DRAFT REPORT OF THE OECD VALIDATION OF THE RAT HERSHBERGER BIOASSAY:

PHASE 3: CODED TESTING OF ANDROGEN AGONISTS, ANDROGEN ANTAGONISTS AND NEGATIVE REFERENCE CHEMICALS BY MULTIPLE LABORATORIES.
SURGICAL CASTRATE MODEL PROTOCOL

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

FOREWORD

This document provides a description and comprehensive summary of the study results for Phase-3 of the OECD validation of the rat Hershberger bioassay. It contains the background on how the validation study was organised and performed, the standardised protocols used, detailed summaries and statistical analyses of the data, and the conclusions drawn from the studies. Phase-3 consisted of coded studies with two androgen agonists (at two dose levels each), three androgen antagonists (one of them at one dose level and the other two at two dose levels), and two negative reference chemicals (at one dose level each). A single protocol was used for the agonists studies based on direct administration of the agonists and statistically significant increases in the target tissues versus an untreated control. Similarly, a single protocol was used for the antagonists based on coadministration with a reference androgen and statistically significant decreases in the target tissues versus the reference androgen group as the control. The negative reference chemicals were tested by both protocols.

The laboratory-testing portion of this phase was conducted between the 2nd and 4th quarter of 2004 and the last reports were obtained in the 1st quarter of 2005. This document was written by Dr. Heinz-Peter Gelbke, consultant for the OECD Secretariat. Comments and input were contributed by the Lead Laboratory, Dr. L. Earl Gray, Jr., (United States Environmental Protection Agency, Research Triangle Park, USA) and members of the Mammalian Validation Management Group, notably Drs. Mike Wade (Health & Welfare, Canada), Gary Timm (US Environmental Protection Agency, Washington, DC), and William Owens (Procter & Gamble, Cincinnati, USA).

The draft report was circulated to the VMG and the EDTA in September 2006. The report has been modified on the basis of the comments received.

Contact for further details
Environment, Health and Safety Division
Environment Directorate
Organisation for Economic Co-operation and Development
2, rue André-Pascal,
75775 Paris Cedex 16, France

Tel: 33-1-45-24-1674 or 9843
Fax: 33-1-45-24-16-75
E-mail: env.edcontact@oecd.org

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

TABLE OF CONENTS

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

ANNEXES

LIST OF TABLES

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

SUMMARY

i. The need to validate test methods for the detection of chemicals interfering with the endocrine system arises from the concerns that ambient levels of natural and industrial chemicals may interact with the endocrine system and as a consequence possibly elicit reproductive, developmental, and other adverse effects in humans and wildlife. The Hershberger bioassay is a leading candidate for a level 3 *in vivo* screening assay of the OECD Conceptual Framework (4) to identify potential androgens and antiandrogens. It is rapid and efficient and there is strong evidence for its sensitivity and specificity from the literature.

ii. The objective for the Hershberger Bioassay validation programme is to validate a test protocol in order to support the development of a Test Guideline for the detection of chemicals having the potential to act as androgen agonists or antagonists in rats. In the preceding validation Phase-1 a protocol was developed for identification of androgen agonists and antagonists by using Testosterone Propionate and Flutamide as potent reference chemicals. In Phase-2 the protocol developed during Phase-1 was used to test two further androgen agonists (17α-Methyl-testosterone and Trenbolone), four androgen antagonists weaker than Flutamide (Procymidone, Vinclozolin, Linuron, and p,p'-DDT) and a potent 5α-reductase inhibitor (Finasteride). The test materials were supplied uncoded at pre-selected dose levels to obtain a dose-response curve by the participating laboratories.

iii. The Phase-2 validation program successfully achieved the goal of demonstrating the reproducibility of the protocol for detecting the weaker androgen agonists and antagonists, as well as the potent 5 α-reductase inhibitor.

iv. After successful completion of the Phase-2 validation testing, the Phase-3 validation was initiated. Coded substances were tested at one or two predetermined dose levels to exclude possible investigator bias. The same dose levels were used by all participants to further substantiate inter-laboratory reproducibility. The dose levels for the agonists and antagonists had already been used in the previous Phase-1 and Phase-2 test series or were derived from the Phase-2 results (DDE). In addition, two chemicals anticipated to act neither as an androgen agonist nor as an androgen antagonist were added to give an indication for the specificity of the Hershberger Bioassay.

v. In Phase 3, the androgen agonists were Testosterone propionate (TP) and Trenbolone (TREN). The test chemical used as a reference androgen agonist was also Testosterone propionate,, which was co-administered with a suspected antagonist. The androgen antagonists were Linuron (LIN) at 2 dose levels (10 and 100 mg/kg/d), p,p'-DDE (DDE) at 2 dose levels (16 and 160 mg/kg/d) and Flutamide (FLU) (3 mg/kg/d). For negative reference chemicals 4-Nonylphenol (mixed isomers) (NP) and 2,4-Dinitrophenol (DNP) were used at dose levels anticipated to approach the maximally tolerable dose (MTD).

vi. The overall goal of the Phase-3 validation study was to further assess the robustness and reproducibility of the Hershberger Bioassay in a blinded manner. The specific goals were to:

Abbreviations: VP: Ventral prostate; SVCG: Seminal vesicles and coagulating glands; LABC: Levator ani and bulbocavernosus muscles; GP: Glans penis; COW: Cowper's gland.
TP: Testosterone Propionate; TREN: Trenbolone; LIN: Linuron; DDE: p,p'-DDE; FLU: Flutamide; NP: 4-Nonylphenol (mixed isomers); DNP: 2,4-Dinitrophenol.
CV: Coefficient of variation; SE: Standard error; SD: Standard deviation

– further demonstrate the reliability and relevance of the Bioassay in response to coded positive and negative substances without investigator bias,
– further evaluate the inter- and intra-laboratory reproducibility of the protocols for identifying androgen agonists and antagonists,
– demonstrate the reproducibility of the Bioassay over time by comparing the data of Phase-3 to that generated in Phase-1 and Phase-2,
– develop data to substantiate the specificity of the protocols both for identifying androgen agonists and antagonists by coded testing of negative reference chemicals
– evaluate the reproducibility of weight changes and relative effectiveness of the 5 target organs
– and thereby support a recommendation of the development of an OECD Test Guideline for the Hershberger Bioassay.

vii. The laboratories participating in Phase-3 had already participated in Phase-1 and -2 thereby enabling a comparison with the data obtained in the previous independent test series.

viii. After finalization of an OECD Test Guideline the Hershberger Bioassay is expected to be widely used in many OECD member countries. Therefore the protocol has to allow for variations in a number of study conditions like the choice of the rat strain, the laboratory diet, housing and husbandry practices such as the use of cage bedding. In addition possible strain differences in the onset of puberty have to be taken into account and the age of castration and the time period of target organ regression after castration have to be kept flexible.

ix. Based on the previous Phase-1 and -2 data the following results were expected for the Phase-3 coded validation study:
a) agonist test protocol:
– TP (0.4 and 0.2 mg/kg/d): statistically significant increase of target organ weights
– TREN (40 mg/kg/d): statistically significant increase of target organ weights
– TREN (1.5 mg/kg/d): no or no consistent increase of target organ weights
– NP (160 mg /kg/d): no increase of target organ weights
– DNP (10 mg/kg/d): no increase of target organ weights
as compared to the vehicle control group.
b) antagonist test protocol:
– FLU (3 mg /kg/d): statistically significant decrease of target organ weights
– LIN (100 mg/kg/d): statistically significant decrease of target organ weights
– LIN (10 mg/kg/d): no or no consistent effect on target organ weights
– DDE (160 mg/kg/d): statistically significant decrease of target organ weights
– DDE (16 mg/kg/d): no or no consistent effect on target organ weights
– NP (160 mg/kg/d): no effect on target organ weights
– DNP (10 mg/kg/d): no effect on target organ weights
as compared to the control group given TP as reference androgen agonist.

x. For statistical analysis the means, standard errors, standard deviations ad 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.

xi. The results obtained with the coded androgen agonists, androgen antagonists, and negative chemicals completely fulfilled the expectations laid down before initiation of the experimental work. By coded testing a possible investigator bias could be ruled out. 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

xii. When comparing the results obtained by different laboratories a high inter-laboratory reproducibility was clearly demonstrated for both protocols of the Hershberger Bioassay by testing androgen agonists, androgen antagonists, and negative reference chemicals.

xiii. To investigate intra-laboratory reproducibility over time the results obtained in Phase-3 were compared to those in Phase-1 or -2 for the laboratories participating in both phases concerned. For each test chemical generally there were only few laboratories available for such a direct comparison. In summary, a good intra-laboratory reproducibility was substantiated although only a limited number of laboratories tested the same chemicals in the different phases of this program.

xiv. Phase-1,-2, and-3 of the validation programme were carried out independently and at different times. The data clearly substantiated a good reproducibility over time.

xv. The specificity of both Hershberger Bioassay protocols for identifying androgen agonists and antagonists was investigated by using two negative reference chemicals, NP and DNP. The doses were selected as to approach or represent the maximally tolerable dose level (MTD). The results clearly substantiated the specificity of the protocols to identify androgen agonists and antagonists.

xvi. The effectiveness of target organs for identification of androgen agonists and antagonists depends to a large extent on the CVs of the weight determinations. Already in the Phase-2 validation report (5) it was noted that the CVs differ among the 5 mandatory tissues. The fluid-filled tissues (VP, SVCG, and COW) generally have 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. Therefore, when developing the final OECD Test Guideline the 5 target organs used in the Phase-2 and Phase-3 validation programmes should all be included.

xvii. During Phase-2 validation, it was already observed that some laboratories encountered a low rate of incomplete preputial separation and this impacted the ability to dissect GP. Therefore laboratories should understand the particular characteristics of their animal strain and supplier relating to the time of preputial separation. Although the Phase-3 protocol was more liberal in respect to the time of castration, there was still one laboratory encountering problems for the evaluation of GP as many animals in this laboratory did not undergo preputial separation. This result underlines the necessity for sufficient flexibility as regards the age of castration for the final Hershberger Bioassay Test Guideline.

xviii. The low dose level of TREN (1.5 mg/kg/d) was expected to represent the NOEL for this androgen agonist or to represent a dose at the very beginning of the ascending part of the dose-response curve. This dose led to a statistically significant weight increase for LABC in two laboratories and to clear numerical weight increases in some others, while for the other 39 target organs investigated, there was only one significant weight increase reported. Thus, LABC seems to be especially responsive to the anabolic action of androgen agonists and this may be of some value for differentiation between androgenic and anabolic activities.

xix. Already in the Phase-2 validation report little or no difference was evident between 0.2 and 0.4 mg/kg/d of TP as androgen agonist reference dose when co-administered with different doses of androgen antagonists. This was substantiated by the data obtained in Phase-3 testing using coded androgen antagonists and negative reference chemicals. Thus, no clear preference can be derived from the data as to the most appropriate dose of the reference androgen agonist (TP), neither for 0.2 nor for 0.4 mg/kg/d.

xx. Overall the following conclusions can be drawn from the results of Phase-3 validation testing:
– A good intra-laboratory reproducibility over time was demonstrated by using coded test samples to avoid investigator bias
– By negative reference chemicals a good specificity for the protocols was demonstrated, both for identification of androgen agonists and antagonists.

And in conjunction with the results of Phase-1 and-2 it is demonstrated that
– The Hershberger protocols are transferable so that they can be used internationally by different laboratories
– The inter-laboratory reproducibility was excellent for agonists, antagonists and for the single 5α-reductase inhibitor tested
– The predictive capacity was very good.

xxi. Within the validation test series only one potent 5α-reductase inhibitor was tested: For this specific mode of action some further validation work might be envisaged with weaker inhibitors to confirm reproducibility and relevance of this assay for such substances.

xxii. Laboratories using the Hershberger Bioassay should provide sufficient training for target organ dissection to their technical staff, especially if this Bioassay is only carried out occasionally.

xxiii. The biological variability of the strains used by the laboratories should be known, especially with regard to onset of puberty and time of preputial separation.

xxiv. On the basis of the successful international OECD validation programme it is proposed to develop an OECD Test Guideline for the identification of androgen agonists and antagonists (Hershberger Bioassay) that may be used as a level 3 screening test within the Conceptual Framework of the OECD (4).

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

INTRODUCTION

1. The need to validate test methods for the detection of chemicals interfering with the endocrine system arises from the concerns that ambient levels of natural and industrial chemicals may interact with the endocrine system and as a consequence possibly elicit reproductive, developmental, and other adverse effects in humans and wildlife. Current reviews have noted that there is very limited evidence for endocrine disruption in humans, but that local, high level exposures to environmental pollutants have probably resulted in endocrine-related effects in wildlife (1)(2)(3).

2. The Hershberger bioassay is a leading candidate for a level 3 *in vivo* screening assay of the OECD Conceptual Framework (4) to identify potential androgens and antiandrogens. It is rapid and efficient and there is strong evidence for its sensitivity and specificity from the literature. Details on the mechanistic background and the most relevant literature are given in the "Introduction" of the OECD Phase-2 validation report (5).

3. The objective for the Hershberger validation programme is to develop and validate a test protocol in order to support the development of a Test Guideline for the detection of chemicals having the potential to act as androgen agonists or antagonists in rats. The Test Guideline, once available, is intended to be used as one element in an overall testing strategy for the detection and assessment of potential endocrine disrupters.

4. Details on the general principles of the validation of test methods and specifically on the history and organization of the OECD validation program for the Hershberger bioassay are given in the section "Test Validation" of the OECD Phase-2 validation report (5). The validation of the Hershberger Bioassay consisted of three separate, consecutive steps:

5. Phase-1: In Phase-1A of the OECD validation study of the Hershberger Bioassay the dose response of several androgen-dependent tissue weights to a series of TP doses was assessed. 17 laboratories in Europe, Japan, Korea, and the United States participated. In Phase-1B the inhibition of this response in these tissues to selected TP doses was assessed against a series of FLU-doses. 7 laboratories from Japan participated. The results showed that the protocol was robust and produced comparable results among laboratories with the reference androgen, testosterone propionate (TP), and with a potent reference androgen antagonist, flutamide (FLU) (3).

6. Phase-2: The protocols used for Phase-2 of the validation study were largely unchanged from the protocol employed in Phases-1A and -1B and had the following characteristics:

- Agonists. Males entering puberty are surgically castrated on pnd 42 or thereafter. The animals are allowed to recover and the tissues of the male reproductive tract to regress for a minimum of 7 days. Treatment would then preferably begin between pnd 49 and 60. The animals are treated for ten consecutive days, once per day, with the test substance by either s.c. injection (TP) or oral gavage. The animals are sacrificed and necropsied 24 hours after the last administration and the target tissues are removed and weighed.
- Antagonists. Males entering puberty are surgically castrated on pnd 42 or thereafter. The animals are allowed to recover and the tissues of the male reproductive tract to regress for a minimum of 7 days. Treatment would then preferably begin between pnd 49 and 60. The animals are treated for ten consecutive days, once per day, with the test substance by oral gavage. The animals are coadministered the reference androgen TP via s.c. injection. The animals are sacrificed and necropsied 24 hours after the last administration and the target tissues are removed and weighed.

7. The results of the Phase-3 validation studies are described in the report presented here. Coded substances were tested at one or two predetermined dose levels to exclude possible investigator bias. The same dose levels were used by all participants to further substantiate inter-laboratory reproducibility. The dose levels for the agonists and antagonists had had already been used in the previous Phase-1 and Phase-2 testing or were derived from the results obtained within Phase-2 (DDE). In addition, two chemicals anticipated to act neither as an androgen agonist nor as an androgen antagonist were added to give an indication for the specificity of the Hershberger Bioassay. Comparison of the results from Phase-3 with those from Phase-1 and -2 will allow to substantiate intralaboratory reproducibility over time. The test substances are listed in table 1.

Table 1. Chemicals used in Phase-3 of the OECD validation of the Hershberger bioassay

Androgen Receptor Agonists	
Testosterone Propionate (TP) [a]	57-85-2
Trenbolone (TREN) 17βHydroxyestra-4,9,11-trien-3-one	10161-33-8
Androgen Receptor Antagonists	
Linuron (LIN)	330-55-2
p,p'-DDE (DDE) 1,1-Dichoro-2,2-bis-(p-chlorophenyl)ethylene	72-55-9
Flutamide (FLU) [b]	1311-84-7
Negative Reference Chemicals	
4-Nonylphenol, mixed isomers (NP)	25154-52-3
2,4-Dinitrophenol (DNP)	51-28-5

[a] Testosterone propionate is the reference androgen agonist, which is coadministered with a suspected antagonist. Standard curves were produced in Phase-1A.

[b] Flutamide is the reference androgen antagonist, which can be coadministered with TP as a positive control antagonist. Standard curves were produced in Phase-1B.

8. The overall goal of the Phase-3 test validation study was to further assess the robustness and reproducibility of the Hershberger bioassay in a blinded manner. Two negative reference chemicals were used to substantiate the specificity of the protocols for testing either androgen receptor agonists or antagonists. The specific goals of Phase-3 were to:

- Further demonstrate the reliability and relevance of the bioassay to respond to and to identify coded positive and negative substances without investigator bias;
- Further evaluate the inter- and intra-laboratory reproducibility of the protocols for identifying strong and weak androgen agonists and antagonists;
- Demonstrate the reproducibility of the bioassay over time by comparing appropriate data to that generated in Phase-1 and Phase-2;
- Develop data to further substantiate the specificity of the protocols for identifying androgen agonists and antagonists by coded testing of negative reference chemicals;
- Evaluate the reproducibility of weight changes and relative effectiveness of the five sex accessory tissues and glands; and
- thereby, support a recommendation for the development of an OECD Test Guideline for the Hershberger Bioassay.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

METHODS

Introduction

9. The rodent Hershberger Bioassay was selected for validation by the OECD following an expert Workshop that was held in Washington DC in 1998. The expert Workshop took account of national recommendations to standardize and validate this procedure as an *in vivo* bioassay in order to identify possible androgen agonists and antagonists (6). In addition, the OECD's Detailed Review Paper on the appraisal of test methods for sex-hormone disrupting chemicals noted the long history of using the Hershberger Bioassay (7), and the Workshop recommended that the OECD proceeds to standardize and validate the Hershberger Bioassay. The EDTA then endorsed the standardisation and validation of the Hershberger Bioassay under the direction of experts in the VMG-mammalian. Further details for the selection criteria are given in § 22-24 of the Hershberger Bioassay Phase-2 report (5).

10. In Phase-1A the dose response of several androgen-dependent tissue weights to a series of TP doses was assessed. For the following organs a satisfactory and reproducible response was demonstrated: ventral prostate (VP), seminal vesicles and coagulating glands (SVCG), the levator ani and bulbocavernosus muscles (LABC), the glans penis (GP) and the Cowper's or bulbourethral glands (COW) (8). In Phase-1B the inhibition of this response in these 5 tissues to selected TP doses was assessed against a series of FLU-doses. Again the response of all 5 tissues was found to be robust and reproducible across laboratories and in the presence of several variations, e.g. strain of rat, diet and modest variations in the age of castration. All laboratories and all protocols were successful in detecting increases in the weights of the accessory sex organs and tissues in response to TP, and there was good agreement among laboratories with regard to the dose-response relationships obtained. There was similar robustness, reproducibility, and agreement between laboratories in their results with regard to the dose response relationships of the anti-androgenic effects of FLU (8).

11. In Phase-2 the performance of the protocol was assessed against additional androgen agonists, antagonists, and a 5α-reductase inhibitor using the same 5 target tissues.

12. The Phase-2 validation program successfully achieved the goal of demonstrating the reproducibility of the protocol for detecting weaker androgen agonists like Methyltestosterone and Trenbolone. All laboratories were able to detect these two substances with all 5 mandatory tissues achieving statistically significant increases. In addition, the Phase-2 validation program successfully demonstrated the reproducibility of the protocol for detecting weaker androgen antagonists like Procymidone, Vinclozolin, Linuron and p,p'-DDE. All laboratories were able to detect p,p'-DDE with all 5 mandatory tissues achieving statistically significant decreases related to the TP reference group. All laboratories were able to detect Procymidone and Vinclozolin (4 laboratories each) with 4 mandatory tissues achieving statistically significant decreases; the weight decrease of GP, however, sometimes failed to achieve statistical significance. 3 of 4 laboratories were able to detect Linuron with at least 4 of the mandatory tissues achieving statistically significant decreases; the 4th laboratory detected Linuron only with SVCG. But this latter laboratory had the largest coefficients of variation (CVs) and encountered some apparent difficulties in dissecting the small unstimulated tissues. Further, the Phase-2 validation program successfully achieved the goal of demonstrating the reproducibility of the protocol for detecting the 5α-reductase inhibitor Finasteride. All laboratories were able to detect this chemical with 4 mandatory tissues by statistically significant decreases. GP did not achieve a statistically significant decrease in one laboratory, and the lack or low level of activity of a 5α-reductase enzyme in this tissue is a plausible reason.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

13. The results of the Phase-2 validation program that were available at the VMG- and EDTA-meeting, May 22-23, 2003, were discussed and led to the recommendation for a Phase-3 validation test series. Here the experimental work should first concentrate on the use of coded positive substances to show reproducibility over time and second on the use of negative substances to get an indication about the specificity of the Hershberger Bioassay (9, 10).

Phase-3 protocols and laboratories

14. After the decision of the VMG and EDTA to initiate Phase-3 of the Hershberger Bioassay validation program a model protocol was developed in cooperation between the Lead Laboratory of the US EPA (Dr. Earl Gray) and the OECD Secretariat. The study manager and director was Dr. J. William Owens on behalf of the OECD Secretariat.

15. The laboratories participating in Phase-3 and the lead investigators are given in annex 1. The OECD model protocol and guidance for the conduct of the study is given in annex 2. The basic protocol requirements are listed in table 2. Mandatory weight determinations were those of total body weight, ventral prostate (fresh) (VP), seminal vesicle and coagulating glands (SVCG), levator ani and bulbocavernosus muscles (LABC), glans penis (GP) and Cowper's glands (COW). Weight determinations of liver, adrenal glands, and kidneys were optional.

Table 2. Protocol summary for OECD Phase-3 validation study

	Factor	**Protocol requirements**
Animals	Species	Rat
	Strain	No preference (not Fischer 344)
	Age at castration	At peripuberty; approx. 6 weeks, but minimum age of 42 days
	Time after castration	1-2 weeks
	Age at initiation of treatment	7-8 weeks
	Weight at time of treatment	Not specified; should be ± 20%
Animal husbandry	Diet	Laboratory preference
	Bedding	Laboratory preference
	Caging	Laboratory preference
Treatment regimen	Animals per dose group	6
	Vehicle	corn oil
	Volume of administration subcutaneous gavage	 0.5 ml/kg/day 5.0 ml/kg/day
	Dosing regimen (mg/kg/day)	10 consecutive daily administrations
	Sacrifice	24-hrs after last treatment
Measurements	Mandatory weights	Total body weight Ventral prostate (fresh) (VP) Seminal vesicle + coagulating glands (SVCG) Levator ani + bulbocavernosus muscles (LABC) Glans penis (GP) Cowper's glands (COWS)
	Optional weights and measurements	Liver Adrenal gland (paired) weight Kidney (paired) weight

16. A total of ten laboratories from Denmark, France, Germany, Japan, Korea, the UK, and the US

volunteered to participate in the Phase-3 studies under the direction of the Lead Laboratory and the Secretariat. All of the laboratories had participated in the previous testing phases. The laboratories were from both the public and private sectors as noted in Table 3.

Table 3. Laboratories participating in the OECD Hershberger Phase-3 validation study

Country	Laboratory	Number of Laboratories
Denmark	Government	1
France	Industry	1
Germany	Industry	2
Japan	Government Private Contract	2 1
Korea	Government	1
United Kingdom	Industry Private Contract	1 1
United States	Government (Lead laboratory) (1)	1
	Total performing studies	**10**

(1) The Lead Laboratory did not perform any study

17. Because the intended purpose of the Hershberger Bioassay is the rapid screening of a potentially large number of chemicals, the judgment was that rigorous and detailed standardisation of all of these variables would constrain the ability to widely use the Hershberger Bioassay in many of the OECD member countries. Therefore, the protocol allowed variations in a number of study conditions, including the choice of rat strain, the laboratory diet, housing and husbandry practices such as the use of cage bedding, and (within the ranges permitted by the model protocol) the age of castration and the time period of regression after castration. The specific conditions in each laboratory at the time that the Hershberger Phase-3 studies were conducted are summarized in Table 4a.

Table 4a. Laboratory parameters and conditions for Phase-3 validation study

Lab	Rat strain	Age at castration	Acclimation time (days)	Vehicle	Age at autopsy	Diet	Animals per cage
1	Wistar rats, CrlGlxBrlHan:Wl	44-46	7	corn oil	61-63	Provimi Kliba SA	1
2	Sprague Dawley	43, 44 [a]	15, 11-13	corn oil	67-68, 65-67	UAR, A04C-10	1
3	SPF-bred Wistar HsdCpb:WU	42	14	corn oil	64-65	Provimi Kliba SA	3
5	HanTac:WH outbred	42-45	14	corn oil	66-70	Proprietary [b]	3
7	CD (SD) IGS BR	41-43, 42-45	11, 10	corn oil	62-64	SDS, R&M1	3
8	Sprague Dawley	44-45	11	corn oil	63-64	PMI 5057	3
9	Alpk:APfSD	42-43	10-11	corn oil	63-64	SDS, R&M1	3
10	Brl Han: WIST Jcl (GALAS)	42	14	corn oil	63	MF, Oriental Yeast	3
11	Crj:CD IGS (SD)	42	7	corn oil	59-61	CE-2, Clea	1
13	Crj:CD IGS (SD)	41-42	10	corn oil	62-63	CRF-1, Oriental Yeast	1

UAR: Usine d'Alimentation Rationnelle; SDS: Special Diet Services

[a] First set of days are for agonist studies, second set of days are for antagonist studies, when different.

[b] Purified (semi synthetic) diet, prepared at laboratory 5

18. For comparison the laboratory parameters and conditions of Phase-2 testing are given in Table 4b. The code numbers of the participating laboratories are the same for Phases-2 and –3.

Table 4b. Laboratory parameter and conditions for Phase-2

Lab	Rat strain	Age at castration	Acclimation time (days)	Vehicle	Age at autopsy	Diet	Animals per cage
1	Wistar rats, CrlGlxBrlHan:Wl	44-46	7	corn oil	61-63	Provimi Kliba SA	1
2	Sprague Dawley	43-45, 47	12, 10-11	0.5% MC	64-67, 67-68	UAR, A04C-10	1
3	SPF-bred Wistar HsdCpb:WU	45-46	12-13	corn oil	67-69	Provimi Kliba SA	3
5	Brl:WIST Han@Mol outbred	42-45	14-15, 18	corn oil	66-70, 70-73	Proprietary[a]	3
7	CD (SD) IGS BR	42-45	7	corn oil	59-63	SDS, RM1	3
8	Sprague Dawley	42	8	corn oil	60	PMI 5057	3
9	Alpk:APfSD	42-43	9-10	corn oil	62-63	SDS, RM1	3
10	Brl Han: WIST Jcl (GALAS)	41-43	7	corn oil	59-61	MF, Oriental Yeast	3
11	Crj:CD IGS (SD) [b]	40-44	8	corn oil	59-63	MF, Oriental Yeast	1
13	Crj:CD IGS (SD) [b]	41-44	11	corn oil	63-66	MF, Oriental Yeast	1

MC: Methyl cellulose; UAR: Usine d'Alimentation Rationnelle; SDS: Special Diet Services

[a] Purified (semi synthetic) diet, prepared at laboratory 5

[b] – Atsugi facility in Kanagawa, Japan.

19. The comparison of table 4a and 4b shows that the specific laboratory parameters and conditions were mostly identical in Phase-2 and Phase-3 with the following exceptions:

- Lab. 2: The vehicle in Phase-3 was corn oil instead of 0.5% Methylcellulose in Phase-2.
- Lab. 5: In Phase-2 and -3 different rat strains were used.
- Lab. 11: Different diets were used in Phase-2 and -3.
- Lab. 13: Different diets were used in Phase-2 and -3.

In addition, in most laboratories there were small differences in age and castration, acclimation time and age at autopsy.

In summary, for most laboratories there was a change of few days for the age at castration, the acclimation time and the age at autopsy. These minor differences are not considered to have a major influence on the results of the Hershberger Bioassay. For one laboratory there was a change of the rat strain and for two laboratories a change of diet that might make some difference for the data obtained in Phase-2 and Phase-3.

Measurements and Data Reporting

20. The mandatory and optional measurements actually performed in each laboratory during the Hershberger- Phase-3 study are summarized in table 5.

Table 5. The mandatory and optional measurements performed by laboratories in Phase-3

	Measurements made in Phase-3 (coded test substances)							
	Mandatory					Optional		
Laboratory	**Ventral prostate**	**Seminal vesicles & coagulating glands**	**LABC muscles**	**Glans penis**	**Cowper's glands**	**Liver**	**Adrenals**	**Kidneys**
1	▲	▲	▲	▲	▲	Y	Y	Y
2	▲	▲	▲	▲	▲	Y	Y	Y
3	▲	▲	▲	▲	▲	Y	N	N
5	▲	▲	▲	▲	▲	Y	Y	Y
7	▲	▲	▲	▲	▲	Y	Y	Y
8	▲	▲	▲	▲	▲	Y	Y	Y
9	▲	▲	▲	▲	▲	Y	Y	Y
10	▲	▲	▲	▲	▲	Y	N	N
11	▲	▲	▲	▲	▲	Y	N	N
13	▲	▲	▲	▲	▲	Y	Y	Y

▲ - Performed mandatory endpoint

Y – Performed optional endpoint; N – Did not perform optional endpoint

21. The Secretariat provided each participating laboratory with a standard Excel spread sheet for recording and transmitting the data in a standard format similar to Phase-2. This standard format provided for rapid e-mail transmission of the results regardless of geographic location to the Secretariat and to the Lead Laboratory. In addition, the work sheets provided the means to quickly calculate basic means, standard deviations and CVs to assist data audits. This proved essential for a rapid assessment of possible entry errors or identification of possible issues by the Secretariat and the Lead Laboratory, e.g., unusually large standard deviations for a group. In addition, the organisation and format of data allowed the rapid extraction into statistical programs by the Secretariat and the Lead Laboratory.

22. The reference androgen agonist was Testosterone Propionate (TP), the reference androgen antagonist was Flutamide (FLU). TP was given by subcutaneous injection, all other chemicals by oral gavage. The test substances for Phase-3 with the exception of TP and FLU were supplied coded to each laboratory. The codes were broken after all laboratories had submitted their results to the OECD Secretariat.

Test Chemicals and Dose Selection for Phase-3

23. For the androgen agonist test protocol, coded samples of Trenbolone (1.5 and 40mg/kg/d) were administered orally. These doses were previously used in Phase-2 by 3 laboratories. A statistically significant increase in target organ weights was to be expected from the earlier studies for the high dose, but no increase for the low dose. Most laboratories also administered TP as additional agonist by subcutaneous injection of either 0.4 or 0.2 mg/kg/d, using the same doses as in their previous investigations. In addition, coded samples of p-Nonylphenol (branched) and 2,4-Dinitrophenol were provided to be orally administered to dose levels of 160 mg/kg/d and 10 mg/kg/d, respectively. These chemicals were expected from the literature to be negative as agonists in the Hershberger Bioassay. The dose levels were selected to avoid acute toxicity either by a separate range finding study (DNP) or by previous data from the Uterotrophic Bioassay results obtained during the validation of this latter test procedure.

24. For the androgen antagonist test protocol, the test chemicals were administered orally at prescribed dosage levels concurrently with TP to determine if they attenuated the effect of growth stimulation by TP in the androgen-dependent tissues. TP was administered subcutaneously at dose levels of 0.4 or 0.2 mg/kg/d according to the previous dosing scheme of the laboratories. Coded samples of the negative reference chemicals (NP and DNP) were provided at the same dose levels as for agonist testing. The coded antagonists and the dose levels were: p,p'-DDE (16 and 160 mg/kg/d) and Linuron (10 and 100 mg/kg/d). Flutamide (3 mg/kg/d) was supplied uncoded and could be tested voluntarily as positive androgen antagonist reference chemical. Dose selection for DDE and LIN was based on the results of the previous Phase-2 studies: For LIN 3 of 4 participating laboratories had found a statistically significant antagonistic effect at a dose level of 100 mg/kg/d (apart from the glans penis). Thus, 100 mg/kg/d should be a dose level showing positive effects in the Phase-3 study. At 10 mg/kg/d no antagonistic effects were observed in the Phase-2 study. For DDE all 4 laboratories using a reference dose of 0.4 mg/kg/d of TP had found statistically significant antagonistic effects at 160 mg/kg/d and no effects at 16 mg/kg/d. All 5 laboratories using a reference dose of 0.2 mg/kg/d of TP had found statistically significant effects at 100 mg/kg/d of DDE (the highest dose tested) for all target organs and no effects at 10 mg/kg/d, while some effects were found at 30 mg/kg/d. For FLU 3 mg/kg/d led to significant antagonistic effects in the Phase-1B of the OECD validation study.

25. In summary based on previous laboratory data, the following results were expected for the Phase-3 coded validation studies:

a) agonist test protocol:

- TP (0.4 and 0.2 mg/kg/d): statistically significant increase of target organ weights
- TRE (40 mg/kg/d): statistically significant increase of target organ weights
- TRE (1.5 mg/kg/d): no increase of target organ weights
- NP (160 mg /kg/d): no increase of target organ weights
- DNP (10 mg/kg/d): no increase of target organ weights

as compared to the vehicle control group.

b) antagonist test protocol:

- FLU (3 mg /kg/d): statistically significant decrease of target organ weights
- LIN (100 mg/kg/d): statistically significant decrease of target organ weights
- LIN (10 mg/kg/d): no effect on target organ weights
- DDE (160 mg/kg/d): statistically significant decrease of target organ weights
- DDE (16 mg/kg/d): no effect on target organ weights
- NP (160 mg/kg/d): no effect on target organ weights
- DNP (10 mg/kg/d): no effect on target organ weights

as compared to the control group given TP as androgen agonist.

The test chemicals and the rationale for dose selection for Phase-3 testing are summarised in table 6.

Table 6. Selected doses for Phase-3 coded test substances

Compound	Doses	Rationale
Testosterone propionate (positive reference agonist)	0.2 mg/kg/d	Selected based on Phase-1 and Phase-2 results.
	0.4 mg/kg/d	Selected based on Phase-1 and Phase-2 results.
Flutamide (positive reference antagonist)	3 mg/kg/d	Selected based on Phase-1 and Phase-2 results.
Nonylphenol (negative)	160 mg/kg/d	Uterotrophic dose
Dinitrophenol (negative)	10 mg/kg/d	Range finding studies to avoid acute toxicity
Trenbolone (previous positive agonist)	1.5 and 40 mg/kg/d	Selected based on Phase-2 results.
Linuron (previous positive antagonist)	10 and 100 mg/kg/d	Selected based on Phase-2 results.
p,p'-DDE (previous positive antagonist)	16 and 160 mg/kg/d	Selected based on Phase-2 results.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Test Chemical Supply

26. The European Chemical Industry Association (CEFIC) has previously supported the uterotrophic validation programme, the enhanced TG 407 validation programme, and Phase-1 and Phase-2 of the Hershberger validation programme by providing financial and managerial responsibility for a centralised chemical repository. CEFIC agreed to continue their support for Phase-3 of the Hershberger validation programme. TNO in the Netherlands continued to serve as the centralised chemical repository as it had for other programmes and phases. Chemicals were purchased, donated, or acquired by synthesis. Where chemicals were included in more than one programme, sufficient quantities of test substances were available so that the same batch could be used in parallel or future studies.

27. After the participation of each laboratory was confirmed, the quantities of test substance that would be needed by each laboratory were calculated. Chemicals were shipped in compliance with regulatory and customs requirements of each nation where participating laboratories were located. Shipments were timed to arrive before the study animals in order to avoid wastage, e.g., expiration of the time window for using immature animals. Other details of the substance supply and handling included:

- The amounts of test chemical needed by each laboratory were calculated and weighed into individually coded, opaque vials. TP as reference androgen agonist was delivered uncoded. Individualised instructions were given to each laboratory, including the volume of test vehicle to be added to provide a test dose solution that could be administered to give the prescribed doses.
- The Lead Laboratory reviewed that the instructions for dissolving, preparing, and making up the solutions were suitable for the test substances. The model protocol for Phase-3 recommended the use of corn oil as most ligands for nuclear receptors such as the androgen receptor have hydrophobic characteristics and corn oil is widely accepted and used by toxicologists for both s.c. and gavage administration.
- Participating laboratories were asked not to break the codes to ensure that all measurements, data acquisition and evaluations were done blindly as to the identity and the doses of the test substances. Codes were disclosed by the OECD Secretariat after the final reports of the Phase-3 work had been received.
- In one case, the sealed code letter was opened at the laboratory. The substances were returned to the repository and a new set of codes and packaged substances was provided to the laboratory.

28. As for TREN there are various specific national regulations, e.g. for importing, a different procedure was used: The material was provided by Sigma Chemicals all from a single lot. The material was shipped from Sigma Chemicals directly to the participating laboratories.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Statistical Analyses of the Data

29. The Lead Laboratory calculated means, standard errors, and the CVs for each endpoint using PROC MEANS on SAS (version 6.08, SAS Institute, Cary, NC, USA). ANOVAs were done using PROC GLM for each laboratory and then the laboratories were pooled for each test substance. Data were then analysed by one-way ANOVA on PROC GLM for each laboratory (with dose as a main effect). Data for each endpoint also were analysed as a two-way ANOVA, with dose and laboratory as main effects, so that the magnitude of the overall dose and laboratory effects, and their interaction, could be determined. In Phase-2 the CV for each androgen-dependent organ weight was fairly constant as the means increased, the standard deviation (SD) being proportional to the mean. This supported the log transformation of the data to provide for a more valid comparison of the effects on organ weights at lower dosage levels.

30. In addition to means and CVs, the R-square values for different effects were calculated. An R-square for an effect was calculated by dividing the sums of squares from the ANOVA for an effect by the total sums of squares in the model. This provides an indication of the strength of the association for an effect with an endpoint. This calculation can be used 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, as indicated by the R-square for the laboratory by dose interaction. The statistical evaluations and the report of the Lead Laboratory are given in annex 3 for androgen agonist testing and in annex 4 for androgen antagonist testing.

31. In the case of testing the antagonist protocol, the organ weights of the vehicle control group, if performed by the participating laboratory, were excluded from the analyses.

32. To enable a direct comparison with the data of Phase-1 and Phase-2 the Secretariat calculated the standard deviations that are given in the tables with the means of the organ weights.

RESULTS OF PHASE-3

33. In this section, the results obtained by the androgen agonist protocol will first be listed, followed by those from the androgen antagonist protocol. The data will be discussed separately for each test chemical. Optional organ weights (liver, adrenals, kidneys), were determined by most laboratories and the data were submitted to the OECD Secretariat. Generally there were no treatment-related effects and these data do not add relevant information to this validation Phase-3 of the Hershberger Bioassay. Therefore, the optional organ weights will only be discussed in specific cases.

34. The results obtained for the androgen agonists and antagonists will first be analysed by pairwise comparison of the target organ weights. The tables relating to this evaluation give the means and standard deviations as calculated by the Secretariat, the results of the pairwise t-test comparisons ($p < 0.05$ and $p < 0.01$) are taken from the report of the Lead Laboratory (Annex 3 and 4). The overall analysis refers to the report of the Lead Laboratory (Annex 3 and 4).

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Androgen agonist data; pairwise comparison of target organ weights

35. The data obtained by the androgen agonist test protocol are given in table 7. The castrated animals are treated for 10 consecutive days with the test chemicals and after the application period the weights of the androgen-dependent tissues (ventral prostate - VP; seminal vesicles and coagulating glands - SVCG; levator ani and bulbocavernosus muscle - LABC; glans penis - GP; Cowper's glands - COW) are determined. An androgen agonist will lead to an increase of weights for these target organs. The absolute target organ weights and the standard deviations for the androgen agonist test protocol are given in table 7. The data for GP of laboratory 5 cannot be evaluated because many animals in this laboratory did not undergo preputial separation, so GP was not dissected or weighed.

Table 7. Hershberger Agonist Data; Target tissue weights by laboratories

NOTE: For TP doses were in Labs 1-9: 0.4 mg/kg/d and in Labs 10-13: 0.2 mg/kg/d
Body weights are given in g; VP, SVCG, LABC, GP, and COW weights are given in mg.

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	Trenbolone 1.5 mg/kg/d	Trenbolone 40 mg/kg/d	TP
Body Weight	**Lab 1**	221.1 ± 6.38	211.7 ± 12.41	221.0 ± 5.15	223.9 ± 6.03	209.3 ± 11.84*	223.2 ± 9.63
	Lab 2	316.5 ± 11.60	302.6 ± 15.45	315.0 ± 13.75	317.8 ± 21.54	274.8 ± 15.95**	Not done
	Lab 3	237.0 ± 20.42	224.7 ± 26.15	254.0 ± 20.84	236.3 ± 11.79	208.8 ± 16.53*	247.8 ± 11.96
	Lab 5	240.3 ± 19.56	223.2 ± 17.22	235.8 ± 17.86	229.8 ± 14.91	214.7 ± 9.99*	242.7 ± 13.75
	Lab 7	275.0 ± 19.77	265.2 ± 9.86	268.6 ± 14.51	283.9 ± 9.73	238.9 ± 11.96**	300.2 ± 9.33
	Lab 8	339.2 ± 8.54	323.7 ± 8.44	323.3 ± 14.09	326.7 ± 20.13	306.6 ± 11.62**	338.5 ± 8.89
	Lab 9	314.0 ± 4.10	308.7 ± 27.80	314.3 ± 14.08	319.5 ± 14.61	289.8 ± 16.07*	338.8 ± 9.91
	Lab 10	262.9 ± 17.03	250.6 ± 12.02	252.4 ± 12.84	249.4 ± 12.89	243.7 ± 9.17**	266.0 ± 12.66
	Lab 11	280.8 ± 7.93	275.9 ± 6.27	285.2 ± 12.40	281.8 ± 15.04	261.5 ± 7.12**	305.6 ± 14.98
	Lab 13	303.2 ± 15.48	297.3 ± 17.61	308.3 ± 14.88	300.8 ± 14.88	264.6 ± 26.37**	320.0 ± 22.73

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	Trenbolone 1.5 mg/kg/d	Trenbolone 40 mg/kg/d	TP
VP	**Lab 1**	18.3 ± 4.95	19.2 ± 4.96	19.1 ± 4.41	19.9 ± 3.14	34.3 ± 9.60**	106 ± 15.48**
	Lab 2	15.9 ± 5.18	14.2 ± 6.68	15.2 ± 6.16	15.3 ± 7.11	71.7 ± 23.23**	Not done
	Lab 3	18.4 ± 4.84	14.3 ± 4.05	20.6 ± 11.19	15.1 ± 4.93	29.6 ± 10.53*	65.9 ± 21.71**
	Lab 5	15.2 ± 1.68	13.4 ± 5.45	15.3 ± 6.68	15.0 ± 4.29	74.2 ± 47.39**	126 ± 28.14**
	Lab 7	48.6 ± 27.65	44.3 ± 20.68	20.5 ± 7.96	48.5 ± 31.78	70.3 ± 43.99(A)	267 ± 134.95**
	Lab 8	20.6 ± 4.99	17.5 ± 4.48	18.0 ± 2.87	22.4 ± 2.63	34.9 ± 3.71**	128 ± 19.11**
	Lab 9	21.9 ± 6.45	16.9 ± 5.44	18.9 ± 4.66	20.8 ± 3.72	30.3 ± 4.12**	146 ± 28.85**
	Lab 10	16.8 ± 1.01	17.0 ± 3.24	15.6 ± 3.43	14.9 ± 1.24	34.1 ± 7.98*	93.6 ± 11.05**
	Lab 11	16.0 ± 5.22	19.8 ± 5.10	15.8 ± 6.53	18.4 ± 3.20	33.4 ± 6.33**	122 ± 25.56**
	Lab 13	22.0 ± 3.08	20.6 ± 1.45	24.0 ± 1.72	26.2 ± 3.81	43.7 ± 11.55**	187 ± 48.39**

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	Trenbolone 1.5 mg/kg/d	Trenbolone 40 mg/kg/d	TP
SVCG	**Lab 1**	37.4 ± 7.71	39.8 ± 9.33	37.2 ± 6.31	38.7 ± 9.24	86.4 ± 14.32**	205.4 ± 38.89**
	Lab 2	40.0 ± 16.85	40.3 ± 6.03	46.6 ± 10.93	58.4 ± 11.52	262.3 ± 78.86**	Not done
	Lab 3	41.5 ± 5.72	38.6 ± 7.24	37.9 ± 11.42	35.3 ± 10.62	108.4 ± 21.33**	177.9 ± 56.36**
	Lab 5	31.1 ± 6.84	30.7 ± 6.81	30.0 ± 5.05	26.9 ± 6.82	169 ± 129.05**	235.8 ± 36.82**
	Lab 7	58.6 ± 14.17	65.1 ± 17.05	60.5 ± 9.30	54.4 ± 6.73	144.4 ± 29.35**	486 ± 135.65**
	Lab 8	57.0 ± 9.74	63.6 ± 10.10	55.2 ± 9.13	70.2 ± 17.55	144.4 ± 26.38**	384.6 ± 57.54**
	Lab 9	72.8 ± 11.16	67.7 ± 19.53	66.0 ± 5.62	72.5 ± 16.46	129.7 ± 15.65**	533.9 ± 53.35**
	Lab 10	27.9 ± 5.56	26.5 ± 1.95	25.7 ± 4.93	29.0 ± 3.50	61.2 ± 9.62*	190.4 ± 19.11**
	Lab 11	42.0 ± 14.50	40.9 ± 11.16	38.8 ± 12.21	41.6 ± 11.51	178.7 ± 60.43**	420.4 ± 32.09**
	Lab 13	61.2 ± 5.94	58.1 ± 6.97	58.4 ± 8.22	61.2 ± 10.92	165.5 ± 37.09**	431.3 ± 55.15**

Table 7. Hershberger Agonist Data (continued)

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	Trenbolone 1.5 mg/kg/d	Trenbolone 40 mg/kg/d	TP
LABC	**Lab 1**	152.4 ± 27.40	161.3 ± 28.62	176.2 ± 29.11	180.0 ± 31.24	358.8 ± 74.25**	369 ± 42.09**
	Lab 2	92.1 ± 20.14	92.2 ± 25.31	95.2 ± 16.69	149.4 ± 41.10**	370.6 ± 46.39**	Not done
	Lab 3	178.9 ± 32.76	160.2 ± 39.65	174.1 ± 25.79	172.0 ± 11.33	358.7 ± 80.40**	356 ± 82.72**
	Lab 5	112.4 ± 29.35	99.8 ± 12.01	101.6 ± 25.84	105.6 ± 15.98	308.4 ± 67.54**	298 ± 44.28**
	Lab 7	232.7 ± 35.76	198.4 ± 39.42	215.2 ± 30.02	214.0 ± 37.03	447.0 ± 50.90**	622 ± 73.24**
	Lab 8	270.1 ± 33.06	240.1 ± 18.63	263.6 ± 27.39	294.5 ± 75.00	548 ± 103.34**	530 ± 25.33**
	Lab 9	162.3 ± 24.43	156.3 ± 18.56	154.5 ± 11.85	176.9 ± 18.31	246.1 ± 41.08**	377 ± 29.43**
	Lab 10	136.8 ± 22.17	128.2 ± 19.24	128.3 ± 11.03	141.5 ± 15.22	298.5 ± 28.15**	312 ± 26.85**
	Lab 11	163.6 ± 38.32	178.6 ± 23.49	189.9 ± 30.38	216.3 ± 17.30**	426.8 ± 46.19**	527 ± 23.47**
	Lab 13	191.3 ± 15.99	178.6 ± 25.15	190.7 ± 6.65	221.1 ± 35.79	452.3 ± 34.53**	544 ± 83.52**

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	Trenbolone 1.5 mg/kg/d	Trenbolone 40 mg/kg/d	TP
GP	**Lab 1**	42.8 ± 3.91	43.2 ± 4.28	43.1 ± 3.99	43.9 ± 2.74	56.5 ± 8.84**	65.2 ± 5.53**
	Lab 2	44.2 ± 7.00	41.9 ± 4.55	43.8 ± 5.89	50.6 ± 6.37	72.1 ± 7.01**	Not done
	Lab 3	39.2 ± 3.99	43.2 ± 4.44	44.0 ± 11.00	40.4 ± 3.25	52.9 ± 2.11(A)	67.4 ± 7.75**
	Lab 5 [a]	[a]	[a]	[a]	[a]	62.9 ± 18.95 [a]	67.5 ± 5.24 [a]
	Lab 7	66.5 ± 5.60	72.9 ± 9.77	66.3 ± 14.79	67.4 ± 3.40	89.6 ± 12.53**	105.3 ± 5.63**
	Lab 8	53.6 ± 6.19	53.1 ± 5.18	54.1 ± 5.23	51.0 ± 7.23	72.4 ± 4.67**	76.6 ± 8.87**
	Lab 9	70.2 ± 6.97	64.1 ± 3.12	66.2 ± 5.25	65.1 ± 6.29	85.1 ± 10.79**	113.7 ± 6.06**
	Lab 10	29.5 ± 5.65	29.9 ± 2.83	28.3 ± 7.21	32.6 ± 5.13	49.8 ± 6.49**	64.4 ± 5.97**
	Lab 11	44.1 ± 4.31	42.9 ± 2.24	41.8 ± 2.25	45.5 ± 2.54	58.5 ± 3.70**	73.9 ± 3.93**
	Lab 13	53.0 ± 7.99	54.6 ± 5.37	54.8 ± 5.59	52.0 ± 2.57	72.9 ± 3.18**	95.1 ± 8.04**

[a] Many animals in this lab did not undergo preputial separation, so glans penis was not dissected or weighed.

