Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

28-Sep-2006

English - Or. English

ENV/JM/MONO(2006)30
Unclassified

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

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

English - Or. English

JT03214577

OECD Environment Health and Safety Publications

Series on Testing and Assessment

No. 62

FINAL OECD REPORT OF THE INITIAL WORK TOWARDS THE VALIDATION OF THE RAT HERSHBERGER ASSAY : PHASE-1, ANDROGENIC RESPONSE TO TESTOSTERONE PROPIONATE, AND ANTI-ANDROGENIC EFFECTS OF FLUTAMIDE

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris

September 2006

Also published in the Series on Testing and Assessment:

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995, revised 2006)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*

No. 20, *Revised Draft Guidance Document for Neurotoxicity Testing (in approval)*

No. 21, *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals (2000)*

No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

No. 24, *Guidance Document on Acute Oral Toxicity Testing (2001)*

No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*

No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*

No 27, *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals Which are Hazardous for the Aquatic Environment (2001)*

No 28, *Guidance Document for the Conduct of Skin Absorption Studies (2004)*

No 29, *Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*

No 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*

No 31, *Detailed Review Paper on Non-Genotoxic Carcinogens Detection: The Performance of In-Vitro Cell Transformation Assays (draft)*

No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)*

No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)*

No. 34, *Guidance Document on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment (in preparation)*

No. 35, *Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (2002)*

No. 36, *Report of the OECD/UNEP Workshop on the use of Multimedia Models for estimating overall Environmental Persistence and long range Transport in the context of PBTS/POPS Assessment (2002)*

No. 37, *Detailed Review Document on Classification Systems for Substances Which Pose an Aspiration Hazard (2002)*

No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)*

No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (in preparation)*

No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures Which Cause Respiratory Tract Irritation and Corrosion (2003)*

No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)*

No. 42, *Guidance Document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure (2003)*

No. 43, *Draft Guidance Document on Reproductive Toxicity Testing and Assessment (in preparation)*

No. 44, *Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)*

No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-range Transport (2004)*

No. 46, *Detailed Review Paper on Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances (2004)*

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 48, New Chemical Assessment Comparisons and Implications for Work sharing (2004)

No. 49, Report from the Expert Group on (Q)SARs on the Principles for the Validation of (Q)SARs (2004)

No. 50, *Report of the OECD/IPCS Workshop on Toxicogenomics (2005)*

No. 51, *Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)*

No. 52, *Comparison of emission estimation methods used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case study of pulp and paper and textile sectors (2006)*

No. 53, *Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (2006)*

No. 54, *Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (2006)*

No. 55, *Detailed Review Paper on Aquatic Arthropods in Life Cycle Toxicity Tests with an Emphasis on Developmental, Reproductive and Endocrine Disruptive Effects (2006)*

No. 56, *Guidance Document on the Breakdown of Organic Matter in Litter Bags (2006)*

No. 57, *Detailed Review Paper on Thyroid Hormone Disruption Assays (2006)*

No. 58, *Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals (2006)*

No. 59, *Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-Day Oral Toxicity Study in Laboratory Rats (2006)*

No. 60, *Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1A) (2006)*

No. 61, *Report of the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1B) (2006)*

1.

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 30 industrialised countries in North America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides and Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and the Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (http://www.oecd.org/ehs/).

This publication was produced within the framework of the Inter-Organisation Programme for the Sound Management of Chemicals (IOMC).

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The participating organisations are FAO, ILO, OECD, UNEP, UNIDO, UNITAR and WHO. The World Bank and UNDP are observers. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

FOREWORD

This document provides a description of Phase-1 of the validation of the rat Hershberger assay. It contains the background on how the validation study was organised and performed, the standardised protocols used, a comprehensive summary of test data and their analyses, and the conclusions drawn from the studies. The study was performed in two parts: a determination of the response and reproducibility of the assay with respect to a reference androgen agonist (Phase-1a), and the response and reproducibility when challenged with a reference androgen antagonist (Phase-1b). In addition, details are provided for the conduct and design of the next phase (Phase-2) in the validation of the assay. Extensive contributions to the report were made by Dr. L. Earl Gray Jr., who provided the initial data summaries and evaluations, and the statistical analyses, and Mr. Mike Walker who served as an independent statistician to confirm and extend the analysis. Dr. Errol Zeiger, consultant to the Secretariat, drafted major parts of the document.

The draft version of this document was submitted for review to the Validation Management Group for the Screening and Testing of Endocrine Disrupters for Mammalian Effects (VMG-mammalian) on 18th April 2002. The Validation Management Group was invited to either make specific comments on the Phase-1 report or to confirm it as sufficient to support progress to Phase-2. The Validation Management Group was further invited to consider the proposed design and protocol for Phase-2 of the validation procedure, and to provide comments and additional recommendations, as appropriate, for the experimental design and protocol(s) to be used. The VMG Mammalian was requested to provide any comments or suggestions before 21st May 2002.

Following the review, the Secretariat took into account all comments received before 28th June and revised the report of the Phase-1 and the recommendations for Phase-2 work accordingly. This document is therefore considered the final report of Phase-1 of the Hershberger assay validation study.

This report provides a comprehensive summary of the Phase-1 testing performed by the participating laboratories, including a detailed presentation and evaluation of their results. It has been automatically declassified, under the responsibility of the Secretary-General, after three years.

TABLE OF CONTENTS

SUMMARY

i) This report summarises the results from an OECD inter-laboratory study conducted in 2000-2001 to examine the reliability and transferability of a standardised protocol for the rat Hershberger assay. This study is considered the first phase in a process to validate the rodent Hershberger assay at the international level. The study was performed in two parts; the first (Phase-1a) involved 17 laboratories that measured the responses of five androgen-dependent tissues to a reference androgen. Other parameters were also measured. In the second part of the study (Phase-1b), seven laboratories that had participated in Phase-1a investigated the ability of the assay to measure the anti-androgenic effects of test substances.

ii) The need to validate the rodent Hershberger assay arises from the concerns that exist that ambient levels of chemicals may interact with the endocrine system to cause adverse effects in humans and wildlife. The evidence for endocrine disruption in humans as a result of exposure to xenobiotic chemicals is limited, but several cases have been reported where local, high level exposures have resulted in adverse effects in wildlife. In 1997, the OECD concluded that existing test methods were insufficient to identify such effects (1). As part of the OECD Test Guidelines Programme a *Special Activity on the Testing and Assessment of Endocrine Disrupters* was initiated to revise existing Guidelines, and develop new OECD Test Guidelines for the testing of potential endocrine disrupting chemicals (Further information concerning the OECD Endocrine Disruptor testing program can be found at http://www.oecd.org/document/62/0,2340, en_2649_34377_2348606_1_1_1_1,00.html). An OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) was subsequently established to provide a focal point within the OECD to identify and recommend priorities for the development and validation of methods for identifying endocrine-disrupting chemicals (2).

iii) The lead laboratory for this Phase-1 validation study was from the U.S. EPA. The participating laboratories were from: Denmark, France, Germany, Japan, Korea, U.K and U.S. Both government laboratories and laboratories from the private sector participated in the work. In addition, Health Canada, provided independent statistical consultation. Further details of participating laboratories are provided in Annex 1.

iv) The rodent Hershberger assay was one of three *in vivo* tests selected by the EDTA to start the international co-operative work. This selection supports the recommendations of the U.S.-EPA Endocrine Disrupter Screening and Testing Advisory Committee (EDSTAC) which, in 1998, recommended the rodent Hershberger assay as a component of its Tier-1 screen for endocrine-disrupting substances (3). The EDSTAC also recognized at that time that although the assay has been in use for many years, that it would have to be validated for use in the Tier-1 screen. The EDSTAC recommendations regarding this assay were subsequently adopted by the EPA for its proposed Endocrine Disruptor Screening Program (EDSP) (4).

v) The principle of the rodent Hershberger assay is that there are organs and tissues in the animal that are under the control of androgens which stimulate and to maintain growth. If the endogenous source of this hormone is not available, either because of immaturity of the animals, or because the animals have been castrated, the animal requires an exogenous source to initiate or restore growth of these tissues, and for normal sexual development. Chemicals that act as agonists may be identified as potential endocrine disrupters if they cause an increase in the weights or these androgen-dependent tissues, or as antagonists if they cause a relative decrease when co-administered with a potent androgen. The rodent Hershberger assay may also serve as a tool for the prioritisation of chemicals for further testing.

vi) The potent androgen, testosterone propionate (CASRN 57-82-5), and androgen antagonist, flutamide (CASRN 1311-84-7), were used as the reference test substances for the Phase-1 study.

Seventeen laboratories from seven Member countries contributed data on the effects of testosterone propionate, and seven of these laboratories also performed the anti-androgenicity studies using flutamide.

vii) As an indication of the androgenic effects of testosterone propionate, weight increases of the accessory sex organs and tissues - ventral prostate (VP), seminal vesicles plus coagulating glands (SV), levator ani and bulbocavernous muscle (LABC), glans penis (GLANS), and Cowper's glands (COWS) - were measured. Additionally, the effects of androgen administration on total body, liver, kidney, and adrenal weights were measured, as well as serum levels of testosterone and lutenizing hormone. Other parameters of the test protocol, regarding the weighing of fresh or fixed tissue, and use of the dorso-lateral prostate weight, were also examined. The same accessory sex organs and tissues were measured as an indication of the anti-androgenic effects of flutamide.

viii) All laboratories and all protocols were successful in detecting increases in the weights of the accessory sex organs and tissues in response to testosterone propionate, and in detecting the anti-androgenic effects of flutamide. There was good agreement among laboratories with regard to the dose responses obtained. There was similar agreement in their ability to identify the anti-androgenic effects of flutamide.

ix) It can be concluded from this first phase of the work that the protocol is robust, reliable and transferable across laboratories for potent androgen agonists and antagonists. Further work needs to be performed to confirm these findings and examine the sensitivity of the various endpoints measured in Phase-1 for identifying less potent androgens and androgen antagonists.

INTRODUCTION

1. The need to validate the rodent Hershberger assay stems from the concerns that exist that ambient environmental levels of chemicals may be causing adverse effects in humans and wildlife due to the interaction of these chemicals with the endocrine system (5)(6)(7)(8)(9). Initial reviews of existing reports have noted limited evidence for endocrine disruption in humans, but there are several reports that local, high level exposures to environmental pollutants have resulted in endocrine-related effects in wildlife (10)(11)(12)(13).

2. The OECD initiative to develop and validate *in vitro* and *in vivo* assays for the detection of chemicals that may interfere with the endocrine response was taken following the recommendations of a number of national, regional and international workshops (6)(7)(8)(10)(14) and following a detailed OECD review of the status of existing test and research methods. This review produced a Detailed Review Paper on test methods for sex hormone disrupting chemicals (1). As part of the OECD Test Guidelines Programme, a *Special Activity on the Testing and Assessment of Endocrine Disrupters* was initiated to revise existing, and develop new OECD Test Guidelines for the testing of potential endocrine disrupters [http://www.oecd.org/ehs/endocrin.htm].

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

4. The OECD framework is designed to develop these assays as individual, multipurpose tools, rather than as part of a rigid testing scheme. The uses of a bioassay for endocrine effects may vary depending on the chemical substance and its available toxicological data. An early screen for one test substance could, for another, be a means to determine the test substance's mode of action (2). The Hershberger assay, once validated, would fit within this framework.

5. The rodent Hershberger assay is based on the principle that a number of accessory sex tissues require androgens to stimulate and to maintain growth. If the endogenous source of this hormone is not available the animal requires an exogenous source to initiate and/or restore the growth of these tissues.

6. The aim of the validation program for the Hershberger assay is to develop a robust, reliable, and relevant test method for the detection of chemicals that have the potential to act like, and consequently interfere with, endogenous male sex hormones. The rodent Hershberger assay will be used to identify chemicals that act as androgen agonists or antagonists (anti-androgens).

TEST VALIDATION

INTRODUCTION TO TEST VALIDATION

7. *Validation* is a specialised term that refers to the scientific process designed to characterise the operational characteristics and limitations of a test method, and to demonstrate its reliability and relevance for a particular purpose.

8. The Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Test Methods (Solna Report) (15) provides the principles of validation which are followed by OECD. Work is underway to incorporate these principles into a revised OECD Guidance Document for the Preparation of Test Guidelines (Guidance Document No.34). The Solna principles are consistent with approaches used in Europe, particularly those of the European Centre for Validation of Alternative Methods (ECVAM) and in the U.S. by the Interagency Co-ordinating Committee on Validation of Alternative Methods (ICCVAM).

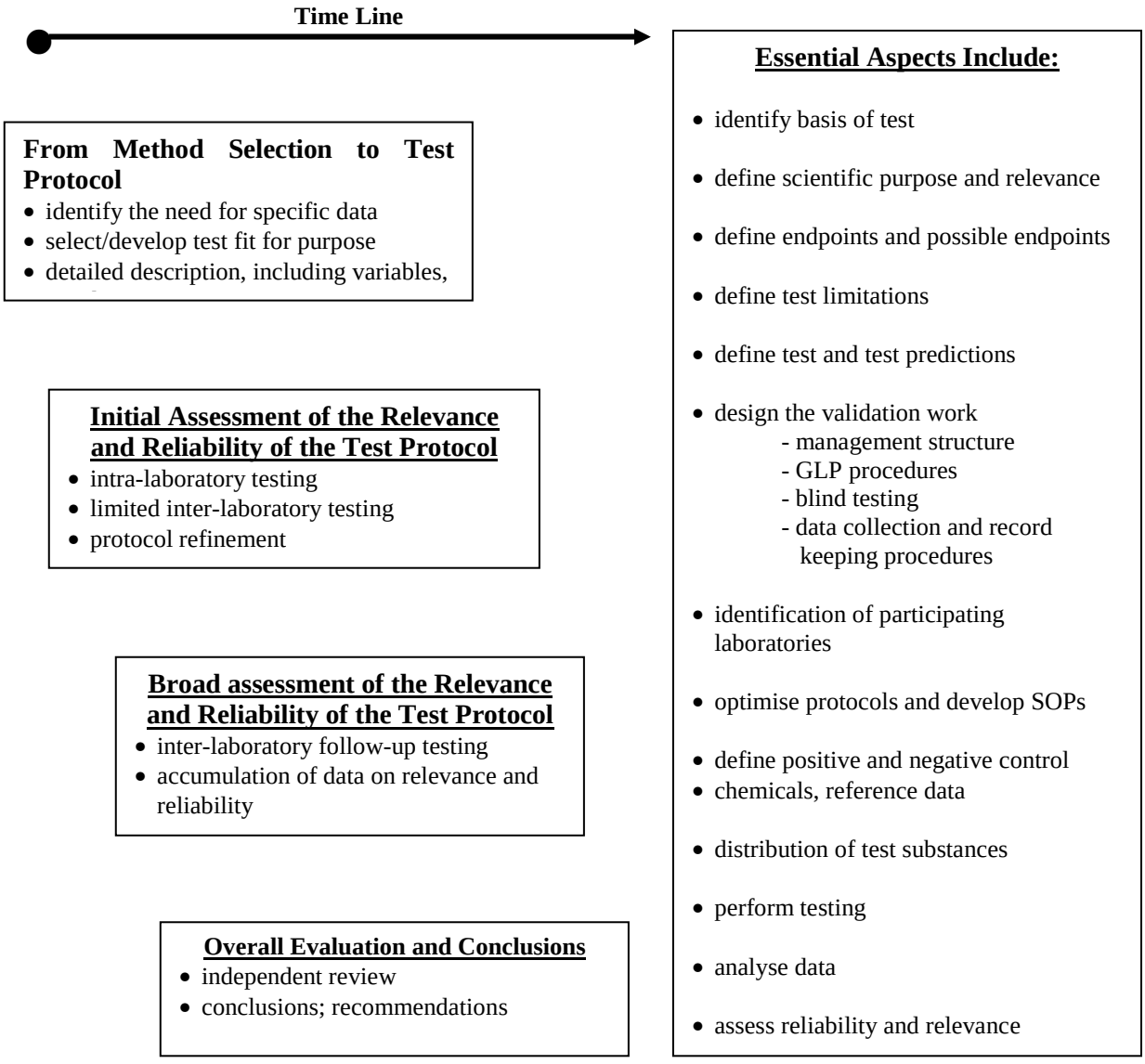
9. Historically, a new test is validated for its proposed use by developing a protocol, standardizing it among one or two laboratories, and then testing a number of potent and weakly acting chemicals under code in a number of laboratories, and evaluating the assay's reliability (i.e., reproducibility within and among laboratories) and relevance (i.e., its ability to accurately measure or predict the effect of concern in the species of concern). The measurement of the responses of accessory sex tissues in castrated or immature rats to administered androgens has been in use since 1932. However, the assay is attributed to Hershberger *et. al*. who in 1953 published the test results for a number of chemicals. There have been a number of protocols used, which vary with regard to whether sexually immature or castrated rats or mice are used; the number of days dosing and the route, and the tissues examined. The assay, denominated the Hershberger assay, has been accepted by testing laboratories, industry, and regulatory authorities for testing pharmaceuticals for androgenic and anti-androgenic effects.