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	Trenbolone 1.5 mg/kg/d	Trenbolone 40 mg/kg/d	TP
COW	**Lab 1**	6.6 ± 1.96	8.7 ± 1.85	9.5 ± 2.79	8.2 ± 2.40	14.1 ± 3.34**	26.0 ± 3.70**
	Lab 2	3.6 ± 1.90	4.6 ± 0.80	5.1 ± 2.14	10.4 ± 3.86*	21.4 ± 10.18**	Not done
	Lab 3	6.1 ± 3.59	5.5 ± 2.49	6.0 ± 2.24	6.1 ± 1.85	10.3 ± 5.66*	16.1 ± 4.95**
	Lab 5	3.5 ± 1.26	4.0 ± 1.04	3.5 ± 0.45	3.5 ± 0.87	13.4 ± 8.60**	22.5 ± 3.45**
	Lab 7	9.8 ± 4.10	7.0 ± 3.48	4.8 ± 2.23	3.2 ± 2.50	11.9 ± 3.63(A)	39.6 ± 4.31**
	Lab 8	8.2 ± 5.18	7.7 ± 1.89	8.2 ± 2.55	11.0 ± 4.81	19.7 ± 5.15**	30.6 ± 7.36**
	Lab 9	7.1 ± 1.04	6.6 ± 1.35	6.6 ± 1.22	7.0 ± 1.59	10.4 ± 2.00**	41.2 ± 4.36**
	Lab 10	4.1 ± 1.30	4.4 ± 1.38	3.9 ± 1.45	4.5 ± 1.06	10.3 ± 2.48**	20.4 ± 3.53**
	Lab 11	5.8 ± 1.31	5.6 ± 1.12	4.4 ± 1.68	5.8 ± 1.82	10.1 ± 2.62**	34.4 ± 8.14**
	Lab 13	8.5 ± 2.15	7.4 ± 1.76	8.0 ± 1.33	8.8 ± 2.37	18.2 ± 5.18**	37.1 ± 6.59**

Pairwise t-test comparisons:**: $p < 0.01$; *: $p < 0.05$

(A): Indicates that an effect that should have been significant was not affected, i.e. a “false negative”.

p-Nonylphenol (branched) (NP)

36. NP at a dose level of 160 mg/kg/d did not lead to an effect on any of the 5 target organ weights. Since in laboratory 5 many animals did not undergo preputial separation, the glans penis could not be dissected or weighed in this laboratory. Thus, as was expected no androgen agonistic response was found for this substance (table 7).

37. When pooling the weights of all optional tissues (data not shown here) there was a significant increase in mean liver and kidney weight ($p<0.01$) (cf Table 4 of the Lead Laboratory report, annex 3). In all laboratories there was a numerical but not a statistically significant decrease in body weight. But pooling body weight data over all laboratories led to a statistically significant decrease ($p<0.01$). These effects on body, liver and kidney weights show that the dose given corresponded to the criteria for a maximally tolerable dose (MTD).

2,4-Dinitrophenol (DNP)

38. DNP, at a dose level of 10 mg/kg/d, did not lead to an effect on body weight nor on any of the 5 target organ weights. Since in laboratory 5 many animals did not undergo preputial separation, the glans penis could not be dissected or weighed in this laboratory. Thus, as was expected no androgen agonistic response was found for this substance (table 7).

Trenbolone (TREN) 1.5 mg/kg/d

39. TREN at a dose level of 1.5 mg/kg/d did not influence body weight and did not increase the weight of VP, SVCG, and GP. Since in laboratory 5 many animals did not undergo preputial separation, the glans penis could not be dissected or weighed in this laboratory.
There was a statistically significant increase in COW ($p<0.05$) reported by laboratory 2. No other laboratory found a statistically significant weight increase of COW, neither clear numerical increases can be seen at this dose level. Laboratory 2 had one of the lowest vehicle control COW weights and one of the highest after treatment. Therefore, this effect may be considered as a chance finding.
Laboratory 2 and laboratory 11 found a statistically significant weight increase for LABC ($p<0.01$). In addition, there is a numerical weight increase for laboratories 1, 8, 9, 10, and 13. This indicates that the androgenic/anabolic effect of TREN starts at about a dose level of 1.5 mg/kg/d with the LABC weight being the most sensitive parameter.
In conclusion as was expected by the Phase-2 data a dose level of 1.5 mg/kg/d of TREN did not lead to an androgenic response of VP, SVCG, GP and COW. But obviously LABC seems to be the most sensitive tissue, the androgenic response starting at about this dose level. This is in line with the results of Phase-2, where laboratory 3 and 7 noted a numerical increase in LABC weight from 195.6 mg to 218.7 mg (laboratory 3) and from 233.5 mg to 260.3 mg (laboratory 7).

40. Three laboratories tested the low dose of TREN both in Phase-2 and-3 (laboratories 1, 3, and 7). In both phases there was no statistically significant weight increase in any of the target organs reported by the participating laboratories (15 target organs in total). Thus, there was a 100 % concordance between Phase-2 and Phase-3.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Trenbolone (TREN) 40 mg/kg/d

41. TREN at a dose level of 40 mg/kg/d led to a statistically significant decrease in body weight in all laboratories. This was expected by the results of Phase-2 where 2 of the 3 participating laboratories found a significant decrease of terminal body weight and the 3rd a numerical decrease.
In Phase-2, 3 laboratories (laboratory 1, 3, and 7) participated in the multiple dose testing of TREN. All 3 of them reported a significant increase in the weights of all target tissues. Correspondingly, the SVCG and LABC weights were found by all laboratories to be significantly increased in Phase-3.
GP weights could not be determined in laboratory 5 since in this laboratory many animals did not undergo preputial separation, so the glans penis could not be dissected. Of the remaining 9 laboratories, all, apart from laboratory 3, achieved statistically significant increased GP weights. But the GP weight in laboratory 3 was numerically clearly increased being in between the control and the TP groups similar to all the other laboratories.
In laboratory 7 VP was not significantly increased as compared to the vehicle control group. On the other hand, in Phase-2 this laboratory had reported a highly significant weight increase after treatment with 40 mg/kg/d TREN. In Phase-2 the ventral prostate weight of the vehicle control group was 15.2 mg, well comparable to the control weights of all the other laboratories. But in Phase-3 the vehicle control weight was 48.6 mg more than by a factor of 3 higher as compared to the weight in Phase-2. Furthermore, the VP weight of this vehicle control group was the highest of all laboratories participating in Phase-3. In addition, in comparison to the VP weight obtained with the negative control substance DNP there was a statistically significant increase after treatment with TREN (40 mg/kg/d). (Note: the weights obtained after administration of the negative reference chemicals, NP and DNP, can also be used for comparison since these negative reference substances did not lead to any effect on the androgen target organs.) Thus, the missing significance can be attributed to insufficient dissecting procedures of this specific laboratory leading to exceptional high and variable VP weights.
A similar situation exists for COW of laboratory 7: While in phase-2 testing laboratory 7 found a significantly increased COW weight there was no significant increase in Phase-3 after application of 40 mg/kg/d TREN. Again the COW vehicle control weight is the highest among all participating laboratories and the weight of the TREN group is in the lower range. On the other hand, in comparison to the COW weights obtained with both of the negative control chemicals (NP and DNP) the TREN treatment group shows a significantly increased COW weight. (Note: the weights obtained after administration of the negative reference chemicals, Nonylphenol and Dinitrophenol, can also be used for comparison since these negative reference substances did not lead to any effect on the androgen target organs). Thus, this missing significance should be regarded as a chance finding at least partially related to the high vehicle control weight possibly due to insufficient dissection procedures.
In conclusion, the target organ weights were increased as was to be expected after treatment with 40 mg/kg/d TREN. The missing significances in one of the 10 participating laboratories for VP and COW may be explained by problems of the laboratory procedures leading to very high control weights.

42. Table 8 gives the CVs for VP and COW. For these tissues laboratory 7 did not achieve a significant organ weight increase in comparison to the vehicle control group after the high dose of TREN. Thereby the hypothesis will be investigated that laboratory 7 belongs to the laboratories with the highest CVs for both of these target organs. In this table only the CVs of the substances are listed that were coded and thereby tested in a blinded manner. The CVs of the uncoded TP doses are not considered. The CVs of laboratory 7 are highlighted in bold figures, the CVs that are higher than those of laboratory 7 are marked in italics.
For VP weights the following picture emerges:

- Vehicle control groups: The CV of lab. 7 (56.91) was not exceeded by any other laboratory.
- Nonylphenol: the CV of laboratory 7 (46.66) was only exceeded by one other laboratory (46.98).
- Dinitrophenol: the CV of laboratory 7 (38.82) was exceeded by 4 of the remaining 9 laboratories.
- Low dose of Trenbolone: the CV of laboratory 7 (65.58) was not exceeded by any other laboratory.
- High dose of Trenbolone: the CV of laboratory 7 (62.55) was only exceeded by one other laboratory (63.87).

In summary, laboratory 7 always belonged to those with the highest CV or actually had the highest CV. This is strong evidence that this laboratory had some deficiencies in dissecting VP.
For COW weights the following picture emerges:

- Vehicle control groups: the CV of laboratory 7 (41.98) was exceeded by 3 of the remaining 9 laboratories.
- Nonylphenol: the CV of laboratory 7 (49.89) was not exceeded by any other laboratory.
- Dinitrophenol: the CV of laboratory 7 (46.88) was not exceeded by any other laboratory.
- Low dose of Trenbolone: the CV of laboratory 7 (79.13) was not exceeded by any other laboratory.
- High dose of Trenbolone: the CV of laboratory 7 (30.59) was only exceeded by 3 of the remaining 9 other laboratories.

In summary, laboratory 7 always belonged to those with the highest CV or actually had the highest CV. This is strong evidence that this laboratory had some deficiencies in dissecting COW.
These findings for VP and COW are a clear indication that a high standard of training of the laboratory personnel for preparation of the small target organs is a major factor to achieve reliable results.

43. The CVs for GP of laboratory 3 did no differ markedly from those of the other laboratories (data not shown here) in contrast to the CVs of laboratory 7 for VP and COW. Thus a laboratory effect for the missing significance of GP weight for laboratory 3 cannot be substantiated. The single missing significance should rather be taken as an indication for a lower responsiveness of the GP. This interpretation corresponds to the results of Phase-2 where GP did not achieve statistical significance in several instances.

44. Three laboratories (1, 3, and 7) tested the high dose of TREN in Phase-2 as well as in Phase-3. In Phase-2 for all organs a significant weight increase was reported by the participating laboratories (15 organs in total). In Phase-3, 12 of these 15 organs again showed a significant weight increase. As mentioned above, 1 of the laboratories obviously had some problems with its dissection procedures for 2 of the target organs that showed a different response as compared to Phase-2. Nevertheless, the overall concordance between Phase-2 and-3 was 80%.

Table 8. Hershberger Agonist Data; Coefficients of variation

Coefficients of variation of VP and COW of all laboratories. CVs of Laboratory 7 that did not achieve statistical significance with the high trenbolone dose are highlighted in bold figures, CVs that are higher than those of laboratory 7 are marked in italics.

Vehicle			**NP**			**DNP**			**TREN 1.5 mg/kg/d**			**TREN 40 mg/kg/d**		
	VP	**COW**		**VP**	**COW**		**VP**	**COW**		**VP**	**COW**		**VP**	**COW**
Lab 1	27.04	29.94	Lab 1	25.80	21.21	Lab 1	23.07	29.45	Lab 1	15.77	29.50	Lab 1	27.98	23.63
Lab 2	32.56	*52.56*	Lab 2	*46.98*	17.19	Lab 2	*40.68*	41.61	Lab 2	32.41	47.52	Lab 2	46.40	*37.13*
Lab 3	26.30	*58.72*	Lab 3	28.25	45.07	Lab 3	*54.27*	37.18	Lab 3	32.61	30.46	Lab 3	35.63	*55.19*
Lab 5	11.12	36.29	Lab 5	40.59	26.22	Lab 5	*43.56*	12.90	Lab 5	28.61	25.19	Lab 5	*63.87*	*64.30*
Lab 7	**56.91**	**41.98**	Lab 7	**46.66**	**49.89**	Lab 7	**38.82**	**46.88**	Lab 7	**65.58**	**79.13**	Lab 7	**62.55**	**30.59**
Lab 8	24.26	*63.32*	Lab 8	25.62	24.65	Lab 8	15.99	31.04	Lab 8	11.76	43.57	Lab 8	10.64	26.14
Lab 9	29.51	14.75	Lab 9	32.20	29.53	Lab 9	24.62	18.51	Lab 9	17.87	22.76	Lab 9	13.59	19.20
Lab 10	6.04	31.83	Lab 10	19.06	31-36	Lab 10	21.91	36.97	Lab 10	8.34	23.49	Lab 10	23.44	24.04
Lab 11	32.62	22.61	Lab 11	25.78	20.19	Lab 11	41.30	38.29	Lab 11	17.45	31.35	Lab 11	18.94	25.97
Lab 13	13.97	25.22	Lab 13	7.03	23.64	Lab 13	7.15	16.56	Lab 13	14.56	27.07	Lab 13	26.45	28.54

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Testosterone propionate (TP) 0.2 and 0.4 mg/kg/d

45. TP was not tested by laboratory 2. Laboratory 1–9 used a dose level of 0.4 mg/kg/d and laboratories 10–13, 0.2 mg/kg/d. Body weights were not affected by treatment with TP. All target organs (VP, SVCG, LABC, GP and COW) showed a significant weight increase in all laboratories ($p<0.01$) as was to be expected by the former results of Phase-1A, regardless whether a dose of 0.2 or 0.4 mg/kg/d was used. In this respect, GP could not be evaluated for laboratory 5 as many animals did not undergo preputial separation, so GP was not dissected or weighed in this laboratory.

46. Nine laboratories tested TP in Phase-1 as well as in Phase-3. For all target organs all laboratories reported a significant weight increase. Thus, there was a 100% concordance for this reference chemical that was supplied uncoded to the Phase-3 participants.

47. Overall, with the androgen agonist test protocol there was no target organ weight increase induced by the negative reference substances. After treatment with androgen agonists the target organs always achieved a statistically significant weight increase with two exceptions, indicating that:

- the efficiency of a laboratory to handle these small target tissues is pivotal to obtain reliable results
- the GP weight increase may be less responsive to androgen agonist treatment and the results for this organ may be less robust.

Androgen agonist data; overall analysis (cf Annex 3 of Lead Laboratory).

48. One of the objectives of the interlaboratory study was to demonstrate that the 10 participating laboratories could get the "correct" responses for all the chemicals on all androgen-dependent tissues. As TP at 0.2 and 0.4 mg/kg/d was used in Phase-1 and -2 and TREN at 1.5 and 40 mg/kg/d in Phase-2 studies of this validation exercise the expectation for these treatment groups was:

- Decreased body weight for TREN at 40 mg/kg/d
- Significant increases in all 5 androgen-dependent organ weights with TREN at 40 mg/kg/d
- Significant increases in all 5 androgen-dependent organ weights with TP at 0.4 and 0.2 mg/kg/d subcutaneously
- Small or only occasional increases in LABC weight with 1.5 mg/kg/d TREN.

49. When all the effects of the androgen agonist treatment groups on the 6 endpoints (body weight and 5 androgen-dependent tissue weights) obtained in Phase-3 were correlated with the effects seen previously in Phase-2 studies, a highly significant positive relationship was found having a correlation coefficient of $r= 0.98$. A graph depicting these observed versus expected effects in Phase-3 is given in the report of the Lead Laboratory (Annex 3). The expected and observed effects were calculated as % of concurrent control values using the vehicle group as control.

50. Pooling the data from all 10 participating laboratories, TREN (40 mg/kg/d) was positive in 96% of the time. At the low dose (1.5 mg/kg bw /d) TREN did not elicit many positive responses (20% of the laboratories reported a stimulatory effect on LABC), and LABC weight was the only tissue with a significantly increased weight after pooling the data from all laboratories (table 9). In Phase-2 this dose did not produce any significant effects. TP at 0.4 and 0.2 mg/kg/d subcutaneously was positive in all instances. The "known negatives" NP (160 mg/kg/d) and DNP (10 mg/kg/d) did not stimulate growth of any of the androgen-dependent tissues neither when analysing the laboratories separately (see above) nor after pooling (table 9).

Table 9. Hershberger Phase-3 agonist data; Means and Standard Errors (SEs) of body, target tissue and optional tissue weights over all participating laboratories

All labs	body weight		VP		SVCG		LABC		GP		COW		Liver		Adrenals		Kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	279	5	21.4	1.7	47	2.2	169	7.4	49.2	1.8	6.3	0.42	11.2	0.3	62	1.5	1970	46
DNP 10	278	4.8	18.3	0.8	45.6	2	169	7	49.2	1.9	6	0.34	11	0.32	63	1.7	1967	44
NP 160	268**	5.3	19.7	1.5	47.1	2.3	159	6	49.5	1.8	6.1	0.3	11.8**	0.36	64.2	1.8	2192**	65
TREN 1.5	277	5.3	21.2	1.7	48.7	2.5	187**	8	49.5	1.5	6.9	0.46	11.3	0.35	59	1.7	2024	51
TREN 40	251**	4.6	48.1**	4.2	152**	11	383**	13	68**	2	14.4**	0.9	11.3	0.32	54.2**	1.9	2003	47
TP.4	286	5.8	138**	10	341**	20	437**	17	81.2**	2.6	30**	1.3	12**	0.4	54.3**	1.9	2128	58

**: $p<0.01$

51. TREN is known to be an androgen agonist *in vivo* and *in vitro* having a more pronounced "anabolic effect" on LABC as compared to the "androgenic effect" on the VP, SV, COW, and GP. This is reflected by the pooled data for the low dose TREN-group (table 9): There was a significant weight increase ($p < 0.01$) for LABC only but not for the other androgen-dependent target tissues.

52. Apart from effects on androgen dependent organs it should be mentioned that NP at a dose level of 160 g/kg/d led to a significant increase of liver and kidney weights and to a decrease of body weight. Thus, the maximally tolerable dose (MTD) was achieved for this anticipated negative reference chemical.

53. Analysis of variance for the 5 androgen-dependent tissues revealed that the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in phase-3 were all statistically highly significant. Importantly, the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by the relatively small F-values for the chemical by laboratory interaction as compared to the size of the 2 main effects (table 10).
There were statistically significant laboratory effects in each ANOVA. This was expected and also seen in Phase-1 and –2 of the validation studies. These effects arose because the laboratories used different rat strains, but even within the same strain the animal body and organ weights would be expected to vary. Such a laboratory effect per se is not problematic, whereas a large laboratory by treatment interaction would be because this would indicate chemicals were not acting similarly in all laboratories. Although the F-values for the interaction were small in comparison to the main effects, the interaction effect is still significant. This could lead to different conclusions in the case of small effects of a substance depending on the laboratory.

Table 10. Hershberger Agonist Data; results of ANOVAs for laboratory interaction and laboratory by treatment interaction
(TP data excluded since not coded)

	All labs			
Body weight	Effect	df	F	p-value
	LAB	9	184	0.0001
	CHEMICAL	4	67	0.0001
	INTERACTION	36	1	> 0.40
CV = 5.6 %				
SVCG	Effect	df	F	p-value
	LAB	9	5.1	0.0001
	CHEMICAL	4	97	0.0001
CV = 54 %	INTERACTION	36	2.2	0.0003
LABC	Effect	df	F	p-value
	LAB	9	61	0.0001
	CHEMICAL	4	377	0.0001
	INTERACTION	36	3.1	0.0001
CV = 18%				
VP	Effect	df	F	p-value
	LAB	9	7.4	0.0001
	CHEMICAL	4	40.2	0.0001
	INTERACTION	36	2.2	0.0003
CV = 59 %				
GP	Effect	df	F	p-value
(no lab 5)	LAB	8	110	0.0001
	CHEMICAL	4	92	0.0001
	INTERACTION	32	1.3	>0.11
CV = 12.1 %				
COW	Effect	df	F	p-value
	LAB	9	8.2	0.0001
CV = 44 %	CHEMICAL	4	63	0.0001
	INTERACTION	36	2.2	0.0002

54. In table 11 the CVs and the R-square values are summarized over the laboratories for the target organs. Thereby the R-square values and the CVs can be compared for laboratory-to-laboratory. The higher the R-square value the better is the strength of association of the chemical effects versus vehicle control (maximum possible R-square value is 100). Laboratory 7 did not detect a statistically significant weight increase for VP and COW by the high TREN dose (40 mg/kg/d). Correspondingly, the R-square value of this laboratory is the lowest for VP and one of the lowest for COW. A similar situation exists for laboratory 3 not detecting a significant increase in GP weight for the high TREN dose. The R-square value of laboratory 3 again is the lowest one of all laboratories evaluating this target tissue.

55. The CVs given in table 11 describe the precision (variability) of the data. A high CV indicates a high variability of the laboratory performance. If all laboratories had the same CV on an endpoint then this would indicate that they were all dissecting the tissues with similar precision. A high CV in a laboratory versus the others might indicate a problem. In a few cases the CVs seemed to be larger than expected: For example, the high CV values for SVCG, VP, and COW in laboratories 5 and 10 arose from a single animal in each laboratory from the high TREN dose group. In these cases the androgenic effect of the chemical treatment was still apparent. However, laboratory 7 data displayed unusually high variability in the vehicle-treated group for VP and COW and this resulted in the only 2 false negatives in the entire data set. These results indicate that this laboratory had some difficulty dissecting the unstimulated small tissues (most likely) in the controls, errors in dosing, incomplete castration or some other technical problem.

56. In summary, the overall analysis of the results of Phase-3 of the Hershberger Assay interlaboratory study using coded samples of known dosage levels of agonists or negative reference chemicals, produced appropriate and consistent responses among all laboratories in more than 90% of the cases (5 androgen-dependent tissues with 4 treatment groups). None of the known androgen agonists were undetected in any laboratory by the total data set for all target tissues.

Table 11. Hershberger Agonist Data; pooled R-square values and CVs by laboratory and target organ

VP	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11	Lab 13
R 2	57	81	37	59	27	77	51	32	64	72
CV	26	45	39	81	62	17	23	110	26	21
SVCG										
R2	82	57	86	53	82	84	77	25	81	57
CV	20	41	24	101	23	21	18	168	43	22
LABC										
R2	80	93	77	86	87	81	70	89	91	95
CV	21	20	21	25	15	19	14	16	14	11
GP										
R2	56	79	45	No data	48	69	60	69	82	72
CV	11	12	13	No data	14	10	10	20	6.7	9
COW										
R2	55	67	23	54	52	59	54	47	59	70
CV	27	56	51	72	44	38	20	76	28	28

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Androgen antagonist data ; pairwise comparison of target organ weights

57. The data obtained by the androgen antagonist protocol are given in table 12. The castrated animals are treated for 10 consecutive days with TP at 0.2 or 0.4 mg/kg/d subcutaneously for stimulation of the androgen responsive organs. In parallel the coded test chemicals (putative androgen antagonists) are administered orally for the same duration: after the application period the weights of the androgen-dependent tissues (VP, SVCG, LABC, GP, COW) are determined. An androgen antagonist will lead to a decrease of weights for these target organs. The absolute target organ weights are given in table 12. The data for GP of laboratory 5 can only be partially evaluated because many animals in this laboratory did not go preputial separation, and then GP was not dissected or weighed.

Table 12. Hershberger antagonist data; Target tissue weights by laboratories

NOTE: For TP doses were in Labs 1-9: 0.4 mg/kg/d and in Labs 10-13: 0.2 mg/kg/d

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol 10 mg/kg/d	DDE 16 mg/kg/d	DDE 160 mg/kg/d	Linuron 10 mg/kg/d	Linuron 100 mg/kg/d	Flutamide 3 mg/kg/d	TP
Body Weight	**Lab 1**	Not done	203.9 ± 7.79*	210.7 ± 9.89	207.3 ± 11.46	199.7 ± 8.30**	204.0 ± 11.13*	195.6 ± 9.45**	201.2 ± 9.26*	215.7 ± 6.78
	Lab 2	322.7 ± 27.24	328.6 ± 23.32	355.2 ± 13.34	345.3 ± 17.05	261.8 ± 22.40**	350.7 ± 18.87	293.1 ± 21.30**	Not done	338.2 ± 33.38
	Lab 3	Not done	247.2 ± 23.37	239.7 ± 15.16	246.0 ± 14.93	227.7 ± 12.50	245.3 ± 20.37	233.3 ± 21.29	246.0 ± 20.24	245.8 ± 19.46
	Lab 5	Not done	230.8 ± 16.99	250.0 ± 11.47	246.2 ± 11.20	221.0 ± 33.35*	241.7 ± 24.11	222.2 ± 15.29*	236.0 ± 19.68	245.3 ± 14.80
	Lab 7	285.6 ± 13.91	283.0 ± 16.29	288.3 ± 17.41	288.4 ± 26.31	259.1 ± 53.09*	313.8 ± 25.62	270.3 ± 18.31	286.6 ± 27.21	301.1 ± 28.69
	Lab 8	343.0 ± 23.61	346.5 ± 14.62	360.2 ± 22.72	353.5 ± 11.78	316.2 ± 37.70*	362.9 ± 19.17	331.2 ± 21.55	348.7 ± 21.09	349.4 ± 23.18
	Lab 9	Not done	336.7 ± 21.68	346.2 ± 11.69	349.3 ± 11.06	328.3 ± 17.14	346.5 ± 19.15	315.5 ± 9.97**	346.7 ± 20.13	347.0 ± 18.20
	Lab 10	260.0 ± 17.41	263.3 ± 17.83	272.2 ± 9.53	266.3 ± 11.68	248.7 ± 17.11*	263.6 ± 11.79	245.7 ± 11.10**	261.5 ± 12.53	267.4 ± 16.51
	Lab 11	313.2 ± 14.67	326.1 ± 16.37	332.2 ± 20.74	329.8 ± 18.69	270.8 ± 44.64**	335.3 ± 15.57	309.5 ± 13.87	327.4 ± 23.06	331.9 ± 14.28
	Lab 13	303.3 ± 16.39	307.0 ± 24.00	317.5 ± 10.92	322.7 ± 23.55	248.1 ± 62.40**	310.6 ± 23.15	292.7 ± 18.07	313.2 ± 24.15	314.7 ± 20.87

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol	DDE 16 mg/kg/d	DDE 160 mg/kg/d	Linuron 10 mg/kg/d	Linuron 100 mg/kg/d	Flutamide	TP
VP	**Lab 1**	Not done	98.9 ± 10.64	109.1 ± 54.56	72.3 ± 12.96	29.4 ± 5.13**	63.9 ± 16.97	43.4 ± 5.35**	25.8 ± 2.43**	87.1 ± 10.53
	Lab 2	15.0 ± 7.15	181.2 ± 28.01	201.2 ± 40.11	177.9 ± 32.79	50.8 ± 18.51**	196.4 ± 16.19	98.9 ± 23.33**	Not done	185.5 ± 67.09
	Lab 3	Not done	72.4 ± 25.10	62.8 ± 13.68	63.8 ± 20.89	23.1 ± 1.51**	72.2 ± 15.34	36.2 ± 9.51**	17.3 ± 4.52**	77.7 ± 20.78
	Lab 5	Not done	116.0 ± 23.24	133.6 ± 37.39	131.9 ± 10.71	36.6 ± 14.89**	143.7 ± 25.07	66.3 ± 16.50**	33.1 ± 5.95**	129.2 ± 20.89
	Lab 7	39.9 ± 16.24	162.3 ± 47.65	178.6 ± 50.35	139.4 ± 41.40*	70.1 ± 37.81**	152.6 ± 38.13	126.9 ± 41.68**	55.9 ± 35.59**	197.3 ± 61.31
	Lab 8	22.6 ± 1.91	131.6 ± 16.68	104.5 ± 14.75**	83.1 ± 17.63**	36.3 ± 8.34**	91.7 ± 21.02**	44.7 ± 11.82**	31.5 ± 4.45**	146.8 ± 13.66
	Lab 9	Not done	156.2 ± 14.37	166.2 ± 23.90*	152.7 ± 9.61	81.0 ± 25.41**	141.7 ± 14.16	97.5 ± 11.08**	55.1 ± 7.29**	145.3 ± 25.66
	Lab 10	14.4 ± 1.55	82.0 ± 7.67**	85.9 ± 19.32*	78.9 ± 14.60**	28.3 ± 5.90**	80.6 ± 10.74**	41.4 ± 5.86**	24.9 ± 3.64**	102.4 ± 17.56
	Lab 11	17.7 ± 2.70	135.0 ± 12.90	129.0 ± 31.54	126.5 ± 39.98	28.0 ± 8.92**	115.5 ± 26.30*	61.6 ± 25.68**	28.2 ± 8.94**	144.9 ± 19.43
	Lab 13	20.7 ± 3.46	142.0 ± 34.49	132.8 ± 20.11	107.2 ± 17.78**	36.2 ± 9.43**	133.6 ± 24.16	67.8 ± 24.48**	39.8 ± 5.18**	159.1 ± 31.18

Table 12 (continued) Antagonist data

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol	DDE		Linuron		Flutamide	TP
					16 mg/kg/d	160 mg/kg/d	10 mg/kg/d	100 mg/kg/d		
SVCG	**Lab 1**	Not done	211.5 ± 42.67	206.4 ± 74.32	112.7 ± 55.46*	41.5 ± 15.55**	136.8 ± 57.04	58.4 ± 10.77**	33.8 ± 7.54**	178.5 ± 39.27
	Lab 2	43.3 ± 8.41	680.8 ± 96.77	593.7 ± 192.80	591.9 ± 297.76	68.8 ± 27.49**	422.7 ± 73.19	277.6 ± 94.17**	Not done	539.7 ± 150.82
	Lab 3	Not done	214.8 ± 74.87	182.2 ± 12.03	180.6 ± 61.40	66.9 ± 18.06**	215.8 ± 53.79	85.2 ± 19.37**	54.6 ± 12.41**	198.2 ± 40.51
	Lab 5	Not done	345.4 ± 159.31	336.4 ± 61.72	328.6 ± 61.13	60.1 ± 16.00**	353.6 ± 54.25	127.6 ± 32.46**	33.1 ± 5.95**	358.4 ± 57.06
	Lab 7	56.5 ± 12.66	468.5 ± 109.96	457.3 ± 43.46	378.4 ± 48.27	159.7 ± 48.70**	447.0 ± 74.37	244.3 ± 25.46**	55.9 ± 35.59**	426.5 ± 76.58
	Lab 8	72.9 ± 17.80	334.0 ± 53.08**	385.3 ± 80.76*	280.6 ± 65.98**	103.4 ± 36.67**	252.7 ± 55.64**	126.8 ± 33.51**	31.5 ± 4.45**	464.3 ± 104.71
	Lab 9	Not done	550.6 ± 64.85	578.0 ± 54.73	559.8 ± 66.36	272.6 ± 51.06**	523.7 ± 45.90	348.1 ± 42.73**	55.1 ± 7.29**	567.4 ± 72.26
	Lab 10	23.8 ± 3.07	191.4 ± 40.05	185.7 ± 33.54	144.5 ± 29.05**	35.5 ± 6.22**	154.7 ± 32.47*	60.6 ± 8.16**	24.9 ± 3.64**	194.8 ± 38.91
	Lab 11	47.5 ± 8.36	405.9 ± 69.84*	439.7 ± 68.67	352.5 ± 80.28**	77.8 ± 26.35**	357.1 ± 40.27**	177.7 ± 43.79**	28.2 ± 8.94**	463.2 ± 70.09
	Lab 13	53.9 ± 5.24	437.1 ± 120.54	411.6 ± 59.71	323.7 ± 58.97**	73.9 ± 7.95**	317.1 ± 62.31**	159.6 ± 44.96**	39.8 ± 5.18**	459.7 ± 97.18

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol	DDE		Linuron		Flutamide	TP
					16 mg/kg/d	160 mg/kg/d	10 mg/kg/d	100 mg/kg/d		
LABC	**Lab 1**	Not done	342.0 ± 51.93	303.2 ± 49.68	273.3 ± 72.18	142.4 ± 26.12**	264.9 ± 40.99	179.0 ± 26.14**	152.4 ± 19.16**	318.3 ± 27.42
	Lab 2	121.8 ± 10.32	454.9 ± 28.50	424.5 ± 23.87	406.5 ± 57.85*	112.7 ± 16.26**	430.3 ± 51.71	252.1 ± 27.12**	Did not run	467.4 ± 43.75
	Lab 3	Not done	320.1 ± 96.19	296.2 ± 35.27	282.2 ± 61.97	166.7 ± 19.69**	320.0 ± 47.48	183.8 ± 20.16**	152.3 ± 16.08**	339.1 ± 42.31
	Lab 5	Not done	331.6 ± 62.60	335.0 ± 28.45	315.9 ± 31.20	117.2 ± 42.13**	322.6 ± 24.16	199.7 ± 30.07**	154.3 ± 34.32**	341.3 ± 70.16
	Lab 7	181.3 ± 34.38	567.9 ± 57.04	569.4 ± 84.62	544.5 ± 93.77	287.1 ± 84.19**	533.5 ± 89.67	398.4 ± 53.77**	262.2 ± 59.86**	543.2 ± 59.40
	Lab 8	293.1 ± 38.26	499.2 ± 32.03**	548.3 ± 42.15	395.8 ± 67.47**	304.6 ± 34.60**	442.9 ± 33.88**	280.5 ± 45.63**	296.2 ± 27.14**	592.2 ± 53.73
	Lab 9	Not done	393.1 ± 16.79	385.0 ± 29.70	400.9 ± 25.76	241.7 ± 41.02**	390.9 ± 22.83	290.0 ± 21.06**	268.8 ± 15.66**	392.6 ± 35.39
	Lab 10	119.0 ± 6.97	307.8 ± 39.96	302.3 ± 29.37	280.7 ± 29.74	125.7 ± 13.39**	302.9 ± 40.44	155.8 ± 21.83**	148.3 ± 18.57**	306.2 ± 39.09
	Lab 11	217.7 ± 25.18	539.6 ± 52.98	529.9 ± 32.25	496.4 ± 71.55	218.1 ± 39.58**	532.4 ± 56.47	336.0 ± 46.34**	235.9 ± 28.96**	575.3 ± 31.81
	Lab 13	193.1 ± 11.30	460.2 ± 84.67	494.4 ± 50.12	396.4 ± 50.95**	165.4 ± 35.48**	452.4 ± 64.24	262.6 ± 47.80**	221.8 ± 51.49**	518.1 ± 74.20

Table 12 (continued) Antagonist data

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol	DDE 16 mg/kg/d	DDE 160 mg/kg/d	Linuron 10 mg/kg/d	Linuron 100 mg/kg/d	Flutamide	TP
GP	**Lab 1**	Not done	70.2 ± 6.40	66.3 ± 6.49	61.7 ± 7.63	44.9 ± 6.87**	61.1 ± 4.22	57.2 ± 5.48*	47.2 ± 9.26**	68.0 ± 10.00
	Lab 2	54.1 ± 10.23	92.6 ± 8.97	93.5 ± 10.06	80.6 ± 5.64	59.3 ± 7.06**	89.5 ± 12.00	76.1 ± 7.29**	Did not run	92.2 ± 14.89
	Lab 3	Not done	69.2 ± 11.74	68.5 ± 8.11	62.0 ± 9.61	45.8 ± 4.57**	67.6 ± 8.86	56.5 ± 4.02**	46.9 ± 6.36**	71.3 ± 12.60
	Lab 5	Not done	66.3 ± 12.58 [a]	68.3 ± 7.34 [a]	58.5 ± 4.74* [a]	[a]	67.1 ± 8.80 [a]	[a]	[a]	75.0 ± 6.97 [a]
	Lab 7	53.2 ± 6.09	89.6 ± 7.80	89.4 ± 6.85	85.4 ± 8.10	67.3 ± 10.11**	89.2 ± 4.46	87.7 ± 4.52	73.5 ± 8.96*	84.4 ± 5.08
	Lab 8	52.6 ± 5.07	81.6 ± 7.86	80.5 ± 7.40	69.5 ± 4.67**	59.6 ± 5.86**	73.2 ± 4.74	61.4 ± 2.59**	57.2 ± 5.79**	80.0 ± 6.29
	Lab 9	Not done	111.3 ± 5.43*	117.3 ± 5.56	112.2 ± 6.03*	97.0 ± 6.21**	116.5 ± 5.94	106.6 ± 3.86**	94.8 ± 3.94**	119.1 ± 5.13
	Lab 10	29.4 ± 2.95	61.9 ± 5.51	61.2 ± 2.10	53.6 ± 3.34*	32.0 ± 3.86**	64.5 ± 6.09	41.7 ± 6.39**	30.3 ± 4.21**	60.9 ± 5.29
	Lab 11	49.9 ± 2.89	75.6 ± 5.90	75.2 ± 2.14	71.7 ± 2.50	49.8 ± 1.65**	73.0 ± 3.29	63.9 ± 6.22	49.7 ± 2.01**	73.0 ± 3.66
	Lab 13	50.9 ± 3.81	86.3 ± 7.94	89.6 ± 7.61	83.9 ± 4.29**	60.2 ± 7.34**	90.2 ± 6.10	73.2 ± 10.01**	60.8 ± 3.36**	93.5 ± 4.59

		Vehicle	Nonylphenol 160 mg/kg/d	Dinitrophenol	DDE 16 mg/kg/d	DDE 160 mg/kg/d	Linuron 10 mg/kg/d	Linuron 100 mg/kg/d	Flutamide	TP
COW	**Lab 1**	Not done	25.7 ± 6.04	21.8 ± 5.79	19.8 ± 7.55	9.3 ± 1.58**	21.6 ± 4.45	12.8 ± 1.23**	9.3 ± 2.96**	21.7 ± 2.89
	Lab 2	4.3 ± 2.02	41.1 ± 7.54	39.9 ± 11.67	42.2 ± 4.70	12.4 ± 2.75**	48.5 ± 7.18	25.6 ± 6.30*	Did not run	40.1 ± 17.73
	Lab 3	Not done	16.2 ± 3.75	15.0 ± 1.76	13.3 ± 2.01	7.7 ± 4.11**	16.6 ± 2.64	10.4 ± 1.52**	5.8 ± 0.69**	15.5 ± 3.71
	Lab 5	Not done	23.1 ± 6.20*	25.6 ± 5.52	23.9 ± 4.12	7.0 ± 2.27**	24.6 ± 5.42	12.3 ± 2.29**	6.0 ± 1.76**	29.8 ± 9.24
	Lab 7	4.7 ± 2.13	42.1 ± 9.47	38.7 ± 7.27	34.9 ± 8.55	18.8 ± 4.71**	32.7 ± 6.29	22.2 ± 2.33**	11.1 ± 5.77**	35.2 ± 5.80
	Lab 8	16.3 ± 2.96	36.2 ± 3.33*	41.1 ± 10.02	26.0 ± 5.82**	17.6 ± 4.47**	31.0 ± 5.78**	18.4 ± 3.92**	12.8 ± 1.80**	43.8 ± 7.88
	Lab 9	Not done	42.0 ± 3.70	43.3 ± 4.51	40.3 ± 6.03	25.3 ± 3.77	40.3 ± 0.98	29.5 ± 3.57	20.1 ± 3.03	39.8 ± 4.81
	Lab 10	3.2 ± 1.29	20.1 ± 1.65	16.2 ± 3.55	16.4 ± 1.88	4.9 ± 0.83**	18.1 ± 3.30	8.3 ± 1.85**	4.7 ± 1.26**	18.4 ± 4.62
	Lab 11	5.7 ± 0.85	28.5 ± 5.61	31.0 ± 2.59	30.9 ± 6.46	10.5 ± 3.78**	25.7 ± 6.21	22.9 ± 9.05	10.4 ± 3.39**	33.0 ± 5.35
	Lab 13	8.3 ± 2.82	37.5 ± 11.13	33.4 ± 5.30	23.5 ± 5.86**	9.3 ± 2.54**	34.8 ± 6.73	20.9 ± 4.60**	12.0 ± 3.80**	35.6 ± 6.51

Many animals in this lab did not undergo preputial separation, so glans penis was not dissected or weighed

Pairwise t-test comparisons:

[1] **: $p < 0.01$; *: $p < 0.05$.

58. Testing for androgen antagonist effects always leads to an interaction of 2 substances, namely the reference androgen agonist (TP) and the test substance. Therefore, changes in body weight or optional organ weight by test substances may be related either to the vehicle or the reference androgen control group. Both of these relationships are of minor biological significance in respect to the overall objective of this Phase-3 validation for identification of androgen antagonistic effects. Therefore a statistical analysis of body and optional organ weights for the TP reference groups was considered meaningless and not conducted.

p-Nonylphenol branched (NP)

59. NP at a dose level of 160 mg/kg/d did not lead to an effect on any of the target organ weights in 5 of the 10 participating laboratories (laboratories 1, 2, 3, 7 and 13). Some of the laboratories reported a statistically significant decrease of target organs for this anticipated negative reference chemical as might have been expected for an anti-androgenic response:

- Laboratory 5: There was a statistically significant decrease in the weight of COW, but it was numerically far less pronounced than the effects observed with DDE (160 mg/kg/d), LIN (100 mg/kg/d) or FLU.
- Laboratory 8: 3 of the 5 target organs showed a statistically significant decrease (SVGC, LABC and COW). Again, these effects were numerically much less pronounced than after administration of DDE (160 mg/kg/d), LIN (100 mg/kg/d) or FLU.
- Laboratory 9 reported a statistically significant decrease for 1 of the 5 target organ weights (GP). This decrease was numerically in the same range as after treatment with LIN (100 mg/kg/d), but clearly less than after treatment with DDE (160 mg/kg/d) or FLU. This laboratory reported by far the highest GP weight for the reference androgen treatment group with TP. This may indicate some problems with the dissection procedure of GP as compared to the other laboratories.
- Laboratory 10 reported a statistically significant decrease in 1 of the 5 target organs investigated (VP). Again, this decrease was numerically much less pronounced than after treatment with DDE (160 mg/kg/d), LIN (100 mg/kg/d) or FLU.
- Laboratory 11 reported a statistically significant decrease in 1 of the 5 target organ weights (SVGC). Again, this decrease was numerically much less pronounced than after treatment with DDE (160 mg/kg/d), LIN (100 mg/kg/d) or FLU.

60. Overall, 1 of the 10 laboratories found a statistically significant organ weight decrease in several of the target organs investigated (laboratory 8). For the other laboratories either none or at most 1 of the target organ weights showed a decrease that might have been interpreted as an androgen antagonistic effect. Of these latter laboratories one might have had some problems with the dissection procedure for GP (laboratory 9).
In total, there is a dataset of 49 determinations of target organ weights and of these only 7 (i.e. 14%) gave an indication for a possible antiandrogenic effect for this negative reference chemical.

61. In conclusion, these data show that the interpretation of the test results with regard to a possible anti-androgenic effect should not rely on only one of the target organs but should take into account the whole dataset.