10. The VMG made the determination to perform the OECD validation work in phases, taking into consideration the long use of the assay and its many variants. The first phase of the validation procedure would be to define a protocol that would be expected to identify potent androgenic and anti-androgenic substances; the second would be to measure the protocol's intra-laboratory variability and inter-laboratory reproducibility with a variety of potent and weakly acting substances, and to determine the relative effectiveness of the different tissues for measuring the effects. The need for subsequent phases for the validation of the assay would be determined following the completion and evaluation of the initial phase. This approach is represented in Figure 1, which shows how the assessment process of the relevance and reliability of new or significantly revised testing methods for hazard characterisation can be undertaken in a stepwise, yet flexible, manner while still providing the information necessary to address the Solna criteria and principles.

11. This report of the first phase of the OECD validation of the rodent Hershberger assay provides the basis for determining its usefulness for the purposes envisioned by the OECD, i.e., the identification of androgenic and anti-androgenic substances among chemicals of interest and in the environment. These results form the basis for the design of the proposed second phase of the validation effort, which examines the effectiveness of the assay for identifying weakly acting substances, the reproducibility of the results of weakly acting substances, and the relative effectiveness of the different target tissues. Phase-2 of the validation of the Hershberger assay will commence upon approval and affirmation of the results of Phase-1, and approval of the recommended Phase-2 design and protocol.

Figure 1

Assessment Process of the Relevance and Reliability of New or Significantly Revised Testing Methods for Hazard Characterisation



Phase-1 of the OECD validation study of the rodent Hershberger assay

12. Phase-1 of this OECD validation study of the Hershberger assay was designed to:

- evaluate the effectiveness of a standardised protocol for identifying androgen agonists and antagonists, by the measurement of the weight increases of five androgen-responsive, accessory sex tissues (ventral prostate; seminal vesicles plus coagulating glands; levetor ani and bulbocavernous muscle; Cowper's glands; glans penis), other organ weights (liver; kidneys; adrenals), body weight, and serum hormone levels in immature castrated male rats to administration of the reference androgen, testosterone propionate;
- measure the ability of the anti-androgen, flutamide, to antagonise the effects of testosterone propionate when administered simultaneously;
- obtain data on intra- and inter-laboratory variability and reproducibility among the investigated endpoints;
- compare weights of fresh and fixed androgen-responsive tissues to determine which procedure, if any, is preferred;
- investigate the sensitivity of the dorso-lateral prostate to androgen administration; identify reference doses of testosterone propionate and flutamide for use in subsequent studies or as positive control substances;
- obtain additional information on the performance characteristics of the recommended protocol;
- enable necessary protocol changes and refinements to be identified; and
- identify a reference dose of testosterone propionate to be used as a positive control in studies for androgen agonists, and as a negative control for the detection of androgen antagonists.

13. The lead laboratory for this Phase-1 validation study was at the U.S.-EPA; Health Canada provided independent statistical consultation to the study (see Annex 1 for details).

14. Seventeen test laboratories from seven Member countries (Denmark, France, Germany, Japan, Korea, the United Kingdom, United States and the Lead Laboratory participated in the Phase-1 study. These laboratories, and their countries, are listed in Table 1. Additional details on the participating laboratories, and their principal investigators, are in Annex 1. All laboratories participated on a voluntary and self-supporting basis. The laboratories included those that were experienced with the assay and those that had not used it prior to this study.

Table 1. Laboratories participating in the OECD Hershberger Phase-1 validation study

Country	Laboratory	Number of Laboratories
Denmark	Government	1
France	Private	2
Germany	Private	2
Japan	Governmant Private	2 5
Korea	Governmant	1
United Kingdom	Private	2
United States	Government (Lead Laboratory) Private	1 2
	Total	18

History and organisation of the OECD endocrine disrupter validation project

15. **In early 1998, the National Co-ordinators of the Test Guidelines Programme established a Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to provide a focal point within OECD to consider and recommend priorities for the development of testing and assessment methods for endocrine disrupters (2). Members of EDTA were nominated by Member countries; industry and environmental groups participated as Invited Experts.**

16. The EDTA subsequently set up two Validation Management Groups (VMG), one for mammalian test methods and one for ecotoxicology test methods. The role of both VMGs is to oversee and manage the conduct of the endocrine disrupter test validation studies. A schematic diagram is provided in Figure 2 which describes the role and structure of the OECD Validation Management Group for mammalian effects.

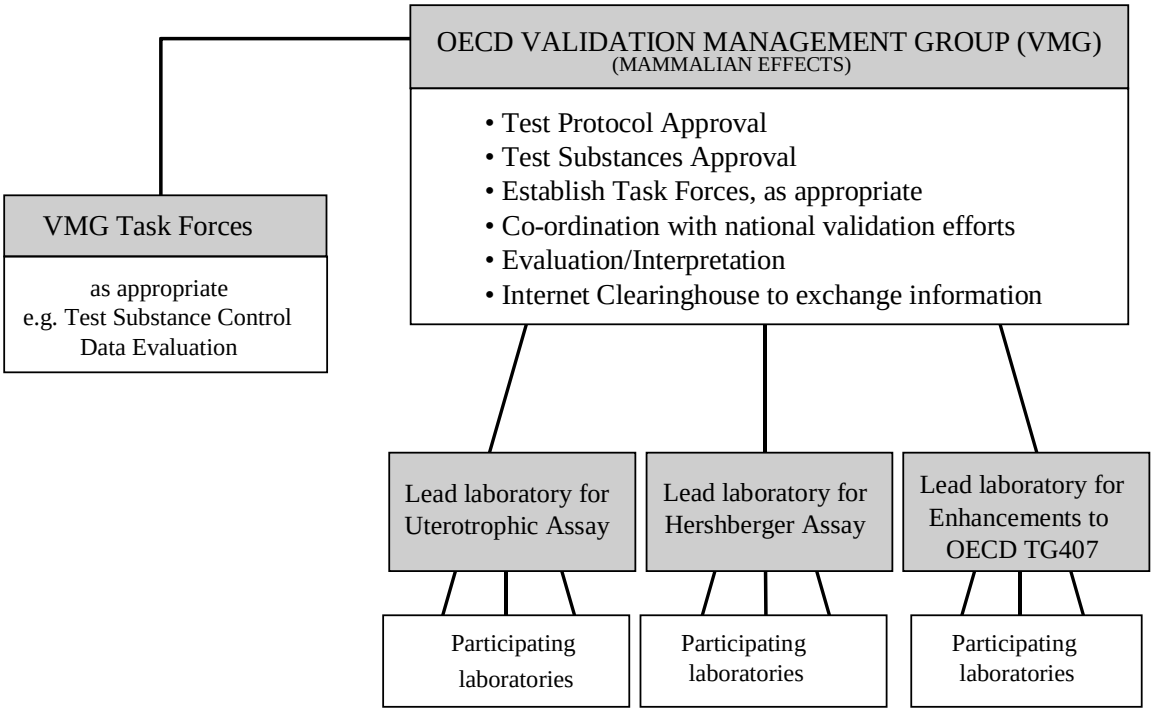
17. The VMG (mammalian) comprises sixteen experts nominated by Member countries and non-government organisations. The membership contains a balance of experts from disciplines including toxicology, endocrinology, and test method development and validation, and is representative of the major OECD regions. Representatives of ICCVAM and ECVAM are members of the Group to provide independent, objective review, to address animal welfare issues, and to provide insight into the requirements for regulatory acceptance of new assays.

18. The VMG developed protocols for the conduct of the Hershberger assay and identified the test chemicals to be used. Expressions of interest were then sought from laboratories wishing to participate in the validation studies. The laboratories that expressed interest were invited to participate in meetings of the VMG, whenever appropriate. The selection of participating laboratories was determined by the willingness of the laboratory to strictly follow the OECD test protocol at their own expense and in accordance with the projected timeline for completion of the study, and provide a formal report containing their experimental data for summary and analysis by the Lead Laboratory and the Secretariat.

19. A U.S. EPA research laboratory, volunteered to assume the responsibility of Lead Laboratory. These responsibilities included drafting the standard experimental protocol on behalf of the VMG; answering day-to-day technical questions from participating laboratories, summarising and evaluating the data, and preparing recommendations for the next validation phase. The Lead Laboratory, however, did not perform the assay protocols. In addition, an independent statistician was asked to evaluate the results and assess the validity of the statistical procedures used by the Lead Laboratory. The OECD Secretariat provided the overall project co-ordination.

Figure 2

The role and structure of OECD Validation Management Group (Mammalian)



20. The participating laboratories each developed their own standard operating procedures (SOP) based on the OECD standardised protocol. The Lead Laboratory reviewed all Phase-1a (testosterone propionate) protocols before the testing work commenced; the flutamide (Phase-1b) protocols were not reviewed prior to testing. All laboratories tested the androgenic activity of testosterone propionate (TP). Seven of these laboratories also tested the antagonist flutamide against TP; four of the laboratories tested against two doses of TP and the other three each tested against one TP dose.

21. A progress report of the Phase-1 validation study, including preliminary summary test results from the testosterone propionate agonism study, was presented to the VMG(mammalian) in March 2001, and to the EDTA at its meeting in May 2001. The EDTA supported the studies performed and initiated planning for subsequent Phase-2 studies, and acknowledged that the VMG (mammalian) would approve the Phase-2 approach following its review and acceptance of the Phase-1 report.

METHODS

INTRODUCTION

22. The rodent Hershberger assay was selected for validation by the OECD following an expert Workshop that was held in Washington, DC in 1998 that recommended that the Hershberger assay be given a high priority for validation. The EDTA considered the recommendations made by the US EDSTAC (3), the OECD's Detailed Review Paper on the appraisal of test methods for sex-hormone disrupting chemicals (1), and the current uses of the assay in OECD countries.

23. At the time of its selection for validation, the assay was in widespread use, following its initial development in the 1930's. The advantages of the assay included that the natural, biological target tissues of endogenous androgens can be examined, the biological response is rapid, and the responses can be quantified and evaluated statistically. The assay can be conducted without the use of specialised facilities, equipment, or techniques.

24. No formal guidelines exist for the assay, and a number of alternative protocol variables have been reported since the assay's initial development. These include:

- the use of rats or mice;
- treatment of sexually immature, mature, or mature castrated animals;
- the tissues examined.

Experiment design

25. The OECD protocol developed by the VMG was provided to the testing laboratories. This protocol is attached as Annex 2. The protocols used for Phase-1a and Phase-1b differed slightly because of the different purposes of the two studies. These protocol details are summarised in Table 2. The rat strains used, ages, suppliers, and husbandry conditions are provided in Annex 3.

Table 2. Protocol design summary for Phases 1a and 1b

	Factor	**Protocol requirements**	**Phase of study**
Animals	Species	Rat	1a and 1b
	Strain	No preference (not Fischer 344)	1a and 1b
	Age at castration	At peripuberty; approx. 5-7 weeks	1a and 1b
	Time after castration	1-2 weeks	1a and 1b
	Age at time of treatment	≤ 7 weeks	1a and 1b
	Weight at time of treatment	Not specified; should be ± 20%	1a and 1b
Animal husbandry	Diet	Lab. preference	1a and 1b
	Food consumption	Lab. preference	1a and 1b
	Bedding	Lab. preference	1a and 1b
	Caging	Lab. preference	1a and 1b
Treatment regimen	Animals per dose group	6	1a and 1b
	Test chemicals	Testosterone propionate Flutamide	
	Route of administration Testosterone propionate Flutamide	 s.c. (dorsal surface) oral gavage	 1a and 1b 1b
	Vehicle Testosterone propionate Flutamide	 corn oil corn oil	 1a and 1b 1b
	Volume of administration Testosterone propionate Flutamide	 0.5 ml/kg/day 5.0 ml/kg/day	 1a and 1b 1b
	Dosing regimen (mg/kg/day) Testosterone propionate Testosterone propionate Flutamide	10 daily administrations 0, 0.1, 0.2, 0.4, 0.8, 1.6 0, 0.2, and/or 0.4 0, 1.0, 5.0, 10.0	 1a 1b 1b
	Sacrifice	24-hrs after last treatment	1a and 1b
Measurements			
	Mandatory weights	Ventral prostate (fresh and fixed) Seminal vesicle + coagulating glands Levetor ani + bulbocavernosus muscles Glans penis Cowper's glands Liver Total body	1a and 1b
	Optional weights and measurements	Adrenal gland (paired) weight Kidney weight Dorso-lateral prostate weight Fresh vs. fixed tissues Negative vs. vehicle control weights Serum testosterone levels Serum lutenizing hormone levels	1a and 1b

Phase-1a: Androgenic (agonist) effect of testosterone propionate

26. Phase-1a of the study was to determine the responses of the androgen-responsive tissues in castrate animals to administered testosterone propionate. All 17 laboratories and the lead laboratory participated in this phase of the study. The standardised OECD protocol used by all of the laboratories is provided in Annex 2.

27. The androgen (TP) was dissolved in corn oil and administered to the animals in the same sequence daily for 10 consecutive days by subcutaneous (s.c.) injection. Body weights were measured daily and the volume of administered test substance adjusted as necessary to maintain the same daily dose. On the day following the tenth dose, the animals were humanely killed in the same sequence and the appropriate tissues and organs were dissected and weighed.

Phase-1b Anti-androgenic (antagonist) effect of flutamide

28. Phase-1b of the study was designed to determine the ability of the assay to measure anti-androgenic effects of administered chemicals. The androgen antagonist (flutamide) was administered to the animals daily for 10 consecutive days by oral gavage. At the same time, the reference androgen, TP, was administered by subcutaneous (s.c.) injection. Two concentrations of TP were used, 0.2 and 0.4 mg/kg/day. The 0.4 mg concentration was selected because it approximates the ED_{70} for LABC, which is one of the moderately sensitive tissues to TP. The 0.2 mg/kg/day concentration was selected because it was half of the ED_{70}. Body weights were measured daily and the volume of test substance adjusted as necessary to maintain the same daily dose. On the day following the tenth dose, the animals were humanely killed in the same sequence as their dosing and the appropriate tissues and organs were dissected and weighed.

Rodent species and strain

29. The assay has been used to evaluate androgenic and anti-androgenic activity using rats and mice. Sexually immature animals, or mature, castrated animals have been used. The species selected by the VMG for the validation work was the rat. Published reports have not shown consistently different assay responses among different animal strains. The VMG considered it undesirable to standardise every aspect of the protocol because if the rodent Hershberger assay was to be adopted as an OECD Test Guideline, as much flexibility as possible should be maintained to ensure wide use while still ensuring that the assay will effectively measure the effects of interest. Therefore, laboratories were encouraged to use the strain of rat that they commonly used and for which they had historical control data. The results from such an approach would provide information on the transferability of the procedure across rat strains, and would allow a guideline to be written that does not specify a single strain, but identifies potential areas of variability. In order to allow for more flexible timing of the test for the laboratories, it was agreed to use mature, castrated animals, rather than immature animals.

30. It was recognized that, for potential incorporation into a Test Guideline, it may be necessary to obtain similar information on immature, rather than castrated, animals. However, in order to limit the protocol variables, these variables were not included in the Phase-1 testing scheme.

Treatment of animals

31. Immature animals were castrated at 33-47 days of age (Phase-1a), and allowed to recover for 7-18 days before starting treatment. In-life testing for Phase-1a (all laboratories) occurred during the period June 2000 through January 2001; the in-life portion of Phase-1b (7 laboratories) was between March 2001

and June 2001. On completion of the experimental work, participating laboratories submitted their individual detailed data on a standardised Microsoft Excel spreadsheet (Annex 4) to the OECD Secretariat and to the Lead Laboratory.

Endpoints examined

32. The Hershberger assay was designed to measure weight increases in androgen-responsive tissues in animals not currently synthesizing endogenous testosterone. The mandatory tissues to be weighed in both the androgen effect procedure (Phase-1a) and the anti-androgen procedure (Phase-1b) were the:

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

Additional mandatory measurements were:

- individual, daily body weights;
- liver weights.

33. In addition to the mandatory measurements identified above, the laboratories had the option to :

- weigh the kidneys and adrenal glands (Phase-1a and 1b);
- weigh the dorso-lateral prostate (Phase-1a);
- compare fresh weights of the androgen-responsive tissues with their fixed weights (Phase-1a)
- measure serum testosterone and lutenizing hormone levels (Phase-1a); and,
- compare the effects of an untreated control with the corn oil vehicle control (Phase-1a).

34. Some participating laboratories had commented that they would prefer to weigh the accessory sex organs after fixation, rather than fresh, because the fixed tissues are easier to dissect and it is less work for the laboratory personnel if they do not have to weigh the fresh tissues at the time of the necropsy. It was also believed that the use of fixed tissues would reduce intra-laboratory variability because the variable drying rates of the fresh tissues would be avoided. To address these concerns, the VMG agreed to expand the protocol to include a comparison of fresh and fixed tissues. The testing laboratories, and the measurements made at each laboratory, are summarised in Table 3.