2,4-Dinitrophenol (DNP)

62. DNP at a dose level of 10 mg/kg/d did not lead to an effect for any of the target organ weights in 7 of the 10 participating laboratories (laboratories 1,2,3,5,7,11,13).
Since in laboratory 5 many animals did not undergo preputial separation, the glans penis could not be dissected or weighed in this laboratory and will not be considered by the statistical analysis. Laboratories 8, 9, and 10 reported the following statistically significant changes of target organ weights in relation to the androgen agonist treatment group (TP):

- Laboratory 9: There was a statistically significant weight increase for VP. Such a weight increase in conjunction with TP treatment is biologically not meaningful and should be regarded as a spurious finding.
- Laboratory 8 reported a statistically significant weight decrease for VP and SVCG. In both of these cases, the weight decrease was numerically much less pronounced than after treatment with DDE (160 mg/kg/d), LIN (100mg/kg/d) or FLU.
- Laboratory 10 found a statistically significant decrease of VP weight. Again, the weight decrease was numerically much less pronounced than after treatment with DDE (160 mg/kg/d), LIN (100mg/kg/d) or FLU.

63. In summary, laboratory 9 reported a spurious weight increase of VP, laboratory 10 a decrease in 1/5 target organ weights (VP), laboratory 8 a decrease in 2/5 target organ weights (VP and SVGC), while the remaining 7 laboratories did not find any effect on the target organ weights as was expected for this negative reference chemical. In total, only 3 of 49 target organ weights investigated (6%) showed a decrease.

64. In conclusion, these data show that the interpretation of the test results with regard to a possible anti-androgenic effect should not rely on only 1 of the target organs but should take into account the whole dataset.

p,p'-DDE (DDE) 16 mg/kg/d

65. The low dose of DDE (16 mg/kg/d) used in the Phase-3 studies was in between two of the dose levels of Phase-2, namely 10 and 30 mg/kg/d, when using a TP reference dose of 0.2 mg/kg/d and corresponded to one of the DDE dose levels with 0.4 mg/kg/d of TP. In the test series with 0.2 mg/kg/d of TP in Phase-2 there was no statistically significant decrease in target organ weights after dosing with 10 mg/kg/d as compared to the androgen reference group treated with TP. On the other hand, with 30 mg/kg/d 4 of the 5 participating laboratories reported an androgen antagonistic effect in some of the target organs (Laboratory 14 did not find any significant weight increase for the target organs after application of 30 mg/kg/d of DDE). In the test series with 0.4 mg/kg/d of TP none of the target organs showed an anti-androgenic response in any of the participating laboratories in Phase-2 with 16 mg/kg/d of DDE. Therefore, for Phase-3 the expectation was that 16 mg/kg/d should either still be in the range of the no effect level or just at the beginning of the ascending dose-response curve for anti-androgenicity.

66. After 16 mg/kg/d of DDE in Phase-3 some laboratories reported for some of the target tissues a significant weight decrease as compared to the TP androgen agonist reference group.

- For VP 4/10 laboratories reported a statistically significant anti-androgenic effect (Laboratories 7, 8, 10 and 13).
- For SVCG 5/10 laboratories reported a statistically significant anti-androgenic effect (Laboratories 1, 8, 10, 11, and 13).
- For LABC 3/10 laboratories reported a statistically significant anti-androgenic effect (Laboratories 2, 8, and 13)

- For GP 4/9 of the evaluated laboratories reported a statistically significant anti-androgenic effect (Laboratories 8, 9, 10, and 13). Laboratory 5 was not included in this evaluation because of problems with dissecting the glans penis.
- For COW 2/10 of the laboratories reported a statistically significant anti-androgenic effect (Laboratories 8 and 13).

In all cases, the target organ weight decrease as related to the TP androgen agonist reference group was numerically clearly less pronounced than the weight decrease observed after application of the high DDE-dose (160 mg/kg/d).
In total, with the low dose of DDE, 2 laboratories found an androgen antagonistic effect in all 5 target organs, 1 laboratory in 3 target organs, 5 laboratories in 1 target organ, and 2 laboratories in none of the target organs.

67. In conclusion, the application of 16 mg/kg/d DDE in conjunction with TP at a dose level of 0.2 or 0.4 mg/kg/d led to a decrease in target organ weights in 18/49 (37%) of the investigated target organs. This indicates that 16 mg/kg/d DDE led to variable effects and that this dose level was just at the beginning of the androgen antagonist dose-response curve.

68. Three laboratories tested DDE at the same dose levels in Phase-2 as well as in Phase-3 using TP as a reference dose of 0.4 mg/kg/d. 16 mg/kg/d of DDE did not lead to weight decrease in any of the target organs for all 3 laboratories in Phase-2. In contrast, in Phase-3 one of the participating laboratories (laboratory 8) reported a statistically significant weight decrease for all target organs. This indicates that for some unknown reason the sensitivity of laboratory 8 for detecting the androgen antagonistic activity of the low dose of DDE had increased. This should be seen in conjunction with the fact that the low dose of DDE was already at the ascending part of the dose-response curve. The other 2 laboratories (3 and 9) reported only for 1 of the 10 target organs (10%) a significant weight decrease in comparison to the TP reference group.

69. Two laboratories (10 and 11) tested DDE in Phase-2 as well as in Phase-3 with an androgen agonist reference dose of 0.2 mg/kg/d. The dose levels selected were different for both phases. The low dose level of 16 mg/kg/d in Phase-3 was in between the dose levels of 10 and 30 mg/kg/d of Phase-2. 10 mg/kg/d in Phase-2 did not show any indication for an androgen antagonistic effect in Phase-2, while 30 mg/kg/d showed an antagonistic response for VP, SVCG, and COW in laboratory 10 and for VP, SVCG, and LABC for laboratory 11. The organ weights in Phase-3 testing at 16 mg/kg/d showed a significant anti-androgenic effect for VP, SVCG, and GP in laboratory 10 and for SVCG in laboratory 11, i.e. 40% of the target organs showed an androgen antagonistic response at the low dose level. Under consideration of the different dose levels tested in Phase-2 and-3 there is only 1 target organ weight in Phase-3 that was in clear disagreement to the results obtained in Phase-2, namely the significant weight decrease of GP in laboratory 10. A weight decrease for GP was not observed at 30 mg/kg/d in both participating laboratories. As in Phase-2 a higher dose level was tested as compared to Phase-3 it is well to be expected that in Phase-2 more target organs showed a treatment related weight decrease, i.e. LABC and COW. Overall there only is one clear disagreement for 1 of 10 organ weights (10%) when considering the results obtained after the three dose levels mentioned above.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

p,p'-DDE (DDE) 160 mg/kg/d

70. In Phase-2 all 5 participating laboratories reported statistically significant anti-androgenic effects for all of the 5 target tissues using 100 mg/kg/d pp'-DDE. Therefore, at the dose level of 160 mg/kg/d in Phase-3 a clear anti-androgenic effect was expected, too.

71. This expectation was clearly fulfilled: only laboratory 9 did not find a statistically significant decrease of COW in comparison to the TP reference group. Interestingly, this laboratory did not find a COW weight decrease after treatment neither with the high LIN dose nor with FLU (see below) indicating that this laboratory may have had some problems in dissecting this target organ.

72. All other laboratories reported a significant organ weight decrease for all target organs after treatment with 160 mg/kg/d DDE in comparison to the TP reference group. In total, a significant anti-androgenic effect was demonstrated for 48/49 (98%) of the target organs investigated.

73. Three laboratories tested the high dose of DDE (160 mg/kg/d) in Phase-2 as well as in Phase-3 using a TP reference dose of 0.4 mg/kg/d. In both phases all 3 laboratories reported a statistically significant weight decrease for all 5 target organs investigated. Thus, there was a 100% concordance.

74. The high dose levels of DDE used in Phase-2 (100 mg/kg/d) and Phase-3 (160 mg/kg/d) were different when using a TP reference dose of 0.2 mg/kg/d. Two laboratories participated in both phases and both of them recorded significant weight decreases for all target organs. When the results obtained for the different dose levels of 100 and 160 mg/kg/d are compared to each other, there is a 100% concordance with regard to the androgen antagonistic target organ effects.

Linuron (LIN) 10 mg/kg/d

75. In Phase-2 testing, 10 mg/kg/d LIN did not lead to a statistically significant target organ weight decrease as compared to the reference TP group in any of the target organs of all of the participating 4 laboratories. Thus, for the Phase-3 testing at the same dose no or at most a first indication for an anti-androgenic effect was to be expected.

76. In fact, this expectation was substantiated by the results of the Phase-3 validation testing for the majority of the target organs investigated:

- For VP only 3/10 of the laboratories found a statistically significant anti-androgenic effect (Laboratories 8, 10, and 11).
- For SVCG 4/10 of the participating laboratories found a statistically significant anti-androgenic effect (Laboratories 8, 10, 11, and 13).
- For LABC 1/10 laboratory found a statistically significant anti-androgenic effect (Laboratory 8).
- For GP none of the 9 evaluated laboratories found a statistically significant anti-androgenic effect.
- For COW 1/10 of the participating laboratories found a statistically significant anti-androgenic effect (laboratory 8).
- It should be noted, that laboratory 8 was always among those (or was the only one) that found a statistically significant anti-androgenic effect. This may indicate that the test conditions of this laboratory had an exceptional high sensitivity for the detection of an anti-androgenic effect at the beginning of the dose response curve.

77. The statistically significant weight decreases in the target organs mentioned above were numerically always far less pronounced than those observed after the high dose of LIN (100 mg/kg/d).

78. In total, 9/49 (18 %) of all target organs investigated by all laboratories showed a statistically significant anti-androgenic effect. Therefore, a dose level of 10 mg/kg/d of LIN is at the beginning of the anti-androgenic dose-response curve, leading to such an effect only in some cases.

79. Two laboratories (1 and 5) tested the low dose of Linuron (10 mg/kg/d) in Phase-2 as well as in Phase-3. Laboratory 5 had problems with dissection of GP due to sometimes incomplete preputial separation, thus the results for GP will not be taken into consideration here. The remaining 9 target organs in both laboratories did not show an anti-androgenic weight decrease. Therefore, overall the concordance is 100%.

Linuron (LIN) 100 mg/kg/d

80. In Phase-2 testing, 3/4 of the participating laboratories reported a statistically significant anti-androgenic effect after treatment with 100 mg/kg/d LIN in all target organs (Note: the GP weight of laboratory 5 was excluded from this evaluation because many animals had not undergone preputial separation). One laboratory (laboratory 6) found a statistically significant weight decrease after LIN treatment in conjunction with TP administration for only 1 target organ (SVCG). Thus, it was to be expected for phase 3 that 100 mg/kg/d LIN should lead to an anti-androgenic effect in most instances.

81. This expectation was substantiated by the data obtained. Six of the 10 participating laboratories found a significant anti-androgenic effect for all 5 target organs, 3 laboratories for 4 target organs and 1 laboratory reported the anticipated weight decrease only for 3 target organs. All of the participants correctly identified the antiandrogenic response in VP, SVCG and LABC.

- For GP, 2/9 of the evaluated laboratories did not find a significant weight decrease as compared to the GP control group: laboratory 7 even reported a slight numerical weight increase for GP indicating to a specific problem with the dissection procedures. The weight decrease of laboratory 11 did not attain statistical significance, but was numerically clearly lower than the GP weight of the TP reference group.
- For COW 2/10 of the participating laboratories did not find statistically significant androgenic effects. As mentioned above, laboratory 9 could not demonstrate an anti-androgenic effect neither with DDE (160 mg/kg/d) nor with FLU (3 mg/kg/d), indicating to some specific problems with this target organ in this laboratory. For laboratory 11 there was a numerical weight decrease of COW after high dose treatment with LIN, but this decrease did not attain statistical significance as compared to the TP control group.

82. In total, only 4/49 (8%) of the target organs investigated by all participating laboratories did not show a statistically significant anti-androgenic effect. Thus, the expectation is clearly fulfilled that LIN at a dose level of 100 mg/kg/d will act as an androgen antagonist.

83. The high dose of Linuron (100 mg/kg/d) was tested by 2 laboratories (1 and 5) both in Phase-2 and Phase-3. As laboratory 5 had problems with dissecting GP due to incomplete preputial separation, this target organ will not be considered here. The remaining 9 target organs in both of the laboratories all showed a statistically significant weight decrease in line with the anti-androgenicity at this LIN dose level. Thus, overall there is a 100% concordance.

Flutamide (FLU) 3 mg/kg/d

84. In Phase-1B of the validation exercise all participating laboratories found a significant decrease of target organ weights after treatment with FLU (3 mg/kg/d) in comparison to the reference group treated with 0.2 or 0.4 mg/kg/d TP. Thus, it was to be expected that this dose level should lead to a clear androgen agonist effect in all participating laboratories. In Phase-3, 9 laboratories used FLU as a reference androgen antagonist. Lab 2 did not participate with FLU and again the data of Lab 5 for GP were excluded from statistical analysis because of incomplete preputial separation.

85. In Phase-3 testing for androgen antagonistic effects, all 9 participating laboratories reported statistically significant decreases for 4 target organs (VP, SVCG, LABC, and GP) in relation to the TP agonist reference group. For COW only laboratory 9 could not find such a decrease in organ weight. As already mentioned above, COW weights in this laboratory were not significantly decreased with all of the anticipated androgen antagonists, namely DDE at 160 mg/kg/d and LIN at 100 mg/kg/d. This is a strong indication for some specific problems with this target organ in this laboratory.

86. In total, only 1/44 (2%) of the target organ weights did not come up with a statistically significant weight increase. Thus, the overall expectation was clearly fulfilled that FLU (3 mg/kg/d) acts as an androgen antagonist.

87. FLU at a dose level of 3 mg/kg/d was tested by 3 laboratories (5, 10 and 13) both in Phase-1B and -3. Due to dissection problems for GP in laboratory 5 this target organ will not be taken into consideration here. All the other target organs (in total 14 from 3 participating laboratories) showed a androgen antagonistic response in Phase-1B and Phase-3. Therefore, the overall concordance is 100%.

Testosterone Propionate (TP) as an androgen agonist at 0.2 and 0.4 mg/kg/d

88. TP was used as reference agonist at a dose level of 0.4 mg/kg/d (Laboratories 1-9) and 0.2 mg/kg/d (Laboratories 10-13). This offers the possibility to again compare the target organ weight increases caused by TP in comparison to the vehicle control group. Weight determinations for the vehicle control group was optional and not carried out by laboratories 1, 3, 5 and 9. In addition, the TP test materials were not supplied in a coded manner. Therefore, a statistical comparison of the organ weights after TP treatment in relation to the vehicle control groups was not done.

89. When comparing the target organ weights of the vehicle control groups and the TP treatment groups, there was always a massive numerical increase that clearly substantiated the androgen agonistic effect of TP, although a detailed statistical analysis was not done. This increase in comparison to the vehicle control group was found by the 3 laboratories using 0.4 mg/kg/d of TP (Laboratories 2, 7 and 8) as well as for the lower dose level of 0.2 mg/kg/d (Laboratories 10, 11, and 13). Thus, an androgen agonist activity of TP again was clearly demonstrated.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Androgen antagonist data; overall analysis (cf. annex 4 of Lead Laboratory)

90. One of the objectives of the inter-laboratory Phase-3 study was to demonstrate that the 10 participating laboratories could get the "correct" responses for all the chemicals on all androgen-dependent tissues. LIN was administered at 2 dose levels of 10 and 100 mg/kg/d that had also been used in Phase-2 testing. The dose levels of DDE in Phase-3 (16 and 160 mg/kg/d) only corresponded to those of Phase-2 when using a TP androgen reference dose of 0.4 mg/kg/d, while with a TP dose of 0.2 mg/kg/d, 3, 10, 30, and 100 mg/kg/d of DDE were tested in Phase-2. But nevertheless from the data obtained during Phase-2 there were clear expectations for the target organ weights of Phase-3 for the androgen antagonist test chemicals:

- No or only occasional target organ weight decreases with DDE at 16 mg/kg/d
- Clear target organ weight decreases with DDE at 160 mg/kg/d
- Small or only occasional target organ weight decreases with LIN at 10 mg/kg/d
- Clear target organ weight decreases with LIN at 100 mg/kg/d
- Clear target organ weight decreases with FLU at 3 mg /kg/d

91. When all the effects of the androgen-dependent treatment groups on 7 endpoints (the 5 target tissue weights and in addition body and liver weights) were correlated with the effects seen previously in Phase-1 or Phase-2 studies, a highly significant positive correlation was found with a correlation coefficient of $r = 0.975$. A graph depicting these observed versus expected effects in Phase-3 is given in the report of the Lead Laboratory (Annex 4). The expected and observed effects were calculated as % of the concurrent control values with the TP 0.4 mg/kg/d group as control (7 laboratories using 0.4 mg TP/kg/d). FLU predictions were from Phase-1B (from 5 laboratories using TP at 0.4 mg/kg/d) and LIN and DDE were derived from Phase-2 (from 4 laboratories using TP at 0.4 mg/kg/d).

92. Pooling the data obtained for the 5 androgen-dependent tissues determined in 10 laboratories, then FLU at 100 mg/kg/d and DDE at 160 mg/kg/d each yielded 96% positive hits (2 % of the misses for these 2 chemicals arose because 1 laboratory could not measure glans penis weight because the prepuce had not separated). Administration of LIN (100 mg/kg/d) produced 90 % positive responses. The low doses of DDE and LIN produced approximately 40 % and 18 % positive responses. Based upon the earlier validation studies it was not expected that the low doses of either of these chemicals would produce consistent positive responses. In fact, the low doses of these 2 test chemicals were more potent as anti-androgens in Phase-3 as compared to the Phase-2 results. DNP und NP were used as putative negative reference chemicals and they yielded only 8 % and 14 % positive responses on the androgen-dependent tissues.

93. Table 13 gives the results obtained after pooling the target tissue weights separately of the laboratories 1-9 using an androgen agonist reference dose of 0.4 mg TP/kg/d and those with a reference dose of 0.2 mg TP/kg/d (laboratories 10-13). With FLU and the high doses of DDE and LIN, a highly significant target organ weight decrease ($p < 0.01$) was observed in all cases. In addition, there was a significant weight decrease for the low dose of DDE for VP, LABC, GP and COW after treatment with 0.4 mg TP/kg/d and for VP and SVCG after treatment with 0.2 mg TP /kg/d demonstrating that 16 mg/kg/d DDE was already at the ascending part of the dose-response curve. Similarly, the low dose of LIN (10 mg/kg/d) showed a significant ($p < 0.01$) target organ weight decrease for VP, SVCG, and LABC after treatment with 0.4 mg TP/kg/d and for VP and SVCG for the androgen agonist reference dose of 0.2 mg TP/kg/d. NP as a negative reference chemical did not lead to a decreased weight of any of the target organs regardless whether a pre-treatment with 0.2 or 0.4 mg TP/kg/d was used (not even at a level of $p < 0.05$, data not given in the table). The same was true for the other negative reference chemical (DNP) with the exception that the VP weight was significantly decreased after pre-treatment with 0.2 mg TP/kg/d. This latter finding may be regarded as spurious as only 3 laboratories were using this TP dose level.

Table 13. Hershberger antagonist data; means and Standard Errors (SEs) of target tissue weights of all participating laboratories

Labs 1-9	VP		SVCG		LABC		GP		COW	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
TP 0.4	138	8.5	390	25	427	17	84.5	2.9	32.3	2
DDE 16	117**	7.2	347	32	374**	16	76.6**	3	28.6**	1.8
DDE 160	49.3**	4.8	119**	15	202**	15	65**	3.6	14.8**	1.3
DNP 10	137	8.8	391	27	409	17	83.3	2.9	32.2	1.9
FLU 3	36.4**	3.4	87**	10	214**	12	64**	3.6	10.8**	0.94
LIN 10	123**	7.7	336**	22	386**	15	81	3	30.8	1.8
LIN 100	73.4**	5.8	181**	17	254**	12	74.2**	3.2	18.8**	1.2
NP 160	131	6.7	401	29	416	16	82.97	2.7	32.3	1.8
OIL	25.8	3.4	57.6	4.2	199	18.4	53.3	1.7	8.4	1.5

Labs 10-13	VP		SVCG		LABC		GP		COW	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
TP 0.2	135	8	378	35	440	37	74	3.9	27.5	2.2
DDE 16	103**	8	270**	27	382	24	70	3.3	23.4	1.9
DDE 160	30.4**	2.2	85.6**	6.7	163**	12.8	45.2**	3.5	7.8**	0.9
DNP 10	116**	7.5	346	30	442	26	75.3	3	26.9	2
FLU 3	31**	2.1	56.8**	5.4	202**	12.2	46.9**	3.1	9**	1
LIN 10	110**	7.1	276**	24	429	26	75.9	2.8	26.2	2.1
LIN 100	56.9**	5.4	133**	14.9	251**	20	59.6**	3.6	17.3**	2.1
NP 160	120	8	345	32	436	27	74.6	2.8	26.7	2.4
OIL	17.6	0.9	41.8	3.4	177	10.8	43.4	2.5	5.7	0.6

** $p<0.001$

94. Analysis of variance for the 5 androgen-dependent tissues revealed that the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in Phase-3 were all statistically highly significant. Importantly, the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by the relatively small F-values for the chemical by laboratory interaction as compared to the size of the two main effects (table 14). When the analysis was conducted separately for the two groups of laboratories that used either TP at 0.4 mg/kg/d (7 laboratories) or 0.2 mg/kg/d (3 laboratories), both protocols yielded similar results.

There were significant laboratory effects in each ANOVA. This was expected and also seen in Phase-1 and Phase-2 of the validation studies. These effects arise because the laboratories used different rat strains, but even within the same strain the animal body and organ weights would be expected to vary. Such a laboratory effect per se is not problematic, whereas a large laboratory by treatment interaction would be because this would indicate chemicals were not acting similarly in all laboratories. Although the F-values for the interaction were small in comparison to the two main effects, the interaction effect is still statistically significant. This could lead to different conclusions in the case of small effects of a substance depending on the laboratory.

Table 14. Hershberger antagonist data; results of ANOVAs for laboratory interaction and laboratory by treatment interaction

Body weight	Labs 1-9				Labs 10--13		
	Effect	df	F	p-value	df	F	p-value
	LAB	6	355	0.0001	2	104	0.0001
	Chemical	7	16	0.0001	7	9	0.0001
	Interaction	41	1.4	NS	14	1.1	NS

SVCG	Labs 1-9				Labs 10- 13		
	Effect	df	F	p-value	df	F	p-value
	LAB	6	133	0.0001	2	158	0.0001
	Chemical	7	100	0.0001	7	95	0.0001
	Interaction	41	3.6	0.0001	14	5.5	0.0001

LABC	Labs 1-9				Labs 10, 11-13		
	Effect	df	F	p-value	df	F	p-value
	LAB	6	146	0.0001	2	109	0.0001
	Chemical	7	151	0.0001	7	54	0.0001
	Interaction	41	3.5	0.0001	14	2	0.03

VP	Labs 1-9				Labs 10, 11-13		
	Effect	df	F	p-value	df	F	p-value
	LAB	6	91	0.0001	2	50	0.0001
	Chemical	7	85	0.0001	7	67	0.0001
	Interaction	41	2.4	0.0001	14	2	0.02

GP	Labs 1-9				Labs 10-13		
	Effect	df	F	p-value	df	F	p-value
	LAB	6	235	0.0001	2	295	0.0001
	Chemical	7	49	0.0001	7	73	0.0001
	Interaction	38	1.1	NS	14	1.8	0.05

COW	Labs 1-9				Labs 10-13		
	Effect	df	F	p-value	df	F	p-value
	LAB	6	111	0.0001	2	82	0.0001
	Chemical	7	83	0.0001	7	40	0.0001
	Interaction	41	2.8	0.0001	14	3	0.002

95. In table 15 the CVs and the R-square values are summarized over the laboratories for the target organs. Thereby the R-square values and the CVs can be compared from laboratory-to-laboratory. The size of the R-square value describes the strength of association of the chemical effects versus TP control and the CV describes the precision (variability) of the data. If all the laboratories had the same CV on an endpoint then this would indicate that they were all dissecting the tissues with similar precision. A high CV in 1 laboratory versus the other laboratories might indicate a problem. However, all CVs seemed to be within a reasonable range, indicating that the dissection had been consistent. If the chemical treatments all had the same magnitude of response in each laboratory then the R-square values were to be similar among laboratories, given that the variances were not markedly different. Indeed, there was only one laboratory in which this was not the case: In laboratory 5 the R-square value for the effect of chemical treatments on GP was lower than for the other 9 laboratories. This

arose because laboratory 5 was unable to obtain GP weights with some of the animals treated with the potent anti-androgens biasing the means as the most-affected animals could not be measured. In fact, for the FLU group in this laboratory, no GP weight data are available. Hence, this low R-square value arises from the incomplete nature of their data set for GP weight.

Table 15. Hershberger antagonist data; pooled R-square values and CVs by laboratory and target organ

VP	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11	Lab 13	Mean(CV)
R2	71	68	68	83	56	90	85	86	80	81	
CV	31	23	30	22	33	17	14	18	25	22	23.5
SVCG											
R2	73	61	71	80	85	84	89	87	89	83	
CV	36	34	29	29	19	24	12	22	19	24	24.8
LABC											
R2	77	91	70	82	75	89	86	88	70	83	
CV	17	10	19	16	16	10	8	13	21	16	14.6
GP											
R2	63	57	55	27	58	75	76	90	74	80	
CV	12	12	15	13	9	8	5	9	9	8	10
COW											
R2	65	58	71	77	73	80	83	86	65	75	
CV	26	26	20	26	22	21	12	20	26	24	22.3

96. In summary, the overall analysis as given in Annex 4 of the results of Phase-3 of the Hershberger Bioassay inter-laboratory study using coded samples of known dosage levels of antagonists or negatives produced appropriate and consistent responses among all laboratories. None of the known positives were undetected in any laboratory. Only a few "false" positive responses were seen with the "known" (or suspected) negative reference chemicals.

DISCUSSION

97. The need to validate test methods for the detection of chemicals interfering with the endocrine system arises from the concerns that ambient levels of natural and industrial chemicals may interact with the endocrine system and as a consequence possibly elicit reproductive, developmental, and other adverse effects in humans and wildlife. The Hershberger Bioassay is a leading candidate for a level 3 *in vivo* screening assay of the OECD Conceptual Framework (4) to identify potential androgens and antiandrogens. It is rapid and efficient and there is strong evidence for its sensitivity and specificity from the literature.

98. The objective for the Hershberger Bioassay validation programme is to validate a test protocol in order to support the development of a Test Guideline for the detection of chemicals having the potential to act as androgen agonists or antagonists in rats. In the preceding validation Phase-1 a protocol was developed for identification of androgen agonists and antagonists by using Testosterone Propionate and Flutamide as potent reference chemicals. In Phase-2 the protocol developed during Phase-1 was used to test two further androgen agonists (17α-Methyl-testosterone and Trenbolone), four androgen antagonists weaker than Flutamide (Procymidone, Vinclozolin, Linuron, and p,p'-DDT) and a potent 5α-reductase inhibitor (Finasteride). The test materials were supplied uncoded at pre-selected dose levels to obtain a dose-response curve by the participating laboratories.

99. The Phase-2 validation program successfully achieved the goal of demonstrating the reproducibility of the protocol for detecting the weaker androgen agonists and antagonists as well as the 5α-reductase inhibitor.

100. After successful completion of the Phase-2 validation testing the Phase-3 validation was initiated. Coded substances were tested at one or two predetermined dose levels to exclude possible investigator bias. The same dose levels were used by all participants to further substantiate inter-laboratory reproducibility. The dose levels for the agonists and antagonists had already been used in the previous Phase-1 and Phase-2 test series or were derived from the Phase-2 results (DDE). In addition, two chemicals anticipated to act neither as an androgen agonist nor as an androgen antagonist were added to give an indication for the specificity of the Hershberger Bioassay.

101. In Phase 3, the androgen agonists were Testosterone propionate and Trenbolone. The test chemical used as a reference androgen agonist was Testosterone propionate, which was co-administered with a suspected antagonist. The androgen antagonists were Linuron at 2 dose levels (10 and 100 mg/kg/d), p,p'-DDE at 2 dose levels (16 and 160 mg/kg/d) and Flutamide (3 mg/kg/d). For negative reference chemicals 4-Nonylphenol (mixed isomers) and 2,4-Dinitrophenol were used at dose levels anticipated to approach the maximally tolerable dose (MTD).

102. The overall goal of the Phase-3 validation study was to further assess the robustness and reproducibility of the Hershberger Bioassay in a blinded manner. The specific goals were to

- further demonstrate the reliability and relevance of the Bioassay in response to coded positive and negative substances without investigator bias,
- further evaluate the inter- and intra-laboratory reproducibility of the protocols for identifying androgen agonists and antagonists,
- demonstrate the reproducibility of the Bioassay over time by comparing the data of Phase-3 to that generated in Phase-1 and Phase-2,
- develop data to substantiate the specificity of the protocols both for identifying androgen agonists and antagonists by coded testing of negative reference chemicals
- evaluate the reproducibility of weight changes and relative effectiveness of the 5 target organs
- and thereby support a recommendation of the development of an OECD Test Guideline for the Hershberger Bioassay.

103. The laboratories participating in Phase-3 had already participated in Phase-1 and -2 thereby enabling a comparison with the data obtained in the previous independent test series.

104. After finalization of an OECD Test Guideline the Hershberger Bioassay should be widely used in many OECD member countries. Therefore the protocol has to allow for variations in a number of study conditions like the choice of the rat strain, the laboratory diet, housing and husbandry practices such as the use of cage bedding. In addition possible strain differences in the onset of puberty have to be taken into account and the age of castration and the time period of target organ regression after castration have to be kept flexible.

105. Based on the previous Phase-1 and -2 data the following results were expected for the Phase-3 coded validation study:

a) agonist test protocol:

- TP (0.4 and 0.2 mg/kg/d): statistically significant increase of target organ weights
- TREN (40 mg/kg/d): statistically significant increase of target organ weights
- TREN (1.5 mg/kg/d): no or no consistent increase of target organ weights
- NP (160 mg /kg/d): no increase of target organ weights
- DNP (10 mg/kg/d): no increase of target organ weights

as compared to the vehicle control group.

b) antagonist test protocol:

- FLU (3 mg /kg/d): statistically significant decrease of target organ weights
- LIN (100 mg/kg/d): statistically significant decrease of target organ weights
- LIN (10 mg/kg/d): no or no consistent effect on target organ weights
- DDE (160 mg/kg/d): statistically significant decrease of target organ weights
- DDE (16 mg/kg/d): no or no consistent effect on target organ weights
- NP (160 mg/kg/d): no effect on target organ weights
- DNP (10 mg/kg/d): no effect on target organ weights

as compared to the control group given TP as reference androgen agonist.

106. For statistical analysis the means, standard errors, standard deviations ad 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.

107. In the following the results of Phase-3 testing will be discussed in relation to the specific goals as mentioned above.

Identification of coded positive and negative substances without investigator bias

108. For the agonist test protocol coded samples were tested at defined dose levels. DNP (10 mg/kg/d) and NP (160 mg/kg/d) were used as negative reference chemicals. None of the 5 target tissues showed a weight increase as would have been expected for an androgen agonist, thereby substantiating the specificity of this test protocol.

109. For TREN a dose level of 1.5 mg/kg/d was not expected to lead to any appreciable androgen agonistic action; again the data showed no target organ weight increases apart from 2 laboratories reporting an androgen agonist action on LABC. In contrast, 40 mg/kg/d of TREN led to the expected target organ weight increases apart from VP and COW in 1 laboratory and GP in another one. Overall the expectations from the results of Phase-2 were fulfilled, i.e. no androgenic activity after application of 1.5 mg/kg/d TREN and a clear androgenic response after 40 mg/kg/d.

110. Similarly for the androgen antagonist test protocol the coded test chemicals led to the expected results: After administration of NP (160 mg/kg/d) there was generally no target organ weight decreases as would have been expected for an androgen antagonist. Only 14% of the target organs investigated by all 10 participating laboratories gave an indication for a possible anti-androgenic effect without any preference of a specific target organ. With DNP (10 mg/kg/d) only 6 % of the investigated target organs showed a weight decrease. Thus, the overall response is clearly in line with the expectation of NP and DNP not acting as an androgen antagonist.

111. For the coded antagonist DDE the expectation was that the low dose (16 mg/kg/d) would either be without an anti-androgenic activity or be close to the start of the dose-response curve. This expectation was fulfilled as only 40 % of the investigated target organs showed a weight decrease in comparison to the TP androgen agonist reference group. And only 1 of the target organs in Phase-3 (10%) at 16 mg/kg/d was in disagreement to the Phase-2 results at 30 mg/kg/d. 160 mg/kg/d in Phase-3 led to an anti-androgenic effect for all target organs in all participating laboratories as expected.

112. For the coded antagonist LIN the expectation was that the low dose (10 mg/kg/d) would either be a NOEL or be close to the start of the dose-response curve. This expectation was fulfilled as only 40 % of the investigated target organs showed a weight decrease in comparison to the TP androgen agonist reference group. The high dose of 100 mg/kg/d in Phase-3 did not lead to an anti-androgenic effect in only 8% of the target organs for all participating laboratories.

113. The data obtained with the coded test chemicals were supported by those with the uncoded positive reference chemicals: TP led to a weight increase for all target organs in all laboratories in the androgen agonist as well as in the androgen antagonist test protocols. Flutamide, as a positive reference chemical, showed androgen antagonist action in all target organs by all participating laboratories with only 1 exception.

114. In summary, the results obtained with the coded negative and positive test chemicals, the latter tested at 2 dose levels, completely fulfilled the expectations laid down before initiation of the experimental work. By coded testing a possible investigator bias could be ruled out.

Inter- and intra-laboratory reproducibility of the protocols for identifying androgen agonists and antagonists

115. Inter-laboratory reproducibility for the androgen agonist test protocol was clearly demonstrated: The coded negative reference chemicals (NP and DNP) did not lead to weight increase in any of the target organs by any of the participating laboratories (49 organs in total). Blinded testing of the low dose of TREN did not show a weight increase for 46 of the 49 target organs investigated while the high dose of TREN led to the expected significant target organ weight increase in 46 of the 49 target organ weights evaluated. For 1 laboratory there was an indication for dissection problems for VP and COW. The uncoded androgen reference chemical (TP) gave a uniform positive response for all target organs by all participating laboratories.

116. For the androgen antagonist test protocol the situation was similar:
For the total of 49 determinations of target organ weights for NP, only 7 (14%) gave an indication for a possible androgen antagonist effect and for DNP only 4 of 49 target organ weights. In total, for NP, 5 of the 10 participating laboratories reported no weight decrease in any of the 5 target organs evaluated and another 4 laboratories found a weight decrease only in 1 of the 5 target organs. Similarly coded testing of DNP did not lead to an effect on any of the target organ weights in 7 of the 10 participating laboratories and 2 of the other 3 laboratories reported a weight change only for one of the 5 target organs.
For the high dose of DDE, 9 of 10 participating laboratories found a statistically significant androgen antagonistic effect in all target organs and one of the participants missed this response only in one of the target organs. With the high dose of LIN, 6 of the 10 laboratories found a significant anti-androgenic effect for all 5 target organs, 3 laboratories for 4 target organs and 1 laboratory reported the anticipated weight decrease only for 3 target organs.
The low doses of DDE and LIN were selected as to be in the very low range of the ascending dose-response curve. Therefore it was expected that either none or only some of the participants would report an androgen antagonistic effect. In total, with the low dose of DDE, 2 laboratories found an androgen antagonistic effect in all 5 target organs, 1 laboratory in 3 target organs, 5 laboratories in 1 target organ, and 2 laboratories in none of the target organs. With the low dose of LIN, one laboratory found an androgen antagonistic effect in 4 target organs, 2 laboratories in 2 target organs, one laboratory in one target organ, and 6 laboratories in none of the target organs.
Finally, when testing the uncoded FLU positive reference antagonist, only 1 of the 44 investigated target organs did not show the anticipated weight decrease as compared to the androgen reference group.

117. These results clearly substantiate a high inter-laboratory reproducibility for the protocols for identifying androgen agonists and antagonists.

118. To investigate intra-laboratory reproducibility over time the results obtained in Phase-3 are compared to those in Phase-1 or -2 for the laboratories participating in the phases concerned. For each test chemical generally there were only few laboratories available for such a direct comparison:

- For the androgenic activity of TP (0.2 and 0.4 mg/kg/d), 9 laboratories with a total of 45 target organs gave an overall concordance of 100% when comparing Phase-1 and -3.
- For TREN at the low dose level of 1.5 mg/kg/d, 3 laboratories with a total of 15 target organs gave an overall concordance of 100 % (no androgenic activity) when comparing Phase-2 and -3.
- For the androgenic activity of the high dose of TREN (40 mg/kg/d), 3 laboratories with a total of 15 target organs gave an overall concordance of 80% when comparing Phase-2 and -3.
- The low dose levels of DDE tested against 0.4 mg/kg/d of TP were the same in Phase-2 and -3. Three laboratories participated in both phases. While in Phase-2 none of the laboratories found a significant target organ weight change, in Phase-3 one of the 3 laboratories reported a significant weight change for all of the target organs. The reason is unclear, but must be seen in conjunction with 16 mg/kg/d being at the beginning of the ascending part of the dose response curve. For the other two laboratories there was deviation for only 10% of the target organs investigated between Phase-2 and -3.

- The low dose levels tested for DDE were different in Phase-2 and-3 using a TP reference dose of 0.2 mg/kg/d. The low dose level in Phase-3 (16 mg/kg/d) is expected to be at the start of the ascending dose-response curve. Therefore a comparison was made with the embracing dose levels of Phase-2, namely 10 and 30 mg/kg/d. When considering the results of the 2 laboratories participating both in Phase-2 and-3 with a total of 10 target organs, there only was a clear disagreement of 10% between Phase-2 and-3.
- With a TP reference dose of 0.4 mg/kg/d the high dose of DDE was the same in Phase-2 and -3. 3 laboratories participated in both phases. For all 3 laboratories all target organs responded with the target organ weight decrease as to be expected for an androgen antagonist. Thus, the intra-laboratory concordance was 100%.
- The high dose levels of DDE were different in Phase-2 (100 mg/kg/d) and Phase-3 (160 mg/kg/d) with a TP reference dose of 0.2 mg/kg/d. Notwithstanding this difference in dosage, for the antiandrogenic activity at both of these high dose levels 2 laboratories with a total of 10 target organs gave an overall concordance of 100% for Phase-2 and -3.
- For the low dose level of LIN (10 mg/kg/d) 2 laboratories with in total 9 target organs gave an overall concordance of 100% (no anti-androgenic activity) when comparing phase-2 and-3.
- For the anti-androgenic activity of the high dose LIN (100 mg/kg/d) 2 laboratories with a total of 9 target organs gave an overall concordance of 100% when comparing Phase-2 and -3.
- The androgen antagonist reference chemical FLU (3 mg/kg/d) was tested by 3 laboratories with a total of 14 target organs; an overall concordance of 100% was found when compared during Phase-2 and -3.

119. In summary by these data, a good intra-laboratory reproducibility over time can be substantiated although only a limited number of laboratories tested the same chemicals in the different phases of this validation exercise.

Reproducibility over time

120. Phase-1, -2, and -3 of the validation program were carried out independently and at different times. Therefore, the data shown above clearly substantiate a good reproducibility over time, too.

Substantiation of the specificity of the protocols for identifying androgen agonists and antagonists

121. NP and DNP at dose levels anticipated to represent a maximally tolerable dose level (160 mg/kg/d and 10 mg/kg/d, respectively) were used as negative reference chemicals both for the androgen agonist and androgen antagonist test protocol.

122. By the androgen agonist test protocol, none of the participating laboratories reported a significant organ weight change for any of the target organs evaluated (in total 49 target organs for each chemical) neither for NP nor for DNP.

123. By the test protocol to detect androgen antagonists the 10 participating laboratories provided a dataset of in total 49 target organ weights for each of the negative reference chemicals. Of these, for NP only 7 (14%) gave a significant although generally small organ weight change as an indication for a possible anti-androgenic effect. Similarly for DNP only 3 of the 49 target organ weights investigated (6%) showed such a decrease.

124. In summary, these results clearly substantiate the specificity for protocols to identifying androgen agonists and antagonists.

Reproducibility of weight changes and relative effectiveness of the 5 target organs

125. The effectiveness of the target organs for identification of androgen agonists and antagonists depends to a large extent on the CVs of the weight determinations. Already in the Phase-2 validation report (5) two important findings were noted:

1. The CVs differ among the 5 mandatory tissues. LABC and GP have lower CVs, suggesting that these tissues may be somewhat easier to consistently dissect and handle. In contrast, the fluid-filled tissues such as the VP, SVCG, and COW generally have greater and more variable CVs from laboratory-to-laboratory.
2. There are differences in the CVs among the laboratories. Certain laboratories were more proficient overall than others and therefore the laboratory procedures and techniques themselves have a major impact on the performance of the assay.

126. These observations by the Phase-2 results are essentially supported by the evaluations of the Lead Laboratory of the Phase-3 data. This is most obvious for the agonist testing subpart. As can be seen from table 10 the CVs of the fluid-filled target organs (SVCG, VP, and COW) were in the range between 44 and 59%, while the CV for GP was 12.1% and that for LABC 18%. According to table 11 laboratories 5 and 10 generally had the highest CVs, especially for the fluid-filled target organs (VP, SVCG, and COW). A detailed analysis by the Lead Laboratory showed that these high CVs arose from a single animal in each laboratory from the TREN 40 mg/kg/d-dose and the antiandrogenic effects of the chemical treatment was still apparent in these laboratories.

127. A similar, but less obvious picture emerges from the Phase-3 antagonist testing. In table 15 the means and ranges of the CVs again are higher for VP (23.5; 14-33), SVCG (24.8; 12-36), and COW (22.3, 12-26) as compared to those of LABC (14.6; 8-21) and GP (10; 5-15).

128. The weight of LABC should be specifically responsive to the anabolic activity of androgen agonists. This is substantiated by the results obtained with the low dose of TREN (1.5 mg/kg/d) that is an androgen agonist with a pronounced anabolic activity (cf. table 7). This dose led to a statistically significant weight increase for LABC in 2 laboratories and to a clear numerical weight increase in some of the other laboratories. For the other 39 target organs investigated there was only one other statistically significant weight increase reported, namely an increase for COW in laboratory 2. Thus, LABC seems to be especially responsive to the anabolic action of androgen agonists and this may be of some value for differentiation between androgenic and anabolic activities.

129. In total, the weight changes of all target organs fulfilled the expectations for chemicals acting as androgen agonists or antagonists. As seen by the CVs, the variability is somewhat higher for the fluid-filled organs (VP, SVCG, and COW) as compared to the solid tissues (LABC and GP). Nevertheless, not always all of the target organs reacted to the test chemicals as was to be expected. Such single "false negative" results were found both for androgen agonists and antagonists for some of the target organs in some laboratories and similarly, occasionally there also were "false positive" organ weight changes after treatment with the negative reference chemicals. Details are given in the RESULTS-section. Therefore, when developing the final OECD Test Guideline, the 5 target organs used in the Phase-2 and Phase-3 validation programs should all be included. As already noted in the Phase-2 report (5) the value of GP still is open to some question.

SURGICAL CASTRATE MODEL PROTOCOL – PHASE-3

Some specific points noted in the Phase-2 report (5)

Incomplete preputial separation of GP

130. During Phase-2 validation it was already observed that some laboratories encountered a low rate of incomplete preputial separation, and this impacted the ability to dissect the GP (cf. summary of Phase-2 report § xiii). Therefore laboratories should understand the particular characteristics of their animal strain and supplier relating to the time of preputial separation. This should be taken into account when selecting the time for castration. Therefore, the Phase-3 protocol was more liberal in respect to the time of castration as can be seen from table 4a. But nevertheless, laboratory 5 still encountered problems for the evaluation of GP as many animals in this laboratory did not undergo preputial separation. This result underlines the necessity for sufficient flexibility as regards the age of castration taking into account the onset of puberty and timing of preputial separation of the experimental animals.

Appropriate dose of TP (reference androgen agonist) when testing for androgen antagonists

131. In the previous validation reports little or no difference became evident between 0.2 and 0.4 mg/kg/d as androgen agonist reference dose of TP when co-administered with different doses of androgen antagonists. This was substantiated by the data obtained in Phase-3 testing. The analysis of variance of Lead Laboratory yielded similar results when conducted separately for the 2 groups of laboratories that used TP either at a dose of 0.4 mg/kg/d (7 laboratories) or 0.2 mg/kg/d (3 laboratories). Details are found in Appendix 3 of the Lead Laboratory report on androgen antagonist testing (Annex 4). Thus, no clear preference can be derived from the data as to the most appropriate dose of the reference androgen agonist (TP), neither for 0.2 nor for 0.4 mg/kg/d subcutaneously.

CONCLUSIONS/RECOMMENDATIONS

132. Phase-3 validation testing has shown a good intra-laboratory reproducibility over time by using coded test samples to avoid investigator bias.

133. The use of negative reference chemicals showed a good specificity of the protocols to identify androgen agonists and antagonists.

134. The Phase-1, -2, and -3 results taken together show
- that the Hershberger protocols are transferable so that they can be used internationally by different laboratories
- that the inter-laboratory reproducibility was excellent for agonists, antagonists and a 5α-reductase inhibitor
- that the predictive capacity was very good.

135. Within the validation test series only one potent 5α-reductase inhibitor was tested. For this specific mode of action some further validation work might be envisaged with weaker inhibitors to confirm reproducibility and relevance of this assay for such a class of substances.

136. Laboratories using the Hershberger Bioassay should provide sufficient training for target organ dissection to their technical staff, especially if this bioassay is only carried out occasionally.

137. The biological variability of the strains used by the laboratories should be known, especially with regard to onset of puberty and time of preputial separation.

138. On the basis of the successful international OECD validation program it is proposed to develop an OECD Test Guideline for the identification of androgen agonists and antagonists (Hershberger Bioassay) that may be used as a level 3 screening test within the conceptual framework of the OECD (4).

REFERENCES

1. Crisp, TM, Clegg ED., Cooper, RL, Wood, WP, Anderson, DG, Baetcke, KP, Hoffmann, JL, Morrow, MS, Rodier, DJ, Schaeffer, JE, Touart, LW, Zeeman, MG, and Patel, YM. (1998). Environmental endocrine disruption: An effects assessment and analysis. Environ. Health Perspect. 106(Suppl. 1):11-56.

2. CSTEE. (1999). CSTEE Opinion on Human and Wildlife Health Effects of Endocrine Disrupting Chemicals, with Emphasis on Wildlife and on Ecotoxicology Test Methods. Report of the Working Group on Endocrine Disrupters of the Scientific Committee on Toxicity, Ecotoxicity and the Environment (CSTEE) of DG XXIV, Consumer Policy and Consumer Health Protection. Brussels, Belgium.

3. IPCS. (2002). Global Assessment of the State-of-the-Science of Endocrine Disruptors. T. Damstra, S. Barlow, A. Bergman, R. Kavlock, G. Van der Kraack, ed. International Programme on Chemical Safety. WHO/PCS/ECD/02.2 Geneva, Switzerland.