Table 3. Testing laboratories, and Phase-1a and Phase-1b measurements made

	Measurements made															
	PHASE-1A										**Phase-1b**					
Laboratory (see Annex 1 for details)	Mandatory*	Kidneys	Adrenals	DL-Prostate	VP (fixed)	DL-P (fixed)	SV (fixed)	COWS (fixed)	Serum hormones	Untreated control	Mandatory*	Kidneys	Adrenals	DL-Prostate	Fixed tissues	Serum hormones
1-UK	▲[a]		▲													
2-FR	▲															
3-GER	▲	▲	▲						▲							
4-GER	▲	▲	▲							▲						
5-JAP	▲	▲	▲							▲	▲	▲	▲			
6-FR	▲	▲	▲													
7-U.S.	▲	▲	▲													
8-JAP	▲	▲									▲	▲				
9-UK	▲															
10-JAP	▲	▲	▲		▲	▲	▲	▲			▲[1)]			▲		
11-DK	▲	▲	▲													
12-JAP	▲	▲	▲					▲	▲		▲	▲	▲		▲[2)]	▲
13-JAP	▲	▲	▲								▲[1)]					
14-KOR	▲	▲	▲													
15-JAP	▲	▲	▲	▲		▲			▲		▲	▲	▲			
16-U.S.	▲											▲	▲			
17-JAP	▲			▲	▲	▲	▲	▲	▲		▲[1)]	▲	▲	▲	▲[3)]	

* Mandatory weights. VP (fresh tissue, and fixed); SV; LABC; GLANS; COWS; daily body weights; liver weights.
[a] no fixed VP
[1] no fixed VP weight
[2] COWS
[3] VP, SV, DL-P, COWS

35. A guide to standardise the dissection procedures and parameters for the tissues of interest was provided by the Lead Laboratory. In addition, two hands-on training sessions were held, at the lead laboratory and at laboratory No. 1. Food consumption was measured by some laboratories. No histopathology was performed.

Test chemicals and routes of administration

36. The test chemicals were supplied by the chemical repository which was formed under OECD auspices and with the financial support from industry, under contract with TNO Nutrition and Food Research, Zeist, The Netherlands. The responsibility of the repository was to purchase test chemicals to be

used for the validation of the *in vivo* assays (see Figure 2) and ship them to the testing laboratories. This assured that all laboratories tested the same purities and lots of test chemicals. The test androgen (agonist), testosterone propionate (TP; CASRN 57-82-5), was from Sigma-Aldrich (99.9% pure). The test anti-androgen (androgen antagonist), flutamide (CASRN 1311-84-7), was from Salutas Pharma, Barleben, Germany (99.9% pure).

37. TP was dissolved in stripped corn oil and administered s.c. Flutamide was dissolved in corn oil and administered by oral gavage. Animals on test were weighed weekly, and the dosing volumes adjusted to compensate for changes in weight.

38. The doses of TP and flutamide to be administered were specified to ensure that results could be compared and the test reproducibility could be assessed. In Phase-1a, TP was to be given, s.c. in corn oil, at 0.1, 0.2, 0.4, 0.8, and 1.6 mg/kg/day for 10 days. In Phase-1b, flutamide was to be administered orally in corn oil at 0.1, 0.3, 1.0, 3.0, and 10.0 mg/kg/day for 10 days to challenge the s.c. administration of 0.2 or 0.4 mg TP/kg/day during the same time period. The doses were selected based on prior experience of the Lead Laboratory with this chemical. The test chemical doses and routes of administration for Phase-1a and 1b experiments are in Table 2.

Other concerns

39. The standard laboratory diets may contain phytoestrogens; these substances are generally not included in the routine diet analyses. As an alternative to specifying specific synthetic diets or diet composition criteria, each laboratory was instructed to record full details of the diet used and retain a sample of the diet for further study and analysis, if necessary. Food and drinking water were available *ad libitum*.

Reporting and analyses of data

40. Laboratories recorded the raw experimental data from their phase one work on a Excel spreadsheet (Annex 4) developed specifically for this validation study. Each laboratory also submitted a written summary report. The individual Excel spreadsheets and reports submitted by each laboratory are available on request to members of the VMG (mammalian) and EDTA from the OECD Secretariat.

41. Data summaries and statistical analyses were prepared by the Lead Laboratory. Subsequent to these analyses, the independent statistician evaluated the results and assessed the validity of the statistical procedures used by the Lead Laboratory, and extended the analyses.

42. The ability of each individual laboratory to detect increased tissue and organ weights at various doses of TP was evaluated by an analysis of variance approach, which included body weight as a co-variable.

43. The results can be analysed using two approaches; one is an analysis of the performance of each laboratory, and a comparison of the individual labs. The second is an analyses of the performance of the test among the laboratories by evaluating the overall test performance. Both approaches were used for the Phase-1 data. In the evaluation of the tissue response data, the emphasis was principally on the coefficient of variation (CV) within and among laboratories. This was because the principal question being addressed was not whether or not TP was androgenic, but the within-laboratory and inter-laboratory variation and agreement in the various measurements.

44. Means, standard errors, and the coefficients of variation were calculated for each endpoint using PROC MEANS on SAS (version 6.08). ANOVAs were done using PROC GLM for each laboratory and

pooled over the 16 laboratories. Examination of the CV among endpoints allows one to compare the statistical precision in the weight of a tissue. Some endpoints are inherently more variable than others, and errors in dissection or weighing can increase the CV. Comparison of means and CVs across laboratories allows one to determine if the technique varies greatly among laboratories, in which case additional efforts may be necessary to standardise dissections and weighing of the tissues.

45. Data were then analysed by ANOVA on PROC GLM for each laboratory (with dose as a main effect) with and without initial body weight (the weight at the start of dosing). 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.

46. The CV for each androgen-dependent organ weight was fairly constant as the means increased, the SD being proportional to the mean, indicated that heterogeneity of variance existed. For this reason, the data were transformed using Log_{10}, because this transformation provides for a more valid comparison of the effects of TP on organ weights at lower dosage levels. Subsequent statistical analyses show that although Log_{10} may not be the most appropriate transformation for each tissue in each laboratory, the results obtained using this transformation are sufficiently robust that it could be used for potent substances, such as those tested here. There is insufficient evidence at this time to determine if the Log_{10} transformation would be uniformly effective with weakly acting substances.

47. These analyses also were conducted with initial body weight as a covariate. Initial body weight at the start of the study was used as the covariate, rather than body weight at necropsy, because the administration of TP affected body weight by increasing body weight gain. Hence, final body weight is not a good covariate because it also is affected by treatment. The use of initial body weight covariate adjusts the analysis for experimental variation from several sources, such as, large differences in the size of the rats from laboratory to laboratory, a large component of which appeared to arise from the use of different aged animals or different strains; and differences in the sizes of the rats on study within a laboratory. Data were not analysed using "relative organ weights" as this manipulation makes several assumptions about the relationship between body size and organ weights (i.e., that a linear relationship exits at all, and that the dose-response line goes through the origin) which often are invalid.

48. In addition to means and CVs, the R^2 values for different effects were calculated. An R^2 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 estimate of the strength of the association for an effect with an endpoint. This calculation can be used to compare the robustness of the TP effect across endpoints, the variation from lab to lab, or to what degree the dose-responses vary among laboratories, as indicated by the R^2 for the lab by dose interaction

49. For the five androgen-dependent sex accessory tissues (VP, SV, LABC, COWS, and GLANS) the data were "normalised" in order to visually compare the shapes of the dose-response curves for each lab such that the data range from 0 to 100%.

50. Additional analyses were performed to further strengthen the conclusions of the original analyses: rigorous validation of ANOVA model assumptions, comparison of LOELs across endpoints and laboratories, comparison of benchmark doses (ED_{05}s) across endpoints and laboratories, and treating the LAB effect as random. In the initial analyses, a Log_{10} transformation was applied to correct the increasing variance seen in many cases, but no formal diagnostics were performed. Normal probability plots of the residuals and applied normality tests (Wilk-Shapiro) were used to assess whether the transformation satisfied the model assumptions. The square root transformation was also tested to see if it properly normalized the data. The best transformation was the one that gave the largest (non-significant) p-value for the normality test statistic.

RESULTS

PHASE 1A: ANDROGENIC (AGONIST) EFFECT

51. The summary report and analyses from the lead laboratory are appended as Annex 5.

Weight increases of accessory sex organs and tissues

52. Results were received from all 17 laboratories. All laboratories provided summaries of the protocols used and Excel spread sheets containing the protocol information and test data. One laboratory notified the Secretariat that it had inadvertently administered TP in μg/kg, rather than in mg/kg doses. There was no dose-related effect on any of the tissue weights in this laboratory. As a result, the data submitted by this laboratory were not included in any of the data analyses or in the summary tables, and are not addressed further in this document.

53. The summary results of the accessory organ and tissue weights are in Table 4. This Table clearly shows that the weights of all tissues increased with increasing TP doses in a dose-responsive manner. The results from the individual laboratories are in Annex 6.

Table 4. Mean summary weights of the accessory sex tissues in castrated, immature rats administered testosterone propionate*

	mg Testosterone Propionate/kg/day					
Tissue	**0**	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
VP (fresh)	22 [48]	71 [54]	110 [38]	172 [32]	233 [30]	262 [32]
VP (fixed)	27 [45]	92 [46]	142 [40]	215 [30]	285 [30]	319 [30]
SV	53 [45]	152 [39]	299 [32]	512 [30]	772 [21]	1029 [25]
LABC	181 [37]	319 [37]	421 [33]	542 [32]	622 [32]	685 [31]
COWS	6.9 [68]	18 [45]	27 [35]	38 [30]	49 [31]	58 [28]
GLANS	48 [27]	70 [24]	78 [19]	86 [19]	89 [22]	92 [18]

* tissue weight in mg; mean of results from 16 laboratories [coefficient of variation]

54. Ventral Prostate (VP): There were significant ($p<0.01$), dose-dependent increases in the weights of the VP in all laboratories (Table 4; Annex 6, Table A). With one exception, (laboratory No. 7) at 0.1 mg TP, all doses differed significantly from the controls. The strain and size of the animals at the time of initiation of dosing did not affect their ability to detect TP-induced changes in VP weight. In most laboratories there was no relationship between body weight and VP weight. The lab-to-lab variability was relatively small (R^2 = 6.6%). With few exceptions, the coefficients of variation were relatively constant among the laboratories at the various TP doses.

55. After the excised VP glands were weighed fresh, they were fixed for 24 hrs and weighed again. Fixation of the VP for 24 h increased the weight of the tissue in all 15 laboratories (Table 4; Annex 6, Table B). All doses differed significantly from the control in all laboratories; the differences seen at 0.1 mg TP were all significant at $p<0.05$. Despite these weight differences, there were no consistent differences in the coefficients of variation of the fresh and fixed weights, showing that fixation did not affect the ability to detect TP-induced increases in tissue weight, or reduce the variability in this measurement.

56. Seminal Vesicles plus Coagulating Glands (SV): There was a dose-dependent increase in the weights of the SV and, with one exception (laboratory No. 7; 0.1 mg TP), all doses differed significantly ($p<0.05$) from the controls (Annex 6, Table C). The lab-to-lab variability was relatively small ($R^2 = 6.2\%$), but the differences among the starting body weights of the animals (see Annex 6, Table G) contributed to 54% of the among-laboratory variability. However, the different animal strains and starting body weights did not affect the ability of this tissue to respond to TP.

57. Levetor Ani and Bulbocavernous Muscle (LABC): There was a dose-dependent increase in the weights of the LABC (Table 4) and, with one exception (laboratory No. 9; 0.1 mg TP) all doses differed significantly ($p<0.05$) from the controls (Annex 6, Table D). Three of the laboratories excised and weighed only the levetor ani muscle, and did not include the bulbocavernous muscle. This did not affect their ability to detect weight increases in response to TP, but introduced a significant lab-to-lab variability. There was, overall, a significant lab-to-lab effect in their responses ($R^2 = 36\%$), although the CV for this endpoint is about half that of the VP and SV.

58. Cowper's Glands (COWS): The COWS are the smallest of all the tissues weighed for this assay (Table 4). There was a dose-dependent increase in the weights in all laboratories and, with one exception (laboratory No. 7; 0.1 mg TP), all doses differed significantly ($p<0.05$) from the controls (Annex 6, Table E). The lab-to-lab effect was highly significant ($R^2 = 14\%$). There was a significant effect of body weight on tissue weight, and the COWS weights were associated with higher coefficients of variation than the other tissues. There was a large range of values in the corn oil control group, and this group had the largest CVs. This suggests that the excision and weighing of these glands in the castrated, immature rat may be technically demanding.

59. Glans Penis (GLANS): The effect on the GLANS was smaller than for the other androgen-dependent tissues examined, and the overall CV for this tissue was similar to that seen for LABC (Table 4). There was a dose-dependent increase in the weights of the GLANS and, with one exception (laboratory No. 4; 0.1 mg TP) all doses differed significantly ($p<0.01$) from the controls (Annex 6, Table F). There was a significant lab-to-lab effect in the responses ($R^2 = 36\%$). Two of the laboratories, laboratory No. 3 and laboratory No. 4, castrated the animals prior to 40 days of age (38 and 31 days, respectively), before preputial separation occurred. This complicates and confounds the accurate measurement of GLANS weight.

Body weight increase

60. Initial body weights, and weight gain characteristics, are functions of the animal strains and their ages at the time they were treated with TP. The mean starting weights of the animals in the different laboratories ranged from 159.8 gm to 280.7 gm (Annex 6, Table G). Regardless of the strain of rats used or the weights of the rats at day 0, the proportional weight gains across laboratories were equivalent (Table 5). There was a small, dose-related weight increase in the TP-treated animals that was consistent across all laboratories.

Table 5. Percent increases in body weight between Day 0 and Day 10 of s.c. administration of testosterone propionate (mean of lab means)

	Avg. wt.*	mg Testosterone Propionate/kg/day					
		0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
mean	232.25	23.68%	26.31%	28.73%	30.53%	31.48%	30.46%
S.D.	33.89	5.35	4.37	4.07	5.19	4.67	4.71
CV	15%	23%	17%	14%	17%	15%	15%

* weight at day 0 prior to treatment for all animals

Other organ weight increases

61. Overall, there were increases in kidney and liver weights as a function of TP dose ($p<0.001$), and a dose-related reduction in adrenal weights ($p<0.001$) (Table 6). Not all of these weight changes were significant in all laboratories. The weight gains for the individual laboratories are in Annex 6, Tables H, I, J. There were insufficient data available to determine if the extent of the variability in responses seen among the laboratories was a function of body weight *per se*, or of the animal strain. The CVs for the liver and kidney weights were similar in the individual laboratories (Annex 6, Tables H, I), and was higher for the paired adrenal glands (Annex 6, Table J). The higher inter-animal variability in the individual laboratories probably reflects the difficulty in excising and trimming the adrenals prior to weighing.

62. Although the combined liver weights showed a dose-related trend (Table 6), not all the individual laboratory increases were dose-related, and a few laboratories did not show an increase (Annex 6, Table H). All laboratories had an overall positive trend in kidney weights, although the responses did not all increase monotonically (Annex 6, Table I). The paired adrenal weights had an overall negative trend in all laboratories, although the responses did not decrease monotonically (Annex 6, Table J).

Table 6. Effects of testosterone propionate administration on liver, kidney, and adrenal weights*

	mg Testosterone propionate/kg/day						
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**	****p =**
Liver (gms) [15 labs]	11.64 [16]	12.21 [18]	12.72 [20]	12.93 [19]	12.91 [19]	12.76 [19]	<0.001
Kidneys (mg) [13 labs]	2003 [14]	2034 [14]	2095 [15]	2122 [16]	2211 [16]	2198 [15]	<0.001
Adrenals (mg) [11 labs]	58.6 [11]	55.7 [13]	56.6 [14]	53.2 [11]	51.9 [9]	49.7 [10]	<0.001

* Mean weights [CV]

** Anova

Effect of fixation on tissue weights

63. All but one of laboratories compared the weights of fresh and subsequently fixed (24-hr) ventral prostate as part of the protocol to measure the TP effects on the accessory sex tissues. Fixation did not affect the CVs of the VP weights (Table 4; Annex 6, Table B). In addition, three laboratories performed additional experiments to examine the fixed weights of the ventral prostate, seminal vesicles plus coagulating glands, and Cowper's glands, and one laboratory weighed the fixed adrenal glands (Table 7). In these experiments, the fresh weights of the tissues were not recorded prior to fixation. Fixation of the tissues did not affect the ability of the laboratories to detect dose-related increases in weight at all TP doses. Fixation also did not consistently lower the CVs in the individual laboratories (compare Table 7 with Appendix 6, Tables A, C, and E).

Table 7. Weights (mg [CV]) of fixed tissues in three laboratories

Tissue	Lab	mg Testosterone Propionate/kg/day					
		0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
VP	**10-1***	21 [24]	69 [24]	104 [13]	132 [21]	173 [17]	202 [13]
	10-2	26 [82]	68 [15]	123 [7]	124 [11]	169 [18]	247 [24]
	15	34 [24]	100 [17]	156 [14]	193 [13]	240 [5]	263 [8]
	17	37 [37]	114 [30]	158 [10]	197 [19]	262 [11]	306 [11]
SV	**10**	45 [44]	128 [28]	289 [17]	449 [9]	603 [19]	780 [9]
	15	83 [15]	271 [11]	444 [13]	695 [10]	912 [11]	1132 [9]
	17	64 [25]	196 [28]	366 [18]	596 [10]	893 [13]	1110 [11]
COWS	**10**	5.1 [19]	15.9 [13]	30.5 [15]	35.1 [10]	44.5 [13]	50.0 [5]
	15	8.6 [18]	25.6 [11]	37.0 [24]	41.7 [10]	64.1 [19]	77.7 [11]
	17	12.3 [15]	29.0 [18]	45.4 [9]	55.8 [12]	72.4 [14]	90.0 [21]
Adrenals	**17**	55 [6]	53 [18]	48.0 [6]	49.0 [9]	51.7 [13]	46.5 [14]

* fixed ventral prostate weights were determined in two sets of animals in laboratory No. 10

All values for VP, SV, and COWS are statistically different from the corresponding 0 mg TP values.