4. OECD. (2002). Summary record of the sixth meeting of the Task Force on Endocrine Disrupters Testing and Assessment (EDTA 6). ENV/JM/TG/EDTA/M(2002)2.

5. OECD. (2003). Draft report of the OECD validation of the rat Hershberger Bioassay: Phase 2: testing of androgen agonists, androgen antagonists and a 5a-reductase inhibitor in dose response studies by multiple laboratories. ENV/JM/TG/EDTA(2003)5/ADD1.

6. EPA. (1998). Endocrine Disrupter Screening and Testing Advisory Committee (EDSTAC) Final Report EPA/743/R-98/003. [http://www.epa.gov/scipoly/oscpendo/history/finalrpt.htm]

7. OECD. (1998). Detailed review paper on the appraisal of test methods for sex hormone disrupting chemicals. OECD Monograph No 21.

8. 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.

9. OECD. (2003). Special activity on endocrine disruptors testing and assessment: Summary of the 4th meeting of the VMG-mammalian and meeting outcome. ENV/JM/EDTA/RD(2003)11.

10. OECD. (2003). Draft report of the 7th meeting of the task force on endocrine disruptors testing and assessment (EDTA). ENV/JM/EDTA/M(2003)5.

ACKNOWLEDGMENTS

The dedicated efforts and work of Dr L. Earl Gray Jr. of the US Lead Laboratory and of Dr. J. William Owen of Procter & Gamble Company are appreciated.

ANNEX 1 − Participating Laboratories – Phase-3 Hershberger Validation Program

This Information is Only Available to Government Representatives of OECD Member Countries

There were ten participating laboratories in Phase-3 testing and all of them had already been involved in Phase-2. Their participation was voluntary. In some laboratories, the work received financial support from their national authorities or from CEFIC. In other cases, the laboratories were bearing the financial costs themselves. This support and assistance to the OECD work is sincerely appreciated and such contributions should be recognized by the VMG-mammalian.

ANNEX 2 − OECD Model Protocol for Phase-3 of the Hershberger Validation Program

This protocol was contained in a document distributed to each of the participating laboratories as the basis for writing their individual laboratory protocols.

OECD MODEL PROTOCOLS AND GUIDANCE FOR THE CONDUCT OF THE RAT HERSHBERGER BIOASSAY WITH THE SURGICALLY CASTRATED PERIPUBERTAL MALE

For Phase-3 of the OECD program to validate the rat Hershberger bioassay

LEAD LABORATORY

United States Environmental Protection Agency
National Health and Environmental Effect Research Laboratory
Endocrinology Branch
2525 Highway 54
Research Triangle Park, North Carolina 27711
UNITED STATES

Lead Investigator: Dr. Earl GRAY

OECD SECRETARIAT

OECD Environment
Health and Safety Division,
2, rue André Pascal
75775 PARIS, Cedex 16
FRANCE

Study Manager and Director: Dr. J. William Owens

OECD PROTOCOL AND GUIDANCE FOR THE SURGICAL CASTRATE PERIPUBERTAL MALE RAT HERSHBERGER BIOASSAY

Model Protocol and Guidance for Phase-3 of the OECD Work to Validate the Rat Hershberger Bioassay

INTRODUCTION

1. The overall aim of the validation program is to demonstrate that the Hershberger bioassay is a reliable and reproducible bioassay that can be considered as the basis for an OECD Test Guideline. This document provides the essential requirements for Phase-3 of the OECD program on the validation of the rat Hershberger bioassay using the surgically castrated pubertal male. Detailed laboratory protocols for the OECD validation program are to be built on the requirements, recommendations, and options contained in this document.

2. The precursor of the rodent Hershberger bioassay was first developed in the 1930s and included various tissues of the male reproductive tract (1) (2) (3), including the ventral prostate, the seminal vesicles with coagulating glands, the Cowper's glands, the glans penis, and preputial glands. The measurement of the levator ani and bulbocavernosus muscles were subsequently investigated in the 1940s (3) (4). After publication of work with an extensive number of compounds by Hershberger *et. al.* in 1953 (5), the procedure has been commonly referred to as the Hershberger bioassay.

3. The Hershberger bioassay is an *in vivo* short-term assay similar in concept to the rodent uterotrophic bioassay. Both procedures measure changes in specific tissues that normally respond to endogenous hormones. The bioassay conditions are designed to achieve low endogenous hormone levels and employ target tissues that are highly responsive to administration of exogenous hormones. The focus of the Hershberger bioassay is on the detection of compounds that may interfere with endogenous male sex hormones. This includes other androgens, androgen antagonists and 5α-reductase inhibitors. The objective in phase-3 of the validation of the surgical castrate is to assess the bioassay's reliability and relevance of the assay using coded substances and negative compounds.

4. The Hershberger and uterotrophic bioassays are both being validated by OECD as potential short-term screens. The information generated by these bioassays can be used to build on that already available, e.g. from relevant *in vitro* screens to narrow the field of chemicals that may need longer-term animal testing. This current protocol is based on the standardisation and optimisation work performed under OECD auspices in Phase-1 and Phase-2 of the Hershberger validation.

PREVIOUS VALIDATION WORK

5. The OECD program on the Hershberger assay has previously demonstrated the response and the reliability of measuring male sex accessory glands and tissues, i.e., ventral prostrate (VP), seminal vesicles with coagulating glands (SVCG), the levator ani and bulbocavernosus muscles (LABC), the Cowper's (or bulbourethral) glands (COWS), and glans penis (GP). Test substances used in Phase-1 and Phase-2 included testosterone propionate (TP) (CAS No. 57-85-2), Flutamide (FLU) (CAS No: 1311-84-7), trenbolone (TREN) (CASRN 10161-33-8), Vinclozolin (VIN) (CASRN 50471-44-8), Procymidone (PRO) (CASRN 32809-16-8), linuron (LIN) (CASRN 330-55-2), *p,p'*-DDE (DDE) (CASRN 72-55-9), Finasteride (FIN) (CASRN 98319-26-7). The assay was robust and reproducible across laboratories in the presence of several variations, e.g., strain of rat, diet, and modest variations in the age of castration. The protocol was successful in detecting increases in the weights of the accessory sex organs and tissues in response to androgens and detecting decreases due antiandrogens. There was good agreement among laboratories with regard to the dose responses obtained (6)(7).

INITIAL CONSIDERATIONS AND PRINCIPLE OF THE ASSAY

6. The rat Hershberger assay is based on changes in the weights of androgen-responsive male sex accessory tissues in peripubertal, surgically castrated male rats. Accessory sex tissues and glands depend upon androgen stimulation to gain and maintain weight during and after puberty. If endogenous sources of androgen (the testes) are removed, the biological activity of exogenous

substances can be assayed by the increase (agonist response) in the weights of these sex accessory tissues or by blocking (antagonist response) the activity of administered androgens and by preventing an increase in the weights of these sex accessory tissues. The rat Hershberger assay then evaluates the ability of a chemical to show biological activities consistent with the agonism or antagonism of androgens (e.g., testosterone and dihydrotestosterone).

7. The surgical castrate protocol uses male rats approaching sexual maturity (peripuberty). The animals are castrated by removing both testes and epididymes (orchidoepididyectomized). In most laboratory strains, such as the Sprague Dawley, Long Evans, or Wistar rats, peripuberty is expected to take place at approximately 6 weeks of age, within an expected age range of 5-7 weeks. Peripuberty is marked by prepuce separation from the glans penis. Prepuce separation is necessary so that the GP can be properly dissected and accurately weighed. At the peripubertal stage of sexual development, the GP and other androgen-dependent sex accessory tissues are sensitive to androgens, having both androgen receptors and appropriate steroidogenic enzymes. The advantage of using this age of rat is that the sex accessory tissues have a high sensitivity and small relative weight, which minimises variation in responses between individual animals. In addition, castration removes the ability to produce and to secrete testosterone (T), breaking the hypothalamic-pituitary-gonadal feedback loop and providing for regression of tissue weights to a lower baseline. As a result, serum T levels are severely reduced (some synthesis may occur in the adrenals).

8. Primary objectives of Phase-3 of the validation program are to demonstrate:

- The reliability of the Hershberger bioassay to respond to and to identify coded positive and negative substances.
- The relative effectiveness of the different sex accessory tissues and glands in the assay.
- The reproducibility of the bioassay over time by comparing appropriate data to that generated in Phase-1 and Phase-2.
- Continue the investigation of the value of the five accessory tissues and glands

9. The test substances will be coded and administered to groups of six animals (n = 6) for 10 consecutive days. The animals will then be necropsied approximately 24 hours later on the 11^{th} day. After dissection, the weights of the designated sex accessory tissues will be measured.

10. In addition to the sex accessory tissues, daily body weights, including at necropsy, are mandatory measures to allow precise does administration, to provide information on the general health and well being of the animals, and so that body weight can be used as a statistical covariable. The liver, adrenal and kidney weights are optional measurements that may provide supplementary information about the systemic toxicity, target organs and other effects of the test substance.

Androgen agonists

11. Biological activity consistent with androgen agonists is tested by administering a test substance to immature castrated rats for 10 consecutive days. The positive control for the tissue responses is TP. The vehicle is the negative control. The weights of the sex accessory tissues of the test chemical groups are compared to the vehicle group for a statistically significant increase in weight.

Androgen antagonists

12. Biological activity consistent with androgen antagonists and 5-alpha reductase inhibitors is tested by administering the test substance to immature castrated rats for 10 consecutive days together with a reference androgen agonist, TP. Administration of TP alone is the negative control. FLU is coadministered with TP to another group as a positive control. The weights of the sex accessory tissues after co-administration of the test chemical and the reference androgen TP together are compared with the weights of tissues of the reference androgen TP alone for a statistically significant decrease in weight.

DESCRIPTION OF METHOD/PREPARATIONS FOR THE TEST

Animal Species and Strain

13. 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. If a laboratory is planning to use an unusual rat strain, or one unique to their own facility, they should determine whether the sexual development criteria noted under the section, *INITIAL CONSIDERATIONS AND PRINCIPLE OF THE ASSAY*, are met.

Castration

14. After an initial acclimatisation period to ensure that the animals are health and thriving, the animals are castrated under anesthesia by placing an incision in the scrotum and removing both testes and epididymes with ligation of blood vessels and seminal ducts. After confirming that no bleeding is occurring, the scrotum should be closed with suture or autoclips. If castrated animals are purchased from an animal supplier, the age of animals and stage of sexual maturity should be assured by the supplier. The time between castration and initiation of dosing will be counted as part of the acclimatisation period. Animals should be castrated on pnd 42 or shortly thereafter, not before, based upon when the animals are expected to achieve full preputial separation by the first day of test substance administration.

Acclimatisation after castration

15. Healthy young animals should be acclimatised to the laboratory conditions 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. The treatment with initiation of dosing (on study) may commence as early as pnd 49 days of age, but preferably not later than pnd 60. This flexibility allows a laboratory to schedule the experimental work efficiently, and for different strains and sources of animals to achieve full preputial separation.

Housing and feeding conditions

16. Temperature in the experimental animal room should be 22 °C (±3°). The relative humidity should be 50 to 60%, but should not exceed limits of 30 to 70% except during room cleaning. Lighting should be artificial, the photoperiod being 12 hours light, 12 hours dark.

17. Laboratories participating in the validation should use the laboratory diet normally used in their chemical testing work. In previous phases, 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. Both diet and drinking water will be supplied *ad libitum.*

18. Animals should be caged in groups of no more than 3 similarly treated rats per cage, giving a minimum of 2 cages of 3 rats/cage per treatment group. Three animals or less per cage will avoid crowding and associated stress with animals of this age that may interfere with the hormonal control of the development of the sex accessory tissue. Individual housing of animals is permitted. Cages should be thoroughly cleaned to remove possible contaminants and arranged in such a way that possible effects due to cage placement are minimised.

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

Body Weight and the selection of animals for the study

20. Increasing differences in body weight may be a source of variability in the weight of tissues of interest within and among groups of animals. Variations in body weight should be both experimentally and statistically controlled, and the statistical analysis should be done both with and without body weight as a covariate. As toxicity may also impact the body weight, the body weight on the first day of administration can be used as the covariate in those cases where significant reductions in body weights have occurred.

21. Experimental control of body weight is accomplished in two steps. The first step involves selection of animals with relatively small variation in body weight for the study cohort from the larger population of animals that have been supplied. Unusually small or large animals should be avoided and should not be placed in the study cohort. A reasonable level of body weight variation within the study cohort should be tolerated. Here, ± 20% of the mean body weight for the cohort population is judged to be reasonable (e.g. 175g ± 35g). The second step involves the assignment of animals to different treatment groups (n = 6) by a randomised complete block approach. Under this approach animals are randomly assigned to treatment groups so that each group has the same mean and standard deviation in weight at the beginning of the study. The procedure used for block randomization should be recorded.

Non-routine health and safety requirements

22. The test substances are possible reproductive and developmental toxicants and, therefore, appropriate precautions should be taken to protect personnel during the validation work, e.g. necessary training, labeling and storage procedures, and protective handling procedures during dose preparation and dose administration.

23. Appropriate precautions such as wearing protective gloves, protective clothing and eye protection will be taken when handling the animals, diets, cages, and wastes (e.g. remaining test solutions, faeces, and carcasses). Waste disposal will be in accordance with good practice and existing regulations applicable to a given laboratory.

PROCEDURE – HERSHBERGER VALIDATION PHASE-3

24. The following procedure is based on the previous validation work in Phase-1 and Phase-2.

Reference substances, vehicle, and route of administration.

25. The reference androgen agonist will be TP. The reference androgen antagonist will be FLU.

26. All participating laboratories should use a vehicle, such as stripped corn oil, that is not easily disposed to potential microbial degradation of the vehicle or the reference and test substances. If the dosing samples are not made daily, care should be taken to preserve and to avoid contamination and spoilage of the samples.

27. The selected route of administration for the reference and coded test substances in Phase-3 are:

Testosterone propionate	subcutaneous injection
Flutamide	per os (oral gavage)
All coded test substances	per os (oral gavage)

Test substances in Phase-3 of the Hershberger validation

28. The test substances for Phase-3 with the exception of TP and FLU will be supplied coded to each laboratory. There will be a set of four coded samples for the agonist studies and six coded samples for the antagonist studies. Each sample will be administered to a group of animals (i.e., it will not be used

for a dose series, but administered as a single specified dose). A minimum of four laboratories will conduct each study.

Test groups in Phase-3 of the Hershberger validation

29. 6 animals of the same age and cohort will be used for the vehicle, TP, any other control group, and each coded treatment group.

30. In one set of studies, the response of the sex accessory tissues and glands to coded agonist samples will be studied. This work will involve four coded test groups having prepared doses of the test substances and one vehicle control group. The doses of positive agonist test substances will be the same as one or more of the doses used in Phase-2 to assess assay reproducibility. The negative test substances will be used to assess the bioassay's false positive rate with agonists. Laboratories may voluntarily choose to include a positive control of the using same TP dose that used by the laboratory in Phase-2 (0.2 mg/kg/d TP was used in seven laboratories and 0.4 mg/kg/d TP was used by nine other laboratories).

31. In a second set of studies, the response of the sex accessory tissues and glands to coded antagonist samples will be studied. This work will involve six coded test groups having prepared doses of the test substances with the coadministration of a selected dose of the reference agonist TP to each group. The reference TP doses will be the same as that used by the laboratory in Phase-2 (0.2 mg/kg/d TP was used in seven laboratories and 0.4 mg/kg/d TP was used by nine other laboratories). The doses of positive test antagonist substances will be the same as one or more of the doses used in Phase-2 to assess assay reproducibility. The negative test substances will be used to assess the bioassay's false positive rate with antagonists. Laboratories may voluntarily choose to include a vehicle control and a positive control of the same TP dose coadministered with 3 mg/kg/d FLU.

Doses in Phase-3 of the Hershberger validation

32. All participating laboratories will use the same dose levels in order to support comparisons of the assays reproducibility with previous and current work. The following table provides the requirements and includes both mandatory and possible voluntary groups.

	Agonist Coded Samples	Antagonist Coded Samples
Group A Vehicle Control	Vehicle	Vehicle – **Voluntary** in antagonists series
Group B Negative Treatment Control	Not applicable - provided by vehicle control (no additional group needed for agonist series)	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d)
Group C	Coded Agonist Test substance #1 [a]	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + Coded Antagonist Test substance #1 [a]
Group D	Coded Agonist Test substance #2 [a]	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + Coded Antagonist Test substance #2 [a]
Group E	Coded Agonist Test substance #3 [a]	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + Coded Antagonist Test substance #3 [a]
Group F	Coded Agonist Test substance #4 [a]	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + Coded Antagonist Test substance #4 [a]
Group G	Not applicable	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + Coded Antagonist Test substance #5 [a]
Group H	Not applicable	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + Coded Antagonist Test substance #6 [a]
Group I Positive Treatment Control	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) **Voluntary in agonist series**	TP dose used by the laboratory in Phase-2 (0.2 or 0.4 mg/kg/d) + FLU 3 mg/kg/d **Voluntary in antagonist series**

[a] The doses of each test substance will be prescribed in order to have comparable data for the analysis of variability among labs and for comparison to previous data.

Administration of doses

33. Subcutaneous injections will be on the dorsal surface of the animal after shaving or trimming of fur. Multiple injections sites may be used. The maximum limit on the volume administered per animal is approximately 0.5 ml/kg body weight per day.

34. Oral gavage will be the delivery of the test substance in vehicle by means such as intubation with an oral gavage syringe. The maximum limit on the volume administered per animal will be 5 ml/kg/day.

35. 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.

Good Laboratory Practice

36. Work should be conducted according to the principles of Good Laboratory Practice (OECD Good Laboratory Practice and Compliance Monitoring) (8). In particular, data should have a full audit trail and be retained on file. Data will be collected in a manner that will allow independent peer review and written records maintained.

OBSERVATIONS

Clinical observations

37. Animals will be evaluated at least once daily for mortality, morbidity, and signs of injury as well as general appearance and signs of toxicity. Any animals in poor health will be identified for further monitoring.

38. 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

39. Individual body weights will be recorded prior to start of treatment (to the nearest 1 g), on each day of administration period and prior to necropsy. Group means and standard deviations will be calculated.

40. Food consumption should be generally observed and any significant changes recorded.

Necropsy

41. Approximately 24 hours after the last administration of the test substance, the rats will be euthanized 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.

42. The order in which the animals are necropsied will be designed such that one or two animals from each of the groups (e.g., one per cage if there are three animals per cage) are necropsied to achieve a randomisation of the groups. In this way, all the animals in the same treatment group are not necropsied at once, and any variation in the procedure over time will not unduly impact any particular group.

43. The five sex accessory tissues (VP, SVCG, LABC, COWS, and GP) are mandatory measurements. The sex accessory 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.

44. Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Previous work has indicated a range of coefficient of variations that appears to differ based upon the proficiency of the laboratory. In a few cases during Phase-2, large differences in the absolute weights of the tissues such as the VP and COWS have been observed within a particular laboratory. 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 has been provided by the Lead Laboratory and was used in Phase-1 and Phase-2. This manual will remain the SOP reference for Phase-3. Careful training according to the SOP guide will minimise a potential source of variation in the study.

45. Each sex accessory tissues will be weighed without blotting to the nearest 0.1 mg, and the weights recorded for each identified animal.

46. 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 weighted to the nearest 0.1 g, and the paired kidneys and paired adrenals will be weighted to the nearest 0.1 mg. All weights will be recorded for each animal.

47. If the evaluation of each chemical requires necropsy of more animals than is reasonable for a single day, the starting date may be staggered on two consecutive days so that the necropsy can be staggered. 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. That is, each group should be split so that half of the animals are necropsied on each day in order to control variability among the groups.

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

REPORTING

Data

49. 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 summarised 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 the test number of animals found showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration and severity.

50. To assist data reporting and compilation, a standardised electronic spreadsheet will be used by participating laboratories to report and transmit data during the validation work to the Secretariat so that it may be easily exchanged and compiled with the Lead Laboratory and independent statisticians. This spreadsheet will be provided by the OECD Secretariat.

Test report

51. The test report must include the following information:

Laboratory identification

- Name of laboratory, location
- Principal investigator and other personnel and their roles in the study
- Dates study began and ended

Test substance:

- Physical nature and, where relevant, physicochemical properties;
- Identification data and source
- Purity

Vehicle identity and supplier:

Test animals and procedures:

- Species/strain used;
- Source or supplier of animals, including full address;
- Number, age and sex of animals;
- Housing conditions (temperature, lighting, and so on), diet used, lot of diet, source of diet, bedding and source of bedding;
- Caging conditions and number of animals per cage;
- Age at castration and time of acclimatisation 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 at the start of the study;
- Necropsy procedures, including means of exsanguination and any anesthesia; and

Results:

- Daily observations during administration, including:
  - Daily body weights (to the nearest 1 g),
  - Clinical signs (if any),
  - TP treatment (Yes or No),

  - Test substance treatment (Yes or No),
  - Dose level and volume administered each day,
  - Time of dosing each day, and
  - Notes on food consumption or measurement of actual food consumption each day.
- On the day of necropsy, individual necropsy data on each animal including absolute sex accessory tissue weights, liver and body weights including the following:
  - Date of necropsy,
  - Animal ID,
  - Home Cage Number or ID,
  - Prosector,
  - Time of day necropsy performed,
  - Animal age, and
  - Order of animal killing and dissection at necropsy,
- Weights of all five mandatory sex accessory tissues and glands.
  - Ventral prostate (fresh weight – to the nearest 0.1 mg),
  - Seminal vesicles plus coagulating glands, including fluid (fresh weight – paired, to nearest 0.1 mg),
  - Levator ani plus bulbocavernosus muscle (fresh weight - 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).
- Weights of optional tissues, if performed.
  - Liver (optional – to nearest 0.1 g),
  - Kidney (optional – paired, to nearest 0.1 mg), and
  - Adrenal (optional – paired, to nearest 0.1 mg).

- General remarks and comments

Discussion

Conclusions

Interpretation of results

52. Statistical comparisons will be made for changes in the weights of the different sex accessory tissues, the body weights, and those optional organs weighed. Statistical significance will be considered as present with $p < 0.05$. For androgen agonism, the test substance groups will be compared to the vehicle control. A statistically significant increase in tissue weight of the mandatory sex accessory tissues versus the same tissue in the vehicle control group will be considered consistent with a positive androgen agonist result. For androgen antagonism, the test substance with co-administered reference androgen groups will be compared to the reference androgen control (TP only group). A statistically significant decrease in tissue weight of the mandatory sex accessory tissues versus the same tissue in the positive control TP group will be considered consistent with a positive antagonist result. Statistically significant changes, positive or negative, in the tissues other than the mandatory sex accessory tissues will be noted and considered to be characteristic for the test substance, but not evidence for androgen agonism or antagonism.

LITERATURE

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

2. Korenchevsky, V, Dennison, M. (1935):
The assay of *trans*dehydroandrosterone and its effects on male and female gonadectomized rats.
Biochem. J. 29:1514-1522.

3. Wainman, P, Shipounoff, GC. (1941):
The effects of castration and testosterone propionate on the striated perineal musculature in the rat.
Endocrinology 29:975-978.

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

5. 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.

6. OECD (2002):
Final OECD report of the 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. October, 2002. Paris, France.

7. OECD (2003)
Draft OECD report of the work towards the validation of the rat Hershberger assay: Phase-2, Dose response studies with androgen agonists, androgen antagonists, and a 5α-reductase inhibitor. ENV/JM/TG/EDTA(2003)12. November, 2003. Paris, France.

8. OECD (1998):
OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, Organisation of Economic Co-operation and Development, Paris 1998. OECD Principles of Good Laboratory Practice (as revised in 1997).

ANNEX 3 − Report of the lead laboratory on experiments with agonists

Please note that information may be repeated as some of the tables from this annex have already been inserted in the report. However, the OECD Secretariat didn't want to change the document prepared by Dr Earl Gray.

Hershberger Interlaboratory Validation Study

Results of Phase III using coded chemicals.

Experiments with Agonists and Negative in Ten Laboratories

Prepared by Leon Earl Gray Jr

Endocrinology Branch, Reproductive Toxicology Division
NHEERL, ORD, USEPA
MD 72, 2525 Highway 54, RTP, NC 27711
gray.earl@epa.gov

Introduction

In our role as the lead laboratory in the OECD Hershberger Assay Validation we received data from Phase 3 experiments from Dr W Owens, acting on behalf of the OECD, and the laboratories for analysis and interpretation. The following are the results of these studies and our analysis of the data from the studies with androgen agonists and negatives. The other report this phase reported on the studies with antiandrogens and negatives.

Phase 3 Objectives

The main objective of this phase was to see if the ten laboratories would obtain results with coded samples that were comparable to the results obtained earlier with identical doses of the chemicals in earlier phases. In this regard, trenbolone was administered orally at 1.5 and 40 mg/kg/d, doses previously used in phase 2 (3 labs). We expected that the statistically significant results from the earlier studies would replicate in phase 3 with coded samples. Most laboratories also administered testosterone propionate (TP) by subcutaneous injection at 0.4 mg/kg/d as well. The phase 3 TP data were compared to results from phase 1 (16 labs). Corn oil (the vehicle) also was administered orally as the control in the agonist studies.

This phase also included dinitrophenol (10 mg/kg/d) and nonylphenol (150 mg/kg/d) which were expected from the literature to be negative as agonists in the Hershberger assay.

In phase 3, ten laboratories from Europe, the UK, the USA and Asia (Japan and Korea) received coded samples from OECD for evaluation in the standardized Hershberger assay for "androgenicity". Chemicals were administered orally at proscribed dosage levels concurrently to determine if they stimulated growth of any or all of the five androgen-dependent tissues (seminal vesicle (SV), ventral prostate (VP), levator ani plus bulbocavernosus (LABC), glans penis (GP) and Cowper's glands). In addition, body weight was monitored throughout the ten day experiment, and body, liver, kidney, and adrenal weights were taken at necropsy (in some but not all laboratories). Like may synthetic norandrogens, trenbolone is a selective androgen agonist, as compared to TP, producing more pronounced "anabolic" effects on muscles than "androgenic" effects on sex accessory tissues. For this reason, growth of the LABC is more sensitive to trenbolone that are the VP and SV.

The objective of the study was to demonstrate that ten laboratories could get the "correct" responses for the agonist and negative chemicals on all androgen-dependent tissues when the chemicals were provided to them as unknown coded samples.

Statistical Methods

Raw data were taken from submitted Microsoft Excel 2003 spreadsheets, converted to text files and modified for statistical analysis using WordPerfect 9 software. The text files were submitted electronically via FTP using the NHEERL intranet to an IBM mainframe computer using SAS proprietary Software Release 8.2 (TS2M0), licensed to US ENVIRONMENTAL PROTECTION AGENCY, Site 0045713001 running on IBM Model 9672 .

The results presented herein are one- way ANOVAs (main effect of the chemicals for each laboratory) and two-way ANOVAs (with laboratory and chemical as the two main effects) from PROC GLM.

Post hoc comparisons were generated using the LSMEANS option of PROC GLM which is a two tailed t-test. T-tests being appropriate for this study because we have a priori expectations about the effects of the different doses of the chemicals. In addition, as a screening

assay, false positives are more tolerable than false negatives that might result from using a more conservative post hoc test. It is also our objective to keep animal use to a minimum so using a t-test enables us to detect weaker effects with a sample size of six rats per group than does a more conservative post hoc test.

Means, standard errors (SE) and coefficients of variation (CV) were generated using PROC MEANS on SAS. The CV is the standard deviation divided by mean and provides an index of the dissection precision of the organ weight data in a laboratory or group. SAS ANOVAs also provides R-square (R2) values (R2 being the variation due to laboratory or chemicals divided by the total variation) which is an index of the robustness of the effects of the chemicals in each laboratory. The expectation is for CV values similar to those seen in Phase 1 and 2 of the validation effort and that all the laboratories would have similar R2 values.

SAS results were then transmitted by FTP to an external hard drive in my laboratory and converted from text files to WordPerfect files for printing. The results from these sheets were then entered into Microsoft Excel spreadsheets to summarize the results of the data analyses. Subsequently, the Excel sheets were printed for this report to pdf files using Adobe Acrobat 5.0 software.

Graphs of histograms of the means and standard errors of the mean for the weight data are presented individually for each laboratory and pooled over the laboratories. The graphs were prepared using Lotus Freelance Graphics 9.7 for Windows. Graphs were arranged by endpoint for the ten laboratories so the repeatability of the effects of the chemicals from lab-to-lab was easily apparent. This file was then also converted to a pdf file for inclusion in the report.

Results

One of the objectives of the interlaboratory study was to demonstrate that ten laboratories could get the "correct" responses for all the chemicals on all androgen-dependent tissues. As trenbolone at 1.5 and 40 mg/kg/d was used in phases 2 studies of this validation exercise we have specific expectations about the quantitative effects of these two coded treatment groups on body weight, liver weight and the five androgen-dependent organ weights. We expected and

observed the following effects:

1. Decreased body weight trenbolone at 40 mg/kg/d
3. Significant increases in all five androgen-dependent organ weights with TP at 0.4 mg/kg/d sc and oral trenbolone at 40 mg/kg/d and small increases in LABC weight at 1.5 mg/kg/d.

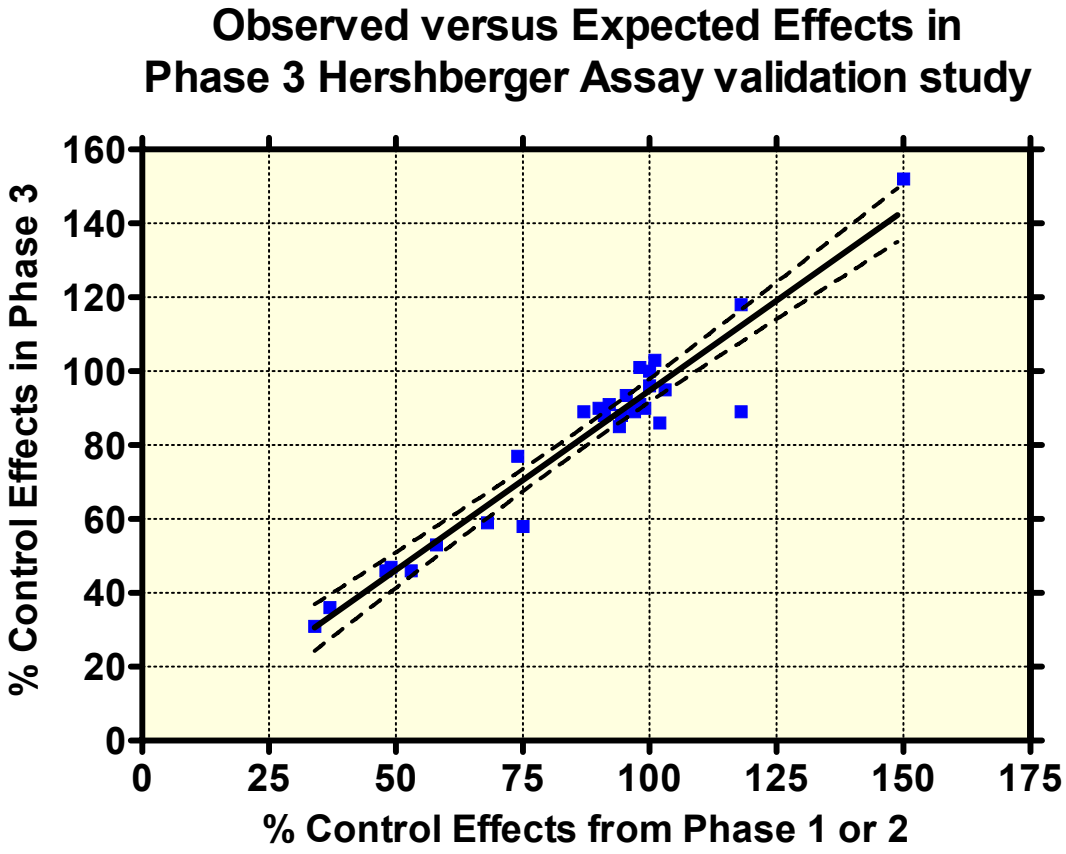
When all the effects of the two agonist treatment groups on the six endpoints (body, and five androgen-dependent tissues) were correlated with the effects seen previously in Phase 2 studies, a highly significant positive relationship was found having a correlation coefficient of r = .98. The expected and observed effects were calculated as percent of concurrent control values using the corn oil as control.

If one considers significant increase (based upon t-tests) in all 5 androgen-dependent tissues from ten laboratories, 17 beta trenbolone was positive in 96 % of the time (Appendix 1). At the low dose, as expected 1.5 mg/kg/d trenbolone did not elicit many positive responses (20% of the labs reported a stimulatory effect on the LABC) LABC weight being the only tissue weight increased significantly when the data from all ten labs was pooled. In phase 2, this dose also did not produce any significant effects. TP at 0.4 mg/kg/d sc was positive in all instances.

The “known negatives” nonylphenol at 150 mg/kg/d and dinitrophenol at 10 mg/kg/d did not stimulate growth of any of the androgen-dependent tissues in any laboratory, as expected.

Figure 1 below displays the consistency of the results of phase 3 with earlier results with the same treatments in earlier phases of the validation effort. In the figure, we compare the effects (as percent of concurrent control) observed in phase 2 for trenbolone at 40 mg/kg/d with the effects in phase 3 whereas the effects of TP at 0.4 mg/kg/d in phase 1 are compared to those seen herein in phase 3. Trenbolone is known to be antiandrogenic in vivo and in vitro, having a more pronounced “anabolic” effect on the LABC than an “androgenic” effect on the VP, SV, COWs and GP and this treatment was readily identified by every laboratory as being an androgen agonist when provided as coded samples to the ten laboratories. The data shown are

for body, ventral prostate, seminal vesicle, glans penis, Cowper's glands, or levator-ani bulbocavernosus muscle weights. We "expected" that the phase 2 effects would be very similar to those "observed" in this study in phases 1 or 2. The figure does not distinguish one endpoint from another with each the observed and expected effect for each endpoint in each treatment being represented by a single blue square. The correlation coefficient for the data is r = 0.98. It is evident that the results of phase 3 are very consistent with the effects in phase 2.



Analyses of variance for the five androgen-dependent tissues revealed that the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in phase 3 were all very statistically significant. Importantly, the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by relatively small F-values for the chemical by laboratory interaction as compare to the size to the two main effects (Appendix 2 and 3). The group means and SEs of the laboratories and the means and SEs of the individual laboratories are presented in Appendix 4 to show the consistency of the antiandrogenic effects quantitatively. The uniformity of the response among the laboratories to each chemical is even more obvious from viewing histograms of the data in Appendix 5. Here the data are presented by endpoint and for the ten individual laboratories. This displays the remarkably uniform profile of effects among laboratories.

The final set of tables summarizes the results by endpoint (Appendix 6). Here it is easy to compare R2 and CV values from laboratory to laboratory for each of the five androgen-dependent organ weights and body weight at necropsy. The size of the R2 describes the strength of association of the chemical effects versus vehicle control and the CV describes the precision (variability) of the data.

If the all laboratories had the same CV on an endpoint then this would indicate that they were all dissecting the tissues with similar precision. A high CV in a laboratory versus the other laboratories might indicate a problem. In a few cases the CVs seemed to be larger than expected, indicating that the labs had some problems with particular endpoints for specific chemicals. For example, the high CV values for the SV, VP and COWs seen in Labs 5 and 10 arose from a single animal in each lab from the trenbolone 40 mg/kg/d dose group. In these cases the androgenic effect of the chemical treatment was still apparent. However, lab 7 data displayed unusually high variability in the control (vehicle-treated) group for the VP and Cowper's glands and this resulted in the only two false negatives in the entire data set. These results indicate that the laboratory had some difficulty dissecting unstimulated small tissues (most likely) in the controls, errors in dosing, incomplete castration or some other technical problem.

In summary, analysis of the results of phase three of the Hershberger assay interlaboratory study using coded samples of known dosage levels of agonists or negatives produced appropriate and consistent responses among all the laboratories in 90% of the cases (5 androgen-dependent tissues with four treatment groups) . None of the known positives were undetected in any laboratory.

Appendix page 1.

Summary of effects for Trenbolone at 40 and 1.5 mg/kg/d. These groups are listed chemical-by-chemical to display the agonist properties of this synthetic steroid. Blue shaded cells differ significantly from control (TP-treated) by $p < 0.01$ whereas yellow shaded cell differ by $p < 0.05$ but t-test from the LSMEANS. The pink color indicates that an effect that should have been significant was not affected, i.e. a "false negative".

p < 0.05 p <0.01 **False negative**

Lab	Chemical	body	ventral prostate	seminal vesicle	LABC	Glans Penis	Cowper's glands	liver
1	TREN 40	209	34.3	86.4	359	57	14.1	7.7
2	TREN 40	275	72	262	371	72	21.4	12
3	TREN 40	209	29.6	108	359	52.9	10.3	10.4
5	TREN 40	215	74.2	169	309	62.9 *	13.4	8.8
7	TREN 40	239	70	144	447	90	12	11.9
8	TREN 40	307	35	144	548	72.3	19.7	11.9
9	TREN 40	290	30.3	130	246	85	10.4	16.4
10	TREN 40	238	58.3	136	317	53.7	14.4	10.7
11	TREN 40	262	33.4	179	427	58.5	10.1	11.5
13	TREN 40	265	43.7	165	452	72.9	18.1	12.2

* GP was signicantly stimulated in this lab as indicated by the fact that these males all displayed PPS and the controls did not

94 % postive/correct responses for androgenicity
6 % False negatives

Lab	Chemical	body	ventral prostate	seminal vesicle	LABC	Glans Penis	Cowper's glands	liver
1	TREN 1.5	224	20	39	180	44	8.2	7.3
2	TREN 1.5	318	15.3	58.4	149	50.6	10.4	12.3
3	TREN 1.5	236	15.1	35.3	172	40.4	6.1	10.8
5	TREN 1.5	230	15	26.9	106	no data	3.5	8.3
7	TREN 1.5	284	48.5	54.4	214	67.4	3.2	13.2
8	TREN 1.5	327	22.4	70.2	295	51	11	11.6
9	TREN 1.5	320	20.8	72.5	177	65.1	7	15.9
10	TREN 1.5	250	14.9	29	141	32.6	4.5	9.7
11	TREN 1.5	282	18.4	41.6	216	45.5	5.8	11.5
13	TREN 1.5	301	26.2	61.2	221	52	8.8	13

LABC by lab is 20% Positive for Androgenicity

Appendix 2.

Summary of the analyses of variance for the five androgen-dependent tissues showing the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in phase 3. Note that the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by relatively small F-values for the chemical by laboratory interaction as compare to the size to the two main effects. There are statistically significant lab effects in each ANOVA. This was expected and also seen in phases 1 and 2 of the validation studies. These effects arise because the laboratories used several different rat strains and even within strain the animal body and organ weights would be expected to vary. In the Hershberger assay protocol, factors that we expected to alter the ability to detect the effects of androgens and antiandrogens on the five androgen-dependent tissues were standardized, but in order to retain flexibility in the protocol so that laboratories could work with animals and conditions that they were familiar with other factors like strain and diet were not standardized because they did not compromise the ability to detect effects. A lab effect is not problematic, whereas a large lab by treatment interaction would be because this would indicate chemicals were not acting similarly in all the labs.

RESULTS OF ANOVAS FOR LAB AND LAB BY TREATMENT INTERACTION

no tp=.4, not an unknown

	All Labs			
	Effect	**df**	**F**	**p-value**
Body Weight	**LAB**	9	184	0.0001
	CHEMICAL	4	67	0.0001
	INTERACTION	36	1	>.40
CV=5.6%				
	Effect	**df**	**F**	**p-value**
SV	**LAB**	9	5.1	0.0001
	CHEMICAL	4	97	0.0001
	INTERACTION	36	2.2	0.0003
CV=54%				
	Effect	**df**	**F**	**p-value**
LABC	**LAB**	9	61	0.0001
	CHEMICAL	4	377	0.0001
	INTERACTION	36	3.1	0.0001
CV=18%				
	Effect	**df**	**F**	**p-value**
VP	**LAB**	9	7.4	0.0001
	CHEMICAL	4	40.2	0.0001
	INTERACTION	36	2.2	0.0003
CV=59%				
	Effect	**df**	**F**	**p-value**
GP	**LAB**	8	110	0.0001
no lab 5	**CHEMICAL**	4	92	0.0001
	INTERACTION	32	1.3	>.11
CV=12.1%				
	Effect	**df**	**F**	**p-value**
COWPER	**LAB**	9	8.2	0.0001
	CHEMICAL	4	63	0.0001
	INTERACTION	36	2.2	0.0002
CV=44%				

Appendix 3.

This table displays the overall effects of the "unknown" coded chemical treatments on organ weights in phase 3. The data are the means from all ten laboratories and the results are from the LSMEANS t-tests from the ANOVAs shown in appendix 2.

		ALL TEN LABS	p-value
BODY WEIGHT			
	OIL	279	
	DNP 10	278	
	NP150	268	
	TREN 1.5	277	
	TREN 40	251	0.0001
SV			
	OIL	47	
	DNP 10	45.6	
	NP150	47.1	
	TREN 1.5	48.8	
	TREN 40	152.5	0.0001
LABC			
	OIL	169	
	DNP 10	169	
	NP150	159	
	TREN 1.5	187	0.011
	TREN 40	383	0.0001
VP			
	OIL	21.4	
	DNP 10	18.3	
	NP150	19.7	
	TREN 1.5	21.6	
	TREN 40	48.1	0.0001
GLANS PENIS 9 labs			
	OIL	49.2	
	DNP 10	49.2	
	NP150	49.5	
	TREN 1.5	49.8	
	TREN 40	68.2	0.0001
COWPER'S			
	OIL	6.3	
	DNP 10	6	
	NP150	6.1	
	TREN 1.5	6.8	
	TREN 40	14.4	0.0001
Liver			
	OIL	11	
	DNP 10	11	
	NP150	11.8	0.0018
	TREN 1.5	11.4	
	TREN 40	11.3	
Kidney			
	OIL	1970	
	DNP 10	1967	
	NP150	2192	0.0001
	TREN 1.5	2029	
	TREN 40	2003	
Adrenals			
	OIL	62	
	DNP 10	63	
	NP150	64.2	
	TREN 1.5	58.7	
	TREN 40	54.2	0.0001

Appendix 4.

Means and standard errors (SEs) of the androgen-dependent tissues presented for the ten individual laboratories. Blue shaded cells differ significantly from control (TP-treated) by $p < 0.01$ whereas yellow shaded cell differ by $p < 0.05$ but t-test from the LSMEANS. The pink color indicates effects that were increased, rather than decreased in an antiandrogenic manner.

Hershberger assay Phase 3 Androgens
values are means and SEs

p < 0.05 p <0.01

All labs	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	279	5	21.4	1.7	47	2.2	169	7.4	49.2	1.8	6.3	0.42	11.2	0.3	62	1.5	1970	46
DNP 10	278	4.8	18.3	0.8	45.6	2	169	7	49.2	1.9	6	0.34	11	0.32	63	1.7	1967	44
NP150	268	5.3	19.7	1.5	47.1	2.3	159	6	49.5	1.8	6.1	0.3	11.8	0.36	64.2	1.8	2192	65
TREN 1.5	277	5.3	21.2	1.7	48.7	2.5	187	8	49.5	1.5	6.9	0.46	11.3	0.35	59	1.7	2024	51
TREN 40	251	4.6	48.1	4.2	152	11	383	13	68	2	14.4	0.9	11.3	0.32	54.2	1.9	2003	47
TP.4	286	5.8	138	10	341	20	437	17	81.2	2.6	30	1.3	12	0.4	54.3	1.9	2128	58

Lab 1	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	221	2.6	18.3	2	37.4	3.1	152	11	42.8	1.6	6.6	0.8	6.5	0.24	65.5	2.6	1632	57
DNP 10	221	2.1	19.1	1.8	37.2	2.6	176	12	43.1	1.6	9.5	1.1	6.4	0.04	71.4	4.7	1690	22.5
NP150	212	5	19.2	2	40	3.8	161	12	43.2	1.7	8.7	0.8	7.1	0.24	78	7	1855	95
TREN 1.5	224	2.5	20	1.3	39	3.8	180	13	44	1.1	8.2	1	7.3	0.4	72.2	4.8	1879	131
TREN 40	209	4.8	34.3	4	86.4	5.8	359	30	57	3.6	14.1	1.4	7.7	0.36	74	2.6	1818	71
TP.4	223	3.9	106	6	205	16	369	17	65	2.3	26	1.5	6.6	0.32	66.2	3.4	1769	60

Lab 2	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	317	4.7	16	2.1	40	6.9	92.1	8.2	44.2	2.9	3.6	0.8	11.1	0.3	69.3	4.6	2012	66
DNP 10	315	5.6	15.2	2.5	46.6	4.5	95	6.7	43.8	2.4	5.1	0.9	11.3	0.4	72.3	5	2054	39
NP150	303	6	14.2	2.7	40.3	2.5	92.2	10.3	41.9	1.9	4.6	0.33	12	0.5	71.8	2.7	2529	258
TREN 1.5	318	8.8	15.3	2.9	58.4	4.7	149	16.8	50.6	2.6	10.4	1.6	12.3	0.6	64.8	2.2	2086	68
TREN 40	275	6.5	72	9.5	262	32	371	19	72	2.9	21.4	4.2	12	0.6	50.1	5.1	2039	63
TP.4	not done																	

Lab 3	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	237	8.3	18.4	2	41.5	2.3	179	13	39.2	1.6	6.1	1.5	10.6	0.32
DNP 10	254	8.5	20.6	4.6	37.9	4.7	174	10.5	44	4.5	6	0.9	11.4	0.6
NP150	225	11	14.3	1.7	38.6	3	160	16	43.2	1.8	5.5	1	10.5	0.48
TREN 1.5	236	4.8	15.1	2	35.3	4.3	172	4.6	40.4	1.3	6.1	0.8	10.8	0.34
TREN 40	209	6.7	29.6	4.3	108	8.7	359	33	52.9	0.9	10.3	2.3	10.4	0.43
TP.4	248	4.8	66	9	178	23	356	34	67.4	3.2	16.1	2	10.96	0.16

Lab 5	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	240	8	15.1	0.7	31.1	2.8	112	12	no data		3.5	0.51	9.3	0.45	61.3	4.7	1526	80
DNP 10	236	7.3	15.3	2.7	30	2.1	102	11	no data		3.5	0.18	8.7	0.33	57.2	4.7	1506	47
NP150	223	7	13.4	2.2	30.7	4.9	100	4.9	no data		3.7	0.46	9.7	0.44	58.2	3.6	1709	47
TREN 1.5	230	6.1	15	1.7	26.9	2.8	106	6.5	no data		3.5	0.35	8.3	0.4	51.6	2.3	1486	65
TREN 40	215	4	74.2	19	169	53	309	28	62.9	13	13.4	3.5	8.8	0.34	49	2.9	1507	51
TP.4	243	5.6	126	11	236	15	298	18	67.5	2.3	22.5	1.4	9.6	0.43	48.3	3.6	1655	68

Lab 7	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	275	8	48.6	11	58.6	5.8	233	15	66.5	2.3	9.8	1.7	13.2	0.4	54.5	3.9	2128	87
DNP 10	269	6	20.5	3.3	60.5	3.8	215	12	66.3	6	4.8	0.9	13	0.46	57.7	3.3	1979	32
NP150	265	4	44.3	8.4	65.1	7	198	16	72.8	4	7	1.4	13.2	0.47	57.5	4	2430	188
TREN 1.5	284	4	48.5	14	54.4	3	214	17	67.4	1.5	3.2	1.1	13.2	0.5	50.7	3.9	2254	115
TREN 40	239	5	70	18	144	12	447	21	90	5	12	1.5	11.9	0.6	43.8	2.6	1962	53
TP.4	300	3.8	267	55	486	55	622	30	105	2.3	39.6	1.8	15.2	0.7	45.2	3	2281	37

Lab 8	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	339	3.5	20.6	2	57	4	270	13	53.6	2.5	8.2	2.1	11.9	0.3	58	2.1	2110	39
DNP 10	323	5.8	18	1.2	55.2	3.7	264	11	54.1	2	8.2	1	11.3	0.38	57.9	3	1993	44
NP150	324	3.4	17.5	1.8	63.6	4	240	7.6	53.1	2.1	7.7	0.8	12.2	0.6	61.2	2.3	2095	59
TREN 1.5	327	8	22.4	1.1	70.2	7	295	30.6	51	3	11	2	11.6	0.38	51.8	1.5	2012	86
TREN 40	307	4.7	35	1.5	144	11	548	42	72.3	1.9	19.7	2.1	11.9	0.35	54.2	2	2167	74
TP.4	338	3.6	128	7.8	385	23	530	3.6	77	3.6	30.6	3	12.4	0.3	51.5	3.3	2252	96

Lab 9	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	314	1.7	22	2.6	73	4.6	162	10	70.2	2.8	7.1	0.4	14.9	0.31	66.3	3.7	2271	54
DNP 10	314	5.7	19	1.9	66	2.3	154	4.8	66.2	2	6.6	0.5	14.8	0.4	63.1	3.4	2261	38
NP150	309	11	17	2.2	67.7	8	156	7.6	64	1.3	6.6	0.6	16.9	0.5	66.2	1.9	2380	91
TREN 1.5	320	6	20.8	1.5	72.5	6.7	177	7.5	65.1	2.6	7	0.6	15.9	0.85	69.5	1.6	2300	48
TREN 40	290	7	30.3	1.7	130	6.4	246	17	85	4.4	10.4	0.8	16.4	0.46	58.3	5.5	2302	79
TP.4	339	4	146	12	534	22	377	12	114	2.5	41.2	1.8	16.7	0.16	63.9	2.8	2373	76

Lab 10	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	263	7	16.8	0.4	27.9	2.3	137	9	29.5	2.3	4.1	0.5	10	0.5
DNP 10	252	5	15.6	1.4	25.7	2	128	4.5	28.3	2.9	3.9	0.6	9.6	0.2
NP150	251	5	17	1.3	26.5	0.8	128	7.9	29.9	1.2	4.4	0.6	10.1	0.5
TREN 1.5	250	5.3	14.9	0.5	29	1.4	141	6.2	32.6	2.1	4.5	0.4	9.7	0.36
TREN 40	238	6.3	58.3	24	136	75	317	21	53.7	4.6	14.4	4.2	10.7	0.48
TP.4	266	5.2	93.5	4.5	190	7.8	312	11	64.4	2.4	20.4	1.4	10.1	0.3

Lab 11	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	281	3.2	16	2.1	42	5.9	164	16	44.1	1.8	5.8	0.5	11.1	0.22
DNP 10	285	5	15.8	2.7	38.8	5	190	12	41.8	0.9	4.4	0.7	10.7	0.43
NP150	276	2.6	19.8	2.1	40.9	4.6	179	9.6	42.9	0.9	5.6	0.5	12.1	0.35
TREN 1.5	282	6	18.4	1.3	41.6	4.7	216	7.1	45.5	1	5.8	0.7	11.5	0.36
TREN 40	262	2.9	33.4	2.6	179	25	427	19	58.5	1.5	10.1	1.1	11.5	0.25
TP.4	306	6.1	122	10.4	420	13.1	527	9.6	73.9	1.6	34.4	3.3	12.8	0.49

Lab 13	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidney	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
OIL	303	6.3	22	1.3	61.2	2.4	191	6.5	53	3.3	8.5	0.9	12.9	0.38	58.8	3.8	2110	29
DNP 10	308	6.1	24	0.7	58.4	3.4	191	2.7	54.8	2.3	8	0.5	13	0.52	61.5	4.3	2290	66
NP150	297	7.2	20.6	0.6	58.1	2.8	179	10	54.6	2.2	7.4	0.7	13.9	0.78	57	3.8	2344	56
TREN 1.5	301	6.1	26.2	1.6	61.2	4.5	221	14.6	52	1	8.8	1	13	0.54	50.2	2.1	2189	78
TREN 40	265	11	43.7	4.7	165	15	452	14	72.9	1.3	18.1	2.1	12.2	0.64	49.9	3.2	2229	92
TP.4	320	9	187	20	431	22	544	34	95.1	3.3	37.1	2.7	13.6	0.83	50.9	4.8	2435	99

Appendix 5.

Histograms of the means and SEs prepared with Lotus Freelance are provided here by endpoint for the two groups of laboratories and for the ten individual laboratories. The large main bar with accompanying value is the mean and the small bar stacked on top is the standard error of the meas. These graphs display the remarkably uniform profile of effects among laboratories.

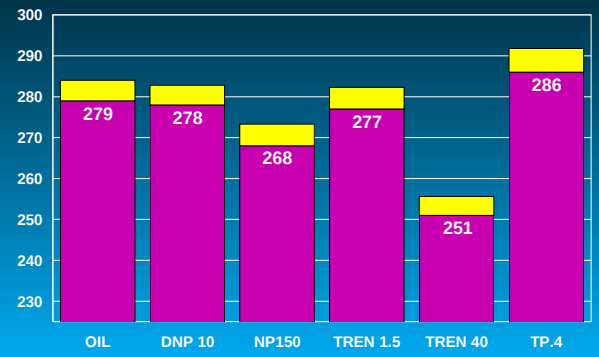
June 2005

prepared by LE Gray Jr

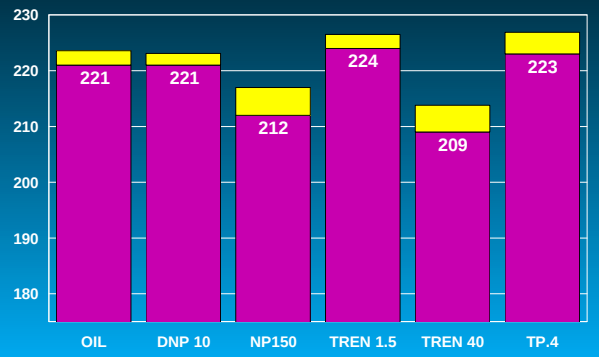
Hershberger Assay phase 3 Agonists and negatives

- Body Weight effects

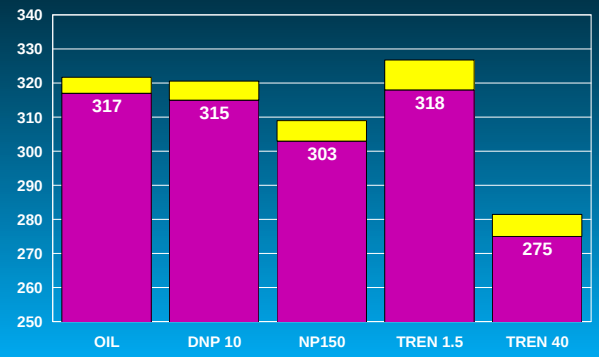
Hershberger assay: Agonist phase 3. All ten laboratories Body Weight Effects



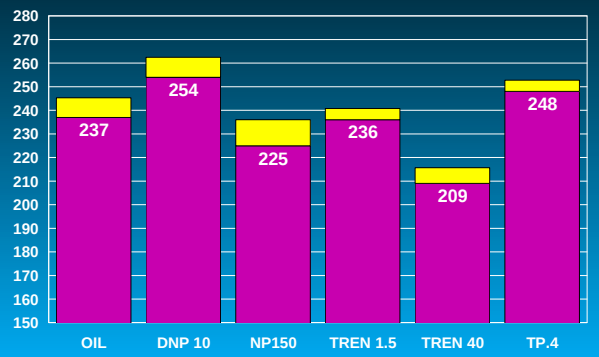
Hershberger assay: Agonist phase 3. Laboratory 1 Body Weight Effects



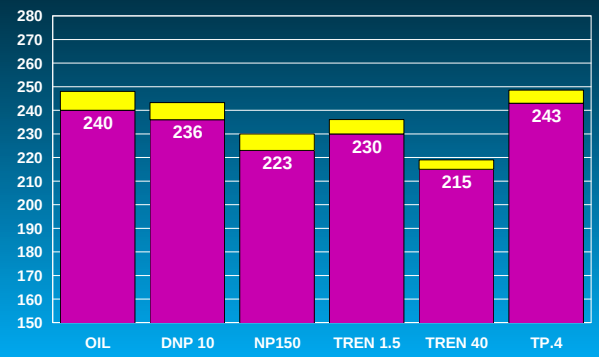
Hershberger assay: Agonist phase 3. Laboratory 2 Body Weight Effects



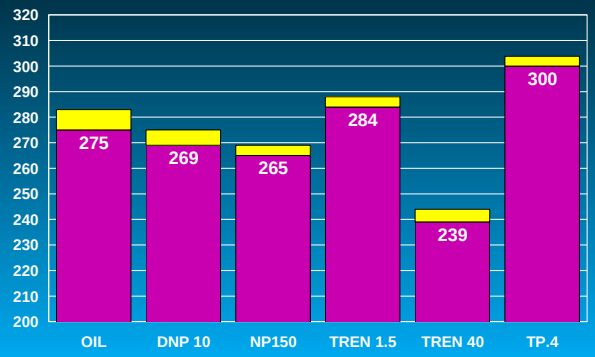
Hershberger assay: Agonist phase 3. Laboratory 3 Body Weight Effects



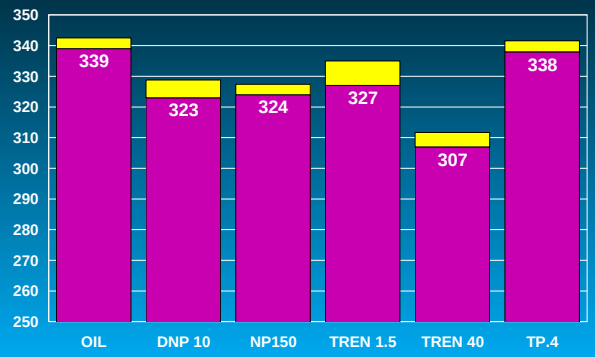
Hershberger assay: Agonist phase 3. Laboratory 5 Body Weight Effects



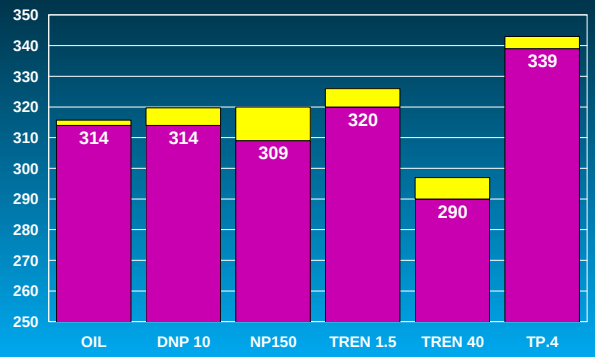
Hershberger assay: Agonist phase 3. Laboratory 7 Body Weight Effects



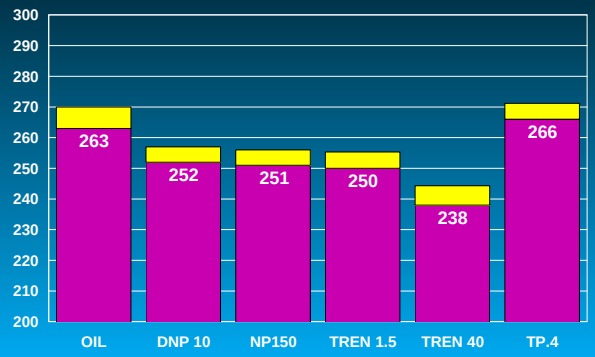
Hershberger assay: Agonist phase 3. Laboratory 8 Body Weight Effects



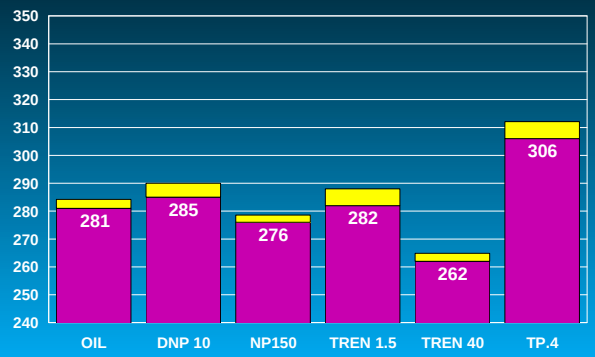
Hershberger assay: Agonist phase 3. Laboratory 9 Body Weight Effects



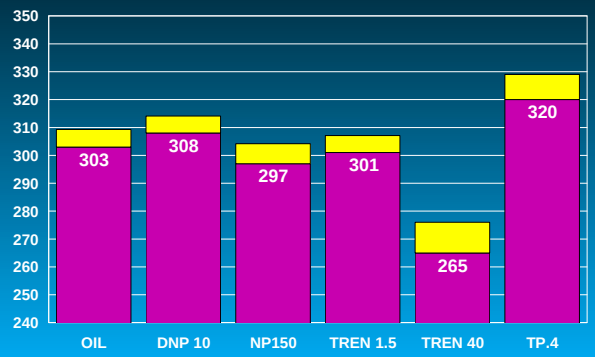
Hershberger assay: Agonist phase 3. Laboratory 10 Body Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11 Body Weight Effects

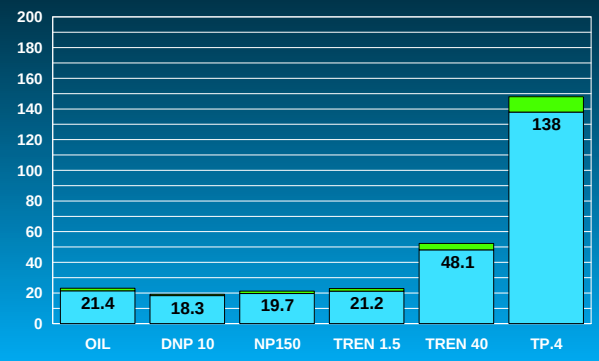


Hershberger assay: Agonist phase 3. Laboratory 13 Body Weight Effects

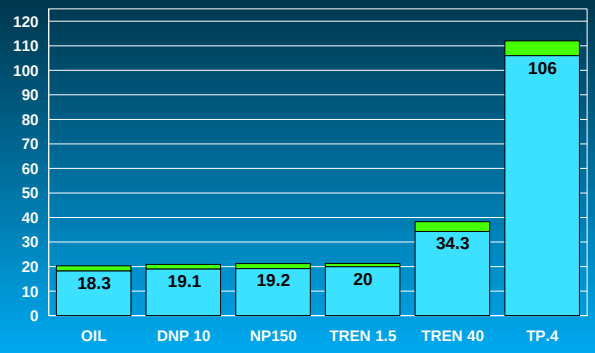


- Ventral Prostate Weight effects

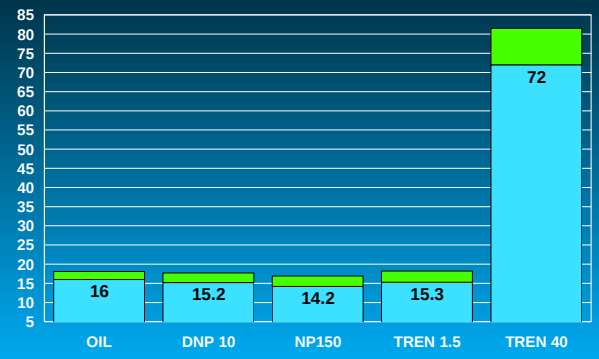
Hershberger assay: Agonist phase 3. All ten laboratories Ventral prostate Weight Effects



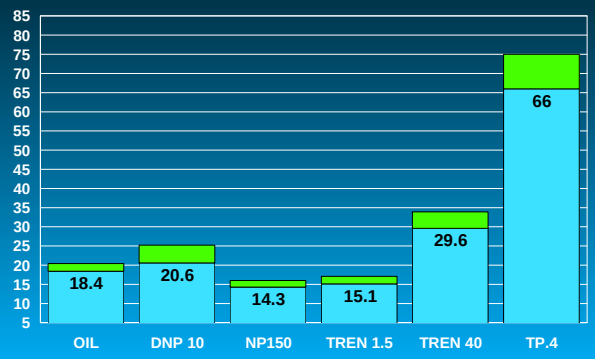
Hershberger assay: Agonist phase 3. Laboratory 1 Ventral Prostate Weight Effects



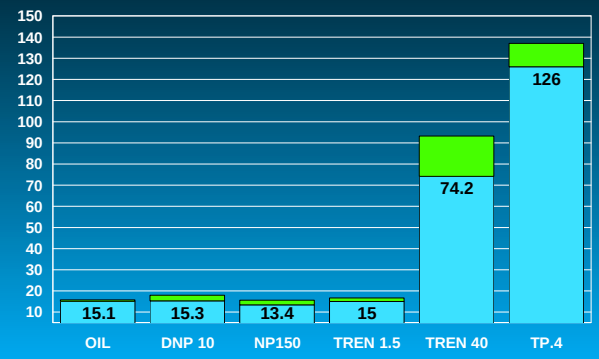
Hershberger assay: Agonist phase 3. Laboratory 2 Ventral Prostate Weight Effects



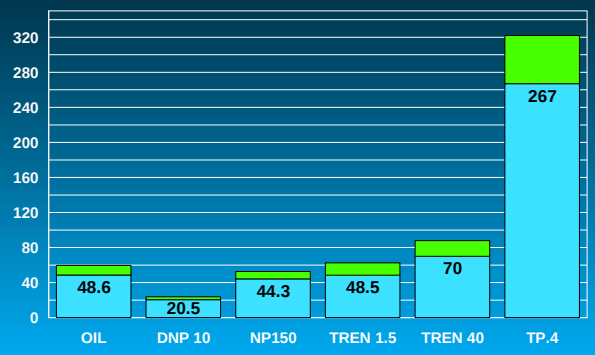
Hershberger assay: Agonist phase 3. Laboratory 3 Ventral Prostate Weight Effects



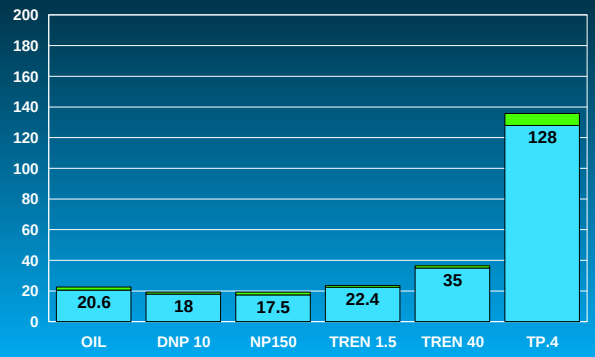
Hershberger assay: Agonist phase 3. Laboratory 5 Ventral Prostate Weight Effects



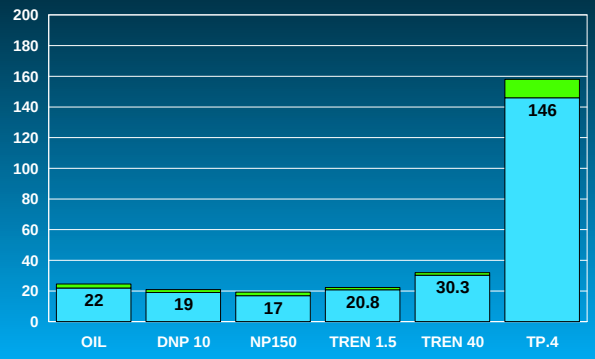
Hershberger assay: Agonist phase 3. Laboratory 7
Ventral Prostate Weight Effects



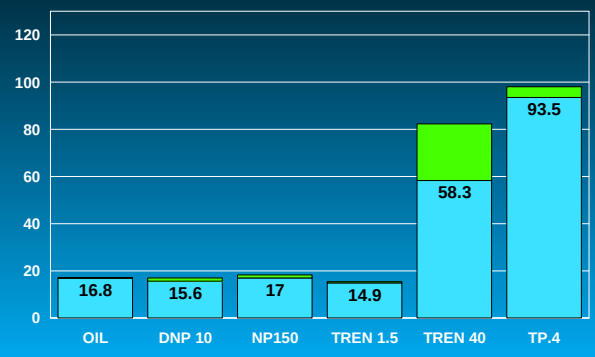
Hershberger assay: Agonist phase 3. Laboratory 8
Ventral Prostate Weight Effects



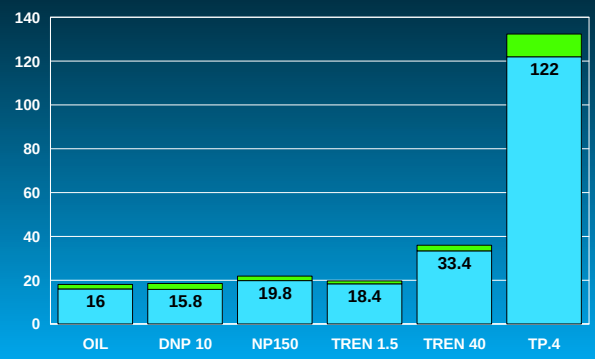
Hershberger assay: Agonist phase 3. Laboratory 9
Ventral Prostate Weight Effects



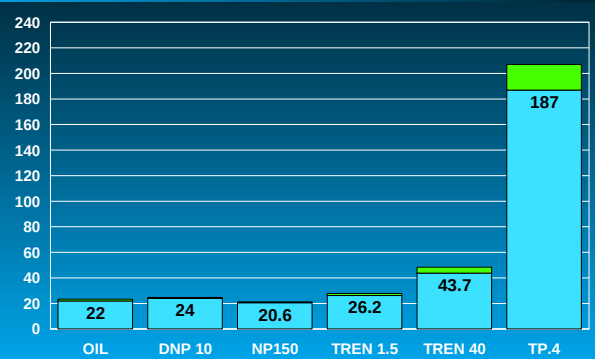
Hershberger assay: Agonist phase 3. Laboratory 10
Ventral Prostate Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11
Ventral Prostate Weight Effects

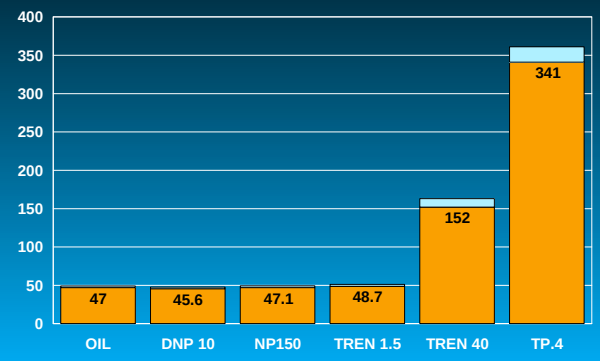


Hershberger assay: Agonist phase 3. Laboratory 13
Ventral Prostate Weight Effects

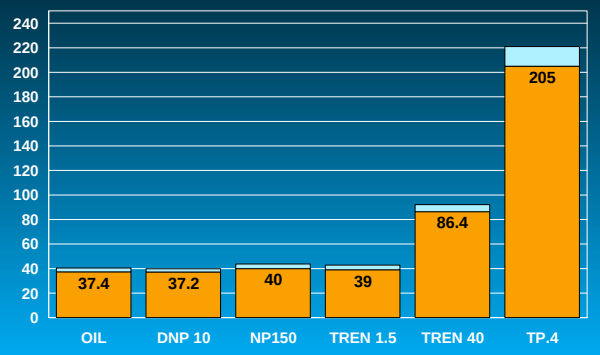


- Seminal Vesicle Weight effects

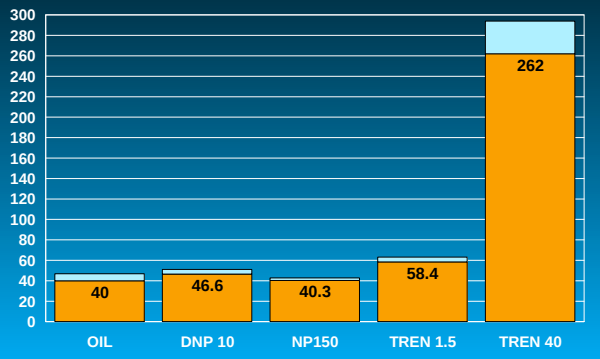
Hershberger assay: Agonist phase 3. All ten laboratories Seminal Vesicle Weight Effects



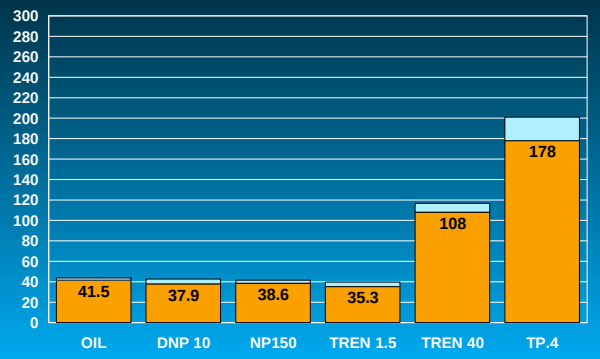
Hershberger assay: Agonist phase 3. Laboratory 1 Seminal Vesicle Weight Effects



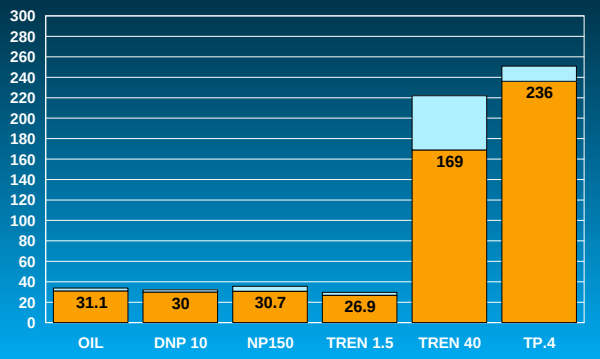
Hershberger assay: Agonist phase 3. Laboratory 2 Seminal Vesicle Weight Effects



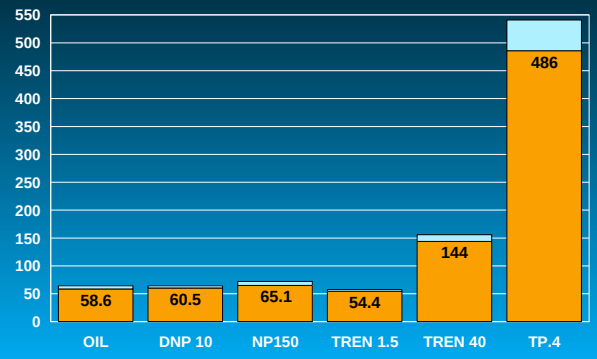
Hershberger assay: Agonist phase 3. Laboratory 3 Seminal Vesicle Weight Effects



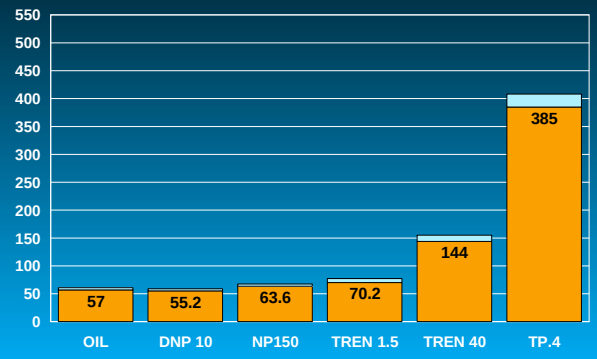
Hershberger assay: Agonist phase 3. Laboratory 5 Seminal Vesicle Weight Effects



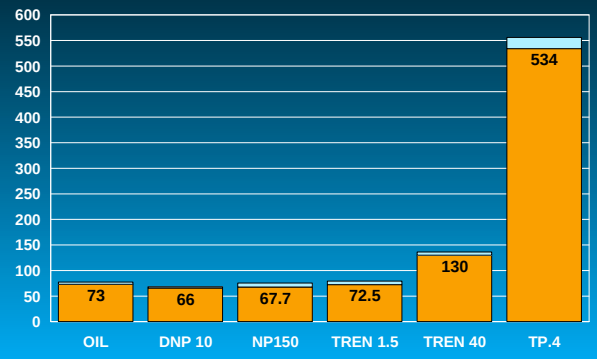
Hershberger assay: Agonist phase 3. Laboratory 7
Seminal Vesicle Weight Effects



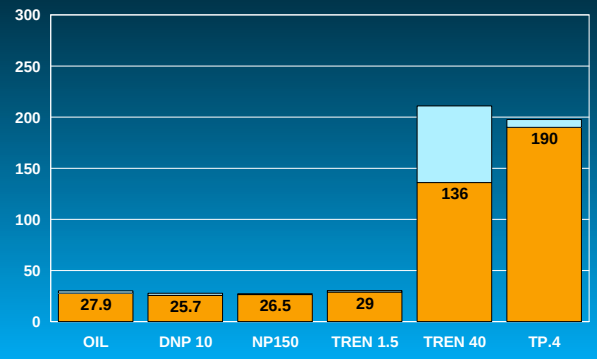
Hershberger assay: Agonist phase 3. Laboratory 8
Seminal Vesicle Weight Effects



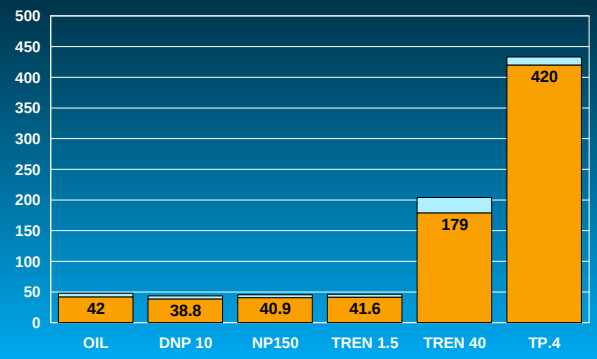
Hershberger assay: Agonist phase 3. Laboratory 9
Seminal Vesicle Weight Effects



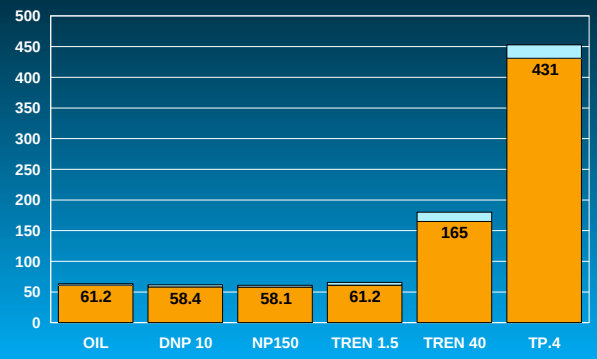
Hershberger assay: Agonist phase 3. Laboratory 10
Seminal Vesicle Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11
Seminal Vesicle Weight Effects

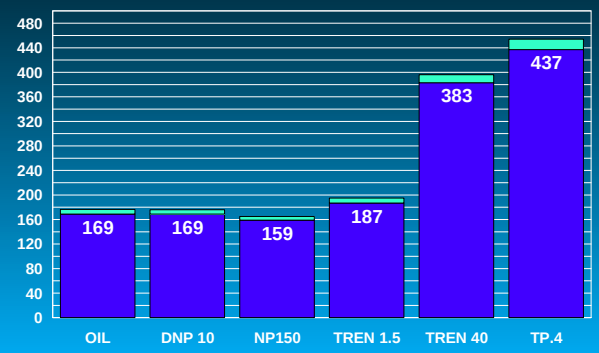


Hershberger assay: Agonist phase 3. Laboratory 13
Seminal Vesicle Weight Effects

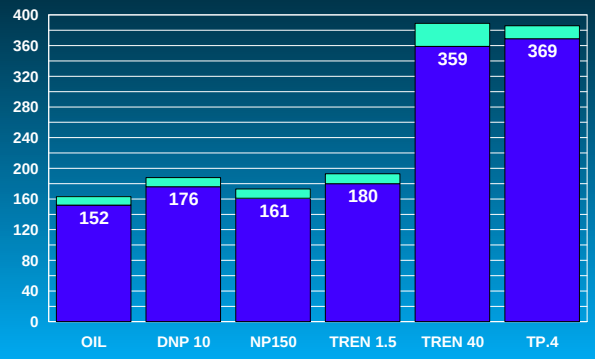


- LABC Weight effects

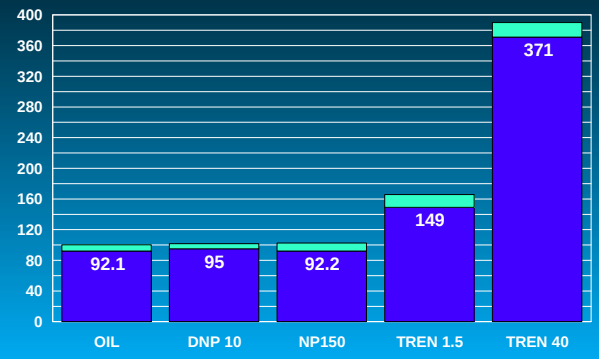
Hershberger assay: Agonist phase 3. All ten laboratories LABC Weight Effects



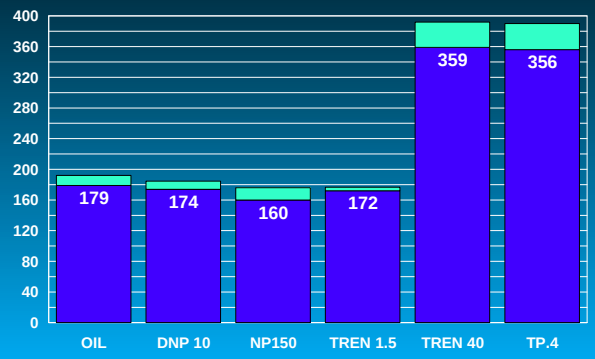
Hershberger assay: Agonist phase 3. Laboratory 1 LABC Weight Effects



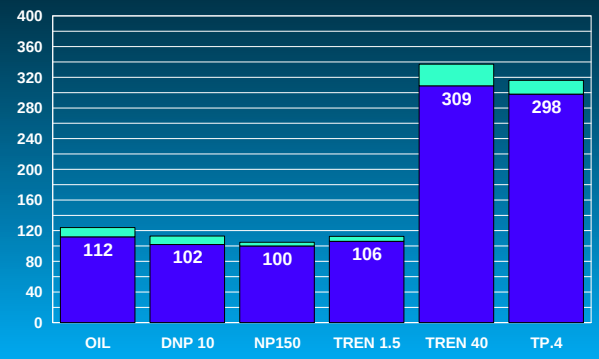
Hershberger assay: Agonist phase 3. Laboratory 2 LABC Weight Effects



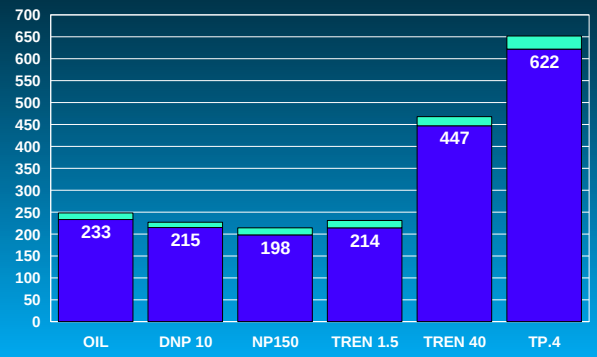
Hershberger assay: Agonist phase 3. Laboratory 3 LABC Weight Effects



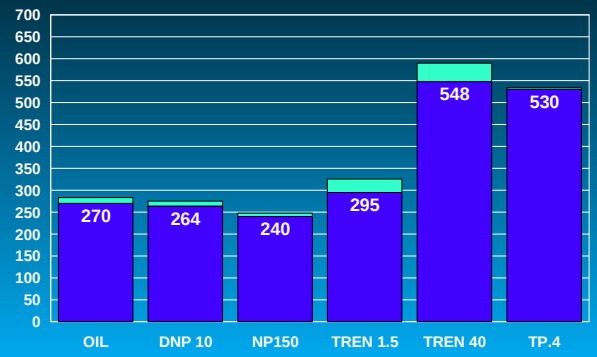
Hershberger assay: Agonist phase 3. Laboratory 5 LABC Weight Effects



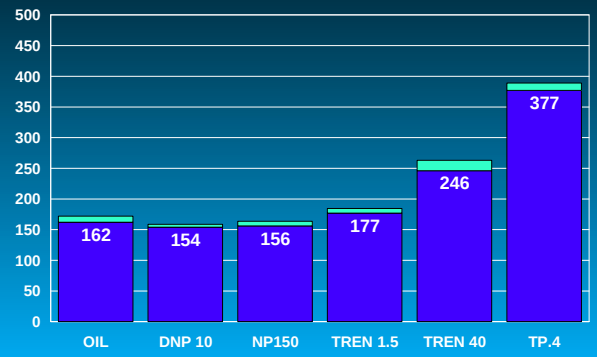
Hershberger assay: Agonist phase 3. Laboratory 7 LABC Weight Effects



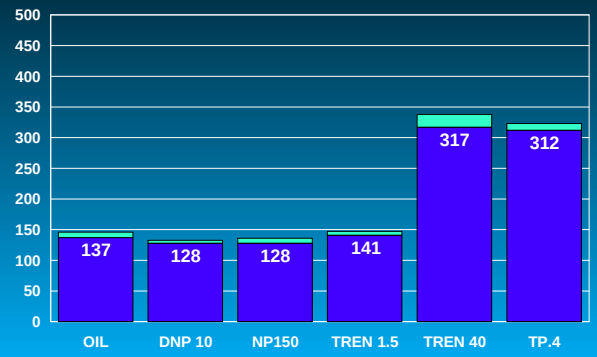
Hershberger assay: Agonist phase 3. Laboratory 8 LABC Weight Effects



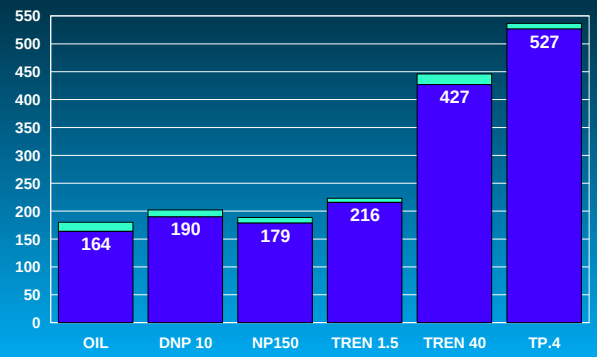
Hershberger assay: Agonist phase 3. Laboratory 9 LABC Weight Effects



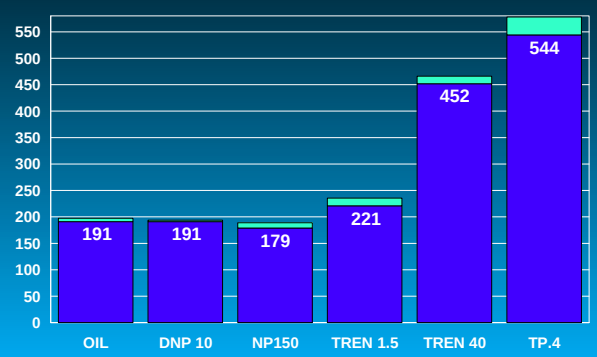
Hershberger assay: Agonist phase 3. Laboratory 10 LABC Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11 LABC Weight Effects

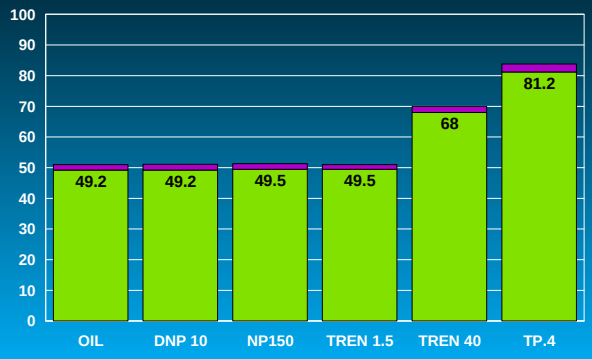


Hershberger assay: Agonist phase 3. Laboratory 13 LABC Weight Effects

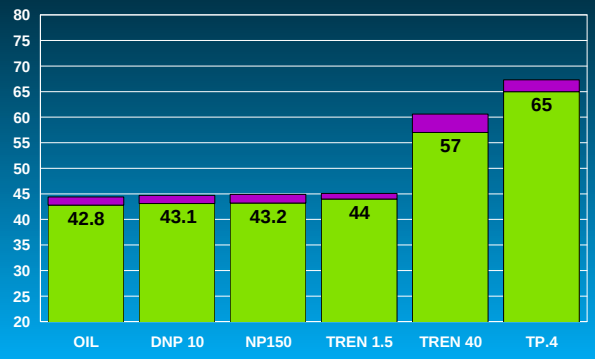


- Glans Penis Weight effects

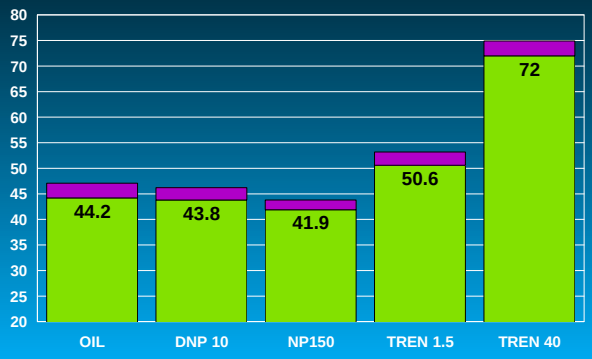
Hershberger assay: Agonist phase 3. All ten laboratories Glans penis Weight Effects



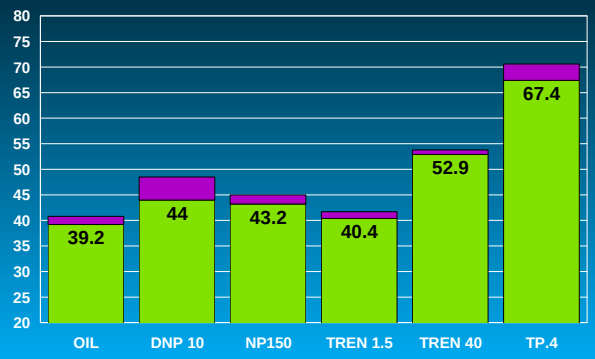
Hershberger assay: Agonist phase 3. Laboratory 1 Glans penis Weight Effects



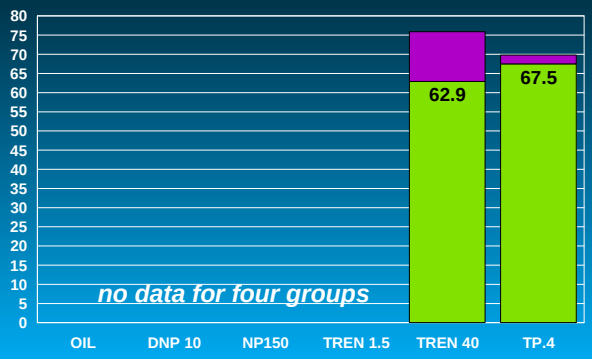
Hershberger assay: Agonist phase 3. Laboratory 2 Glans penis Weight Effects



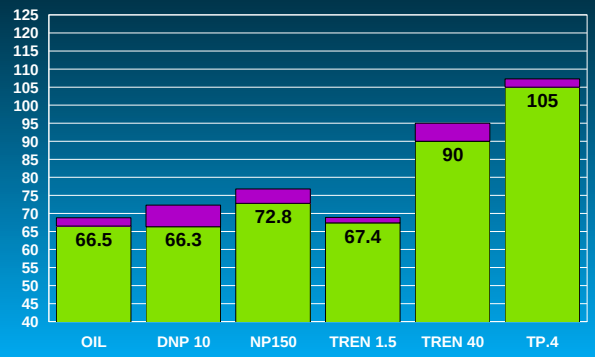
Hershberger assay: Agonist phase 3. Laboratory 3 Glans penis Weight Effects



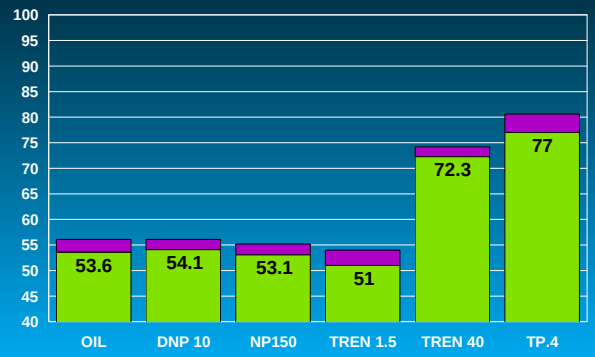
Hershberger assay: Agonist phase 3. Laboratory 5 Glans penis Weight Effects



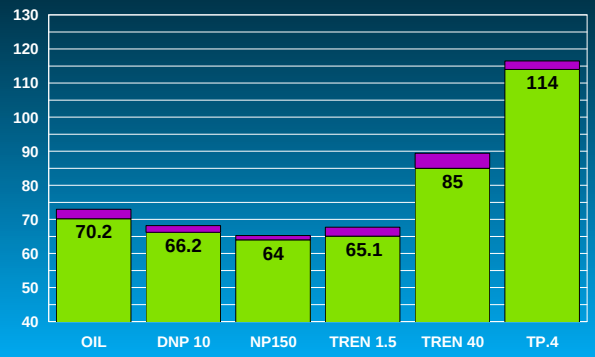
Hershberger assay: Agonist phase 3. Laboratory 7 Glans penis Weight Effects