Use of the dorso-lateral prostate (DL-P) gland

64. To address the question of the suitability of the dorso-lateral prostate (DL-P) as an indicator of androgen effects, two laboratories also excised and weighed the DL-P, and three laboratories also weighed the fixed tissues. The results from the individual laboratories are presented in Table 8. There was a significant dose-related increase in the weights of the fresh and fixed DL-P at all TP doses. Although the weights of the gland after fixation were higher, the fixation process did not appear to affect the CVs in each of the laboratories.

Table 8. Weights of the fresh and fixed dorso-lateral prostate gland

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
Fresh tissues						
15	32 [32]	78 [32]	123 [23]	141 [13]	195 [9]	195 [25]
17	49 [31]	99 [15]	138 [14]	171 [13]	223 [19]	246 [11]
Fixed tissues						
10	21 [26]	69 [26]	104 [15]	132 [17]	173 [11]	202 [13]
10	26 [90]	68 [17]	123 [7]	124 [12]	169 [20]	247 [27]
15	34 [26]	100 [18]	156 [15]	193 [14]	240 [6]	263 [8]
17	37 [41]	114 [33]	158 [11]	197 [20]	262 [12]	306 [12]

mean mg [coefficient of variation]

Serum testosterone and lutenizing hormone levels

65. Four laboratories measured serum testosterone and lutenizing hormone levels. Two of these laboratories performed 2 separate sets of measurements; in the animals that were used for the fresh tissue weights, and in a different set of animals that were used to compare fresh and fixed tissue weights.

66. The serum testosterone measurements are in Table 9; the lutenizing hormone measurements from

the same animals are in Table 10.

67. Serum testosterone levels. The s.c. injection of TP to castrated animals would be expected to result in relatively constant levels of serum testosterone. The serum testosterone concentrations at all levels of TP dosing are highly variable. The data from three of the laboratories were equivalent, however the fourth laboratory (laboratory No. 3) produced results that were 1 to 2.5 orders of magnitude higher, and highly variable (Table 9). In all laboratories, the sensitivities of the analytical methods used were not sufficient to detect increases in serum testosterone in the animals dosed s.c. with 0.1 mg TP/day, which was sufficient to induce significant weight changes in SV, VP, LABC, and COWS. There were statistically significant, dose-related increased mean levels of testosterone in the serum of animals at the higher TP doses.

68. Serum luteinizing hormone (LH) levels. There were dose-related decreases in LH levels at increasing TP doses. in all of the laboratories (Table 10).

Table 9. Serum testosterone concentrations (ng/ml) in Phase-1a testosterone propionate studies*

		mg Testosterone propionate/kg/day					
LAB*		**0**	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
3**	mean	0.41	46.39	102.27	141.20	337.78	310.27
	S.D.	0.46	41.74	121.25	119.79	299.04	349.62
	range	0.05-1.20	2.13-100.0	22.58-339.83	1.33-307.36	27.57-617.54	3.35-850.42
	C.V.	112	90	119	85	89	113
12	mean	<0.10	<0.17	.45	1.01	2.17	4.12
	S.D.	--	>0.05	0.05	0.19	0.38	0.69
	range	<0.10-<0.10	<0.1-0.2	0.4-0.5	0.8-1.3	1.6-2.7	3.4-5.3
	C.V.	--	>12	11	19	18	17
15-1	mean	0.10	0.10	0.27	0.63	1.58	3.67
	S.D.	--	--	0.08	0.08	0.48	0.35
	range	0.10-0.10	0.1-0.1	0.2-0.4	0.5-0.7	1.3-2.3	2.9-5.0
	C.V.	--	--	30	13	6	10
15-2	mean	0.10	0.15	0.32	0.67	1.58	3.75
	S.D.	--	0.05	0.12	0.23	0.31	0.83
	range	0.10-0.10	0.1-0.2	0.2-0.5	0.4-1.0	1.3-2.0	2.9-5.0
	C.V.	--	33	38	34	20	22
17-1	mean	<0.20	<0.20	<0.20	0.67	1.12	2.78
	S.D.	--	--	>0.004	0.08	0.27	0.59
	range	<0.20-<0.20	<0.20-<0.20	<0.2-0.2	0.5-0.7	0.7-1.5	2.2-3.5
	C.V.	--	--	>2	12	24	21
17-2	mean	<0.20	<0.20	<0.20	0.43	1.27	3.17
	S.D.	--	--	>0.004	0.12	0.45	0.84
	range	<0.20-<0.20	<0.20-<0.20	<0.2-0.2	0.3-0.6	0.7-1.7	2.1-4.2
	C.V.	--	--	>2	28	35	26

* Two laboratories (Nos. 15 and 17) performed more than one analysis. The first (15-1; 17-1) used the animals that were used for the wet tissue weight determinations. The second (15-2; 17-2) used a separate set of animals, tested at a later date, that were used for fixed tissue weight determinations.

** Laboratory No. 3 used a Diagnostic Systems Laboratories radio-immunoassay kit. One possible cause for the high testosterone values in the laboratory No. 3 study is that a different procedure was used to prepare the serum specimens. Rather than use the immunoassay to detect testosterone directly in rat serum in the same manner as is recommended in the kit for human samples, on advice of experts of Diagnostic Product Corp., laboratory No. 3 had extracted the rat sera with diethyl ether. The extracts were evaporated to dryness under nitrogen and reconstituted the residue for testosterone determination (16)

Laboratory No. 12 used a Diagnostic Systems Laboratories Active TM Testosterone enzyme immunoassay (EIA)system.

Laboratory No. 15 used a Diagnostic Products Corp. Coat-A-Count Total Testosterone kit, based on a solid-phase 1251 radio-immunoassay.

Laboratory No. 17 used a Diagnostic Products Corp., DPC total testosterone, radio-immunoassay kit.

Table 10. Serum lutenizing hormone concentrations (ng/ml) in Phase-1a testosterone propionate studies

		mg Testosterone propionate/kg/day					
LAB*		**0**	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
3	mean	18.78	16.66	8.46	1.70	1.44	0.49
	S.D.	6.14	4.75	6.20	1.02	0.75	0.12
	range	10.99-26.72	12.94-22.11	1.9-18.33	0.95-3.66	0.15-2.19	0.29-0.64
	C.V.	33	29	73	60	52	24
12	mean	38.43	42.20	32.92	20.62	15.62	10.60
	S.D.	7.13	12.67	11.95	4.92	1.30	2.59
	range	32.4-52.4	25.6-62.4	17.5-46.1	12.6-27.0	13.5-17.0	11.9-19.1
	C.V.	19	30	36	24	8	24
15-1	mean	8.12	10.08	7.63	2.30	1.33	1.50
	S.D.	2.54	2.25	2.25	1.03	0.23	0.15
	range	5.2-11.9	6.4-13.5	4.0-10.9	1.3-3.8	1.1-1.7	1.3-1.7
	C.V.	31	22	29	45	17	10
15-2	mean	7.60	11.25	9.08	2.48	1.48	1.70
	S.D.	1.26	2.44	2.46	1.19	0.31	0.13
	range	5.7-9.0	8.2-15.6	6.5-11.8	0.9-4.2	0.9-1.7	1.5-1.8
	C.V.	17	22	27	48	21	8
17-1	mean	7.82	10.45	9.70	2.78	<1.08	<1.03
	S.D.	2.04	2.40	1.72	1.24	>0.60	>0.57
	range	5.2-11.3	8.3-14.9	7.5-11.7	1.7-4.8	<0.8-2.3	<0.8-2.2
	C.V.	26	23	18	45	>56	>55
17-2	mean	8.82	8.87	8.05	4.67	<0.88	<0.93
	S.D.	2.57	2.82	1.41	1.47	>0.13	>0.12
	range	6.4-12.4	4.4-12.4	6.1-9.1	2.6-6.8	<0.8-1.1	<0.8-1.1
	C.V.	29	32	18	31	>15	>13

* Two laboratories (laboratory Nos. 15 and 17) performed more than one analysis. The first (laboratory No. 15-1; 17-1) used the animals that used for the wet tissue weight determinations. The second (laboratory No. 15-2; 17-2) used a separate set of animals, tested at a later date, that were used for fixed tissue weight determinations.

Untreated control vs. vehicle (corn oil) control

69. TP was administered dissolved in corn oil; the volume administered was adjusted to the weight of the animal. As a result, the administered volumes varied from approx. 0.06 ml/animal to 0.2 ml/animal. To determine if the corn oil induced weight changes in the accessory sex tissues, two laboratories included an untreated control group for comparison with the vehicle control group that was administered 0.5 ml corn oil/kg body weight. There were no effects on androgen-responsive tissues, in other organ weights, or total body weights, as an effect of corn oil injection (Table 11).

Table 11. Effects of corn oil on tissue weights*

Lab		VP-fresh	SV	LABC	COWS	GLANS	VP-fixed	Total body	Liver**	Adrenals	Kidneys
4 un-treated											
	mean	30.5	46.8	69.6	2.42	46.6	35.7	224.3	11.1	55.8	1664.8
	S.D.	15.50	8.93	33.36	0.77	15.50	19.20	6.15	0.91	8.75	106.91
	C.V.	51	19	48	32	33	54	2.7	8	16	6
4 corn oil											
	mean	28.0	46.7	65.9	1.62	45.2	35.3	226.0	11.3	55.2	1626.7
	S.D.	10.26	8.71	21.56	1.53	19.26	14.33	10.73	1.25	6.15	118.17
	C.V.	37	19	33	94	43	41	5	11	11	7
5 un-treated											
	mean	16.7	44.2	186.1	4.9	52.7	19.5	292.6	11.8	55.8	1925.0
	S.D.	3.16	8.38	8.30	1.20	6.38	3.93	12.66	0.69	6.79	83.61
	C.V.	19	19	4	24	12	20	4	6	12	4
5 corn oil											
	mean	16.3	53.2	200.2	8.2	53.5	18.6	291.1	11.4	48.7	1960.0
	S.D.	4.22	7.15	18.29	5.79	6.50	4.59	6.71	0.68	7.54	125.06
	C.V.	26	13	9	71	12	25	2.3	6	15	6

* weights expressed as mg tissue
** expressed as gm
S.D., Standard deviation of the mean
C.V., Coefficient of variation

PHASE-1B: ANTI-ANDROGENIC (ANTAGONIST) EFFECT OF FLUTAMIDE

70. The Phase-1b summary report and analyses from the Lead Laboratory are appended as Annex 7.

Flutamide effects on TP-induced weight increases of accessory sex organs and tissues

71. Seven laboratories examined the ability of flutamide to block the androgenic responses to TP (Annex 1). Four of the laboratories used 0.2 and 0.4 mg TP/kg/day, two used only 0.2 mg/kg TP, and one used only 0.4 mg/kg TP. All laboratories provided summaries of the protocols used and detailed Excel spread sheets containing the protocol information and test data.

72. The rat strains used, animal ages, and diets for this study are described in Annex 3. The protocol is summarised and compared with the Phase-1a protocol in Table 2. The summary effects of flutamide administration on the TP-induced weights of the androgen-responsive tissues are in Table 12. The data from the individual laboratories are in Annex 6.

73. Flutamide, administered by oral gavage in corn oil in 10 consecutive, daily doses, at the same time as TP administration by s.c. injection, blocked the androgenic effects of TP in a dose-responsive manner in all of the laboratories ($p<0.001$). Although there was variation among laboratories with respect to the lowest flutamide dose that produced significant decreases in tissue weights, they all showed the same patterns of decreases (Annex 6, Tables K, L, M, N, O). In all cases, the flutamide effect in the pooled analysis was larger in the 0.4 mg TP group than in the 0.2 mg TP group. Both TP groups were equivalent with respect to their CVs.

Table 12. Antagonism by flutamide of testosterone propionate (TP) activity in accessory sex tissues

	TP	mg Flutamide/kg/day					
		0	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
VP	0	18.4 (39)					
	0.2 mg/kg	127 (20)	127 (26)	102 (23)*	69 (30)*	35 (27)*	25 (25)*
	0.4 mg/kg	218 (24)	184 (27)*	170 (24)*	122 (34)*	59 (35)*	30 (37)*
SV	0	47.2 (34)					
	0.2 mg/kg	346 (25)	287 (27)*	242 (29)*	133 (33)*	70 (27)*	56 (25)*
	0.4 mg/kg	580 (18)	488 (23)*	447 (24)*	279 (33)*	117 (38)*	59 (30)*
LABC	0	206.2 (22)					
	0.2 mg/kg	490 (12)	475 (15)	432 (16)*	351 (21)*	260 (17)*	229 (17)*
	0.4 mg/kg	645 (15)	607 (13)*	553 (12)*	456 (17)*	327 (17)*	243 (21)*
COWS	0	7.9 (36)					
	0.2 mg/kg	29.5 (24)	28.5 (19)	22.9 (24)*	20 (28)*	11.5 (33)*	8.8 (32)*
	0.4 mg/kg	44.6 (16)	40.6 (21)*	35.7 (14)*	29 (24)*	18.2 (30)*	10.4 (35)*
GLANS	0	51.4 (25)					
	0.2 mg/kg	87.4 (19)	81.7 (13)*	80.1 (15)*	73.7 (14)*	62.7 (18)*	58.6 (18)*
	0.4 mg/kg	93.0 (15)	94.4 (15)	88.3 (14)*	81.8 (18)*	69.1 (20)*	58.0 (23)*

in mg, mean across laboratories [Coefficient of variation]
* Significantly decreased from the 0 mg flutamide value.

74. Ventral prostate (VP): Flutamide significantly inhibited the effects of 0.2 and 0.4 mg TP/kg/day in each laboratory ($p<0.001$) (Table 12; Annex 6, Table K). The 0.2 mg/kg TP-induced weight gain of VP was significantly reduced at doses of 0.3 mg/kg and above in two of the laboratories, and at 1.0 mg and above in all laboratories. Weight gain induced by 0.4 mg/kg TP was reduced by all flutamide doses, with two of the laboratories showing significant reductions at 0.1 mg flutamide, and all laboratories having significant responses at 1.0 mg. At 10 mg/kg flutamide, the VP weights approached the untreated (no TP) weights. In the pooled analysis, VP weight was significantly related to animal body weight in the 0.2 mg/kg TP group.

75. Seminal vesicle (SV): Flutamide significantly inhibited the effects of 0.2 and 0.4 mg TP/kg/day in each laboratory ($p <0.001$) (Table 12; Annex 6, Table L). The lowest dose of flutamide, 0.1 mg/kg/day, produced highly significant ($p<.0.001$), dose-related weight decreases at both concentrations of TP when the individual laboratory results were combined. When either TP dose was used, the lowest dose of flutamide, produced significant weight gain decreases in two of the laboratories, and the overall weight decreases were significant. Five of the six laboratories showed significant weight decreases at 0.3 mg flutamide in the 0.2 mg TP group, and 3 of the laboratories showed significant decreases in the 0.4 mg TP group. At 10 mg/kg flutamide, the SV weights approached the untreated (no TP) weights. There was significant lab-to-lab variability in SV weights, but the dose-responses were similar.

76. Levator ani plus bulbocavernosus muscle (LABC): Flutamide significantly inhibited the effects

of 0.2 and 0.4 mg TP/kg/day in each laboratory ($p < 0.001$) (Table 12; Annex 6, Table M). Flutamide doses of 0.3 mg/kg and above reduced the 0.2 mg/kg TP-induced weight gain of LABC in three of the laboratories, and in two of the laboratories using 0.4 mg/kg TP. Weight gain induced by 0.4 mg/kg TP was reduced by all flutamide doses when the individual laboratory results were combined. At 10 mg/kg flutamide, the LABC weights approached the untreated weights. The CVs for LABC were lower than the CVs seen for the VP and SV weights.

77. Cowper's glands (COWS): Flutamide significantly inhibited the effects of 0.2 and 0.4 mg TP/kg/day on COWS in each laboratory ($p < 0.001$) (Table 12; Annex 6, Table N). The 0.2 mg/kg TP-induced weight gain of COWS was significantly reduced at doses of 0.3 mg/kg flutamide and above in the combined laboratory data, and in three of the individual laboratories. Weight gain induced by 0.4 mg/kg TP was reduced by all flutamide doses when the laboratory results were combined, although none of the individual laboratories had significant decreases at 0.1 mg flutamide, and two laboratories had significant responses at 0.3 mg/kg flutamide. At 10 mg/kg, flutamide completely inhibited the TP-induced weight gain. The lab-to-lab variability was highly significant, and was larger than that seen with the other tissues.

78. Glans penis (GLANS): Flutamide significantly inhibited the effects of 0.2 and 0.4 mg TP/kg/day on GLANS in each laboratory ($p < 0.001$) (Table 12; Annex 6, Table O). The weight gain reduction was significant when combined over laboratories in the 0.2 mg/kg TP animals administered 0.1 mg/kg flutamide, although only one laboratory produced a significant decrease. All laboratories showed significant decreases at 3.0 mg/kg flutamide and above. Flutamide significantly reduced the 0.4 mg/kg TP-induced weight gain of GLANS at doses of 0.3 mg/kg and above when the data were combined across laboratories, however individual laboratories did not show significant decreases until 1.0 mg/kg flutamide (three laboratories). This is the opposite effect as was seen with VP, LABC, and COWS, where flutamide was most, or equally, effective against 0.4 mg/kg TP. The CVs for GLANS were lower than the CVs seen for the VP, SV, and COWS weights.

Effect of fixation on ventral prostate weights

79. Some of the laboratories participating in the flutamide study compared the weights of fresh and fixed (24-hrs) ventral prostate (Table 13; Annex 6, Table P). There was an overall weight gain as a result of the fixation, but not all tissues in all laboratories showed the effect. Although the tissue weights were heavier, the statistical analyses of the fixed VP weights yields the same results as the fresh VP weights. Fixation of the tissue did not consistently lower the CV of the measurement within or among laboratories (Table 13; Annex 6, Table P).

Table 13. Comparison of fresh and fixed ventral prostate weights in animals treated with testosterone propionate and flutamide

		mg flutamide/kg/day					
Tissue	**mg TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
fresh	0	18 [34]					
fixed		20 [33]					
fresh	0.2	120 [19]	120 [28]	98 [22]	64 [29]	32 [22]	23 [24]
fixed		140 [22]	139 [31]	117 [21]	77 [34]	37 [25]	25 [32]
fresh	0.4	230 [20]	187 [22]	184 [14]	122 [20]	60 [32]	28 [22]
fixed		257 [24]	210 [26]	203 [18]	137 [24]	67 [36]	31 [26]

TP = 0.2, 4 labs; TP=0.4, 3 labs; mean [CV]

Effects on body weight

80. The initial body weights, and weight gain characteristics, are functions of the animal strains and their ages at the time they were placed on test. Treatment of the rats with 0.2 or 0.4 mg TP/day for 10 days resulted in low, but consistent, weight gains (Table 14). The administration of flutamide to the TP-treated rats led to slight, but not significant, decreased weight gain over the 10-day treatment period. The weight gain changes in the individual laboratories are in Annex 6, Table Q.

Table 14. Percent increases in body weights between Day 0 and Day 10 of administration of flutamide to castrated, immature rats receiving testosterone propionate.

	Avg. wt. at day 0*	**Untreated control ****	**mg flutamide/kg/day + 0.2 mg TP/kg/day**					
			0	**0.1**	**0.3**	**1.0**	**3.0**	**10**
mean	233.7	23.4	27.6	28.3	27.6	26.9	21.1	25.5
S.D.	16.3	5.2	4.9	4.4	5.8	5.1	4.9	4.8
CV	7	22	18	16	21	19	23	19
			mg flutamide/kg/day + 0.4 mg TP/kg/day					
			0	**0.1**	**0.3**	**1.0**	**3.0**	**10**
mean	238.0	23.0	30.6	30.5	29.3	27.1	27.1	26.0
S.D.	18.7	5.0	6.1	5.4	5.8	6.6	5.5	5.4
CV	8	22	20	13	20	24	20	21

* avg. weight in gms at day 0 for all untreated animals; mean (S.D.)

** Animals not receiving testosterone propionate or flutamide

Flutamide effects on TP-induced weight changes in other organs

81. Body and organ weights: Not all laboratories performing Phase-1b measured all the non-reproductive organs. The flutamide dose did not significantly affect the liver and kidney weight increases induced by TP over the 10-day treatment period (Table 15). Administration of flutamide mitigated the TP-induced decrease in adrenal weights. All increases in adrenal weight were significant, and showed a dose-relationship over the 0 dose flutamide control. The data from the individual laboratories are in Annex 6, Table R (0.2 mg TP) and Table S (0.4 mg TP).

Table 15. Effects of testosterone propionate and flutamide administration on liver, kidney, and adrenal weights

Organ	Untreated control	mg flutamide/kg/day + 0.2 mg TP/kg/day					
		0	**0.1**	**0.3**	**1.0**	**3.0**	**10**
Liver (gm) (6 labs)	12.3 [9]	13.2 [9]	13.1 [11]	12.8 [11]	13.1 [10]	12.8 [8]	13.2 [11]
Kidneys (mg) (5 labs)	1986 [7]	2051 [7]	2047 [9]	2029 [6]	2085 [8]	1995 [9]	2014 [8]
Adrenals (mg) (4 labs)	55.9 [15]	48.3 [12]	53.4 [14]	52.2 [13]	53.7 [12]	50.4 [16]	56.9 [14]
		mg flutamide/kg/day + 0.4 mg TP/kg/day					
		0	**0.1**	**0.3**	**1.0**	**3.0**	**10**
Liver (gm) (4 labs)	12.2 [12]	13.6 [13]	13.5 [13]	13.5 [13]	12.9 [14]	13.2 [11]	13.1 [11]
Kidneys (mg) (3 labs)	1998 [5]	2175 [6]	2172 [7]	2137 [9]	2079 [8]	2081 [4]	2106 [6]
Adrenals (mg) (3 labs)	55.3 [12]	46.4 [13]	48.2 [12]	52.1 [9]	52.6 [15]	51.1 [11]	54.0 [16]

mean [CV]

DATA EVALUATION

DATA ANALYSES

82. Data analyses were performed by the Lead Laboratory. The conclusions of the Lead Laboratory, and graphic depictions of the data, are in Annex 5 and 7. The data, and these analyses, were subsequently forwarded to the independent statistical consultant for confirmation and further analysis, if necessary. The independent statistician's analyses supported the analyses by the Lead Laboratory, and additional analyses were performed. The report and the results of these additional analyses are in Annex 8.

Phase-1a data: Androgenic (agonist) effect of testosterone propionate

83. All laboratories obtained dose-responsive increases in weights of the five androgen-dependent tissues beginning with the lowest TP concentration tested, 0.1 mg/kg (Table 4). The data from the individual laboratories, and the combined laboratory data, were transformed to achieve a normal distribution. No single transformation adequately normalized the data across all laboratories and endpoints in the Phase-1a studies. Although the Log_{10} transformation was sufficient for most of the data, there were a number of data points where such a transformation would have led to incorrect determinations of LOEL's. The transformation model that best fit the combined data from all laboratories is presented in the following tables as the "Overall model."

84. Table 16 indicates which transformation was most appropriate for each laboratory and endpoint. For some laboratories and combined laboratory data, no obvious transformation was available to transform the data to normality; these transformations are presented in the Tables as "no obvious transformation." For most laboratory-tissue combinations, the Log_{10} transformation yields the same LOEL as the "correct" transformation. However, in six cases, the correct transformation leads to a higher LOEL than the Log_{10} transformation. These cases are identified in Table 17.

Table 16. Statistical transformations needed to normalise the Phase-1a data

Tissue	Overall model (all labs)	Individual labs	Most Appropriate Transformation
VP (fresh)	No obvious transformation	1, 4, 14, 15	Log_{10}
		2, 3, 7, 8, 17	Untransformed
		5, 6, 10, 12, 16	Square root
		9, 13	No obvious transformation
SV	No obvious transformation	1, 3, 4, 5, 7, 8, 9,10, , 15	Log_{10}
		2, 6, 12, 13, 16, 17	Square root
LABC	No obvious transformation	1, 3, 4, 10, 12, 13	Log_{10}
		6, 7, 16, 17	Untransformed
		8, 15	Square root
		2, 5, 9,14	No obvious transformation
COWS	No obvious transformation	1,8, 12,13, 16	Log_{10}
		2, 5, 7, 9, 15, 17	Untransformed
		3, 4, 6, 10, 14	Square root
GLANS	No obvious transformation	1, 9, 13, 15	Log_{10}
		2, 3, 5, 6, 7, 14, 16, 17	Untransformed
		8, 10, 12	Square root
		4	No obvious transformation

Table 17. LOEL (mg TP/kg/day) changes as an effect of data transformation used

Tissue	Lab	LOEL, Log_{10} transformation	Most appropriate transformation	LOEL; Most appropriate transformation
SV	6	0.1	Untransformed	0.4
LABC	2	0.1	Untransformed	0.2
COWS	2	0.1	Untransformed	0.2
COWS	8	0.1	Untransformed	0.2
COWS	14	0.1	Untransformed	0.2
GLANS	2	0.1	Untransformed	0.2

85. The ED_{70} determinations (i.e., the dose at which 70% of the maximum response in that tissue was seen) were in the order of GLANS (ED_{70} = 0.2 mg TP/kg/day) ≥ LABC > COWS, VP > SV (ED_{70} = 0.8 mg/kg/day). The ED_{70} calculations were used to select TP concentrations for use in Phase-1b; the ED_{70} of the median reactive tissue, LABC (0.4 mg/kg/day), was used, and one-half that value (0.2 mg/kg/day).

86. The results were also analyzed to determine the Benchmark Dose (BMD) across laboratories and endpoints. This value, which is used by regulatory agencies, is an estimate of the dose that causes the mean response to increase or decrease by double the standard deviation of the control group. Unlike the LOELs that are relatively uniform across tissues (0.1 mg TP/kg/day), the BMD values show a wider variation among tissues, from a low of 0.054 (VP; Log_{10} transformed) to 0.295 (GLANS; untransformed) (Table 18).

Table 18. Benchmark doses (BMDs) of testosterone propionate (mg/kg/day) for the Phase-1a study combined across all laboratories and listed in order of decreasing potency

Tissue	BMD, Log_{10} transformed (BMDL*)	Most appropriate transformation	BMD, Most appropriate transformation (BMDL*)
VP	0.054 (0.044)	No obvious transformation	-
SV	0.065 (0.055)	Log_{10}	0.065 (0.055)
LABC	0.195 (0.154)	Unknown	-
COWS	0.076 (0.058)	Square root	0.126 (0.103)
GLANS	0.199 (NA**)	No obvious transformation	0.295 (0.199)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

87. The BMD values indicate that the relative order of activity of the tissues (i.e., VP>SV>COWS>LABC=GLANS) are in almost the opposite order from the relative activities as measured by the ED_{70}. This is because the ED_{70} is a function of what dose is needed with respect to the maximum response for that tissue, whereas the BMD identifies the lowest dose that produces an effect.

Phase-1b data: Anti-androgenic (antagonist) effect of flutamide

88. The ED_{70} values for the anti-androgenic effect of flutamide against TP are opposite in potency to the ED_{70}'s for TP agonism, i.e., SV > LABC = COWS = VP > GLANS. As with the Phase-1a data, no single transformation adequately normalized the data across all laboratories and endpoints in the Phase-1b studies. Although the Log_{10} transformation was sufficient for most of the data, there were a number of data sets where such a transformation was not appropriate, and would have led to incorrect determinations of LOEL's. Table 19 indicates which transformation was most appropriate for each laboratory and endpoint, and for the combined laboratory data. For some laboratory data, no obvious transformation was available to transform the data to normality.

Table 19. Transformations needed to normalize the Phase-1b data

Tissue	mg TP/ kg/day	Overall model (all labs)	Individual labs	Most Appropriate Transformation
VP (fresh)	0.2 mg	Square root	8, 12, 15	Square root
			5, 15	Log_{10}
			13	Untransformed
	0.4 mg	Log_{10}	5, 10, 12	Square root
			13, 15	Log_{10}
SV	0.2 mg	Log_{10}	5	Untransformed
			8, 12, 13, 15	Log_{10}
			17	Square root
	0.4 mg	Square root	5	Untransformed
			10, 13,15	Square root
			12	Log_{10}
LABC	0.2 mg	Square root	5, 13	Square root
			8,13	Log_{10}
			12, 15	Untransformed
	0.4 mg	Untransformed	5, 15	Untransformed
			10	Square root
			12, 13	Log_{10}
COWS	0.2 mg	Log_{10}	5, 8,10, 12, 15,17	Log_{10}
			13	Untransformed
	0.4 mg	Log_{10}	5	Square root
			10, 13	Untransformed
GLANS	0.2 mg	No obvious transformation	5	No obvious transformation
			8, 12, 17	Untransformed
			13	Log_{10}
			15	Square root
	0.4 mg	Square root	5	Square root
			10, 13, 13	Untransformed
			12	Log_{10}

89. The results were also analyzed to determine the BMD for flutamide activity across laboratories and endpoints for each level of TP used. When the most appropriate transformation is used, the BMD values show a wide variation among tissues, The relative orders of response of the different tissues when the most appropriate transformations are used are, for 0.2 mg/kg TP, GLANS>VP>SV>LABC>COWS. At 0.4 mg/kg TP, the relative order of response is LABC> SV>VP>COWS>GLANS (Table 20).

Table 20. Benchmark Doses (BMDs) of flutamide (mg/kg/day) for the Phase-1b study combined across laboratories and listed in order of decreasing potency

Tissue	mg TP/ kg/day	BMD, Log_{10} transformed (BMDL*)	Most appropriate transformation	BMD, Most appropriate transformation (BMDL)
VP	0.2	0.603 (0.512)	Square root	0.499 (0.418)
	0.4	0.609 (0.525)	Log_{10}	0.609 (0.525)
SV	0.2	0.542 (0.477)	Log_{10}	0.542 (0.477)
	0.4	0.510 (NA**)	Square root	0.311 (0.271)
LABC	0.2	1.115 (1.007)	Square root	0.917 (0.790)
	0.4	0.501 (NA)	Untransformed	0.293 (0.240)
COWS	0.2	1.333 (NA)	Log_{10}	1.333 (NA)
	0.4	0.948 (0.737)	Log_{10}	0.948 (0.737)
GLANS	0.2	0.502 (NA)	Untransformed	0.332 (0.218)
	0.4	1.308 (NA)	Square root	1.067 (0.825)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

OVERALL CONCLUSIONS

90. A number of general and specific conclusions can be drawn from Phase-1 of this validation study.

a) The OECD's Hershberger assay protocol selected for Phase-1 is sufficient to detect testosterone propionate-induced weight gains in male rat accessory sex tissues, and the mitigation of that weight gain by the anti-androgen flutamide.

i) Phase-1a: The protocol allowed the detection of the androgenic effects of testosterone propionate. All five of the androgen-sensitive tissues sampled showed dose-related weight increases as a function of testosterone propionate dose in all laboratories. The differences in rat strain used, and the differences in the ages at which the animals were castrated, did not affect the ability of the animals to respond to testosterone propionate.

ii) Phase-1b: The protocol allowed the detection of the anti-androgenic effects of flutamide in all five androgen-sensitive tissues by all laboratories. Flutamide antagonised both reference doses of testosterone propionate (0.2 and 0.4 mg/kg/day) in a dose-responsive manner. The magnitudes of reductions in tissue weights were similar across all laboratories.

b) The need for use of all 5 tissues remains to be tested against weaker androgens and anti-androgens than testosterone propionate and flutamide, against lower concentrations of strong actors, and against substances that exert their hormonal effects via different mechanisms or pathways.

c) The responses seen appeared to be unaffected by the strain of rat that was used.

d) On average, fixed tissues were heavier than fresh tissues. However, there were no affects of fixation on the ability to detect the responses of testosterone propionate or flutamide, or on the CVs of the responses. This suggests that either procedure could be used.

e) The dorso-lateral prostate showed equivalent sensitivity to the ventral prostate to the androgenic and anti-androgenic effects of testosterone propionate and flutamide.

f) Testosterone propionate induced a significant, dose-related increase in serum testosterone levels, and decrease in lutenizing hormone levels. Serum testosterone level was a less sensitive endpoint than accessory sex tissue weight gain because 0.1 mg/kg testosterone propionate, which was sufficient to induce a significant weight gain in the accessory sex tissues, did not produce increased serum hormone levels. Similarly, lutenizing hormone levels were not significantly, or consistently, affected at the 0.1 mg/kg testosterone propionate dose.

g) Body weight change was associated with increasing androgen levels. Although the trend was significant, the weight change was relatively weak, and may not be sufficient as an endpoint for identifying androgens.

h) Testosterone propionate induced significant and dose-related decrease in adrenal weights. This weight loss was partially mitigated by treatment with the anti-androgen, flutamide at concentrations up to 3.0 mg/kg, and was fully mitigated at 10 mg flutamide/kg.

i) Testosterone propionate produced small, but significant, dose-related increases in liver and kidney weights that were not affected by flutamide treatment.

j) The use of 6 animals per dose group was sufficient to detect the androgenic and anti-androgenic activity of potent substances. Additional analyses are needed to determine if 6 animals per group will be sufficient for detecting weaker androgens and anti-androgens, or strong ones at lower concentrations whether fewer animals could be used.

k) The androgen-sensitive tissues are not all easily excised. The differences in the tissue weights, and coefficients of variance in the different laboratories, demonstrate that not all laboratories dissected the tissues with equal skill.

l) The animals should be castrated after preputial separation occurs (usually after 42 days). If they are castrated earlier, the glans penis (GLANS) has not fully separated, making it difficult to dissect in the control animals. This can affect the accuracy of the tissue weight measurements.

RECOMMENDED PHASE-2 VALIDATION STUDIES

91. The Lead Laboratory, following discussions of the VMG, developed a proposal for the design of Phase-2 of the inter-laboratory validation study (Annex 9). The recommended protocol includes the testing of known strong and weak androgens and strong and weak anti-androgens. A recommendation is also made that some of the tests be performed "blind," without the laboratory knowing the identity of the test chemical. At this time, 15 laboratories, from Europe, the US, Japan, and Korea, have agreed to participate in the Phase-2 studies.