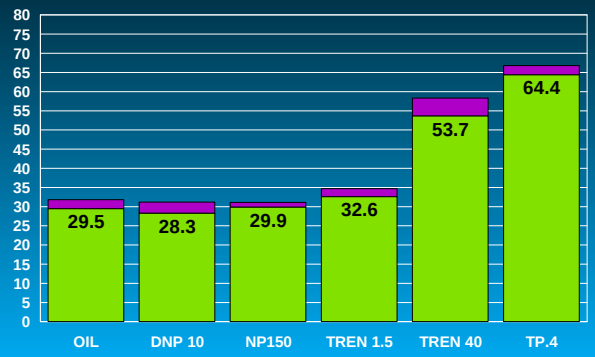
Hershberger assay: Agonist phase 3. Laboratory 8 Glans penis Weight Effects



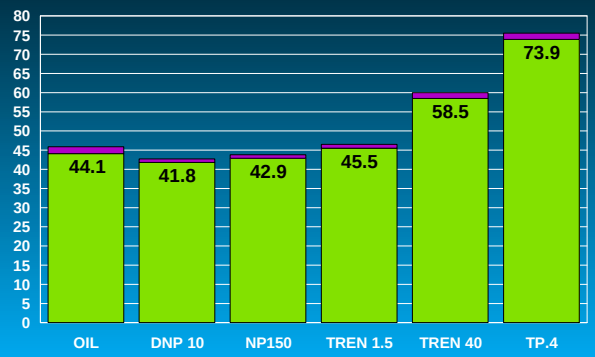
Hershberger assay: Agonist phase 3. Laboratory 9 Glans penis Weight Effects



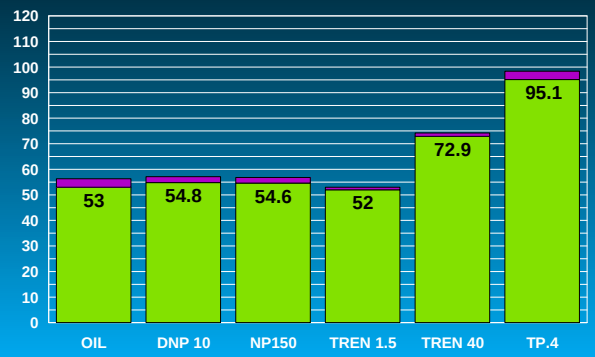
Hershberger assay: Agonist phase 3. Laboratory 10 Glans penis Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11 Glans penis Weight Effects

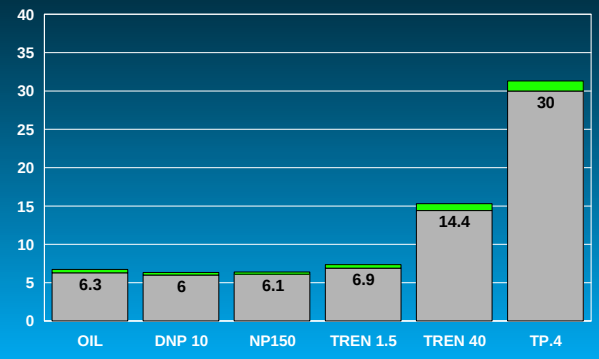


Hershberger assay: Agonist phase 3. Laboratory 13 Glans penis Weight Effects

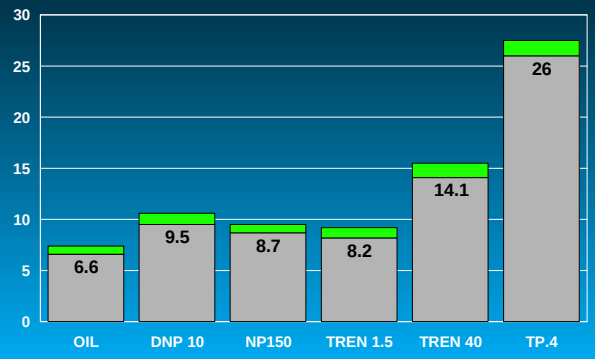


- Cowper's Gland Weight effects

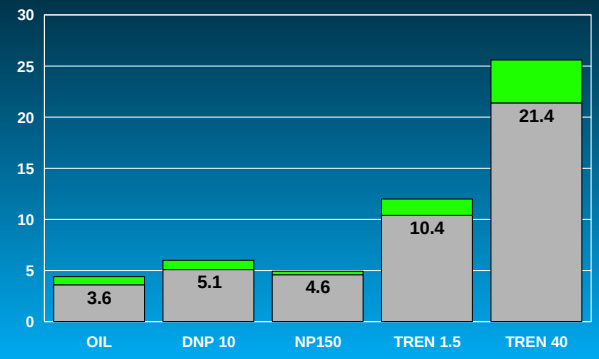
Hershberger assay: Agonist phase 3. All ten laboratories Cowper's gland Weight Effects



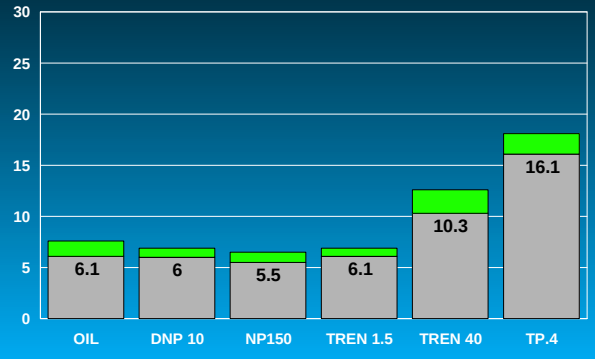
Hershberger assay: Agonist phase 3. Laboratory 1 Cowper's gland Weight Effects



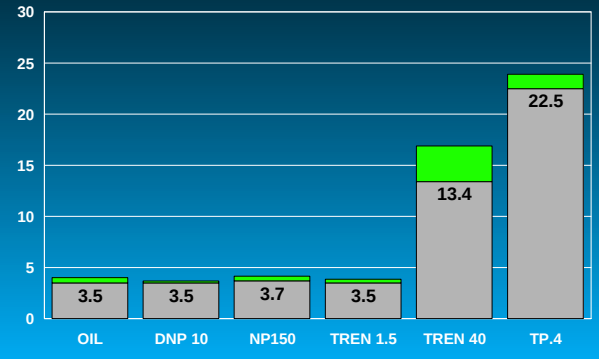
Hershberger assay: Agonist phase 3. Laboratory 2 Cowper's gland Weight Effects



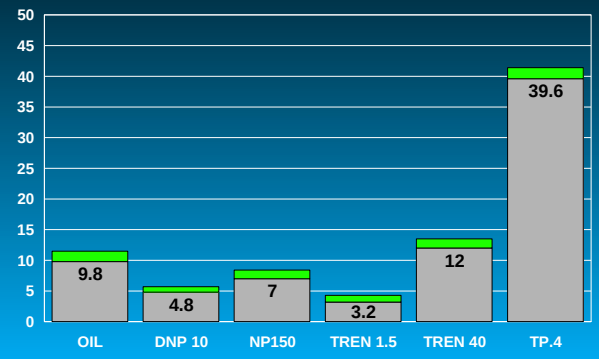
Hershberger assay: Agonist phase 3. Laboratory 3 Cowper's gland Weight Effects



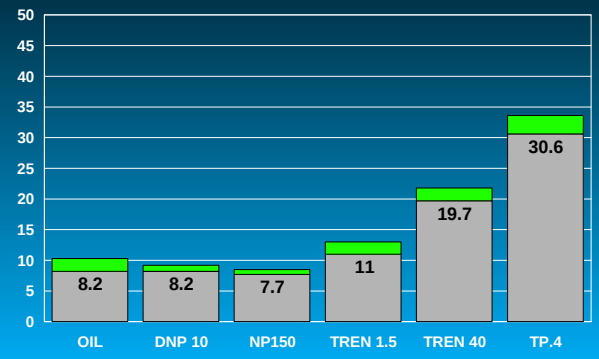
Hershberger assay: Agonist phase 3. Laboratory 5 Cowper's gland Weight Effects



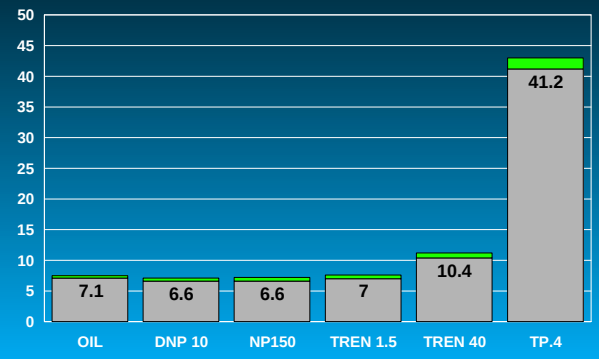
Hershberger assay: Agonist phase 3. Laboratory 7 Cowper's gland Weight Effects



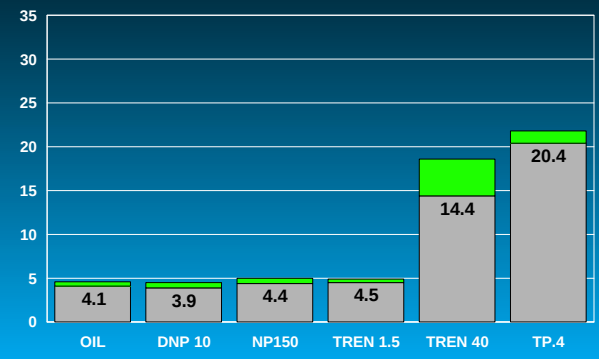
Hershberger assay: Agonist phase 3. Laboratory 8 Cowper's gland Weight Effects



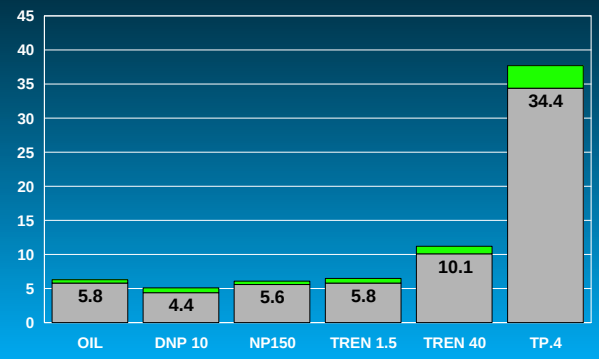
Hershberger assay: Agonist phase 3. Laboratory 9 Cowper's gland Weight Effects



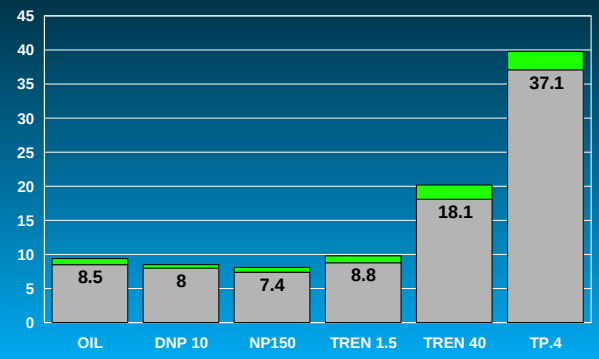
Hershberger assay: Agonist phase 3. Laboratory 10 Cowper's gland Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11 Cowper's gland Weight Effects

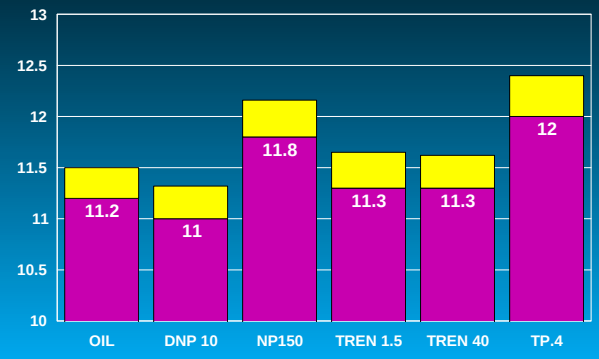


Hershberger assay: Agonist phase 3. Laboratory 13 Cowper's gland Weight Effects

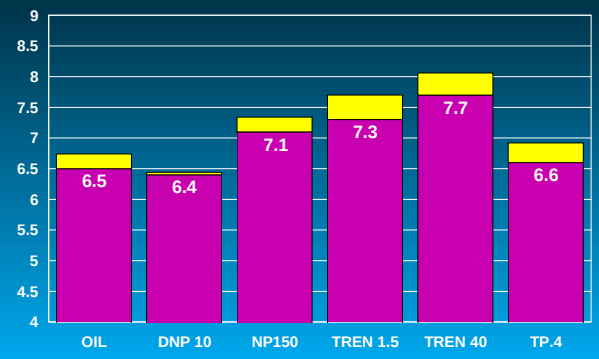


- Liver Weight effects

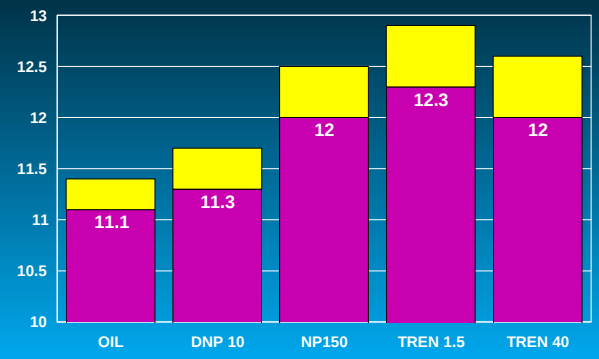
Hershberger assay: Agonist phase 3. All ten laboratories Liver Weight Effects



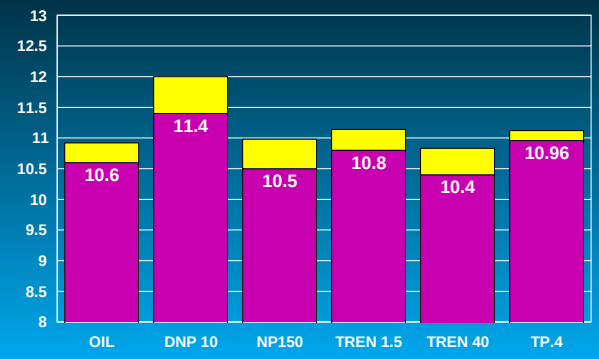
Hershberger assay: Agonist phase 3. Laboratory 1 Liver Weight Effects



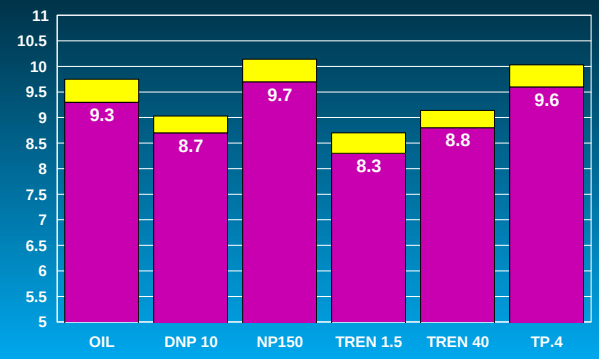
Hershberger assay: Agonist phase 3. Laboratory 2 Liver Weight Effects



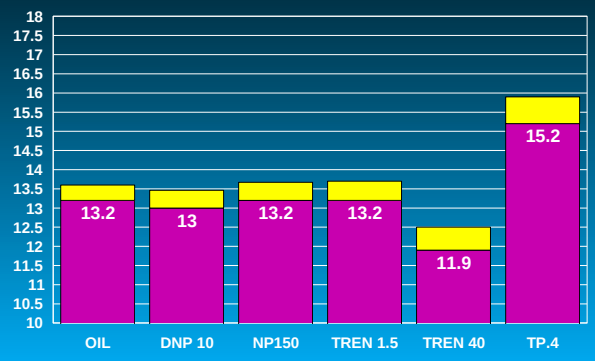
Hershberger assay: Agonist phase 3. Laboratory 3 Liver Weight Effects



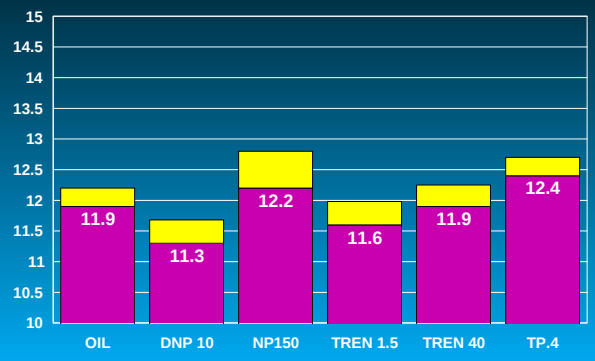
Hershberger assay: Agonist phase 3. Laboratory 5 Liver Weight Effects



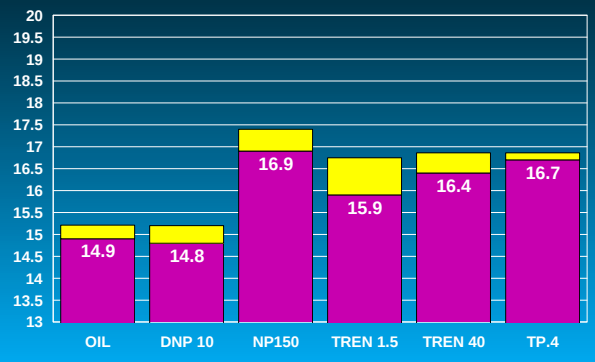
Hershberger assay: Agonist phase 3. Laboratory 7 Liver Weight Effects



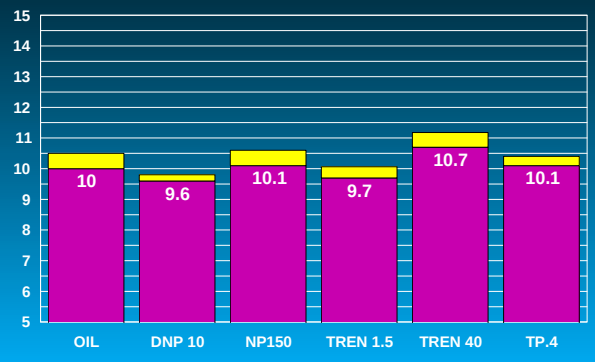
Hershberger assay: Agonist phase 3. Laboratory 8 Liver Weight Effects



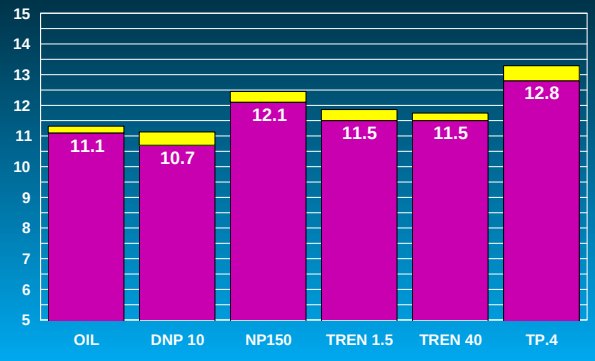
Hershberger assay: Agonist phase 3. Laboratory 9 Liver Weight Effects



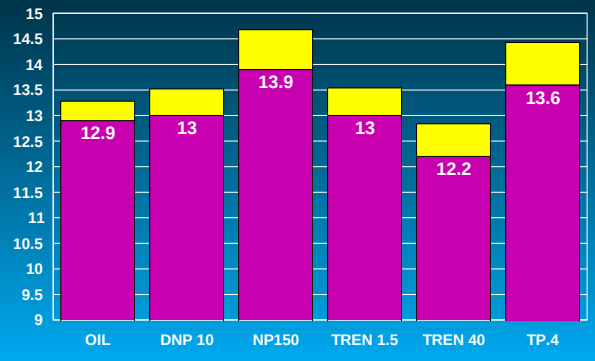
Hershberger assay: Agonist phase 3. Laboratory 10 Liver Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 11 Liver Weight Effects

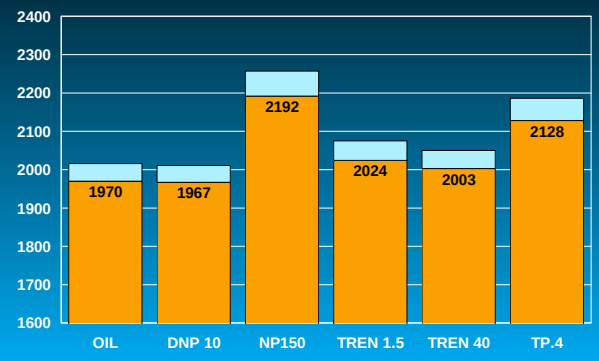


Hershberger assay: Agonist phase 3. Laboratory 13 Liver Weight Effects

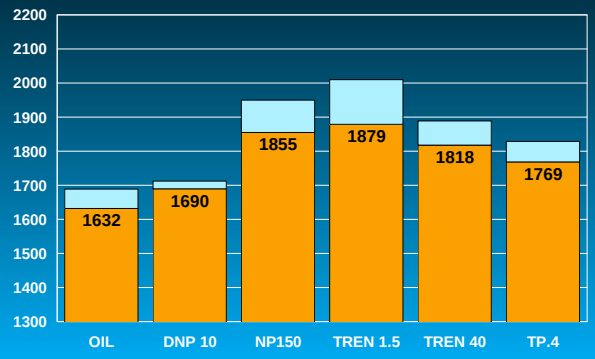


- Kidney Weight effects

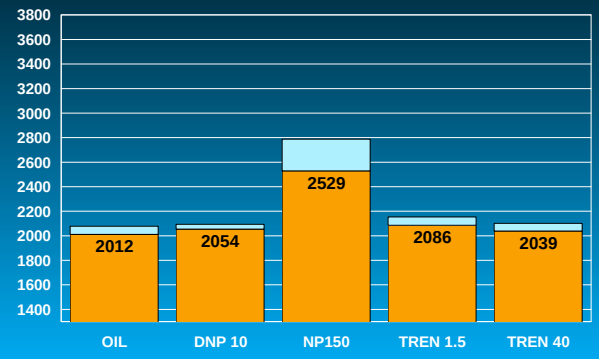
Hershberger assay: Agonist phase 3. All ten laboratories Kidney Weight Effects



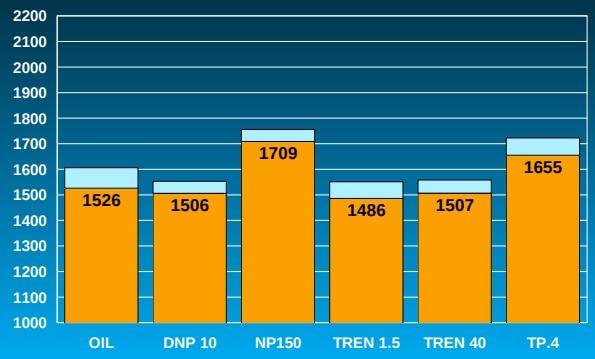
Hershberger assay: Agonist phase 3. Laboratory 1 Kidney Weight Effects



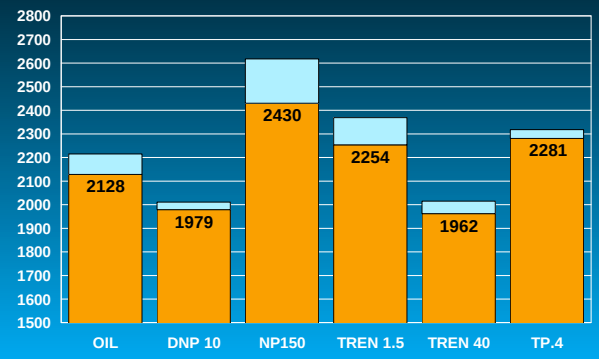
Hershberger assay: Agonist phase 3. Laboratory 2 Kidney Weight Effects



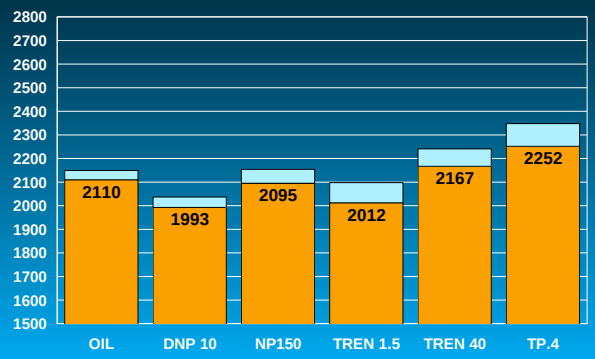
Hershberger assay: Agonist phase 3. Laboratory 5 Kidney Weight Effects



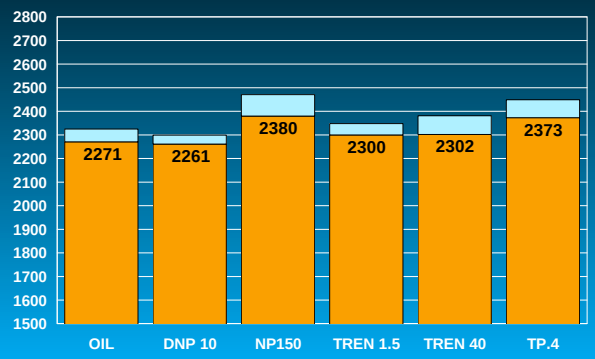
Hershberger assay: Agonist phase 3. Laboratory 7 Kidney Weight Effects



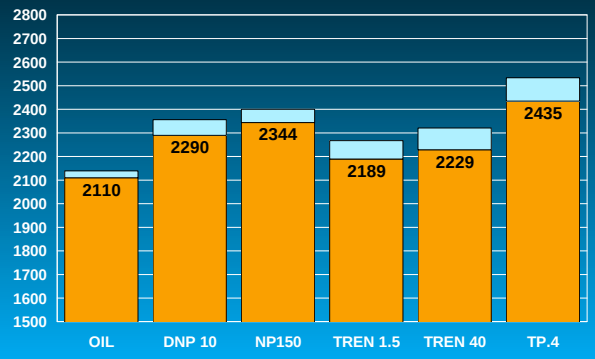
Hershberger assay: Agonist phase 3. Laboratory 8 Kidney Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 9 Kidney Weight Effects



Hershberger assay: Agonist phase 3. Laboratory 13 Kidney Weight Effects



Appendix 6.

The data are summarized here for each endpoint for the ten laboratories individually in order to display the R2 and CV values in a manner that the strength of the chemical effects (R2) and precision (CV) of the data can be easily compare among the laboratories for consistency. There were a few unusually large CVs. These do not always indicate that a lab was less proficient in their dissections than the other laboratories.

body weight Means of Coded Chemicals p < 0.05 p <0.01 False negative

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	lab 10	Lab 11	lab 13
OIL	221	317	237	240	275	339	314	263	281	303
DNP 10	221	315	254	236	269	323	314	252	285	308
NP150	212	303	225	223	265	324	309	251	276	297
TREN 1.5	224	318	236	230	284	327	320	250	282	301
TREN 40	209	275	209	215	239	307	290	238	262	265
R2	33	55	41	27	59	42	30	27	44	46
CV	4.1	5.2	8.5	7	5.2	4.1	5.5	5.7	3.7	6.2

ventral prostate

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	lab 10	Lab 11	lab 13
OIL	18.3	16	18.4	15.1	48.6	20.6	22	16.8	16	22
DNP 10	19.1	15.2	20.6	15.3	20.5	18	19	15.6	15.8	24
NP150	19.2	14.2	14.3	13.4	44.3	17.5	17	17	19.8	20.6
TREN 1.5	20	15.3	15.1	15	48.5	22.4	20.8	14.9	18.4	26.2
TREN 40	34.3	72	29.6	74.2	70	35	30.3	58.3	33.4	43.7
R2	57	81	37	59	27	77	51	32	64	72
CV	26	45	39	81	62	17	23	110	26	21

Seminal vesicle

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	lab 10	Lab 11	lab 13
OIL	37.4	40	41.5	31.1	58.6	57	73	27.9	42	61.2
DNP 10	37.2	46.6	37.9	30	60.5	55.2	66	25.7	38.8	58.4
NP150	40	40.3	38.6	30.7	65.1	63.6	67.7	26.5	40.9	58.1
TREN 1.5	39	58.4	35.3	26.9	54.4	70.2	72.5	29	41.6	61.2
TREN 40	86.4	262	108	169	144	144	130	136	179	165
R2	82	57	86	53	82	84	77	25	81	57
CV	20	41	24	101	23	21	18	168	43	22

LABC

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	lab 10	Lab 11	lab 13
OIL	152	92.1	179	112	233	270	162	137	164	191
DNP 10	176	95	174	102	215	264	154	128	190	191
NP150	161	92.2	160	100	198	240	156	128	179	179
TREN 1.5	180	149	172	106	214	295	177	141	216	221
TREN 40	359	371	359	309	447	548	246	317	427	452
R2	80	93	77	86	87	81	70	89	91	95
CV	21	20	21	25	15	19	14	16	14	11

Glans Penis

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	lab 10	Lab 11	lab 13
OIL	42.8	44.2	39.2	no data	66.5	53.6	70.2	29.5	44.1	53
DNP 10	43.1	43.8	44	no data	66.3	54.1	66.2	28.3	41.8	54.8
NP150	43.2	41.9	43.2	no data	72.8	53.1	64	29.9	42.9	54.6
TREN 1.5	44	50.6	40.4	no data	67.4	51	65.1	32.6	45.5	52
TREN 40	57	72	52.9	62.9	90	72.3	85	53.7	58.5	72.9
R2	56	79	45		48	69	60	69	82	72
CV	11	12	13		14	10	10	20	6.7	9

Cowper's glands

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	lab 10	Lab 11	lab 13
OIL	6.6	3.6	6.1	3.5	9.8	8.2	7.1	4.1	5.8	8.5
DNP 10	9.5	5.1	6	3.5	4.8	8.2	6.6	3.9	4.4	8
NP150	8.7	4.6	5.5	3.7	7	7.7	6.6	4.4	5.6	7.4
TREN 1.5	8.2	10.4	6.1	3.5	3.2	11	7	4.5	5.8	8.8
TREN 40	14.1	21.4	10.3	13.4	12	19.7	10.4	14.4	10.1	18.1
R2	55	67	23	54	52	59	54	47	59	70
CV	27	56	51	72	44	38	20	76	28	28

Liver										
	Lab 1	**Lab 2**	**Lab 3**	**Lab 5**	**Lab 7**	**Lab 8**	**Lab 9**	**lab 10**	**Lab 11**	**lab 13**
OIL	6.5	11.1	10.6	9.3	13.2	11.9	14.9	10	11.1	12.9
DNP 10	6.4	11.3	11.4	8.7	13	11.3	14.8	9.6	10.7	13
NP150	7.1	12	10.5	9.7	13.2	12.2	16.9	10.1	12.1	13.9
TREN 1.5	7.3	12.3	10.8	8.3	13.2	11.6	15.9	9.7	11.5	13
TREN 40	7.7	12	10.4	8.8	11.9	11.9	16.4	10.7	11.5	12.2
TP.4	6.6	NO DATA	10.96	9.6	15.2	12.4	16.7	10.1	12.8	13.6
R2	**35**	13	11	25	19	11	32	15	28	15
CV	10	11	10	11	9	8.4	8.2	10	7.1	11

Note regarding the CVs for the SV, VP and Cowper's gland weights

High CV arises in Labs 5 and 10 from a single high Tren 40 value in each data set

note that this did not affect the ability to detect the inceases in VP, SV or Cowper's gland weights.

High CV in lab 7 is due to high variance overall in control and trenolone 40 groups for VP and Cowper's glands

note that this did preclude detection of effects, causing false negatives

Appendix 7.

The following table lists the data used in the calculation of the correlation coefficient describing the strength of association between the treatment effects from trenbolone at 1.5 and 40 mg/kg/d and 0.4 mg TP/kg/d from phases 2 and 1, respectively, 2 with the effects seen herein in phase 3. These were used to develop the graph shown within the text of the document. The effects observed in phase 3 were remarkably similar to those seen earlier in phase 2, demonstrating that the androgens produced reproducible responses.

Correlation of effects from phase 1 or 2 studies with phase 3
Shown as percentages of concurrent control values

Tissue		Expected from Phase 1 or 2	Observed in Phase 3
VP	TREN 1.5	118	99
VP	TREN 40	211	225
VP	TP .4	792	645
SV	TREN 1.5	97	100
SV	TREN 40	211	323
SV	TP .4	962	725
LABC	TREN 1.5	109	111
LABC	TREN 40	197	226
LABC	TP .4	299	259
BWT	TREN 1.5	102	99.2
BWT	TREN 40	91	90
BWT	TP .4	106	103
GP	TREN 1.5	100	100
GP	TREN 40	130	138
GP	TP .4	178	165
COWS	TREN 1.5	103	110
COWS	TREN 40	201	229
COWS	TP .4	545	476

R2 LINEAR =0.96 r=.98

ANNEX 4 − Report of the lead laboratory on experiments with antagonists

Please note that information may be repeated as some of the tables from this annex have already been inserted in the report. However, the OECD Secretariat didn't want to change the document prepared by Dr Earl Gray.

Hershberger Interlaboratory Validation Study

Results of Phase III using coded chemicals.

Experiments with Antagonists in Ten Laboratories

Prepared by Leon Earl Gray Jr

Endocrinology Branch, Reproductive Toxicology Division
NHEERL, ORD, USEPA
MD 72, 2525 Highway 54, RTP, NC 27711
gray.earl@epa.gov

Introduction

In our role as the lead laboratory in the OECD Hershberger Assay Validation we received data from Phase 3 experiments from Dr W Owens, acting on behalf of the OECD, and the laboratories for analysis and interpretation. The following are the results of this effort.

In phase 3, ten laboratories from Europe, the UK, the USA and Asia (Japan and Korea) received coded samples from OECD for evaluation in the standardized Hershberger assay for "antiandrogenicity".

Chemicals were administered orally at proscribed dosage levels concurrently with testosterone propionate (TP) to determine if they attenuated the effect of growth stimulating effect of TP in five androgen-dependent tissues (seminal vesicle (SV), ventral prostate (VP), levator ani plus bulbocavernosus (LABC), glans penis (GP) and Cowper's glands). In addition, body weight was monitored throughout the ten day experiment, and body, liver, kidney, and adrenal weights were taken at necropsy (in some but not all laboratories). Seven of the laboratories administered TP sc at 0.4 mg/kg/d whereas the three from Japan used 0.2 mg

TP/kg/d.

The chemicals included p,p' DDE at 16 and 160 mg/kg/d, linuron at 10 and 100 mg/kg/d, nonylphenol at 150 mg/kg/d, dinitrophenol at 10 mg/kg/d, and flutamide at 3 mg/kg/d. Corn oil (the vehicle, in some laboratories) also was included as a treatment in some of the laboratories.

All of these antiandrogenic chemicals (p,p' DDE, linuron and flutamide) were evaluated earlier in Phase II of this validation effort and nonlyphenol and dinitrophenol were administered as known negatives. The objective of the study was to demonstrate that ten laboratories could get the "correct" responses for all the chemicals on all androgen-dependent tissues when the chemicals were provided to them as unknown coded samples.

Statistical Methods

Raw data were taken from submitted Microsoft Excel 2003 spreadsheets, converted to text files and modified for statistical analysis using WordPerfect 9 software. The text files were submitted electronically via FTP using the NHEERL intranet to an IBM mainframe computer using SAS (r) proprietary Software Release 8.2 (TS2M0), licensed to US ENVIRONMENTAL PROTECTION AGENCY, Site 0045713001.running on IBM Model 9672 .

The results presented herein are one- way ANOVAs (main effect of the chemicals for each laboratory) and two-way ANOVAs (with laboratory and chemical as the two main effects) from PROC GLM. In addition, because on group of seven laboratories used 0.4 mg TP/kg/d and the second group of three laboratories used 0.2 mg TP/kg/d the data were analyzed by group and also will all ten laboratories pooled.

Post hoc comparisons were generated using the LSMEANS option of PROC GLM which is a two tailed t-test. T-tests being appropriate for this study because we have a priori expectations about the effects of the different doses of the chemicals. In addition, as a screening assay, false positives are more tolerable than false negatives that might result from using a more conservative post hoc test. It is also our objective to keep animal use to a minimum so using a t-test enables us to detect weaker effects with a sample size of six rats per group than does a more

conservative post hoc test.

Means, standard errors (SE) and coefficients of variation (CV) were generated using PROC MEANS on SAS. The CV is the standard deviation divided by mean and provides an index of the dissection precision of the organ weight data in a laboratory or group. SAS ANOVAs also provides R-square (R2) values (R2 being the variation due to laboratory or chemicals divided by the total variation) which is an index of the robustness of the effects of the chemicals in each laboratory. The expectation is for CV values similar to those seen in Phase 1 and 2 of the validation effort and that al the laboratories would have similar R2 values.

SAS results were then transmitted by FTP to an external hard drive in my laboratory and converted from text files to WordPerfect files for printing. The results from these sheets were then entered into Microsoft Excel spreadsheets to summarize the results of the data analyses. Subsequently, the Excel sheets were printed to for this report to pdf files using Adobe Acrobat 5.0 software.

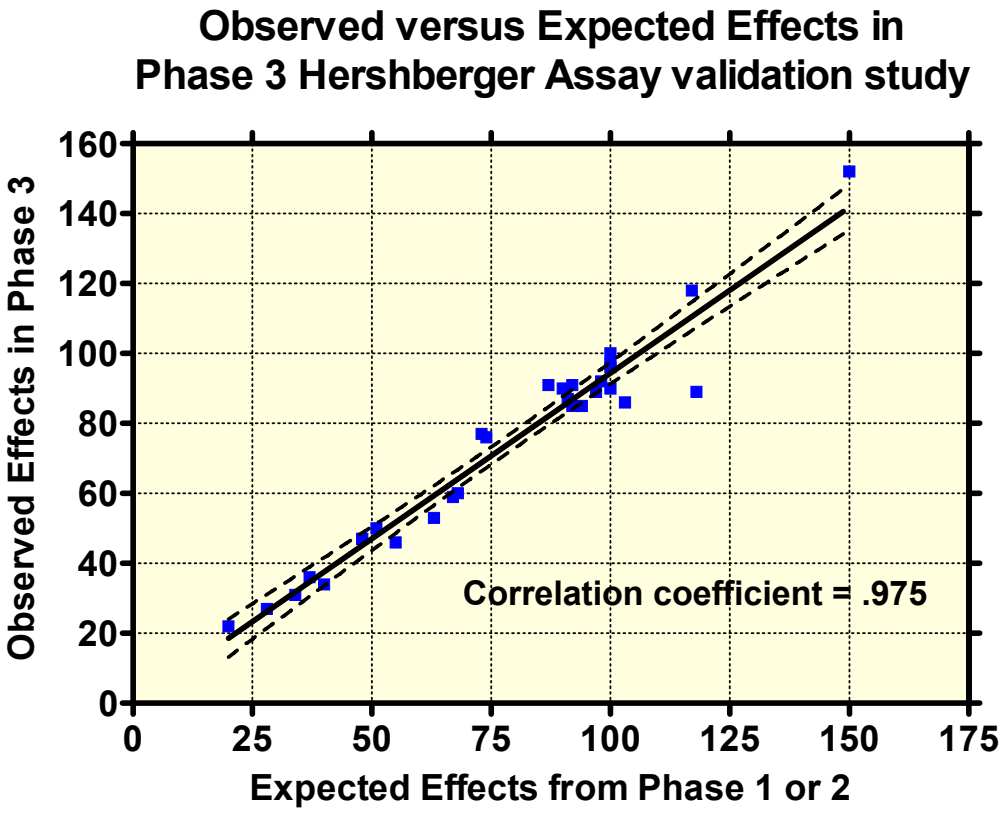
Graphs of histograms of the means and standard errors of the mean for the weight data are presented individually for each laboratory and pooled over the laboratories. The graphs were prepared using Lotus Freelance Graphics 9.7 for Windows. Graphs were arranged by endpoint for the ten laboratories so the repeatability of the effects of the chemicals from lab-to-lab was easily apparent. This file was then also converted to a pdf file for inclusion in the report.

Results

One of the objectives of the interlaboratory study was to demonstrate that ten laboratories could get the "correct" responses for all the chemicals on all androgen-dependent tissues. As p,p' DDE at 16 and 160 mg/kg/d, linuron at 10 and 100 mg/kg/d and flutamide at 3 mg/kg/d were used in either phases 1 or 2 studies of this validation exercise we have specific expectations about the quantitative effects of these five treatment groups on body weight, liver weight and the five androgen-dependent organ weights. We expected and observed the following effects:

1. Liver weight increases with p,p' DDE at 16 and 160 mg/kg/d
2. Decreased body weight with p,p' DDE at 160 and linuron at 100 mg/kg/d
3. Reduced androgen-dependent organ weights with p,p' DDE at 160 mg/kg/d, linuron at 100 mg/kg/d and flutamide at 3 mg/kg/d.

When all the effects of these five treatment groups on the seven endpoints (body, liver and five androgen-dependent tissues) were correlated with the effects seen previously in Phase 1 or 2 studies a highly significant positive relationship was found having a correlation coefficient of r = .975. The expected and observed effects were calculated as percent of concurrent control values with the TP 0.4mg/kg/d group as the control (seven laboratories using 0.4 mg TP/kg/d). Flutamide predictions were from phase 1(from five labs using TP at 0.4 mg/kg/d) and linuron and p,p' DDE were derived from phase 2 data (from four labs using TP at 0.4 mg/kg/d).



If one considers that there 5 androgen-dependent tissues were measured in ten laboratories than Flutamide and p,p' DDE at 160 mg/kg/d yielded 96% positive hits (2% of the misses for these two chemicals arose when one lab could not measure glans penis weight because the prepuce had not separated) (Appendix 1). Administration of linuron at 100 mg/kg/d produced 90% positive responses. The low doses of p,p' DDE and linuron produced 40% and 18% positive responses. We did not expect the low dose of either of these chemicals to produce consistent positive responses based upon earlier validation studies and, in fact, there were more potent as antiandrogens herein than in phase 2 studies. Dinitrophenol and nonylphenol were negative yielding only 8% and 14% positive responses on the androgen-dependent tissues.

Analyses of variance for the five androgen-dependent tissues revealed that the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in phase 3 were all very statistically significant. Importantly, the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by relatively small F-values for the chemical by laboratory interaction as compare to the size to the two main effects (Appendix 2). When the analysis was conducted separately for the two groups of laboratories that used either TP at 0.4 mg/kg/d (7 labs) or 0.2 mg/kg/d (3 labs), both protocols yielded similar results (Appendix 3). The groups means and SEs of the two groups of laboratories and the means and SEs of the individual laboratories are presented in Appendix 4 to show the consistency of the antiandrogenic effects quantitatively. The uniformity of the response among the laboratories to each chemical is even more obvious from viewing histograms of the data in Appendix 5. Here the data are presented by endpoint for the two groups of laboratories and for the ten individual laboratories. This displays the remarkably uniform profile of effects among laboratories.

The final set of tables summarizes the results by endpoint (Appendix 6). Here it is easy to compare R2 and CV values from laboratory to laboratory for each of the five androgen-dependent organ weights and body weight at necropsy. The size of the R2 describes the strength of association of the chemical effects versus TP control and the CV describes the precision (variability) of the data. If the all laboratories had the same CV on an endpoint then this would indicate that they were all dissecting the tissues with similar precision. A high CV in one laboratory versus the other laboratories might indicate a problem. However, all the CVs seemed

to be within a reasonable range, indicating that the dissections had been consistent. If the chemical treatments all had the same magnitude of response in each laboratory then the R2 values were be similar among laboratories, given that the variances were not markedly different. Indeed, there was only one case in which this was not the case. In laboratory 5 the R2 value for the effect of chemical treatments on the glans penis was lower than the other 9 laboratories. This arose because they were unable to obtain glans penis weights with some of the animals treated with the potent antiandrogens biasing the means as the most affected animals could not be measured. In fact, for the flutamide group in the laboratory, no glans penis weight data are available. Hence, this low R2 value arises from the incomplete nature of their data set for glans penis weight.

In summary, analysis of the results of phase three of the Hershberger assay interlaboratory study using coded samples of known dosage levels of antagonists or negatives produced appropriate and consistent responses among all the laboratories. None of the known positives were undetected in any laboratory. In fact the positive responses were quite strong. Only a few "false" positive responses were seen with the "known" (or suspected negatives).

A similar report on the studies with the coded agonists is in preparation.

Appendix page 1.

Summary of effects listed chemical-by-chemical. Blue shaded cells differ significantly from control (TP-treated) by $p < 0.01$ whereas yellow shaded cell differ by $p < 0.05$ but t-test from the LSMEANS.

Appendix 2.

Summary of the analyses of variance for the five androgen-dependent tissues showing the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in phase 3. Note that the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by relatively small F-values for the chemical by laboratory interaction as compare to the size to the two main effects.

Appendix 3.

Responses of the androgen-dependent tissues were very similar in the two groups of laboratories that varied the dose of TP injected from 0.4 to 0.2 mg/kg/d. The p - value is provided for each endpoint by chemical showing the consistency of the results using two protocol modifications.

Appendix 4.

Means and CVs of the androgen-dependent tissues presented for the two groups of laboratories (using different doses of TP) and for the ten individual laboratories. Blue shaded cells differ significantly from control (TP-treated) by $p < 0.01$ whereas yellow shaded cell differ by $p < 0.05$ but t-test from the LSMEANS.

Appendix 5.

Histograms of the means and SEs prepared with Lotus Freelance are provided here by endpoint for the two groups of laboratories and for the ten individual laboratories. These graphs display the remarkably uniform profile of effects among laboratories.

Appendix 6.

The data are summarized here for each endpoint for the ten laboratories individually in order to display the R2 and CV values in a manner that the strength of the chemical effects (R2) and precision (CV) of the data can be easily compare among the laboratories for consistency. There were no unexplained low R2 values and no unusually large CVs.

Appendix page 1.

Summary of effects listed chemical-by-chemical. Blue shaded cells differ significantly from control (TP-treated) by $p < 0.01$ whereas yellow shaded cell differ by $p < 0.05$ but t-test from the LSMEANS.