Recommended Phase-2 test protocol

92. The recommended test protocol for the Phase-2 study would be essentially the same as the protocol used in Phase-1. The Phase-2 study design includes the testing of two chemicals per laboratory, so that each chemical would be tested in four laboratories.

Recommended Phase 2 test chemicals

93. The following chemicals are recommended for use in the Phase-2 testing:

Androgens	Anti-androgens
Testosterone propionate (TP)	Vinclozolin
Trenbolone	Procymidone
Methyl testosterone	Linuron
	p,p'-DDE
	Finasteride

a. Trenbolone would be evaluated for androgenic effects using at least four dose levels using both s.c. and oral administration.
b. Methyl testosterone would be evaluated for androgenic effects using the oral route at three dose levels.
c. p,p'-DDE and finasteride would be evaluated for anti-androgenic effects against TP.
d. Linuron would be evaluated for anti-androgenic effects against TP.
e. Vinclozolin and/or procymidone would be evaluated for anti-androgenic effects against TP using the oral route and three dose levels.
f. At least two of the laboratories would run two of the above anti-androgens chemicals "blind" against TP. These laboratories would also be responsible for performing dose-range studies to determine the anti-androgen doses to be used.

94. The criteria for chemical selection were: it is a known androgen or anti-androgen and has a mechanism of action that should be detected using the Hershberger assay protocol; the androgenic effects *in vivo* are well-documented and known to be detectable at dose levels that do not induce systemic toxicity; it is known or suspected to affect reproductive development or pregnancy by an androgen receptor-mediated mechanism; and, it is commercially available at a reasonable cost.

Recommended androgens

95. Testosterone propionate was the androgen used in Phase-1. For this reason, it is recommended for use in Phase-2 as the control against which the anti-androgen responses will be measured. The 0.4 mg/kg/day dose should be used based on the responses to this dose in Phase-1.

96. Trenbolone is used to promote muscle growth in cattle, and is expected to have a greater effect on

the LABC than on the other tissues. It is more effective by s.c. injection than by the oral route. Doses have been recommended for both routes of administration (see Annex 9). Trenbolone is difficult to ship because of export restrictions. Five laboratories in Europe have volunteered to procure samples for testing directly from the supplier.

97. Methyl testosterone differs from testosterone and trenbolone in that it is relatively potent when administered orally. Dose levels similar to trenbolone are recommended.

Recommended anti-androgens

98. Based on the responses in the flutamide Phase-1b experiments, it is recommended that 0.4 mg TP/kg/day be used as the reference androgen for the anti-androgenicity studies. At this dose, the tissues are larger than at the 0.2 mg TP/kg/day dose and are therefore easier to dissect and weigh. It is recommended that all other aspects of the Phase-1b protocol remain the same.

99. Vinclozolin and procymidone have similar mechanisms of action as flutamide, but are about 1-2 orders of magnitude less potent in both the Hershberger assay and as developmental toxicants. Dose ranges for their used have been recommended based on prior studies in the Lead Laboratory.

100. Linuron is a herbicide that is a weak agonist and also acts as an antagonist *in vivo*. The activity of this chemical in the Hershberger assay is dependent on the assay protocol. Linuron exhibits anti-androgenic activity when administered to castrate, immature males for 10 days, but not for 5 days, and it was negative in a 7-day castrate adult assay, and in intact adult rats. Dose levels have been recommended based on prior studies in the Lead Laboratory.

101. Finasteride was selected because it specifically inhibits 5α-reductase, a mechanism of anti-androgenicity that is not addressed by the other chemicals. It should show activity only in those tissues that contain 5α-reductase. Therefore, it should produce greater responses in tissues like the VP and SV, which have the enzyme, than in the LABC which do not have high levels of the enzyme.

102. p,p'-DDE is an effective antagonist in castrate, immature rats, but produced smaller effects in castrate adults. It was not effective in intact adult rats. However, in a pubertal rat assay, p,p'-DDE significantly delays puberty in the absence of androgen-dependent tissue changes. Recommended dose levels are based on published studies.

Tissues to be examined

103. The five androgen-responsive tissues used in Phase-1 (VP, SV, LABC, COWS, and GLANS) are recommended because they each have a unique responsiveness to different chemicals or sensitivities at different dose ranges. Additional rationales for the use of the five tissues is that the test is less likely to produce a false negative in response to weakly acting substances than if only one tissue were examined and, similarly, a false positive is less likely if the evaluation is based on multiple tissues than on the response of a single tissue. It is also recommended that the paired adrenal gland weights be included because this organ is affected by androgen-receptor antagonists and inhibitors of steroidogenic P450 enzymes, and responded to the TP and flutamide concentrations used in the Phase-1 studies.

REFERENCES

1) OECD. Detailed Review Paper on the Appraisal of Test Methods for Sex Hormone Disrupting Chemicals. OECD Monograph No 21. Paris: OECD, 1998.

2) OECD. Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Working Group, 10th-11th March 1998, ENV/MC/CHEM/RA(98)5. Paris: OECD, 1998.

3) EPA 1998a. Endocrine Disrupter Screening and Testing Advisory Committee (EDSTAC) Final Report EPA/743/R-98/003.

4) [http://www.epa.gov/scipoly/oscpendo/history/finalrpt.htm]

5) EPA 1998b. Endocrine Disruptor Screening Program: Proposed Statement of Policy. 63 FR 71542-71568.

6) Colborn T, Clement C, eds. Chemically Induced Alterations in Sexual and Functional Development. Princeton, NJ: Princeton Scientific Publishing, 1992.

7) EC/EEA/OECD/WHO. European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife. European Union Report EUR 17549. Workshop held at Weybridge, UK, Dec. 1996. Brussels: European Union, 1997.

8) SETAC–Europe/OECD/EC Expert Workshop on Endocrine Modulators and Wildlife Assessment and Testing (EMWAT) Report of an expert workshop held in Veldhoven, the Netherlands 10-13 April 1997. (published by SETAC Europe, September 1997)

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

10) NRC. Hormonally Active Agents in the Environment. Washington, DC: National Academy Press, 1999.

11) Toppari J, Larsen JC, Christiansen P, Giwercman P, Grandjean P, Guillette LJ Jr, Jegou B, Jensen TK, Jouannet P, Keiding N, Leffers H, McLachlan JA, Meyer O, Muller J, Rajpert-DE Meyts E, Scheike T, Sharpe R, Sumpter J, Skakkebaek NE. Male reproductive health and environmental xenoestrogens. Environ. Health Perspect. 104(Suppl. 4): 741-803 (1996).

12) Adami H, Bergstrom R, Mohner M, Zaronski W, Storm H, Ekbom A, Tretli S, Teppo L, Ziegler H, Rahu M. Testicular cancer in nine northern European countries. Int. J. Cancer 59: 33-38 (1994)

13) Carlsen E, Giewrcman A, Keiding N, Skakkabaek NE. Evidence for decreasing quality of semen during the last 50 years. Brit. Med. J. 305: 609-613 (1992).

14) Sharpe RM, Skakkabaek NE. Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? Lancet 341: 1992 (1993).

15) Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Gray LE, Kaatari S, Lucier G, Luster M, Mac MJ, Maczka C, Miller R, Moore J, Rolland R, Scott G, Sheehan DM, Sinks T, Tilson H. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: A report of the U.S.

EPA-sponsored workshop. Environ. Health Perspect. 104(Suppl. 4): 715-740 (1996).

16) OECD. Final Report of the OECD Workshop on the Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods (Solna Report, 1996) as presented to the Seventh Meeting of the National Co-ordinators of the Test Guidelines Programme, 18th-19th September 1996. ENV/MC/CHEM/TG(96)9. Paris: OECD, 1996.

17) Email from W. Mellert (BASF) to E. Zeiger (OECD Secretariat), January 29, 2002.

ACKNOWLEDGMENTS

The dedicated efforts and work of Dr. L. Earl Gray Jr. of the U.S. Lead Laboratory, and the statistical support of Mr. Mike Walker are appreciated.

ANNEX - 1

Participating Laboratories in Phase 1 of the OECD Validation of the Rodent Hershberger Assay

This information is only available to Government representatives of OECD member countries.

ANNEX-2

OECD PROTOCOL AND GUIDANCE FOR THE CONDUCT OF THE PHASE-1 VALIDATION OF THE RODENT HERSHBERGER ASSAY

Initial OECD work on the validation of the rodent Hershberger assay[1]

1. As agreed at the Second meeting of the OECD Validation Management Group (VMG) for the Screening and testing of Endocrine Disrupters (20-21 January 2000) and subsequently revised further by teleconference of the VMG on 6 March 2000.

INITIAL OECD WORK ON THE VALIDATION OF THE RODENT HERSHBERGER ASSAY[2]

INTRODUCTION

1. The overall aim of the validation work is to develop a robust assay that can be considered as the basis for an OECD Test Guideline. This document provides the essential requirements for the initial OECD work on the validation of the rodent Hershberger assay. More detailed practical laboratory protocols for the OECD Validation work may be built on the essential requirements contained in this document.

2. The rodent Hershberger assay was first described in 1953 (Hershberger *et. al.*, 1953). Since that time it has been used primarily in the pharmaceutical industry. A standardised and validated protocol has not been available for consideration internationally. This protocol provides the initial protocol for further standardisation and optimisation within the OECD.

3. The Hershberger assay is an *in vivo* short-term assay for chemicals that have the potential to act like endogenous sex hormones. The rodent Hershberger assay is similar in concept to the rodent uterotrophic assay - both measure as endpoints changes in specific tissues that normally respond to endogenous hormones. The focus of the Hershberger assay is on male sex hormone interactions while the uterotrophic assay's focus is on female sex hormone interactions.

4. The Hershberger and uterotrophic assays are both being considered by OECD as potential short term screening assays. The information generated by use of the assay 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.

INITIAL VALIDATION WORK

5. The aims of the initial OECD work on the Hershberger assay are to:

- Demonstrate the reliability of measuring sex accessory tissues among participating laboratories;
- Demonstrate the responses of the different sex accessory tissues to the reference androgen agonist and, testosterone propioante (TP) (CAS No. 57-85-2) and the reference androgen antagonist - Flutamide (FT) (CAS No: 1311-84-7);
- Enable any sources of variables to be investigated further, e.g., different strains of animals and to enable the protocol to be modified further as appropriate; and
- Enable a standard reference dose of TP to be calculated for use as positive control when detecting androgen agonists and as the negative control in case of detection of antagonists.

INITIAL CONSIDERATIONS

6. The rodent Hershberger assay evaluates the ability of a chemical to show biological activities

2. As agreed at the Second meeting of the OECD Validation Management Group (VMG) for the Screening and testing of Endocrine Disrupters (20-21 January 2000) and subsequently revised further at the teleconference of the VMG on 6 March 2000

consistent with the agonism or antagonism of natural hormones, that have masculinising effects. These hormones are known as androgens (e.g., testosterone).

7. Accessory sex glands and accessory sex tissues are dependent upon androgen stimulation to gain and maintain weight during and after puberty. If endogenous sources of androgen are removed, exogenous sources of androgen are necessary to increase or maintain the weights of these sex accessory tissues.

8. The sex accessory glands and tissues for this protocol are the: Ventral Prostrate (VP); Seminal Vesicles (SV); Coagulating Glands (CG); Levator ani plus Bulbocavernosus muscles (LABC); Glans Penis (GP) and Cowpers (or bulbourethral) Glands (CP) .

9. This protocol uses sexually immature male rats, castrated at peripuberty by removal of testes and epididymi (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. TP will initiate prepuce separation so that the Glans Penis (GP) can be 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 rodent is that the sex accessory tissues have a high sensitivity and small relative weight which both help to minimise variation in responses between individual animals.

10. Little is known about the response of individual sex accessory tissues to exogenous chemicals that may cause androgenic effects, although it has been shown that the male sex accessory tissues have different sensitivities to androgens and other steroid hormones. [Ashby *et al* (in press).] This differential sensitivity has been used historically and continues to be used to this day in the pharmaceutical industry by companies searching for chemicals that are anabolic but not either androgenic or oestrogenic. One example of differential sensitivity is the LABC muscles that lack the enzyme 5-alpha reductase. These muscles lack the ability to convert testosterone to its active form dihydrotestosterone. Weight increases of the LABC without concomitant weight increases in the VP, CG and SV glands (which contain 5-alpha reductase) may reflect an anabolic rather than an androgenic response.

11. As part of the development of this protocol, study variables have been standardised as far as possible based on historical experience and current research. The key variables not standardised in this protocol are the strain of rodent, diet, and housing conditions.

PRINCIPLE OF TEST

12. The rodent Hershberger assay is based on changes in weight of male sex accessory tissues in sexually immature castrated male rats.

13. Test substances may stimulate or, in the presence of a reference androgen, inhibit the stimulated development of sex accessory tissues.

14. The test substance is administered in graduated doses to several groups of male rodents for a number of consecutive days. Measurement of the weight of sex accessory tissues provides information on the androgenic nature of a chemical, however it can also provide additional information on whether effects are due to the effects on the androgen hormone receptor *in vivo* or on other relevant biochemical mechanisms, e.g., effect on other enzymes involved in the production of sex hormones such as 5-alphareductase.

15. In addition to the sex accessory tissues, body weight gain is a mandatory measurement to provide information on the general health and wellbeing of the animals. In the initial validation work, liver weight at necropsy is also a mandatory endpoint as some test substances may appear to be anti-androgenic by inducing an increased metabolism of TP by the liver. This may be indicated by an increase in liver size. Necropsy of the adrenals and kidneys may provide supplementary information about the effects of the test substance on other related biochemical pathways and are therefore optional supplementary endpoints. Measurement of serum testosterone and leutinising hormone may also be investigated in this context.

Androgen agonists

16. To test for androgen agonists a test substance is administered to immature castrated rats for ten consecutive days. TP is administered by daily sub-cutaneous injection. TP provides the positive control in studies with substances of unknown androgenic activity. The vehicle provides the negative control.

Androgen antagonists

17. To test for androgen antagonists, the test substance is administered to immature castrated rats for ten consecutive days together with a reference androgen agonist (TP). Administration of TP alone is used as the negative control which treatments are compared to for antiandrogenic activity. The weights of the sex accessory tissues after co-administration of the test chemical and reference androgen are compared with the weights of tissues from this control group.

DESCRIPTION OF METHOD/PREPARATIONS FOR THE TEST

Animal Species and Strain

18. 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, which has a different schedule of sexual development compared to other more commonly used strains such as Sprague Dawley, Long Evans or Wistar strains. 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* are met. The strain of rat used will be recorded in the report.

Acclimatisation

19. Healthy young animals that have been acclimatised to the laboratory conditions for 1-2 weeks following castration will be used. Animals will be observed daily, and any animals with evidence of disease or physical abnormalities will be removed. If castrated animals are purchased from an animal supplier the age of animals and stage of sexual maturity should be assured by the supplier and the time between castration and initiation of dosing will be counted as part of the acclimatisation period. In such cases the animals will be no more than 8 weeks of age at the initiation of dosing. A period of between one and two weeks acclimatisation has been chosen to allow sufficient period of acclimatisation while also allowing a laboratory to schedule the experimental work efficiently.

Housing and feeding conditions

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

21. There is currently insuffcient information showing the influence of laboratory diets on the identification of androgenic substances *in vivo*. Laboratories participating in the validation should use the laboratory diet normally used in their chemical testing work. 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*.

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

23. Each animal will be identified individually (e.g., ear mark or tag).

24. 6 animals of the same age and cohort will be used per treatment group.

Body Weight and the selection of animals for the study

25. Variations in body weight may be a source of variation in the weight of tissues of interest (especially the liver). This variation, if present, will increase variability within a group or among groups of animals. This may interfere with assay sensitivity, and possibly lead to false positives or false negatives.

26. Body weights will vary from study to study and different rodent strains. Each participating laboratory should establish its own procedure for limiting the variability in body weight. These procedures will be recorded in the report and should ensure that all groups of animals reflect normal variations expected for healthy animals.

27. As a precautionary measure, any effect of body weight on sex accessory tissue weight will be controlled in both the experimental design and data analysis phases of the study.

28. Within the experimental design the variation in body weight will be both experimentally and statistically controlled. Within the data analysis phase, body weight will be used as a covariate in the overall analysis.

29. Experimental control is accomplished in two steps. The first step involves selection of animals with relatively small variation in body weight from the larger population. Avoiding unusually small or large animals achieves this. A reasonable level of body weight variation within a study should be tolerated to ±20% of the mean body weight (e.g. 175g ± 35g). While this degree of variability may seem large it is not expected to alter the outcome of the study, as long as the animals are healthy, and will reduce the numbers of animals that would be rejected.

30. The second part of "experimental" control of body weight involves the assignment of animals to different treatment groups by a randomised complete block approach rather than by completely randomisation. 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. This variable is then included in the data analysis to adjust for differences in body weight.

Non-routine health and safety requirements

31. The test substances are known as possible reproductive and developmental toxicants and therefore appropriate precautions should be taken to protect personnel during the validation work, e.g.

necessary training, labelling and storage procedures, and protective handling procedures during dose preparation and dose administration.

32. 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.

PROCEDURE - INITIAL VALIDATION

33. The following procedure is focused on the initial validation work where the only test substances used are a reference androgen and a reference anti-androgen.

Reference substances and vehicle

34. The reference androgen will be Testosterone Propionate (TP), CAS No 57-85-2. The reference anti-androgen (or androgen antagonist control) will be Flutamide (FT) CAS No 1311-84-7.

35. In the initial validation work TP and FT will be administered in a specified laboratory grade stripped corn oil. All participating laboratories will use stripped corn oil to eliminate potential differences in absorption as a source of variation. Participating laboratories will be supplied with both TP and FT from the central chemical repository.

The number of test groups

36. Participating laboratories will first examine the response of the sex accessory tissues to the reference androgen-TP. This work will involve five test groups and one vehicle control group.

37. In a second step, three test groups will be studied where FT is co-administered with TP to examine the effect of the androgen antagonist on the sex accessory tissues. The negative control group will be the reference dose of TP. The second step will be conducted after the overall analysis of results from step1 from all participating laboratories has been completed.

Doses

38. All participating laboratories will use the same dose levels. The following table provides the requirements:

	Agonist response	**Antagonist response**
Vehicle Control	Vehicle	Vehicle
Negative Control	Provided by vehicle control	TP (ref)[3]
Group A	TP: 0.1 mg/kg/day	TP (ref)[3] FT 1 mg/kg/day
Group B	TP: 0.2mg/kg/day	TP (ref)[3] FT 5mg/kg/day
Group C	TP: 0.4 mg/kg/day	TP (ref)[3] FT 10 mg/kg/day
Group D	TP: 0.8 mg/kg/day	
Group E	TP: 1.6 mg/kg/day	

3. TP (ref) is the reference dose of TP established from the first stage of the study approximating to a ED70 on the ascending part of the dose-response curve

39. The dose of TP that provides the negative control will be established based on the evaluation of the initial TP dose response work (paragraph 36).

Administration of doses

40. TP is administered by s.c injection. TP is not administered orally as this is known to be less effective and to produce more variable results because absorption via the gut is influenced by many factors such as diet and gut flora.

41. FT will be administered by oral gavage.

42. For subcutaneous administration, all treatments are administered by s.c. injections on the dorsal surface of the animal. The maximum limit on the volume administered per animal is approximately 0.5 ml/kg body weight per day.

43. For oral administration, all treatments are administered by gavage. The maximum limit on the volume administered per animal will be 5 ml/kg/day.

44. The animals will be dosed in the same manner and time sequence for ten consecutive days at approximately 24 hour intervals. The dosage level will be adjusted for changes in body weight. The volume of dose and time that it is administered will be recorded on each day of exposure.

Good Laboratory Practice

45. Work will be conducted according to the principles of Good Laboratory Practice (OECD Good Laboratory Practice and Compliance Monitoring (OECD, 1998). In particular data will have a full audit trail and be retained on file. Data will be collected in a manner that will allow independent peer review. Calibration data for all balances used should be determined a part of the study and written records maintained.

OBSERVATIONS

Clinical observations

46. Animals will be evaluated 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.

47. 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 the reasons.

Body weight and food consumption

48. Individual body weights will be recorded prior to start of treatment (to the nearest 0.1g), on each day of administration period and prior to necropsy. Group means and standard deviations will be calculated.

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

Necropsy

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

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

52. The sex accessory tissue and liver weights are mandatory measurements. Adrenal and kidney weights are optional additional measurements.

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

54. The sex accessory tissues will be excised and their weights determined, for comparison with the weights of sex accessory tissues from the vehicle control group, or reference TP group (in the case of antagonist response). If serum hormones are to be measured as an option, the rodents will be anaesthetised prior to necropsy and blood taken by cardiac puncture. If serum hormones are to be measured, the method of anaesthesia should be chosen with care so that it does not affect hormone measurement.

55. It is important that persons carrying out the dissection of the sex accessory tissues are familiar with standard dissection procedures for these tissues. This will minimise a potential source of variation in the study. Ideally the same prosector should be responsible for the weighing a given tissue to eliminate inter-individual differences in tissue processing. If this is not possible, the necropsy should be designed such that each prosector weighs a given tissue from all treatment groups as opposed to one individual weighing all tissues from a control group, while someone else is responsible for the treated groups.

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

Measurement of sex accessory tissues

57. After necropsy, the sex accessory tissues will be removed and weighed without blotting (to the nearest 0.1mg). The excised tissues will be trimmed of any fat. Participating laboratories should ensure that the excision procedures used are reproducible over time and pay particular care to prevent variations in fluid losses from tissues during processing. A standard operating procedure will be followed for the excision of sex accessory tissue. This procedure will be provided by the Lead Laboratory.

58. After excision and weighing of the ventral prostate it will be fixed for 24 hours in 10% neutral buffered formalin (4% formaldehyde) and weighed again.

59. The following weight of the following sex accessory tissues will be measured:

- Ventral Prostate (VP) – fresh and fixed tissue weight (24 hours)
- Seminal vesicles together with coagulating gland (SV and CG) – fresh tissue weight
- Levator ani and bulbocavernous muscles (LABC) – fresh tissue weight
- Glans penis (GP) - fresh tissue weight
- Cowpers (or bulbourethral) Glands (CG) – fresh tissue weight

60. The weight of the adrenal glands and the kidneys and levels of serum leutinising hormone and testosterone may be measured as optional endpoints.

REPORTING

Data

61. Data will be reported individually and for each group of animals (i.e. body weights, liver weight, accessory sex tissue weights, optional measurements and other responses and observations). 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. To assist data reporting and compilation a standardised electronic spreadsheet will be used by participating laboratories to report data during the initial validation work.

Test report

62. The test report must include the following information:

Laboratory identification

Test substance:

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

Vehicle:

Test animals:

- Species/strain used
- Number, age and sex of animals
- Source, housing conditions, diet, and bedding
- Individual weights of animals at the start of the study (to nearest 0.1 g)

Test conditions:

- Housing conditions
- Number of animals per cage
- Necropsy procedures
- Diet

Results:

- Daily observations

- 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
- Animal age
- Order of animal in the necropsy

- TP treatment (Yes or No and dosage level)
- FT treatment (Yes or No and dosage level)

- Body weight at start of dosing (to nearest 0.1g)
- Body weight at necropsy (to nearest 0.1g)

- Weights of sex accessory tissues[4] (to the nearest 0.1g)
  - Ventral prostate (fresh weight and weight after fixation)
  - Seminal vesicle plus coagulating gland, including fluid (fresh weight)
  - Levator ani plus bulbocavernosus muscle (fresh weight)
  - Glans penis (fresh weight)
  - Cowpers Gland (fresh weight)
  - Liver (fresh weight)
  - Kidney weight (optional)
  - Adrenal weight (optional)
  - Serum LH (optional)
  - Serum T (optional)

- General remarks and comments

Discussion

Conclusions

Interpretation of results

64. Statistical comparisons in individual laboratories will be made for the different sex accessory by analysis of variance. For androgen agonism, the test substance groups will be compared to the vehicle control. A statistically significant increase in tissue weight will be considered 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. A statistically significant decrease in tissue weight will be considered a positive antagonist result. If more than one set of comparisons is required, all comparisons will be conducted separately for each test group against its control.

4. In a parallel protocol, identical in all aspects to this, some laboratories may generate data by fixing the sex accessory tissues before separation and weighing. This is an optional additional protocol for comparative purposes.

ANNEX – 3

Table A. Rat strains used, ages, suppliers, and husbandry conditions.

LAB	Rat Strain and Supplier	Bedding material	Rats per cage	Diet	Age at Castration (Days)	% Controls With PPS	Starting Weight (gms)
1	Alpk:APfSD; AstraZeneca Breeding	Paper (coffee filter)	3	R&M No. 1 Special Diet Services Ltd., batch 6458	42-44	100	224
2	Sprague-Dawley; Iffa Credo, France	Suspended steel wire mesh	1	Pietrement aliment type M20; lot 991125	33-47	100	271
3	CRL:WI(GLX/BRL/HAN) IGS BR; Charles River, Germany	Suspended steel wire mesh	1	9433LL Meal,Eberle Nafag AG,CH - Lot 44/00	38	0	160
4	HSD/CPB-WU			Eberle Nafag AG, CH Gossau, NAFAG no.9349	31	0	163
5	Crj:CD(SD) IGS; Hino Breeding Ctr.	Sunflake®	3	Oriental Yeast Co., Ltd, Tokyo, Japan	42-44	100	231
6	Crl CD®(SD) IGS BR; Charles River, France	Autoclaved sawdust	3	AO4 C pelleted maintenance diet, batch No: 00331	42	100	254
7	CRL:CD(SD) IGS BR; Charles River	Suspended steel wire mesh	3	PMI Certified 5002 Mash Lot # May 22 00 3A	42	67	311
8	Crj:CD(SD)IGS; Charles River Japan	Suspended steel wire mesh	1	CE-2; Clea Japan Co. Ltd.; Lot#, E2050-P8	40	100	213
9	CD Sprague-Dawley; Charles River UK	Suspended steel wire mesh	3	Special Diet Services RM1(E) SQC expanded pellet (lot 6706)	42	100	237
10	Jcl:Wistar; Fuji Farm, Clea, Japan	Suspended steel wire mesh	3	CE-2, Lot no. E2050-P8 (Clea Japan, Inc.)	41-43	100	214
12	Crj CD(SD)IGS SPF/VAF; Charles River Japan	Suspended steel wire mesh	1	Oriental Yeast Co., Ltd., Lot No. 000405	41-44	100	256
13	Crj:CD(SD)IGS; Tsukuba Facility, Charles River Japan	Autoclaved hardwood chips (Beta Chip)	2	Oriental Yeast, lot no. 000412A1	44-46	100	265
14	Sprague-Dawley; Korea FDA	Autoclaved elm wood	3	PMI LabDiet, 5014	40	83	224
15	Crj:CD(SD)IGS; Charles River Japan	Autoclaved “White flake®”	3	CRF-1, Lot No. 00.05.09	42	100	223
16	CD; Charles River, Raleigh, NC, US	Sani-chips cage litter	1	Pelleted Purina Certified Chow # 5002	44-46	83	349
17	Crj:CD(SD)IGS; Charles River Japan	Beta chip	3	Oriental Yeast Company, CRF-1, Lot # 000208, 000602	44	100	234

Table B. Animals used for the Phase-1b studies; ages and weights

LAB	Rat Strain	Age at Castration (Days)	Age on Study (Days)	% Controls With PPS	Starting Weight (gms)	Bedding material	Rats per cage	Diet
5	Crj:CD(SD) IGS	42-44	51-53	100	229	Sunflake®	3	Oriental Yeast Co., Ltd, Tokyo, Japan
8	Crj:CD(SD)IGS	40-42	46-48	100	209	Suspended steel wire mesh	1	CE-2; Clea Japan Co. Ltd.; Lot# E2050-P8
10	Jcl:Wistar	41-43	47-49	n.r.	230	Suspended steel wire mesh	3	CE-2; Clea Japan, Inc., Lot # E2050-P8
12	Crj CD(SD)IGS SPF/VAF	41-44	52-55	100	256	Suspended steel wire mesh	1	Oriental Yeast Co., Ltd., Lot No. 000405
13	Crj:CD(SD)IGS	44-46	51-53	100	255	Autoclaved hardwood chips (Beta Chip)	2	Oriental Yeast, lot no. 000412A1
15	Crj:CD(SD)IGS	41-43	48-50	n.r.	207	Autoclaved “White flake®”	3	CRF-1, Lot No. 00.05.09
17	Crj:CD(SD)IGS	43-45	50-52	100	231	Beta chip	2	Oriental Yeast Company, CRF-1, Lot # 000208, 000602

PPS: Preputial separation
n.r.: not reported

ANNEX - 4

Necropsy Data Sheet

REQUIRED **Body and organ Weights (fresh)**

Date of Necropy

Age of animals:Days

Order of Necropsy	Animal ID	Time of Day	Dose Group	PPS Yes or No	Body Weight grams	Glans Penis mg	Ventral Prostate mg	Seminal Vesicle Coag. Gland and Fluid mg	LABC mg	Cowper's Glands mg	Liver grams	Ventral Prostate Fixed mg
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
Prosectior Initials												

Note: PPS refers to preputial separation, the glans penis weight cannot be measured if the animal is PPS negative, i.e. the prepuce has not separated from the glans penis

Key Necropsy Personnel

OPTIONAL ENDPOINTS

Adrenals mg	Kidneys mg	Serum Testosterone ng/ml	Serum LH ng/ml	Other 1	Other 2	Other 3	Other 4	Other 5	Other 6
Prosectior Initials									

ANNEX – 5

LEAD LABORATORY SUMMARY REPORT OF INITIAL WORK TOWARDS THE VALIDATION OF THE RAT HERSHBERGER ASSAY: PHASE-1A, ANDROGENIC DOSE RESPONSE EFFECTS OF TESTOSTERONE PROPIONATE

Lead Laboratory Summary Report of the OECD Interlaboratory Study on the Hershberger Assay: Phase 1A: Dose Reposnse Effects of Testosterone Propionate (sc)

Prepared by Leon Earl Gray Jr., Endocrinology Branch, Reproductive Toxicology Division, National Health and Environmental Effects Research Laboratory, US EPA, NC 27711

Reviewed by Mike Walker and William Owens
Their comments are incorporated

INTRODUCTION

1. In the fall of 2000, 17 laboratories participated in an interlaboratory investigation of the Hershberger assay, using a protocol developed by the OECD earlier in the year. This report presents the lead laboratory's summary of the results of the data analysis from these tests of the protocol.

PRIMARY STUDY: THE OECD PHASE IA INTERLABORATORY STUDY: TESTOSTERONE PROPIONATE DOSE RESPONSE

2. In this investigation, each laboratory examined the effects of graded doses of testosterone propionate (TP) (0, 0.1, 0.2, 0.4, 0.8, and 1.6 mg/kg/d, sc (in 0.5 mL corn oil/kg/d, for ten consecutive days) on organ weights in the immature castrated male rat. Sample sizes were designed as six rats per group. Two laboratories included "untreated" controls in addition to the "vehicle control".

3. In an attempt to start off with a population of animals of fairly uniform size, the animals on study were typically selected from a larger population of about 50 animals, with the "outliers" not being included in the investigation. It was suggested that the range in weight at the start of the study should not exceed 20%. In addition, animals were randomly assigned to treatments in a manner that provided equivalent initial body weights in each group within a laboratory.

4. A range in age at castration and the initiation of treatment were recommended (and followed). However, it was deemed unnecessary, based upon the results of the OECD uterotropic assays study and the literature, to standardise the strain of rat used or the diet in each laboratory. Fifteen of 17 laboratories reported data indicating that all recommendations were followed (Table 1). One laboratory (lab 1) did not submit all of the data in a usable format, while another provided ages that ranged over 14 days rather than giving a specific age (lab 2). In the latter case, it is unlikely that the animals actually differ in age by 2 weeks and are the same age, but this age is unknown.

5. Sixteen of 17 laboratories successfully executed the protocol, as designed. One laboratory (lab 11), however, administered TP at micrograms/kg/d rather than mg/kg/d.

6. Most laboratories measured all of the "required" endpoints: ventral prostate (fresh), ventral prostate (weighed after 24 hours of fixation), seminal vesicle (plus coagulating glands, presumably with their fluids), glans penis, Cowper's gland and levator ani plus bulbocavernosus muscle weights. One laboratory (lab 1) did not report the fixed weight of the ventral prostates, and it appears that at least three labs weighed the LA rather than the LABC muscles, as required. This variation in dissection of the LABC affected the absolute weights of this tissue, but did not alter the response to TP. Body, paired kidney and liver weights were reported by all of the laboratories. In addition to the "required" endpoints, several laboratories weighed the adrenals, three laboratories (labs 10, 15, 17) weighed the dorsolateral prostate, and three laboratories (labs 12, 15, 17) measured serum testosterone and LH levels.

AN ADDITIONAL COMPARATIVE STUDY

7. Three laboratories (labs 10, 15, 17) executed a study in parallel to the primary study to determine how altering the method of tissue dissection and weighing affected the weights of the ventral and dorsolateral prostates, the seminal vesicles and Cowper's glands. In these studies, the sex accessory glands were dissected as a unit, not weighed fresh as above and the glands were preserved in fixative for 24 hours after which they were separated and weighed.

DATA ANALYSIS

8. Means, standard errors and the coefficient of variation (CV or the standard deviation divided by the mean, as an estimate of relative variability) were calculated for each endpoint using PROC MEANS on SAS (available with SAS version 6.08 on the USEPA IBM mainframe). ANOVAs were done using PROC GLM for each laboratory and pooled over the 16 laboratories (excluding lab 11 that administered TP in μg/kg/day). The ventral prostate (VP) (fresh and fixed), seminal vesicles plus coagulating glands with fluids (SV), levator ani plus bulbocavernosus muscles (LABC), Cowper's glands (COWS), glans penis (GLANS) and dorsolateral prostate (DLP) data are presented in attached Tables. Examination of the CV among endpoints allows one to compare the statistical precision in the weight of a tissue. Some endpoints are inherently more variable than others while in some cases error in dissection or weighing can increase the CV. Comparison of means and CVs across labs allows one to determine if the technique varies greatly from lab to lab, in which case additional efforts may be necessary to standardise dissections and weighing of the tissues.