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	DDE 16	207	4.7	72.3	5.3	113 a	23	273	29	62	3.1	19.8	3.1	9.4 b	0.16	55.6	3.8	1785	85
2	DDE 16	345	7	178	13	592	122	406	24	81	2.3	42.2	1.9	16	0.46	60.2	3.3	2392	47
3	DDE 16	246	6.1	64	8.5	181	25	282	25	62	3.9	13.3	0.81	13.7	0.54				
5	DDE 16	246	4.6	132	4.4	329	25	316	13	58.4	2.4	24	1.7	11.9	0.12	50.1	1.8	1658	51
7	DDE 16	288	11	139	17	378	20	544	38	85 .4	3.3	35	3.5	16	0.8	52.1	3.2	2215	52
8	DDE 16	353	4.8	83	7.2	281	27	396	27.5	69.5	1.9	26	2.4	16.1	0.6	61.1	2.4	2337	69
9	DDE 16	349	4.5	153	3.9	560	27	401	11	112	2.5	40.3	2.5	19.9	0.24	61.3	1.8	2534	46
10	DDE 16	266	4.8	79	6	144	12	281	12	53.6	1.4	16.35	0.8	12.5	0.34				
11	DDE 16	331	9.3	128	20	357	40	486	34	72	1.2	31.8	3	17.7	0.8				
13	DDE 16	323	10	107	7	324	24	396	21	84	1.8	23.5	2.4	15.95	0.53	56.6	1.4	2329	81

40% positive for antiandrogen

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	DDE 160	200	3.7	29.4	2.3	41.5	7	142	12	44.96	3.1	9.3	0.8	12.3	0.29	56.1	3.96	1788	27.5
2	DDE 160	262	11	51	9	69	8	113	8	59	3.5	12.4	1.4	19	1.2	69.6	7.2	2050	72
3	DDE 160	228	7.2	23.1	0.87	66.9	10.4	167	11.3	45.8	2.6	7.74	2.4	17.31	1.1				
5	DDE 160	221	14	36.6	6.1	60.1	6.5	117	17			7	0.9	16.34	1.3	51.1	1.3	1665	85
7	DDE 160	259	22	70	15	160	20	287	34	67	4.1	18.8	1.9	20.3	2	56.4	2	2164	152
8	DDE 160	316	17	36	3.7	103	16	305	15	60	2.6	17.6	2	19.7	1.3	64.3	0.6	2306	82
9	DDE 160	328	7	81	10	273	20.8	242	17	97	2.5	25.3	1.5	26.4	0.9	67.4	2	2591	66
10	DDE 160	249	7	28.3	2.4	35.5	2.5	126	5.5	32	1.6	4.9	0.34	17.25	0.49				
11	DDE 160	271	22	28	4.4	78	13	218	20	50	0.8	10.5	1.9	22.8	1.8				
13	DDE 160	248	31	36	4.7	74	4	165	18	60	3.7	9.3	1.3	18.6	1.3	55.5	1.8	2152	108

96% positive for antiandrogen

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	DNP 10	211	4	109	24	206	30	303	20	66	2.7	21.8	2.3	6.3	0.2	59.2	2.1	1595	34
2	DNP 10	355	5.5	201	16	594	79	424	9.7	93	4.1	40	4.8	13.8	0.34	62	1.8	2339	99
3	DNP 10	240	6.2	63	5.6	182	5	296	14	68.5	3.3	15	0.7	10.5	0.23				
5	DNP 10	250	4.7	134	15	336	25	335	12	68.3	3	25.55	2.3	11	0.3	50	3.2	1715	50
7	DNP 10	288	7	179	21	457	18	569	35	89	2.8	39	3	16.4	0.4	60	2.6	2344	56
8	DNP 10	360	9.3	104	6	385	33	548	17	80.5	3	41.1	4.1	13.5	0.43	57.6	1.9	2242	60
9	DNP 10	346	4.8	166	9.8	578	22	385	12	117	2.3	43.3	1.8	17.05	0.41	59	2.7	2401	61
10	DNP 10	272	3.9	85.9	1.9	186	14	302	12	61	0.9	16.2	1.4	10.92	0.35				
11	DNP 10	332	8.5	129	13	440	28	530	13	75.2	0.9	31	1	15.1	0.35				
13	DNP 10	317	4.5	133	8	412	24	494	20	90	3.1	33.4	2.2	13.7	0.22	58	2.4	2342	14

8% positive

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	FLUT 3	201 a	3.8	26	1	34 b	3.1	152 b	7.8	47.2 b	3.8	9.3	1.2	6.16	0.2	58.2	2.4	1609	101
3	FLUT 3	246	8.2	17.3	1.8	54.6	5.1	152	6.6	46.9	2.6	5.8	0.3	11.1	0.61				
5	FLUT 3	236	8	33	2.4	43.3	7.7	154	14			6	0.72	8.9	0.7	83	5.1	1515	58
7	FLUT 3	287	11	56	15	101	15	262	24	74	3.7	11	2.4	15.93	0.8	65.6	4.9	2335	66
8	FLUT 3	349	8.6	31.5	1.8	88.4	3.1	296	11.1	57.2	2.4	12.8	0.7	13.5	0.4	61	1.9	2243	87
9	FLUT 3	347	8.2	55.1	3	201	6.4	269	6.4	94.8	1.6	20.1	1.2	18	0.6	60.7	3.1	2540	58
10	FLUT 3	261	5.1	24.85	1.49	28.2	1.5	148	7.6	30.3	1.7	4.7	0.5	10	0.33				
11	FLUT 3	327	9	28.2	3.6	65.9	4.6	236	12	50	0.8	10.3	1.4	15.1	0.8				
13	FLUT 3	313	10	40	2	76.3	4.4	222	21	61	1.4	12	1.5	13.5	1	55.5	3.1	2194	82

96% positive

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	LINURON 10	204 a	4.5	64	7	137	23	265 a	17	61	1.7	21.6	1.8	6.48	0.29	55.9	2.5	1628	37
2	LINURON 10	351	7.7	196	6.6	423	30	430	21	90	4.9	48	2.9	14.2	0.6	62.2	1.9	2410	32
3	LINURON 10	245	8.3	72	6.3	216	22	320	19.4	67.7	3.6	16.6	1.1	10.9	0.28				
5	LINURON 10	242	10	144	10	354	22	323	10	67	3.6	24.6	2.2	9.1	1.1	45.5	2	1567	42
7	LINURON 10	314	10	153	16	447	30	533	37	89	1.8	32.7	2.6	17.3	0.65	55.7	2.3	2429	73
8	LINURON 10	363	7.8	92	8.6	253	23	443	14	73.2	1.9	31	2.4	14.5	0.6	58.5	2.6	2407	93
9	LINURON 10	347	7.8	142	5.8	524	19	391	9.3	117	2.4	40.3	0.4	17.6	0.51	57.5	2.3	2526	94
10	LINURON 10	264	4.8	81	4.4	155	13	303	17	64.5	2.5	18.1	1.3	10.4	0.2				
11	LINURON 10	335	6	115	11	357	16	532	23	73	1.3	25.7	2.5	15.56	0.55				
13	LINURON 10	311	9	134	10	317	25	452	26	90	2.5	35	2.7	12.8	0.6	56.5	3.2	2184	81

18% positive

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	_INURON 10	196 b	3.9	43.4	2.2	58.4 b	4.4	179 b	11	57.2 a	2.2	12.8	0.5	6.33	0.17	57.5	3.5	1563 a	49
2	_INURON 10	293	8.7	99	9.5	278	38	252	11	76.1	3	26	2.6	11	0.48	65	1.5	2082	92
3	_INURON 10	233	8.7	36	4	85	8	184	8.2	56	1.6	10.4	0.6	10.74	0.48				
5	_INURON 10	222	6.2	66	7	128	13	200	12			12.28	0.9	8.6	0.33	58.5	5.9	1618	57
7	_INURON 10	270	7.5	127	17	244	10	398	22	88	1.8	22.2	1	15.4	0.75	54.4	4.4	2283	64
8	_INURON 10	331	8.8	44.7	4.8	127	13.7	280	18.6	61.4	1.1	18.4	1.6	13.77	0.55	68.6	3.2	2227	66
9	_INURON 10	316	4.1	97.5	4.5	348	17	290	8.6	107	1.6	29.5	1.5	16.1	0.5	61.5	4	2444	52
10	_INURON 10	246	4.5	41.4	2.4	60.6	3.3	156	8.9	41.7	2.6	8.3	0.7	9.55	0.35				
11	_INURON 10	310	5.7	61.6	10	178	18	336	19	64	2.5	22.9	3.7	14.14	0.33				
13	_INURON 10	293	7.4	68	10	160	18	263	20	73.2	4	20.9	1.9	13	0.35	63.2	2.8	2236	54

90% positive

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	NP 150	204	3.2	99	4.3	212	17	342	21	70	2.6	26	2.5	6.9	0.21	58	2.6	1746	40
2	NP 150	329	10	181	11	681	40	455	12	93	3.7	41.1	3.1	13.65	0.6	61.3	1.8	2474	78
3	NP 150	247	9.5	72	10	215	31	320	39	69	4.8	16.2	1.5	11.8	0.7				
5	NP 150	231	7	116	9.5	345	65	332	26	66	5	23.1	2.5	11.7	0.7	49.3	2	1731	70
7	NP 150	283	6.6	162	19	468	45	567	23	89.6	3.2	42.1	3.9	15.9	0.55	53.1	3.5	2471	104
8	NP 150	346	6	132	7	334	22	499	13	82	3.2	36.2	1.35	15.1	0.55	57.6	1.6	2426	26.5
9	NP 150	337	8.8	156	6	551	26	393	7	111	2.2	42	1.5	18.8	0.9	59.4	1.5	2616	94
10	NP 150	263	7.3	82	3.1	191	16	308	16	61.9	2.25	20.1	0.67	11.3	0.5				
11	NP 150	326	6.7	135	5.3	406	29	540	22	75.6	2.4	28.5	2.3	15.8	0.9				
13	NP 150	307	10	142	14	437	49	460	35	86.3	3.2	37.5	4.5	14.2	0.7	54.3	5.1	2316	111

14 % positive

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
2	OIL	323	11	15	2.9	43	3.4	122	4.2	54.1	4.2	4.25	0.8	11.8	0.84	78.4	5.2	2135	87.3
7	OIL	286	5.7	40	6.6	56.5	5.2	181	14	53.2	2.5	4.7	0.9	12.7	0.5	56.1	1.5	2147	32
8	OIL	343	10	22.6	0.85	72.9	7.3	293	15.6	52.6	2.1	16.3	1.2	13.3	0.4	61.8	2.6	2179	87
10	OIL	260	7.1	14.4	0.6	23.8	1.25	119	2.8	29.4	1.2	3.2	0.5	10.02	0.4				
11	OIL	313	6	17.7	1.1	48	3.4	218	10.3	50	1.2	5.65	0.35	13.76	0.29				
13	OIL	303	7	21	1.4	54	2.1	193	4.6	50.9	1.6	8.3	1.1	12.9	0.5	57	4.4	2106	87

Lab	Chemical	body		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
1	TP.4	216	3.1	87.1	4.3	178	16	318	11	68	4.1	21.7	1.2	6.58	0.18	61.8	2.4	1756	25
2	TP.4	338	14	185	27	540	62	467	18	92.2	6.1	40.1	7.2	13.5	0.7	57.6	3.2	2268	62
3	TP.4	246	8	78	8.5	198	17	339	17	71.3	5.1	15.5	1.6	11.1	0.4				
5	TP.4	245	6	129	8.5	358	23	341	29	75	3.1	30	4.1	9.65	0.28	52.5	3.8	1623	46
7	TP.4	301	12	197	25	426	31	543	24	85.4	2.1	35.2	2.4	16	1	52.3	5.2	2450	122
8	TP.4	349	6.6	147	5.6	464	43	592	22	80	2.6	43.8	3.2	13.3	0.4	64.7	1.7	2241	75
9	TP.4	347	7.4	145	10	567	29	393	14	119	2.1	40	2	17.3	0.5	61.5	1.9	2538	77
10	TP.2	267	6.7	102	7	195	16	306	16	60.9	2.1	18.4	1.9	10.4	0.4				
11	TP.2	332	5.8	145	8	479	30	495	85	68	5.6	28. 5	3 .3	15.94	0.9				
13	TP.2	315	8.5	159	13	460	40	518	30	93.5	1.9	35.6	2.7	13.65	0.7	49.5	3.4	2207	93

Appendix 2.

Summary of the analyses of variance for the five androgen-dependent tissues showing the magnitude of effect of the chemicals and the laboratory-to-laboratory variability in phase 3. Note that the chemical effects were fairly consistent from laboratory-to-laboratory as indicated by relatively small F-values for the chemical by laboratory interaction as compare to the size to the two main effects.

RESULTS OF ANOVAS FOR LAB AND LAB BY TREATMENT INTERACTION

	Group 1 Labs 1-9				Group 2 Labs 10, 11 and 13		
	Effect	**df**	**F**	**p-value**	**df**	**F**	**p-value**
Body Weight	**LAB**	6	355	0.0001	2	104	0.0001
	CHEMICAL	7	16	0.0001	7	9	0.0001
	INTERACTION	41	1.4	NS	14	1.1	NS
	Effect	**df**	**F**	**p-value**	**df**	**F**	**p-value**
SV	**LAB**	6	133	0.0001	2	158	0.0001
	CHEMICAL	7	100	0.0001	7	95	0.0001
	INTERACTION	41	3.6	0.0001	14	5.5	0.0001
	Effect	**df**	**F**	**p-value**	**df**	**F**	**p-value**
LABC	**LAB**	6	146	0.0001	2	109	0.0001
	CHEMICAL	7	151	0.0001	7	54	0.0001
	INTERACTION	41	3.5	0.0001	14	2	0.03
	Effect	**df**	**F**	**p-value**	**df**	**F**	**p-value**
VP	**LAB**	6	91	0.0001	2	50	0.0001
	CHEMICAL	7	85	0.0001	7	67	0.0001
	INTERACTION	41	2.4	0.0001	14	2	0.02
	Effect	**df**	**F**	**p-value**	**df**	**F**	**p-value**
GP	**LAB**	6	235	0.0001	2	295	0.0001
	CHEMICAL	7	49	0.0001	7	73	0.0001
	INTERACTION	38	1.1	NS	14	1.8	0.05
	Effect	**df**	**F**	**p-value**	**df**	**F**	**p-value**
COWPER	**LAB**	6	111	0.0001	2	82	0.0001
	CHEMICAL	7	83	0.0001	7	40	0.0001
	INTERACTION	41	2.8	0.0001	14	3	0.002

Appendix 3.

Responses of the androgen-dependent tissues were very similar in the two groups of laboratories that varied the dose of TP injected from 0.4 to 0.2 mg/kg/d. The p - value is provided for each endpoint by chemical showing the consistency of the results using two protocol modifications.

		GROUP 1 LABS 1-9	GROUP 2 LABS 10-13	ALL TEN LABS
BODY WEIGHT	**DDE 16**			
	DDE 160	**0.0001**	**0.0001**	**0.0001**
	DNP 10			
	FLUT 3			
	LIN 10			
	LIN 100	**0.0001**	**0.003**	**0.0001**
	NP 150			**0.03**
SV	**DDE 16**	**0.04**	**0.0001**	**0.0001**
	DDE 160	**0.0001**	**0.0001**	**0.0001**
	DNP 10			
	FLUT 3	**0.0001**	**0.0001**	**0.0001**
	LIN 10	**0.007**	**0.0001**	**0.0001**
	LIN 100	**0.0001**	**0.0001**	**0.0001**
	NP 150			
LABC	**DDE 16**	**0.0001**	**0.03**	**0.0001**
	DDE 160	**0.0001**	**0.0001**	**0.0001**
	DNP 10			
	FLUT 3	**0.0001**	**0.0001**	**0.0001**
	LIN 10	**0.0001**		**0.004**
	LIN 100	**0.0001**	**0.0001**	**0.0001**
	NP 150			
VP	**DDE 16**	**0.001**	**0.0001**	**0.0001**
	DDE 160	**0.0001**	**0.0001**	**0.0001**
	DNP 10		**0.009 ?**	
	FLUT 3	**0.0001**	**0.0001**	**0.0001**
	LIN 10	**0.02**	**0.0007**	**0.0002**
	LIN 100	**0.0001**	**0.0001**	**0.0001**
	NP 150		**0.04 ?**	**0.05**
GLANS PENIS	**DDE 16**	**0.0001**	**0.04**	**0.0001**
	DDE 160	**0.0001**	**0.0001**	**0.0001**
	DNP 10			
	FLUT 3	**0.0001**	**0.0001**	**0.0001**
	LIN 10	**0.04**		
	LIN 100	**0.0001**	**0.0001**	**0.0001**
	NP 150			
COWPER'S	**DDE 16**	**0.02**	**0.05**	**0.002**
	DDE 160	**0.0001**	**0.0001**	**0.0001**
	DNP 10			
	FLUT 3	**0.0001**	**0.0001**	**0.0001**
	LIN 10			
	LIN 100	**0.0001**	**0.0001**	**0.0001**
	NP 150			

Appendix 4.

Means and CVs of the androgen-dependent tissues presented for the two groups of laboratories (using different doses of TP) and for the ten individual laboratories. Blue shaded cells differ significantly from control (TP-treated) by $p < 0.01$ whereas yellow shaded cell differ by $p < 0.05$ but t-test from the LSMEANS.

Hershberger assay Phase 3 Antiandrogens
values are means and SEs

		p < 0.05	p <0.01

	Labs 1-9 Group 1	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
G1	TP.4	292	9.7	138	8.5	390	25	427	17	84.5	2.91	32.3	2	12.51	0.57	58.4	1.47	2146	64
G1	DDE 16	291	9	117	7.2	347	32	374	16	76.6	3	28.6	1.77	14.72	0.52	56.7	1.3	2154	59
G1	DDE 160	262	9.3	49.3	4.8	119	15	202	15	65	3.6	14.84	1.3	19.03	0.85	60.6	1.65	2099.7	68
G1	DNP 10	293	9.1	137	8.8	391	27	409	17	83.3	2.9	32.2	1.9	12.65	0.55	57.2	1.1	2106	60
G1	FLUT 3	278	10	36.4	3.4	87	10	214	12	64	3.6	10.8	0.94	12.25	0.72	59.7	1.7	2048	82
G1	LIN 10	295	9.6	123	7.7	336	22	386	15	81	3	30.76	1.75	12.87	0.65	55.9	1.2	2161	72
G1	LIN 100	266	7.7	73.4	5.8	181	17	254	12.25	74.2	3.2	18.75	1.2	11.69	0.54	60.9	1.7	2036	61
G1	NP 150	282	8.6	131	6.7	401	29	416	16	82.97	2.7	32.3	1.8	13.4	0.59	56.4	1.1	2244	67.3
G1	OIL	317	7.6	25.8	3.4	57.6	4.2	199	18.4	53.3	1.7	8.4	1.45	12.58	0.36	65.4	3	2154	40

	Labs 10-13 Group 2	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver					
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE				
G2	TP.2	305	7.7	135	8	378	35	440	37	74	3.9	27.48	2.2	13.17	0.67				
G2	DDE 16	305	8.4	103	8	270	27	382	24	70	3.3	23.4	1.9	15.25	0.6				
G2	DDE 160	255	11	30.4	2.2	85.6	6.7	163	12.8	45.2	3.5	7.8	0.9	19.2	0.87				
G2	DNP 10	307	7	116	7.5	346	30	442	26	75.3	3	26.9	2	13.25	0.45				
G2	FLUT 3	301	8.2	31	2.1	56.8	5.4	202	12.2	46.9	3.1	9	1	12.87	0.68				
G2	LIN 10	303	8.2	110	7.1	276	24	429	26	75.9	2.8	26.2	2.1	12.9	0.57				
G2	LIN 100	283	7.3	56.9	5.4	133	14.9	251	20	59.6	3.6	17.3	2.1	12.2	0.5				
G2	NP 150	299	7.7	120	8	345	32	436	27	74.6	2.8	26.7	2.4	13.75	0.6				
G2	OIL	292	6.7	17.6	0.9	41.8	3.4	177	10.8	43.4	2.5	5.7	0.64	12.21	0.44				

Group 1

Lab 1	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1 TP.4	216	3.1	87.1	4.3	178	16	318	11	68	4.1	21.7	1.2	6.58	0.18	61.8	2.4	1756	25
1 DDE 16	207	4.7	72.3	5.3	113	23	273	29	62	3.1	19.8	3.1	9.4	0.16	55.6	3.8	1785	85
1 DDE 160	200	3.7	29.4	2.3	41.5	7	142	12	45	3.1	9.3	0.8	12.3	0.29	56.1	3.96	1788	27.5
1 DNP 10	211	4	109	24	206	30	303	20	66	2.7	21.8	2.3	6.3	0.2	59.2	2.1	1595	34
1 FLUT 3	201	3.8	26	1	34	3.1	152	7.8	47.2	3.8	9.3	1.2	6.16	0.2	58.2	2.4	1609	101
1 LIN 10	204	4.5	64	7	137	23	265 a	17	61	1.7	21.6	1.8	6.48	0.29	55.9	2.5	1628	37
1 LIN 100	196	3.9	43.4	2.2	58.4	4.4	179	11	57.2	2.2	12.8	0.5	6.33	0.17	57.5	3.5	1563	49
1 NP 150	204	3.2	99	4.3	212	17	342	21	70	2.6	26	2.5	6.9	0.21	58	2.6	1746	40

Lab 2	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
2 TP.4	338	14	185	27	540	62	467	18	92.2	6.1	40.1	7.2	13.5	0.7	57.6	3.2	2268	62
2 DDE 16	345	7	178	13	592	122	406	24	81	2.3	42.2	1.9	16	0.46	60.2	3.3	2392	47
2 DDE 160	262	11	51	9	69	8	113	8	59	3.5	12.4	1.4	19	1.2	69.6	7.2	2050	72
2 DNP 10	355	5.5	201	16	594	79	424	9.7	93	4.1	40	4.8	13.8	0.34	62	1.8	2339	99
2 LIN 10	351	7.7	196	6.6	423	30	430	21	90	4.9	48	2.9	14.2	0.6	62.2	1.9	2410	32
2 LIN 100	293	8.7	99	9.5	278	38	252	11	76.1	3	26	2.6	11	0.48	65	1.5	2082	92
2 NP 150	329	10	181	11	681	40	455	12	93	3.7	41.1	3.1	13.65	0.6	61.3	1.8	2474	78
2 OIL	323	11	15	2.9	43	3.4	122	4.2	54.1	4.2	4.25	0.8	11.8	0.84	78.4	5.2	2135	87.3

Lab 3	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
3 TP.4	246	8	78	8.5	198	17	339	17	71.3	5.1	15.5	1.6	11.1	0.4
3 DDE 16	246	6.1	64	8.5	181	25	282	25	62	3.9	13.3	0.81	13.7	0.54
3 DDE 160	228	7.2	23.1	0.87	66.9	10.4	167	11.3	45.8	2.6	7.74	2.4	17.31	1.1
3 DNP 10	240	6.2	63	5.6	182	5	296	14	68.5	3.3	15	0.7	10.5	0.23
3 FLUT 3	246	8.2	17.3	1.8	54.6	5.1	152	6.6	46.9	2.6	5.8	0.3	11.1	0.61
3 LIN 10	245	8.3	72	6.3	216	22	320	19.4	67.7	3.6	16.6	1.1	10.9	0.28
3 LIN 100	233	8.7	36	4	85	8	184	8.2	56	1.6	10.4	0.6	10.74	0.48
3 NP 150	247	9.5	72	10	215	31	320	39	69	4.8	16.2	1.5	11.8	0.7

Lab 5	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
5 TP.4	245	6	129	8.5	358	23	341	29	75	3.1	30	4.1	9.65	0.28	52.5	3.8	1623	46
5 DDE 16	246	4.6	132	4.4	329	25	316	13	58.4	2.4	24	1.7	11.9	0.12	50.1	1.8	1658	51
5 DDE 160	221	14	36.6	6.1	60.1	6.5	117	17			7	0.9	16.34	1.3	51.1	1.3	1665	85
5 DNP 10	250	4.7	134	15	336	25	335	12	68.3	3	25.55	2.3	11	0.3	50	3.2	1715	50
5 FLUT 3	236	8	33	2.4	43.3	7.7	154	14			6	0.72	8.9	0.7	83	5.1	1515	58
5 LIN 10	242	10	144	10	354	22	323	10	67	3.6	24.6	2.2	9.1	1.1	45.5	2	1567	42
5 LIN 100	222	6.2	66	7	128	13	200	12			12.28	0.9	8.6	0.33	58.5	5.9	1618	57
5 NP 150	231	7	116	9.5	345	65	332	26	66	5	23.1	2.5	11.7	0.7	49.3	2	1731	70

Group 1 (continued)

Lab 7	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
7 TP.4	301	12	197	25	426	31	543	24	85.4	2.1	35.2	2.4	16	1	52.3	5.2	2450	122
7 DDE 16	288	11	139	17	378	20	544	38	85 .4	3.3	35	3.5	16	0.8	52.1	3.2	2215	52
7 DDE 160	259	22	70	15	160	20	287	34	67	4.1	18.8	1.9	20.3	2	56.4	2	2164	152
7 DNP 10	288	7	179	21	457	18	569	35	89	2.8	39	3	16.4	0.4	60	2.6	2344	56
7 FLUT 3	287	11	56	15	101	15	262	24	74	3.7	11	2.4	15.93	0.8	65.6	4.9	2335	66
7 LIN 10	314	10	153	16	447	30	533	37	89	1.8	32.7	2.6	17.3	0.65	55.7	2.3	2429	73
7 LIN 100	270	7.5	127	17	244	10	398	22	88	1.8	22.2	1	15.4	0.75	54.4	4.4	2283	64
7 NP 150	283	6.6	162	19	468	45	567	23	89.6	3.2	42.1	3.9	15.9	0.55	53.1	3.5	2471	104
7 OIL	286	5.7	40	6.6	56.5	5.2	181	14	53.2	2.5	4.7	0.9	12.7	0.5	56.1	1.5	2147	32

Lab 8	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
8 TP.4	349	6.6	147	5.6	464	43	592	22	80	2.6	43.8	3.2	13.3	0.4	64.7	1.7	2241	75
8 DDE 16	353	4.8	83	7.2	281	27	396	27.5	69.5	1.9	26	2.4	16.1	0.6	61.1	2.4	2337	69
8 DDE 160	316	17	36	3.7	103	16	305	15	60	2.6	17.6	2	19.7	1.3	64.3	0.6	2306	82
8 DNP 10	360	9.3	104	6	385	33	548	17	80.5	3	41.1	4.1	13.5	0.43	57.6	1.9	2242	60
8 FLUT 3	349	8.6	31.5	1.8	88.4	3.1	296	11.1	57.2	2.4	12.8	0.7	13.5	0.4	61	1.9	2243	87
8 LIN 10	363	7.8	92	8.6	253	23	443	14	73.2	1.9	31	2.4	14.5	0.6	58.5	2.6	2407	93
8 LIN 100	331	8.8	44.7	4.8	127	13.7	280	18.6	61.4	1.1	18.4	1.6	13.77	0.55	68.6	3.2	2227	66
8 NP 150	346	6	132	7	334	22	499	13	82	3.2	36.2	1.35	15.1	0.55	57.6	1.6	2426	26.5
8 OIL	343	10	22.6	0.85	72.9	7.3	293	15.6	52.6	2.1	16.3	1.2	13.3	0.4	61.8	2.6	2179	87

Lab 9	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidneys	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
9 TP.4	347	7.4	145	10	567	29	393	14	119	2.1	40	2	17.3	0.5	61.5	1.9	2538	77
9 DDE 16	349	4.5	153	3.9	560	27	401	11	112	2.5	40.3	2.5	19.9	0.24	61.3	1.8	2534	46
9 DDE 160	328	7	81	10	273	20.8	242	17	97	2.5	25.3	1.5	26.4	0.9	67.4	2	2591	66
9 DNP 10	346	4.8	166	9.8	578	22	385	12	117	2.3	43.3	1.8	17.05	0.41	59	2.7	2401	61
9 FLUT 3	347	8.2	55.1	3	201	6.4	269	6.4	94.8	1.6	20.1	1.2	18	0.6	60.7	3.1	2540	58
9 LIN 10	347	7.8	142	5.8	524	19	391	9.3	117	2.4	40.3	0.4	17.6	0.51	57.5	2.3	2526	94
9 LIN 100	316	4.1	97.5	4.5	348	17	290	8.6	107	1.6	29.5	1.5	16.1	0.5	61.5	4	2444	52
9 NP 150	337	8.8	156	6	551	26	393	7	111	2.2	42	1.5	18.8	0.9	59.4	1.5	2616	94

Group 2

Lab 10	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
10 TP.2	267	6.7	102	7	195	16	306	16	60.9	2.1	18.4	1.9	10.4	0.4
10 DDE 16	266	4.8	79	6	144	12	281	12	53.6	1.4	16.35	0.8	12.5	0.34
10 DDE 160	249	7	28.3	2.4	35.5	2.5	126	5.5	32	1.6	4.9	0.34	17.25	0.49
10 DNP 10	272	3.9	85.9	1.9	186	14	302	12	61	0.9	16.2	1.4	10.92	0.35
10 FLUT 3	261	5.1	24.85	1.49	28.2	1.5	148	7.6	30.3	1.7	4.7	0.5	10	0.33
10 LIN 10	264	4.8	81	4.4	155	13	303	17	64.5	2.5	18.1	1.3	10.4	0.2
10 LIN 100	246	4.5	41.4	2.4	60.6	3.3	156	8.9	41.7	2.6	8.3	0.7	9.55	0.35
10 NP 150	263	7.3	82	3.1	191	16	308	16	61.9	2.25	20.1	0.67	11.3	0.5
10 OIL	260	7.1	14.4	0.6	23.8	1.25	119	2.8	29.4	1.2	3.2	0.5	10.02	0.4

Lab 11	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
11 TP.2	332	5.8	145	8	479	30	495	85	68	5.6	28. 5	3 .3	15.94	0.9
11 DDE 16	331	9.3	128	20	357	40	486	34	72	1.2	31.8	3	17.7	0.8
11 DDE 160	271	22	28	4.4	78	13	218	20	50	0.8	10.5	1.9	22.8	1.8
11 DNP 10	332	8.5	129	13	440	28	530	13	75.2	0.9	31	1	15.1	0.35
11 FLUT 3	327	9	28.2	3.6	65.9	4.6	236	12	50	0.8	10.3	1.4	15.1	0.8
11 LIN 10	335	6	115	11	357	16	532	23	73	1.3	25.7	2.5	15.56	0.55
11 LIN 100	310	5.7	61.6	10	178	18	336	19	64	2.5	22.9	3.7	14.14	0.33
11 NP 150	326	6.7	135	5.3	406	29	540	22	75.6	2.4	28.5	2.3	15.8	0.9
11 OIL	313	6	17.7	1.1	48	3.4	218	10.3	50	1.2	5.65	0.35	13.76	0.29

red shading indicates a significant response in the wrong direction for an antiandrogen

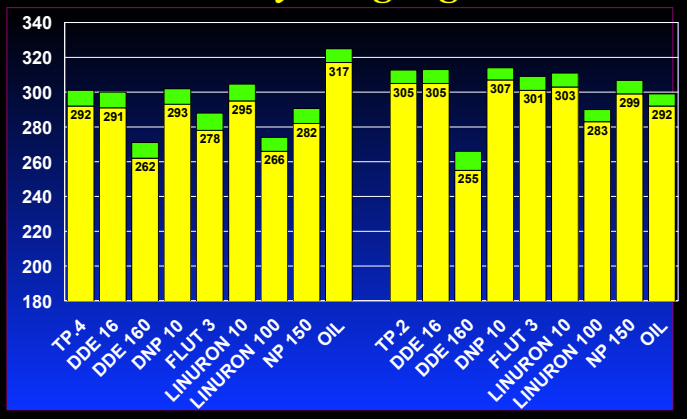
Lab 13	body weight		ventral prostate		seminal vesicle		LABC		Glans Penis		Cowper's glands		liver		adrenals		kidney	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
13 TP.2	315	8.5	159	13	460	40	518	30	93.5	1.9	35.6	2.7	13.65	0.7	49.5	3.4	2207	93
13 DDE 16	323	10	107	7	324	24	396	21	84	1.8	23.5	2.4	15.95	0.53	56.6	1.4	2329	81
13 DDE 160	248	31	36	4.7	74	4	165	18	60	3.7	9.3	1.3	18.6	1.3	55.5	1.8	2152	108
13 DNP 10	317	4.5	133	8	412	24	494	20	90	3.1	33.4	2.2	13.7	0.22	58	2.4	2342	14
13 FLUT 3	313	10	40	2	76.3	4.4	222	21	61	1.4	12	1.5	13.5	1	55.5	3.1	2194	82
13 LIN 10	311	9	134	10	317	25	452	26	90	2.5	35	2.7	12.8	0.6	56.5	3.2	2184	81
13 LIN 100	293	7.4	68	10	160	18	263	20	73.2	4	20.9	1.9	13	0.35	63.2	2.8	2236	54
13 NP 150	307	10	142	14	437	49	460	35	86.3	3.2	37.5	4.5	14.2	0.7	54.3	5.1	2316	111
13 OIL	303	7	21	1.4	54	2.1	193	4.6	50.9	1.6	8.3	1.1	12.9	0.5	57	4.4	2106	87

Appendix 5.

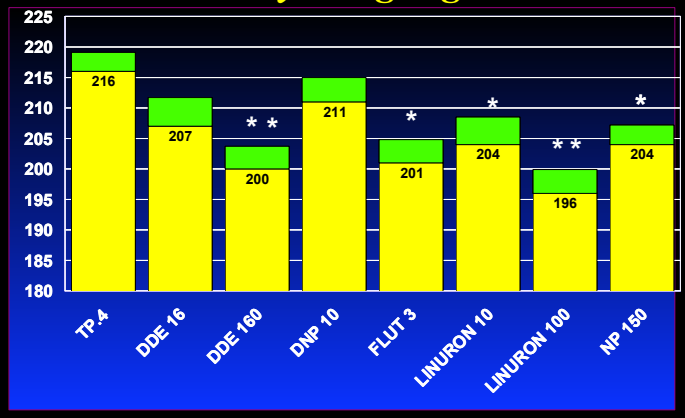
Histograms of the means and SEs prepared with Lotus Freelance are provided here by endpoint for the two groups of laboratories and for the ten individual laboratories. These graphs display the remarkably uniform profile of effects among laboratories.

Body weight

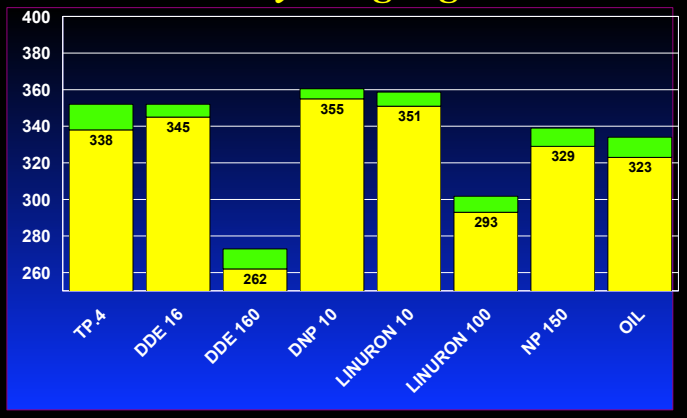
Phase 3 Antiandrogen All Labs
Body Weight g



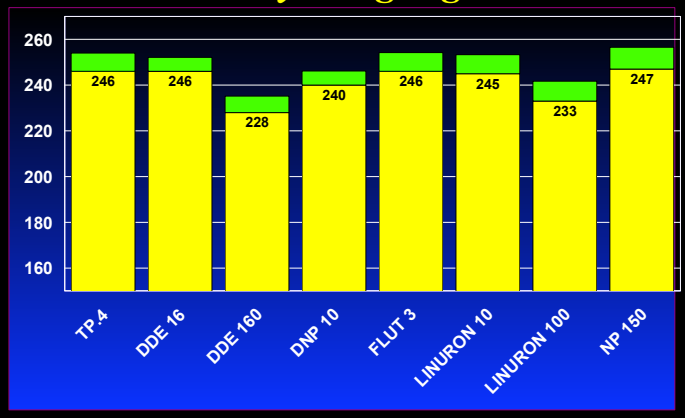
Phase 3 Antiandrogen Lab 1 data
Body Weight g



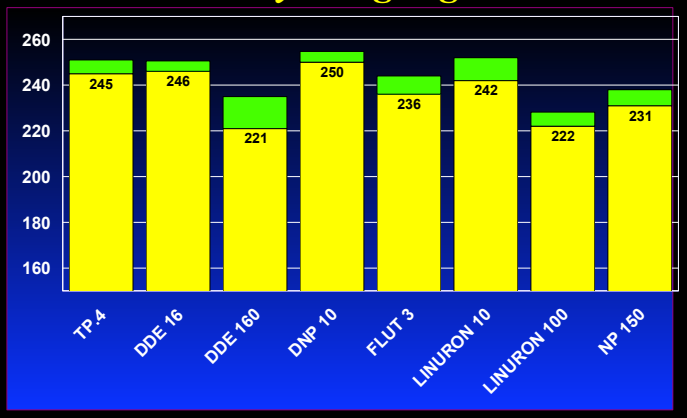
Phase 3 Antiandrogen Lab 2 data
Body Weight g



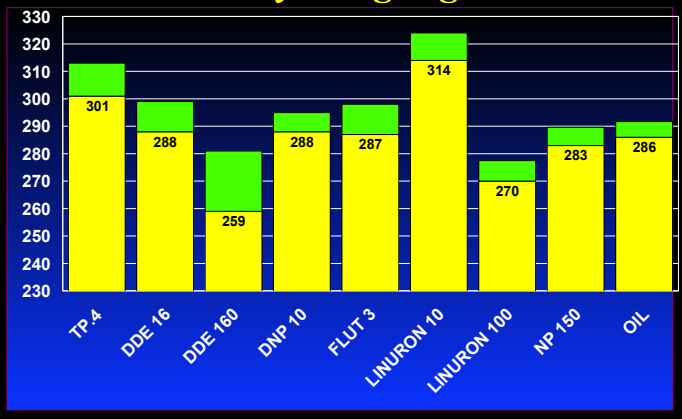
Phase 3 Antiandrogen Lab 3 data
Body Weight g



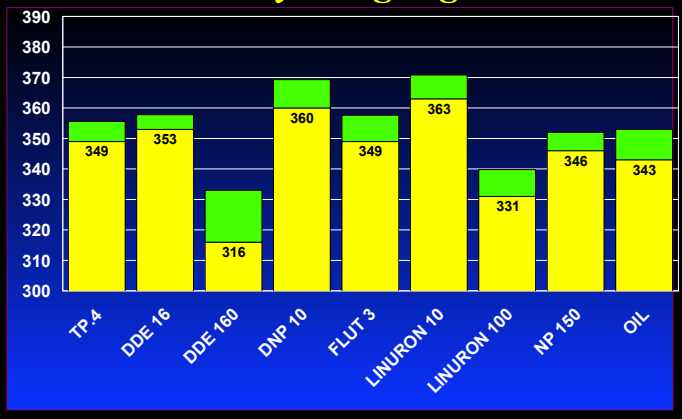
Phase 3 Antiandrogen Lab 5 data
Body Weight g



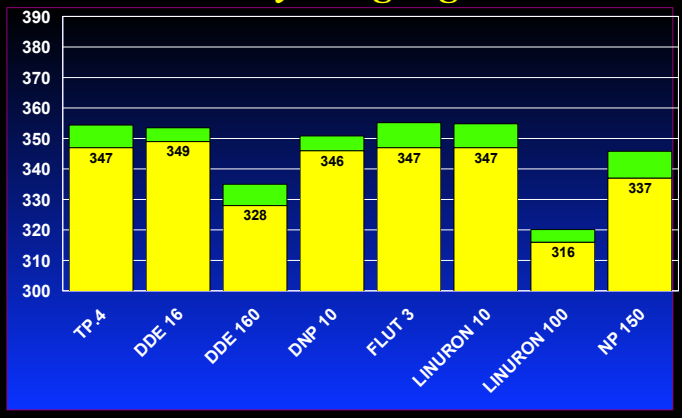
Phase 3 Antiandrogen Lab 7 data
Body Weight g



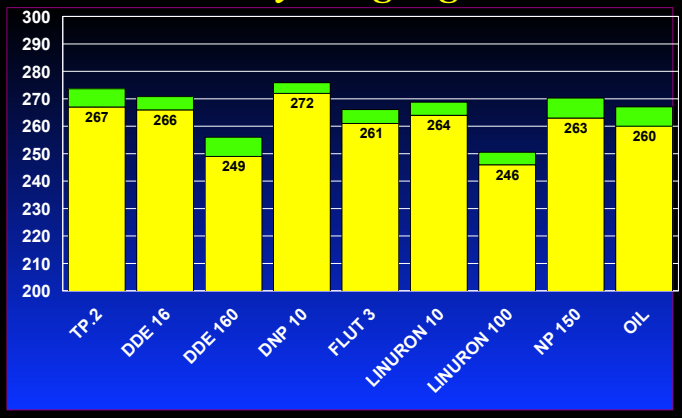
Phase 3 Antiandrogen Lab 8 data
Body Weight g



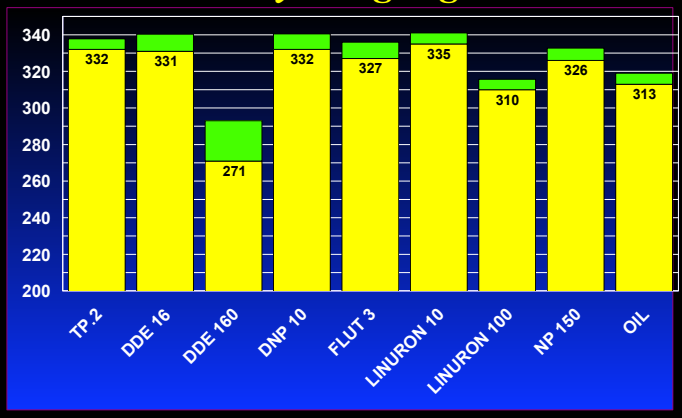
Phase 3 Antiandrogen Lab 9 data
Body Weight g



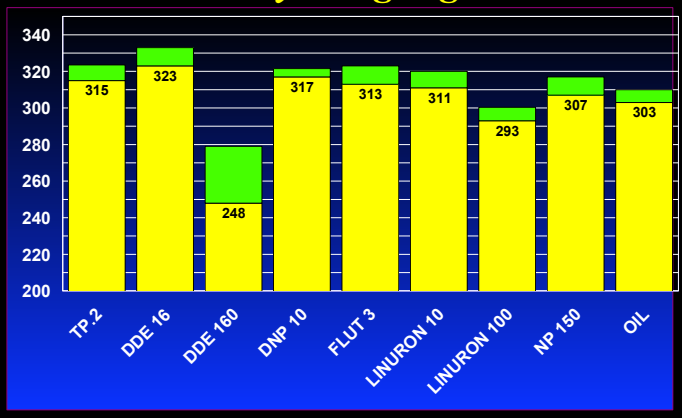
Phase 3 Antiandrogen Lab 10 data
Body Weight g



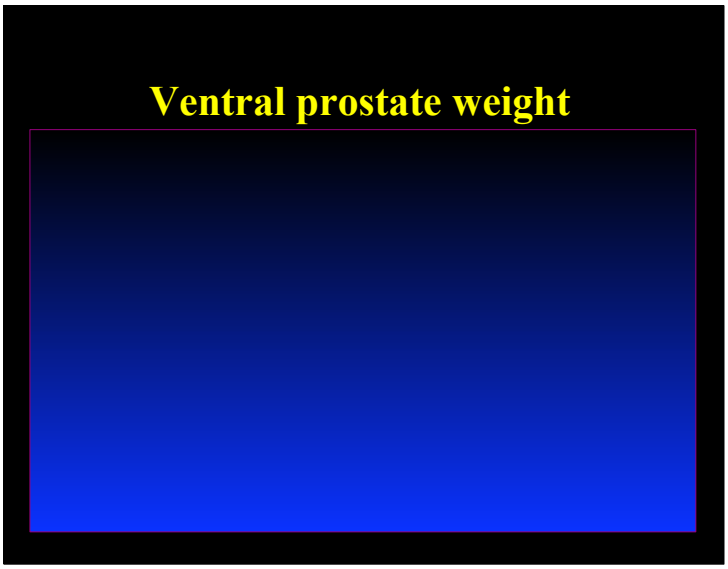
Phase 3 Antiandrogen Lab 11 data
Body Weight g



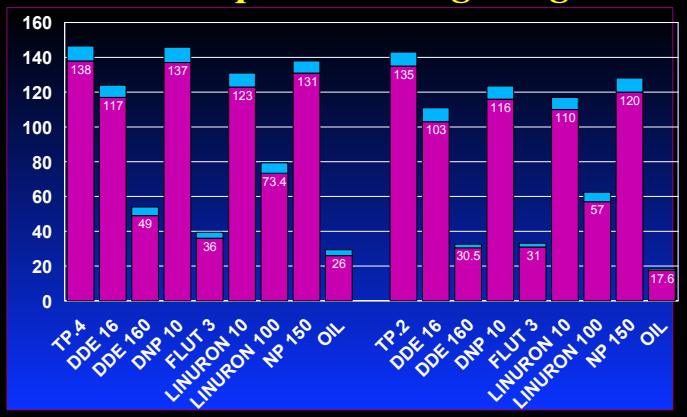
Phase 3 Antiandrogen Lab 13 data
Body Weight g



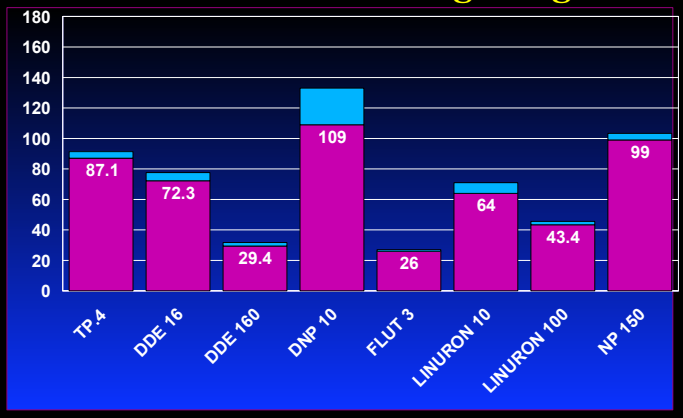
Ventral prostate weight



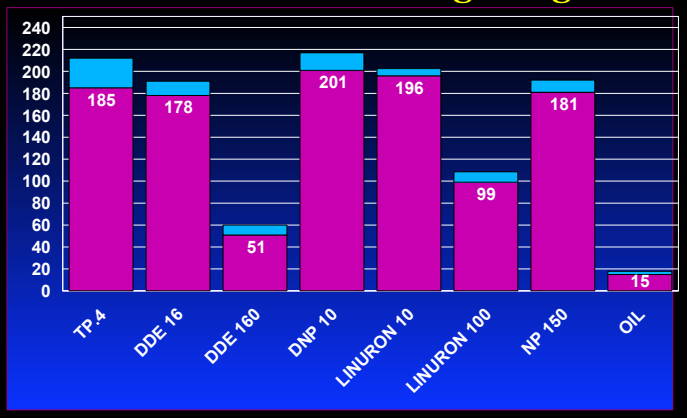
Phase 3 Antiandrogen All Labs
Ventral prostate Weight mg



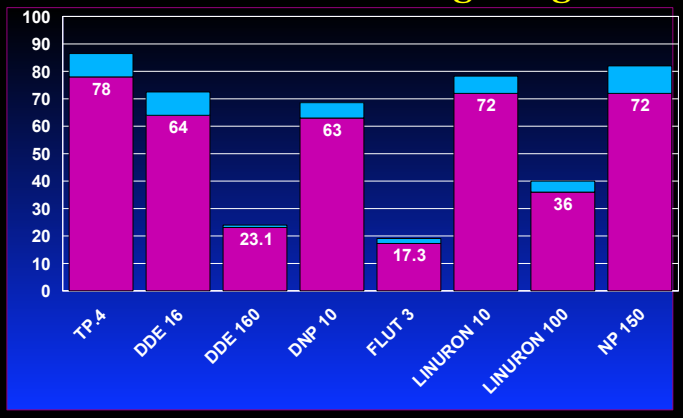
Phase 3 Antiandrogen Lab 1 data
Ventral Prostate Weight mg



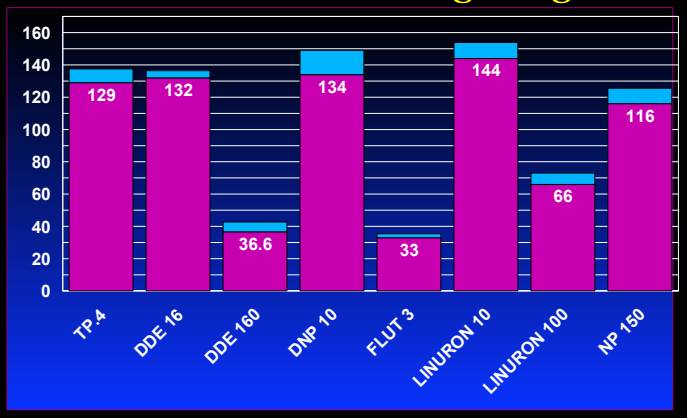
Phase 3 Antiandrogen Lab 2 data
Ventral Prostate Weight mg



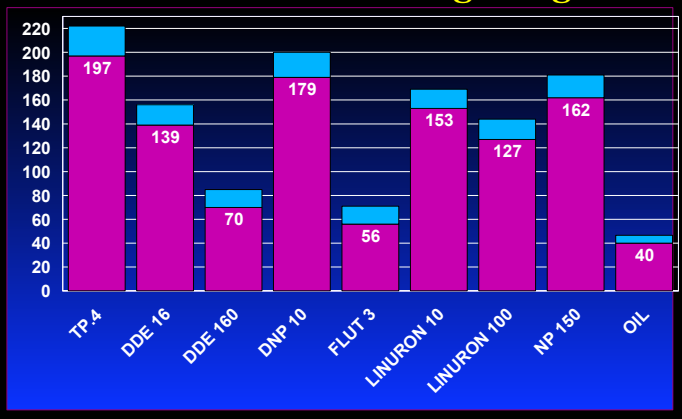
Phase 3 Antiandrogen Lab 3 data
Ventral Prostate Weight mg



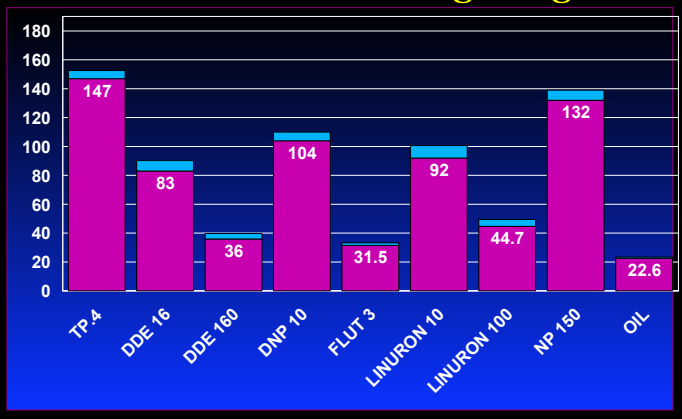
Phase 3 Antiandrogen Lab 5 data
Ventral Prostate Weight mg



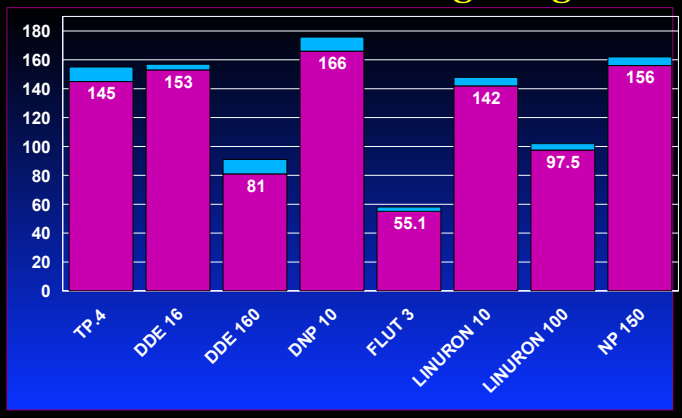
Phase 3 Antiandrogen Lab 7 data
Ventral Prostate Weight mg



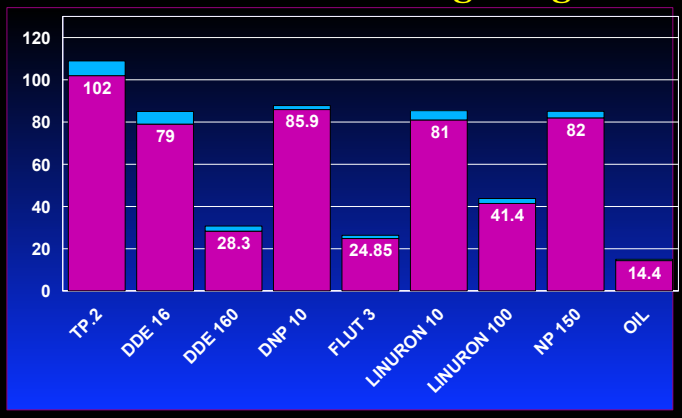
Phase 3 Antiandrogen Lab 8 data
Ventral Prostate Weight mg



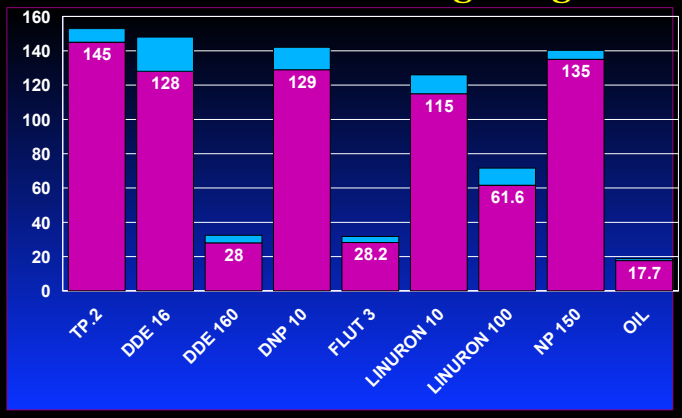
Phase 3 Antiandrogen Lab 9 data
Ventral Prostate Weight mg



Phase 3 Antiandrogen Lab 10 data
Ventral Prostate Weight mg



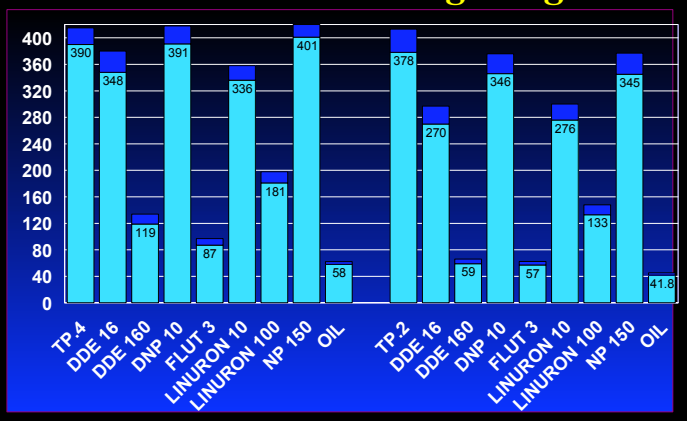
Phase 3 Antiandrogen Lab 11 data
Ventral Prostate Weight mg



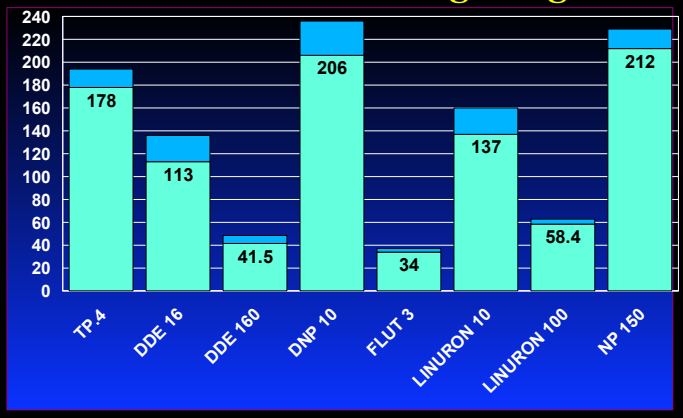
Phase 3 Antiandrogen Lab 13 data
Ventral Prostate Weight mg

Seminal Vesicle weight data

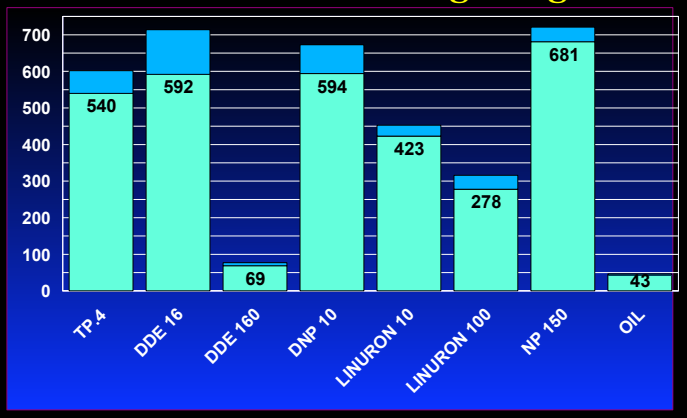
Phase 3 Antiandrogen All Labs
Seminal vesicle Weight mg



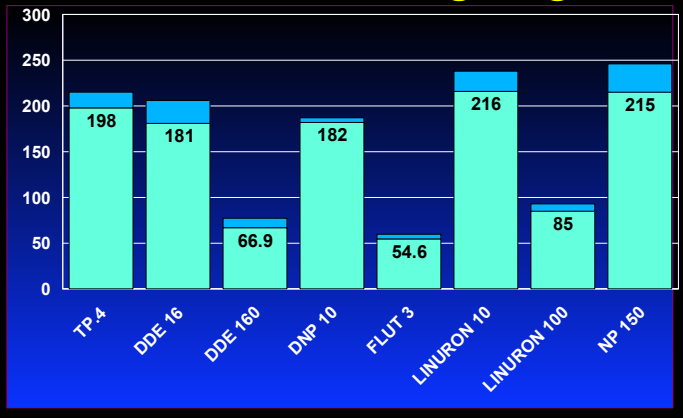
Phase 3 Antiandrogen Lab 1 data
Seminal Vesicle Weight mg



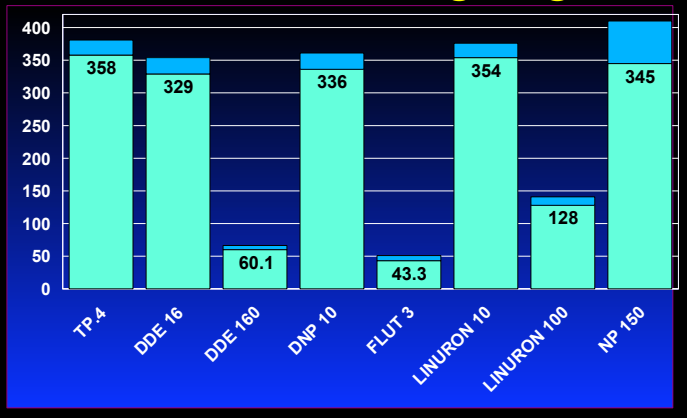
Phase 3 Antiandrogen Lab 2 data
Seminal Vesicle Weight mg



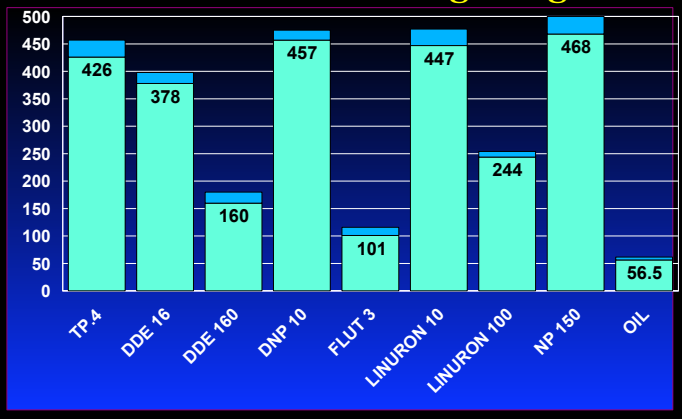
Phase 3 Antiandrogen Lab 3 data
Seminal Vesicle Weight mg



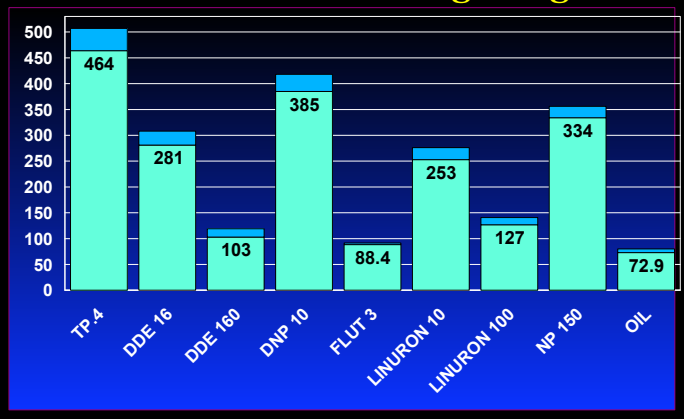
Phase 3 Antiandrogen Lab 5 data
Seminal Vesicle Weight mg



Phase 3 Antiandrogen Lab 7 data
Seminal Vesicle Weight mg

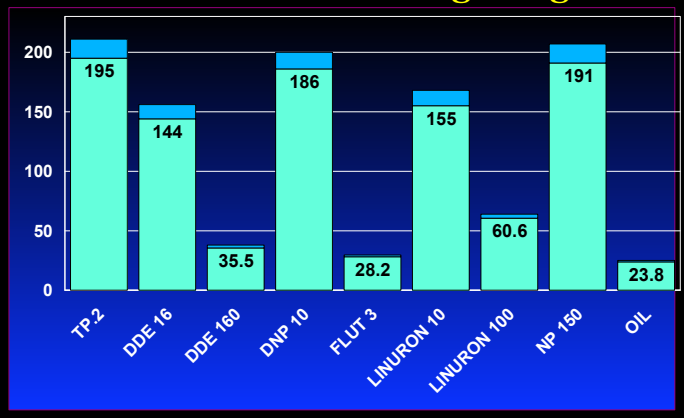


Phase 3 Antiandrogen Lab 8 data
Seminal Vesicle Weight mg

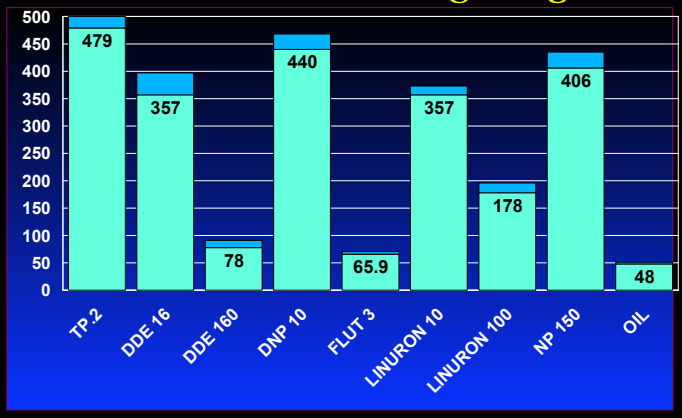


Phase 3 Antiandrogen Lab 9 data
Seminal Vesicle Weight mg

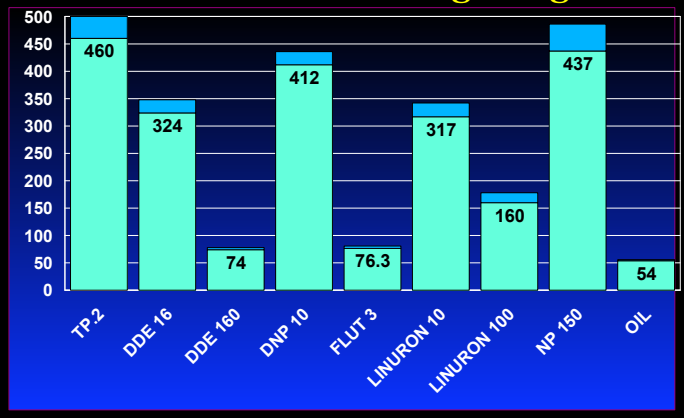
Phase 3 Antiandrogen Lab 10 data
Seminal Vesicle Weight mg



Phase 3 Antiandrogen Lab 11 data
Seminal Vesicle Weight mg



Phase 3 Antiandrogen Lab 13 data
Seminal Vesicle Weight mg

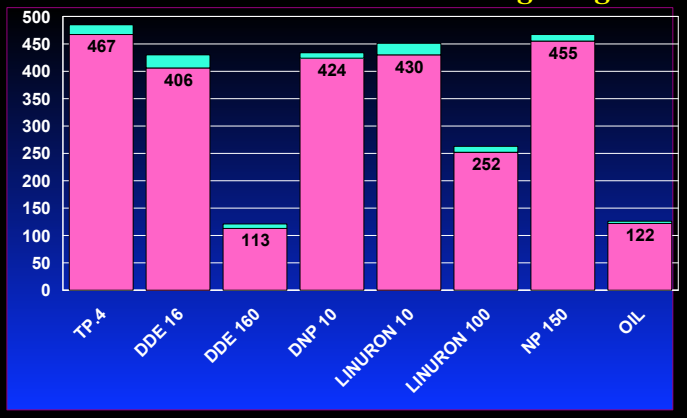


LABC weight data

Phase 3 Antiandrogen All Labs
LABC Weight mg

Phase 3 Antiandrogen lab 1 data
Levator Ani-bulbocavernosus Weight mg

Phase 3 Antiandrogen lab 2 data
Levator Ani-bulbocavernosus Weight mg



Phase 3 Antiandrogen lab 3 data
Levator Ani-bulbocavernosus Weight mg

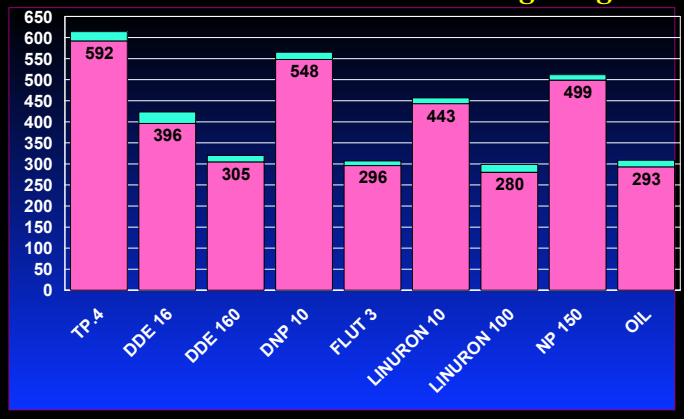
Phase 3 Antiandrogen lab 5 data
Levator Ani-bulbocavernosus Weight mg

Phase 3 Antiandrogen lab 7 data

Levator Ani-bulbocavernosus Weight mg

Phase 3 Antiandrogen lab 8 data

Levator Ani-bulbocavernosus Weight mg



Phase 3 Antiandrogen lab 9 data

Levator Ani-bulbocavernosus Weight mg

Phase 3 Antiandrogen lab 10 data

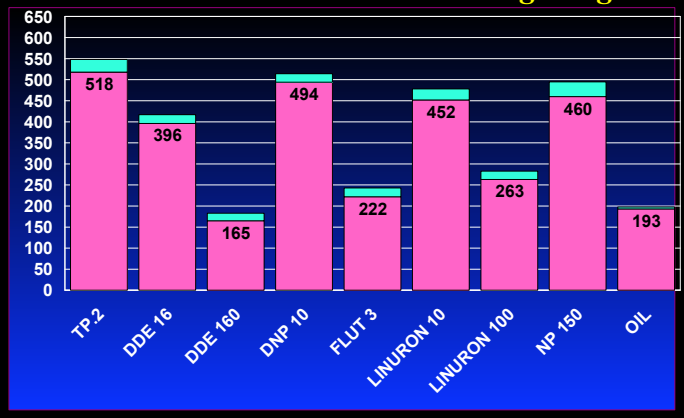
Levator Ani-bulbocavernosus Weight mg

Phase 3 Antiandrogen lab 11 data

Levator Ani-bulbocavernosus Weight mg

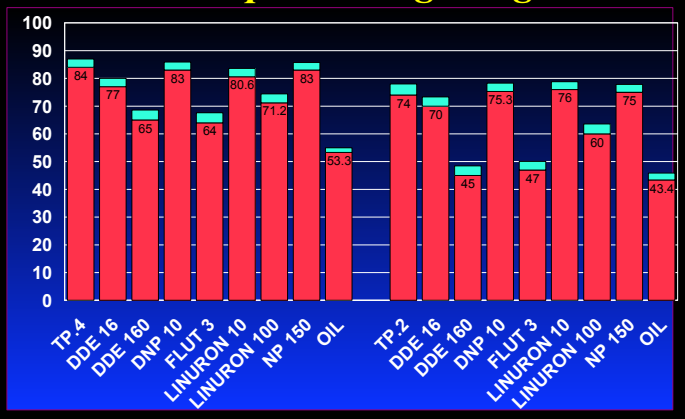
Phase 3 Antiandrogen lab 13 data

Levator Ani-bulbocavernosus Weight mg

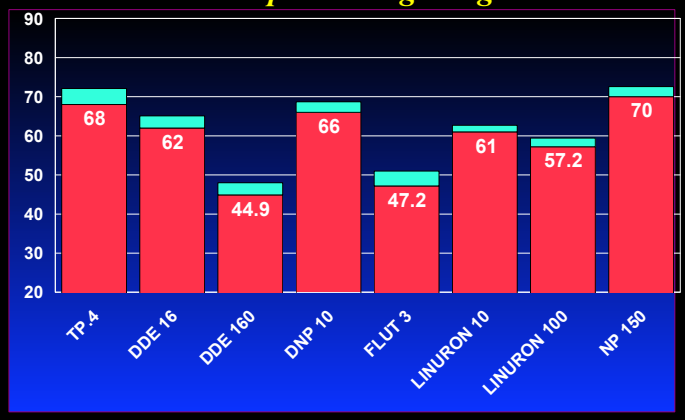


Glans penis weight data

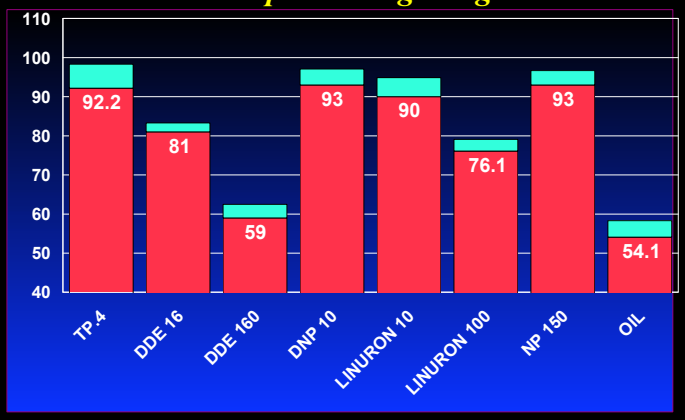
Phase 3 Antiandrogen All Labs
Glans penis Weight mg



Phase 3 Antiandrogen Lab 1 data
Glans penis Weight mg

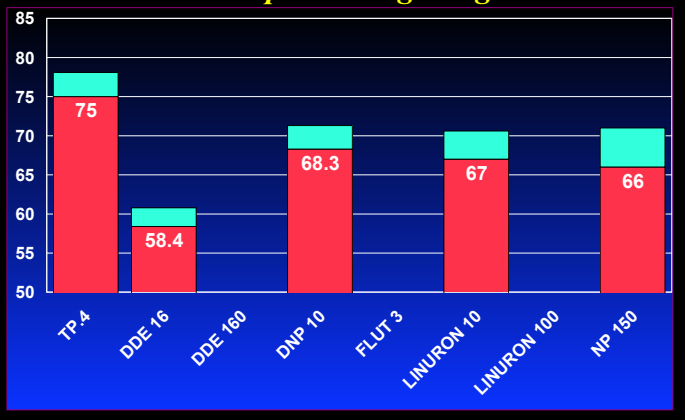


Phase 3 Antiandrogen Lab 2 data
Glans penis Weight mg



Phase 3 Antiandrogen Lab 3 data
Glans penis Weight mg

Phase 3 Antiandrogen Lab 5 data
Glans penis Weight mg

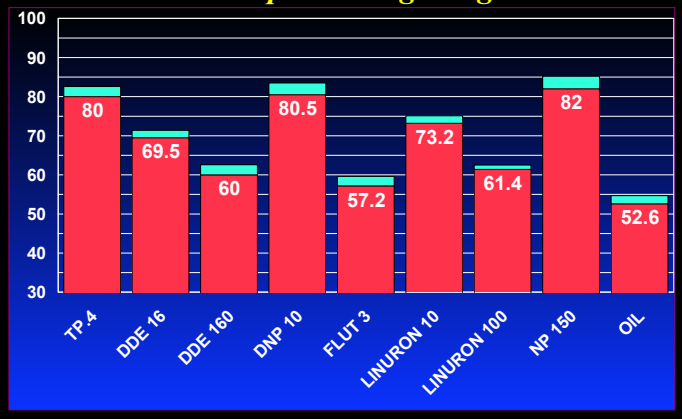


Phase 3 Antiandrogen Lab 7 data

Glans penis Weight mg

Phase 3 Antiandrogen Lab 8 data

Glans penis Weight mg

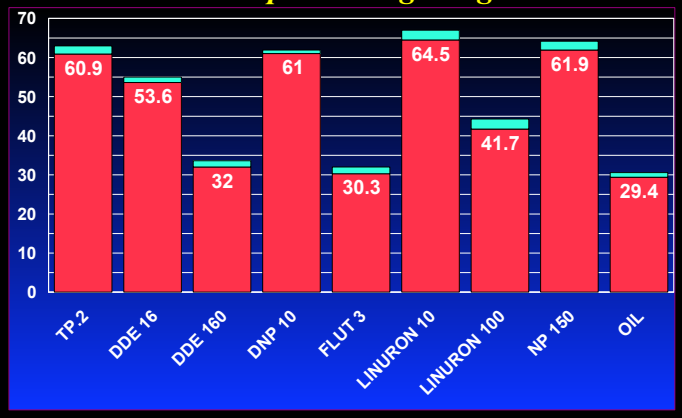


Phase 3 Antiandrogen Lab 9 data

Glans penis Weight mg

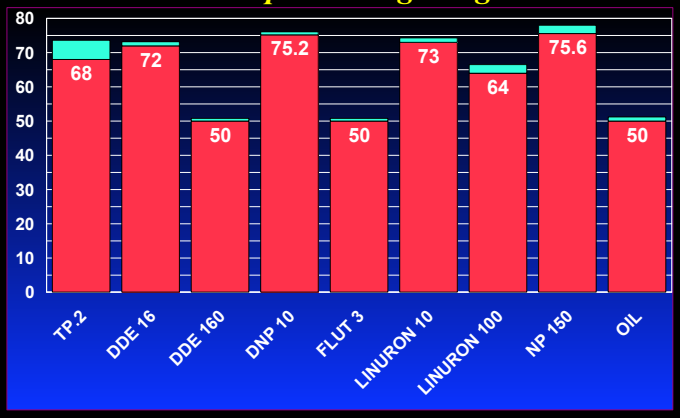
Phase 3 Antiandrogen Lab 10 data

Glans penis Weight mg



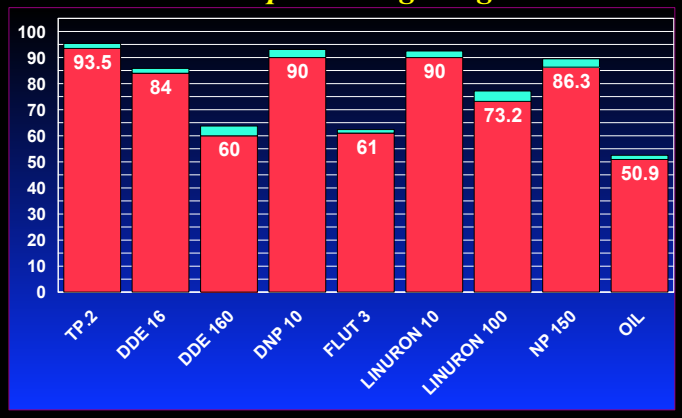
Phase 3 Antiandrogen Lab 11 data

Glans penis Weight mg



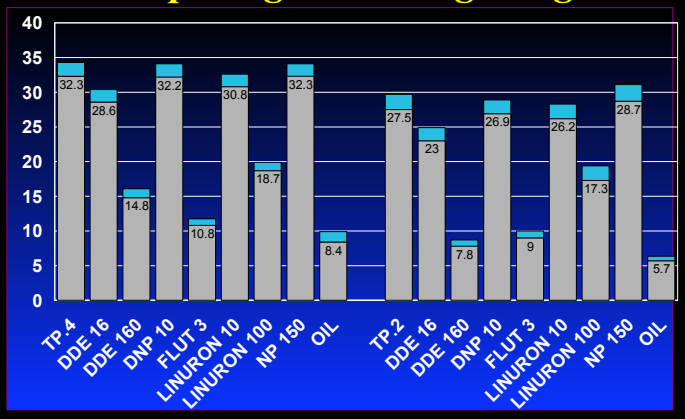
Phase 3 Antiandrogen Lab 13 data

Glans penis Weight mg

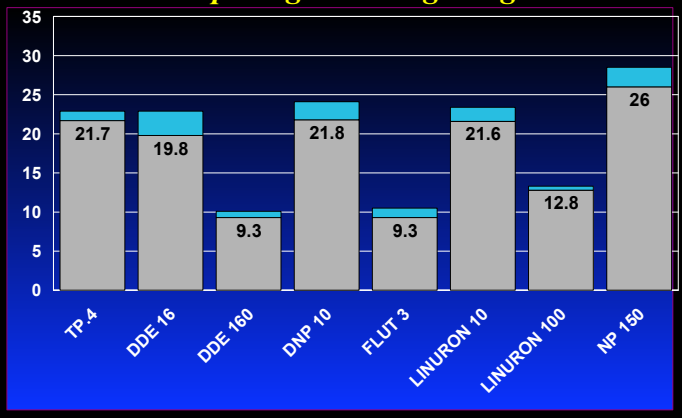


Cowper's gland weight data

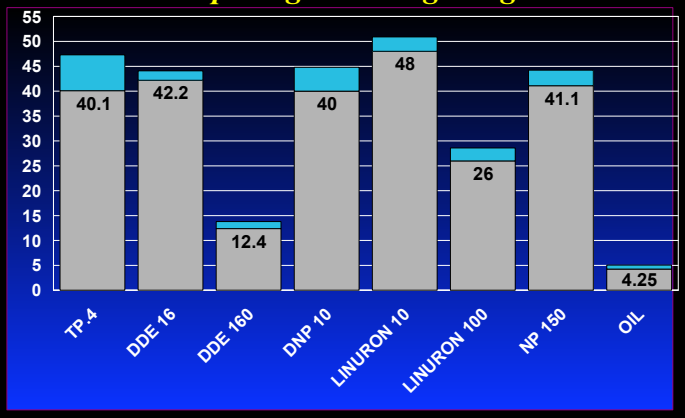
Phase 3 Antiandrogen All Labs
Cowper's glands Weight mg



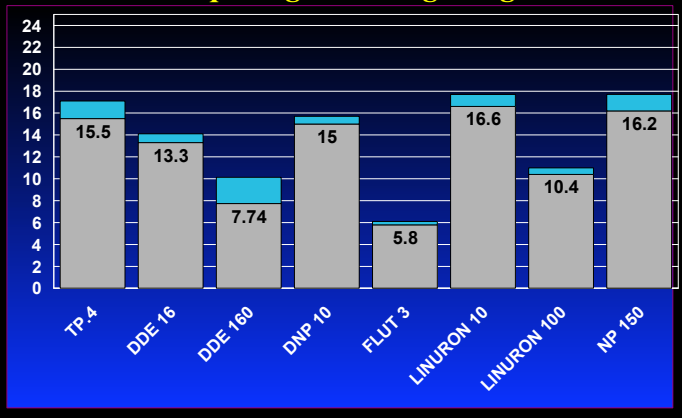
Phase 3 Antiandrogen Lab 1 data
Cowper's gland Weight mg



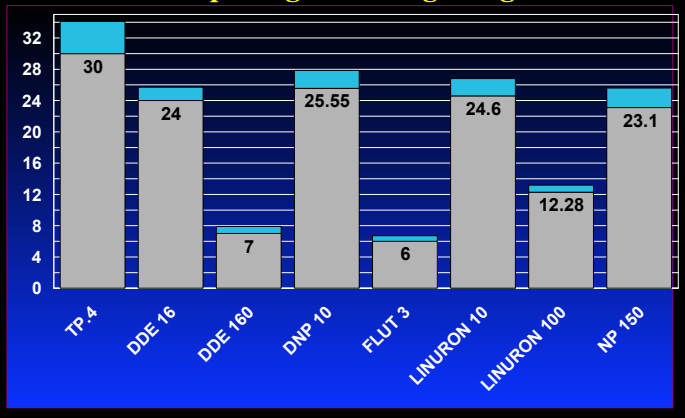
Phase 3 Antiandrogen Lab 2 data
Cowper's gland Weight mg



Phase 3 Antiandrogen Lab 3 data
Cowper's gland Weight mg

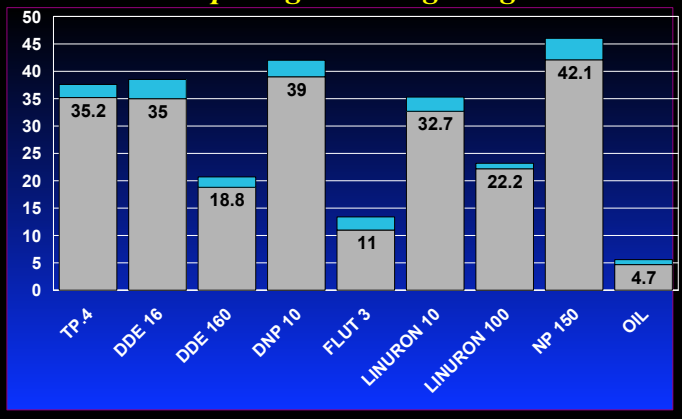


Phase 3 Antiandrogen Lab 5 data
Cowper's gland Weight mg



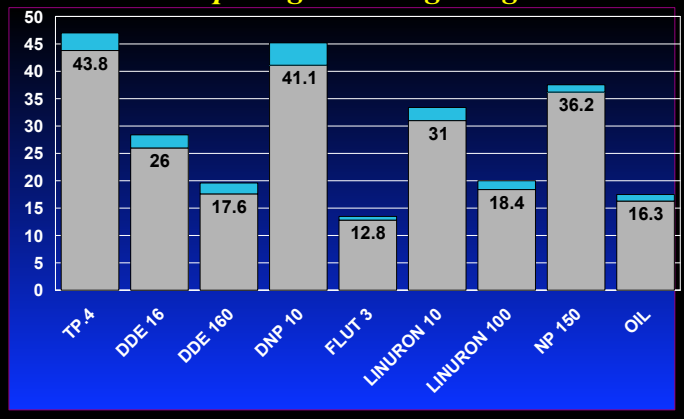
Phase 3 Antiandrogen Lab 7 data

Cowper's gland Weight mg



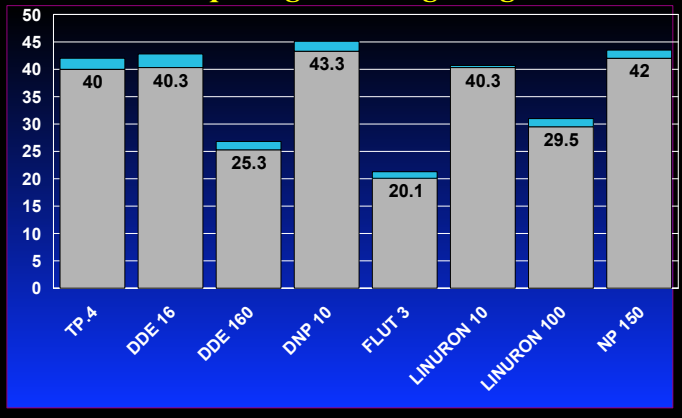
Phase 3 Antiandrogen Lab 8 data

Cowper's gland Weight mg



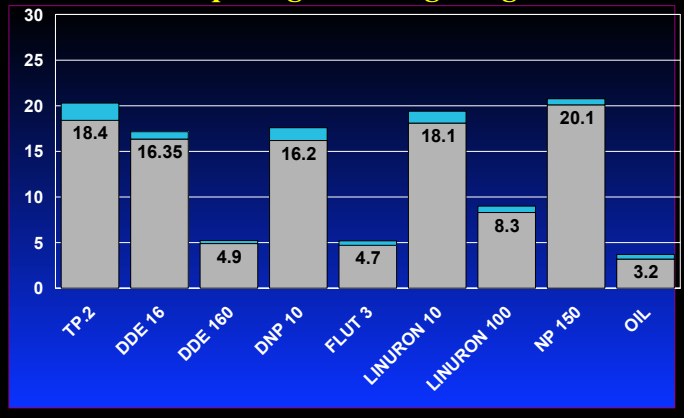
Phase 3 Antiandrogen Lab 9 data

Cowper's gland Weight mg



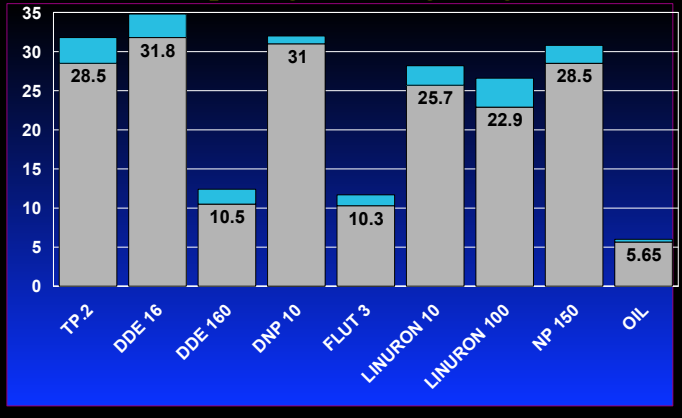
Phase 3 Antiandrogen Lab 10 data

Cowper's gland Weight mg



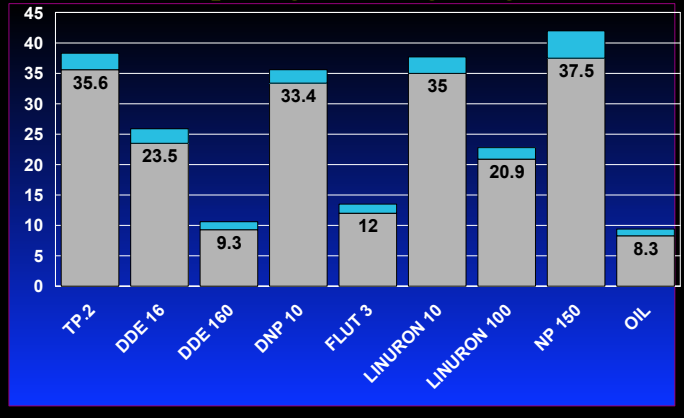
Phase 3 Antiandrogen Lab 11 data

Cowper's gland Weight mg



Phase 3 Antiandrogen Lab 13 data

Cowper's gland Weight mg



Appendix 6.

The data are summarized here for each endpoint for the ten laboratories individually in order to display the R2 and CV values in a manner that the strength of the chemical effects (R2) and precision (CV) of the data can be easily compare among the laboratories for consistency. There were no unexplained low R2 values and no unusually large CVs.

body weight

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9		lab 10	Lab 11	lab 13
TP.4	216	338	246	245	301	349	347	**TP.2**	267	332	315
DDE 16	207	345	246	246	288	353	349	**DDE 16**	266	331	323
DDE 160	200 b	262	228	221	259	316	328	**DDE 160**	249	271	248
DNP 10	211	355	240	250	288	360	346	**DNP 10**	272	332	317
FLUT 3	201 a	nd	246	236	287	349	347	**FLUT 3**	261	327	313
INURON 1	204 a	351	245	242	314	363	347	**LINURON**	264	335	311
NURON 10	196 b	293	233	222	270	331	316	**LINURON**	246	310	293
NP 150	204	329	247	231	283	346	337	**NP 150**	263	326	307
Oil	nd	323	nd	nd	286	343	nd	**OIL**	260	313	303
R2	33	68	11	25	26	32	35	**R2**	32	46	39
CV	4.6	6.8	7.9	8.3	10	6	5	**CV**	5	7	9

ventral prostate

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9		lab 10	Lab 11	lab 13
TP.4	87.1	185	78	129	197	147	145	**TP.2**	102	145	159
DDE 16	72.3	178	64	132	139	83	153	**DDE 16**	79	128	107
DDE 160	29.4	51	23.1	36.6	70	36	81	**DDE 160**	28.3	28	36
DNP 10	109	201	63	134	179	104	166	**DNP 10**	85.9	129	133
FLUT 3	26	nd	17.3	33	56	31.5	55.1	**FLUT 3**	24.85	28.2	40
INURON 1	64	196	72	144	153	92	142	**LINURON**	81	115	134
NURON 10	43.4	99	36	66	127	44.7	97.5	**LINURON**	41.4	61.6	68
NP 150	99	181	72	116	162	132	156	**NP 150**	82	135	142
Oil	nd	15	nd	nd	40	22.6	nd	**OIL**	14.4	17.7	21
R2	71	68	68	83	56	90	85	**R2**	86	80	81
CV	31	23	30	22	33	17	14	**CV**	18	25	22

Seminal vesicle

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9		lab 10	Lab 11	lab 13
TP.4	178	540	198	358	426	464	567	**TP.2**	195	479	460
DDE 16	113 a	592	181	329	378	281	560	**DDE 16**	144	357	324
DDE 160	41.5 b	69	66.9	60.1	160	103	273	**DDE 160**	35.5	78	74
DNP 10	206	594	182	336	457	385	578	**DNP 10**	186	440	412
FLUT 3	34 b		54.6	43.3	101	88.4	201	**FLUT 3**	28.2	65.9	76.3
INURON 1	137	423	216	354	447	253	524	**LINURON**	155	357	317
NURON 1(	58.4 b	278	85	128	244	127	348	**LINURON**	60.6	178	160
NP 150	212	681	215	345	468	334	551	**NP 150**	191	406	437
Oil		43			56.5	72.9		**OIL**	23.8	48	54
R2	73	61	71	80	85	84	89	**R2**	87	89	83
CV	36	34	29	29	19	24	12	**CV**	22	19	24

LABC

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9		lab 10	Lab 11	lab 13
TP.4	318	467	339	341	543	592	393	**TP.2**	306	495	518
DDE 16	273	406	282	316	544	396	401	**DDE 16**	281	486	396
DDE 160	142 b	113	167	117	287	305	242	**DDE 160**	126	218	165
DNP 10	303	424	296	335	569	548	385	**DNP 10**	302	530	494
FLUT 3	152 b		152	154	262	296	269	**FLUT 3**	148	236	222
INURON 1	265 a	430	320	323	533	443	391	**LINURON**	303	532	452
NURON 1(	179 b	252	184	200	398	280	290	**LINURON**	156	336	263
NP 150	342	455	320	332	567	499	393	**NP 150**	308	540	460
Oil		122			181	293		**OIL**	119	218	193
R2	77	91	70	82	75	89	86	**R2**	88	70	83
CV	17	10	19	16	16	10	8	**CV**	13	21	16

Glans Penis

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9		lab 10	Lab 11	lab 13
TP.4	68	92.2	71.3	75	85.4	80	119	TP.2	60.9	68	93.5
DDE 16	62	81	62	58.5	85 .4	69.5	112	DDE 16	53.6	72	84
DDE 160	44.96	59	45.8	.	67	60	97	DDE 160	32	50	60
DNP 10	66	93	68.5	68.3	89	80.5	117	DNP 10	61	75.2	90
FLUT 3	47.2		46.9	.	74	57.2	94.8	FLUT 3	30.3	50	61
INURON 1	61	90	67.7	67	89	73.2	117	LINURON	64.5	73	90
NURON 1(	57.2	76.1	56	.	88	61.4	107	LINURON	41.7	64	73.2
NP 150	70	93	69	66.3	89.6	82	111	NP 150	61.9	75.6	86.3
Oil		54.1			53.2	52.6		OIL	29.4	50	50.9
R2	63	57	55	27	58	75	76	R2	90	74	80
CV	12	12	15	13	9	8	5	CV	9	9	8

Cowper's glands

	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9		lab 10	Lab 11	lab 13
TP.4	21.7	40.1	15.5	30	35.2	43.8	40	TP.2	18.4	28. 5	35.6
DDE 16	19.8	42.2	13.3	24	35	26	40.3	DDE 16	16.35	31.8	23.5
DDE 160	9.3	12.4	7.74	7	18.8	17.6	25.3	DDE 160	4.9	10.5	9.3
DNP 10	21.8	40	15	25.55	39	41.1	43.3	DNP 10	16.2	31	33.4
FLUT 3	9.3		5.8	6	11	12.8	20.1	FLUT 3	4.7	10.3	12
INURON 1	21.6	48	16.6	24.6	32.7	31	40.3	LINURON	18.1	25.7	35
NURON 1(	12.8	26	10.4	12.28	22.2	18.4	29.5	LINURON	8.3	22.9	20.9
NP 150	26	41.1	16.2	23.1	42.1	36.2	42	NP 150	20.1	28.5	37.5
Oil		4.25			4.7	16.3		OIL	3.2	5.65	8.3
R2	65	58	71	77	73	80	83	R2	86	65	75
CV	26	26	20	26	22	21	12	CV	20	26	24