9. Data were then analysed by ANOVA on PROC GLM for each laboratory (with dose as a main effect) with and without initial body weight (the weight at the start of dosing). Data for each endpoint also were analysed as a two-way ANOVA, with dose and laboratory as main effects, so the magnitude of the overall dose and laboratory effects and their interaction could be determined.

10. In addition, the fact that CV for each androgen-dependent organ weight was fairly constant as the means increased, the SD being proportional to the mean, indicated that heterogeneity of variance existed. For this reason, the data were transformed using LOG10. In particular, this transformation provides for a more valid comparison of the effects of TP on organ weights at lower dosage levels.

11. These analyses also were conducted with initial body weight as a covariate. Initial body weight at the start of the study was used as the covariate rather than body weight at necropsy, because the administration of TP significantly affected body weight by increasing body weight gain (expected for an anabolic steroid). Hence, final body weight is not a good covariate because it also is affected by treatment. This covariate adjusts the analysis for experimental variation from several sources (i.e., first, large differences in the size of the rats from lab to lab, a large component of which appeared to arise from the use of different aged animals or different strains, and, second, differences in the size of the rats on study within a lab). Data were not analysed using "organ weights relative to body weight" as this manipulation makes several assumptions about the relationship between body size and organ weights, which often are invalid (e.g., first, that a linear relationship exits at all and, second. that any relationship is linear and its line goes through the origin). Instead, the increase in tissue weights of treated animals was expressed relative to the control tissue weight.

12. In addition to means and CVs, R-square (R2) values for different effects were calculated and presented in the tables. An R2 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 estimate of the strength of the association for an effect with an endpoint. This can be used to compare the robustness of the TP effect across endpoints, the variation from lab to lab, or to what degree the dose response curves vary from lab to lab (as indicated by the R2 for the lab by dose interaction). It is also useful to note how the R2 for

an effect declines after adjustment of the data with analysis of covariance using initial body weight as a covariate.

13. For the five androgen-dependent sex accessory tissues (SV, VP, GLANS, COWS, and LABC) the data were "normalised" in order to visually compare the shapes of the dose-response curves for each lab such that the data range from 0 to 100%. In this normalisation, the vehicle control value was set to zero, while the response seen in the high TP dose group was set at 100% (example of how the ED-70 calculation is arrived at is shown in Figure 1).

RESULTS

Overall "normalised" dose-response curves for the five androgen-dependent tissues

14. Figures 2 and 3 display the dose-response curves for the five tissues, pooled across all the labs and normalised such that the values for each organ range from zero to 100%. One objective of the current investigation was to determine an approximate ED70-value for TP to be used in subsequent phases of the OECD Hershberger assay standardisation and validation exercise. As seen in Figure 1, the visually estimated ED70 values range from ~0.2 for the GLANS to ~0.8 mg TP/kg/d for the SV. It appears from these figures that three distinct dose-response relationships exist for the five endpoints. The glans penis reaches a maximum response at a lower dosage level than the other four tissues, while, in contrast, the SV shows a more gradual and more linear response over the dose range used in the current studies. The LABC, COWS and VP appear to respond to the same degree to different doses of TP and are intermediate between the GLANS and SV curves. These normalised values were not analysed statistically and the details of the statistical analyses of each organ are discussed below.

Effect of TP on Ventral Prostate Weight in the OECD Phase I Interlaboratory Study

15. The ventral prostate (fresh) weight data will be discussed first. However, many of the results of the data analysis are very similar for the VP (Fixed), SV, LABC, GLANS and COWS. Hence, the VP analysis will be covered in more detail and then referred to in subsequent sections of the document.

16. The ventral prostate (fresh) weight data from 16 labs over the 6 TP dosage groups are shown in Table 2 and Figures 4-6. Figure 4 displays the mean values and standard error of the mean (SEM) (from PROC MEANS, with an SEM, unadjusted for lab to lab variability). These means and the sample sizes also are shown at the bottom of Table 2. In the overall analysis of the VP data, even the 0.1 mg/kg/d TP dose differed from control by $p < .0001$.

17. The overall CV for VP weight was 25%, but ranged from 16% in one lab to 42% in another. In general, as the CV increased, the F and R2-values for the effect of TP declined. F-values ranged from 12 to 210 and R2 ranged from 67 to 97% for the effect of TP from lab to lab. In 15 of 16 of the labs, the control value differed from the 0.1 TP dose by $p < 0.01$. One would expect that labs with higher CVs would have more difficulty detecting less robust effects than those described here. Pooled (across dose) VP values ranged from 90 to 220 mg (Table 2). As discussed below, some of this variability arises from the use of different size animals in different labs. An examination of the magnitude (R squared or R2) of the effect of TP across the labs indicates that the size of the animal is not related to the ability to detect TP-induced changes in ventral prostate weight, i.e. the assay is robust across a broad range of animals with mean body weights which range from 160 g in one lab to 349 g in another.

18. In the overall analysis of the log-transformed VP-weight data, the F-value for treatment was 1143 and the R2 was about 83.

19. The lab to lab variability, termed a lab effect, was highly significant and accounts for about 7% of the variance (R2 lab=6.6%). Due to the large numbers of animals and precision of the data, the lab by dose interaction (which has an R2 of 2.5%) is also highly significant ($p < 0.001$).

20. When initial body weight is included in the analysis of the log-transformed data the difference between the labs is reduced by 28%. In most labs there was no relationship between body weight and ventral prostate weight within the data. Hence, the overall significance in body weight versus ventral prostate weight in the combined analysis (all labs together) results from large differences between labs in the size, and age, of animals on study.

21. The lack of relationship between VP and body size seen in this study results in part from the design of the study and this conclusion cannot be extended to other protocols and should not be taken to indicate that controlling body weight is not an important consideration. Although within each lab, body weight was unrelated to VP, SV, LABC, COWS and GLANS weights, body weight did covary significantly with liver, kidney and adrenal weights.

Effect of preservation of the ventral prostate in fixative for 24 hours on the tissue weight

22. Figure 7 and Table 3 compare the VP weight data from 16 labs, weighed fresh and then reweighed 24 hours later after preservation in fixative. VP tissue weighed significantly more ($p < .0001$) after fixation than before. In addition, the increase in weight was affected by the size of the tissue, increasing with TP dose (Figure 8). It seems reasonable to conclude from this study, that fixation of VP alters tissue weight such that it is less "accurate" (i.e. the weight deviates from the "true" value, assuming that the fresh weight is the "true" value).

23. Although fixation of the VP alters the weight of the tissue, there is no indication from the analysis of these data that this method altered the ability to detect the effects of TP or altered the variability from lab to lab (Table 3). If one compares the overall F-values, R2s for treatment (TP) and lab, or the CV, from tables 2 (fresh) and 3 (fixed) the statistical analyses are almost identical. On a lab-by-lab basis, the weight of the VP after fixation differed significantly between the vehicle control and 0.1 mg TP/kg/d by $p < 0.05$ for all 15 labs, which measured the VP after fixation.

Results of the comparative study examining fresh versus fixed tissues (VP, COWS and SV) in different animals after TP treatment (0, 0.1, 0.2, 0.4, 0.8, and 1.6 mg/kg/d sc)

24. Three laboratories extended the study of the effect of tissue fixation to other tissues. As shown in Figure 9, fixation of the VP in these labs again significantly increased ($F=27.5$, $p < 0.001$) the weight of this tissue (as analysed by a 3 way ANOVA with TP dose (6 doses), lab (3 labs) and method (fresh versus fixed) as main effects. The effects of fixation on the COWS and the SV in these labs is also shown in Figure 9.

Effect of TP on Seminal Vesicle Weight in the OECD Phase I Interlaboratory study

25. The overall effect of TP on SV weight is shown in Figure 10 and at the bottom of Table 4. The raw data from the 16 individual labs are shown in Figures 11, 12 and Table 4. In general the results are very similar to those reported above for the VP in the Interlaboratory study. TP-treatment significantly increased SV size (overall F-value=2535, R2 for treatment = 88%) at all dosage levels. The labs differed significantly ($R2=6.2$, $p < 0.001$) and the dose response curves also differed (slightly, but significantly) by lab ($R2=2.1\%$, $p < 0.001$). In 15 of 16 labs, the SV in the 0.1 mg TP/kg/d group was significantly greater than the control SV group ($P < 0.05$). Adjusting the data for initial body weight reduced the lab-to lab effect by 54%. On a within lab basis, body weight was generally not significantly correlated with SV size.

26. The overall CV (the CV of the data pooled from all laboratories) = 22%, similar to the individual CVs of the VP, which ranged from 13% to 37%. SV values, by lab pooled over dose (Table 4), ranged from 219 in one lab to 588 mg in another. As indicated above, about half of this variability was related to the size of the animals used in the study. Although initial size of the animals contributes to the variability in the results from lab to lab, there is no indication that this alters the ability to detect the effect of TP within a lab or in the overall analysis. Hence, the robustness of the assay is not compromised by the use of animals of varying sizes in different studies, as long as the methods are precise and the size of the animals is controlled within a laboratory within the prescribed limits.

Effect of TP on LABC Weight in the OECD Phase I Interlaboratory study

27. The effects of TP on growth of the LABC are shown in Figures 13-15 and Table 5. TP increased the growth of the LABC (F-value=861, R2=56%), and a significant lab effect also was detected (R2=36%). It is noteworthy that, as compared to the SV and VP, the lab effect is increased for the LABC. This occurs in spite of the fact that the overall CV of 12% of the LABC is about half that of the SV and VP. As indicated above, some of the lab effect resulted from the fact that three of the labs weighed only the LA muscle rather than the complete LABC. Adjusting the data using initial body size reduces the lab effect by 24%. Taken together, these results suggest that some of the labs weighed the LABC differently from others. This does not appear to be a "strain" effect because two (labs 2, 4) of the four (labs 1, 2, 3, 4) labs with low overall LABC values are using SD rats, the strain used by a majority of the labs. We have discussed this with some of the PIs in these labs, and they have confirmed that they did not weigh the entire LABC. While, the use of smaller animals might contribute to the lab-to-lab variability as two of these labs have used smaller and/or younger rats (labs 3, 4), two of the labs with low LABC values used animals of "average" size (labs 1, 2). Although this difference exists, it had no impact on the ability to detect the effects of TP in any lab. In 15 of 16 labs, including all four of the above with relatively low LABC values, the control LABC weight was significantly lower than that of the 0.1 mg TP/kg/d dose group (only lab 8 was not significant, with a $p < 0.25$ for 0 versus 0.1 using log transformed data with body weight as a covariate). For the LABC weight, different labs had fairly similar dose response curves (Figure 15). The R2 for the interaction of lab * dose was 1.3% ($p < 0.03$), being about half that of the VP or SV.

Effect of TP on Glans Penis (GLANS) Weight in the OECD Phase I Interlaboratory study

28. The effects of TP treatment on GLANS weight is shown in Figures 16-18 and Table 6. TP treatment significantly increased glans penis weights (F-value 222, $P < 0.001$, R2 for treatment = 45%). This endpoint was of similar precision to the LABC having a CV of 14%, about half that of the VP or SV. The effect of TP was smaller on this tissue in relative magnitude than the VP, SV and LABC.

29. On a lab by lab basis, 15 of 16 labs found that the lowest dose of TP caused a significant increase in GLANS size ($P < 0.01$ except for BAY, $p < 0.1$). The R2 varied from a low of 29% in one lab to a high of 93% in another, while the CVs varied inversely with R2 from 6 to 26%. The lab effect was relatively large (as compared to VP or SV). The lab effect was reduced by 25% by inclusion of initial body weight as a covariate in the analysis of variance. The dose response curves were similar for this endpoint; the lab by dose interaction being significant (R2=4.5%, $P < 0.03$).

30. As indicated in Table 1, castration of the immature male rat prior to 40 days of age, which is around the normal age at preputial separation (PPS), precluded PPS in the vehicle controls. Three labs (labs 3, 4, 11) castrated animals at 38, 31 and 35 days of age, respectively and none of the controls in these labs displayed PPS (Figure 34). The failure of PPS likely complicates and confounds and accurate measurement of GLANS weight at necropsy in the control group due to the possible necessity to excise the prepuce from the GLANS during dissection.

31. If this endpoint is to be included in subsequent OECD evaluations of the Hershberger assay, it appears necessary to require that the animals be castrated at 40 days of age, or later, depending upon the natural timing of PPS. Glans penis weight does have considerable utility because it is sensitive to low doses of TP and can easily be dissected in castrate-no TP (if PPS positive). This contrasts with the VP, SV and COWS, which are hard to dissect in the vehicle control animals, being small and embedded in fat.

Effect of TP on Cowper's gland (COWS) Weight in the OECD Phase I Interlaboratory study

32. The effect of TP on the growth of the COWS is shown in Fig 19-21 and Table 7 and is statistically significant at all dosage levels ($p < 0.001$). The F-value for TP-treatment is 773 with an overall R2 of about 74%, with a CV of 22%. The lab effect is significant ($p < 0.001$) with an R2 of about 14%. Addition of initial body weight as a covariate reduces the lab effect by 44%.

33. The CVs range from 13% to 38% and the R2 for TP treatment range from 60% to 97%. Overall means, pooled across dose, range from 15 mg in one lab to 48 mg in another. Of note, is the range in values and CVs for COWs weight in the vehicle control group. One lab reports weights of 1.6 mg, while another reports values of 16 mg. CVs are much higher in this dose group for this endpoint than in any of the TP-treated groups. It is obvious that weighing these glands in the castrated immature male rat is technically difficult and the control values may be highly variable within and across laboratories, depending on the skill of the dissecting technician. In the overall ANOVA the dose by lab interaction has an R2 of 3.4% ($p < 0.001$).

Effect of TP on body weight and body weight gain in the OECD Phase I Interlaboratory study

34. Body weights of the animals on study varied greatly from lab - to - lab (Figures 22 - 24). One source of variation in the size of the animals is obviously related to the age at which they assigned to treatments and necropsied. However, the small size of the rats from one lab (lab 3) may be related to the strain used (described as lab 8). These rats were much smaller than SD rats of similar age, and they gained much less weight than did rats from other labs throughout the experiment (Figure 23). It must be emphasised that the fact that these rats appeared to differ in terms of weight for age, and growth during the study, did not compromise the robust nature of the responses of the tissues to TP in these animals.

35. Although the body weights did not show a statistically significant impact from the TP doses, body weight gain during the study was significantly enhanced by TP-treatment in a dose related manner (Table 8 and Figure 25) (Fs for TP=60 and lab=107). Weight gain was significantly related to initial body weight (F=25), indicating that larger animals gained more weight.

Effect of TP on nonreproductive organs measured by some or all laboratories in the OECD Phase I Interlaboratory study

36. All tissues contain androgen-receptors and, to some degree, respond to TP-treatment. In the current study TP treatment significantly increased the means of the pooled kidney (Figure 26) and liver (Figure 27) weights (both, $p < 0.0001$). TP-treatment also caused a dose-related reduction in the means of pooled adrenal weights (Figure 28, $p < 0.0001$). Some of these effects were not significant in all labs when tested individually because they are much less robust than the effects seen in the sex accessory tissues, described above.

Effects of TP on serum testosterone and LH levels measured by three laboratories in the OECD Phase I Interlaboratory study

37. In the current study, immature male rats were castrated and injected with TP. This should result in relatively constant levels of serum testosterone (T), which in turn should suppress the levels of LH in the serum by inhibiting pituitary LH secretion. However, the serum hormone data from the three labs is extremely

variable. For serum testosterone the CV=76 % and apparently the radioimmunoassays (RIAs) are not of sufficient sensitivity to accurately measure serum T levels in the 0.1 mg TP/kg/d dose group, so it cannot be statistically distinguished from the corn oil control even though this is a highly effective dose of TP. For serum LH, the CV=42%, there is a large lab effect, and LH is not reduced by TP-treatment until the dose reaches 0.4 mg/kg/d.

38. In contrast to the inability to measure serum testosterone by RIA in the low dose group, the uniform response of the sex accessory tissues across the 14 laboratories in the low dose group and all the higher dose groups clearly indicates that sc injection of TP produced uniform, dose-related internal exposures to testosterone. For four of the five androgen-dependent organs (SV, VP, COWS, and LABC), each administered dose of TP produced effects that differed significantly from every other dose group (analysed using Duncan's Multiple Range test). This implies that the actual internal testosterone levels also did not overlap between the groups.

39. Proponents of the use of various serum hormone levels, or even more sophisticated measures, should consider the apparent technical difficulties presented. It should be noted that the RIA kits themselves are not standardised. In addition, it appears that such measures will require a significant increase in the numbers of animals to attain sufficient power. Similar difficulties were encountered in phase 1 of the program to enhance the 407 Test Guideline. This is consistent with the difficulties here in standardising and validating such measures.

Lack of Effects of corn oil administration (sc) on organ and body weights as measured by two laboratories in the OECD Phase I Interlaboratory study: Comparison of injected versus untreated controls

40. In the current interlaboratory study, two laboratories included untreated controls in addition to the vehicle-injected controls. The analysis of these data are included because they demonstrate that administration of corn oil, at 0.5 ml/kg/d for ten days did not induce any changes in reproductive organs, did not induce changes in liver weight, and, based on adrenal weight, did not "stress" the animals by the daily handling and injection.These results are similar to those between untreated controls and vehicle treated controls in phase 1 of the uterotrophic validation program.

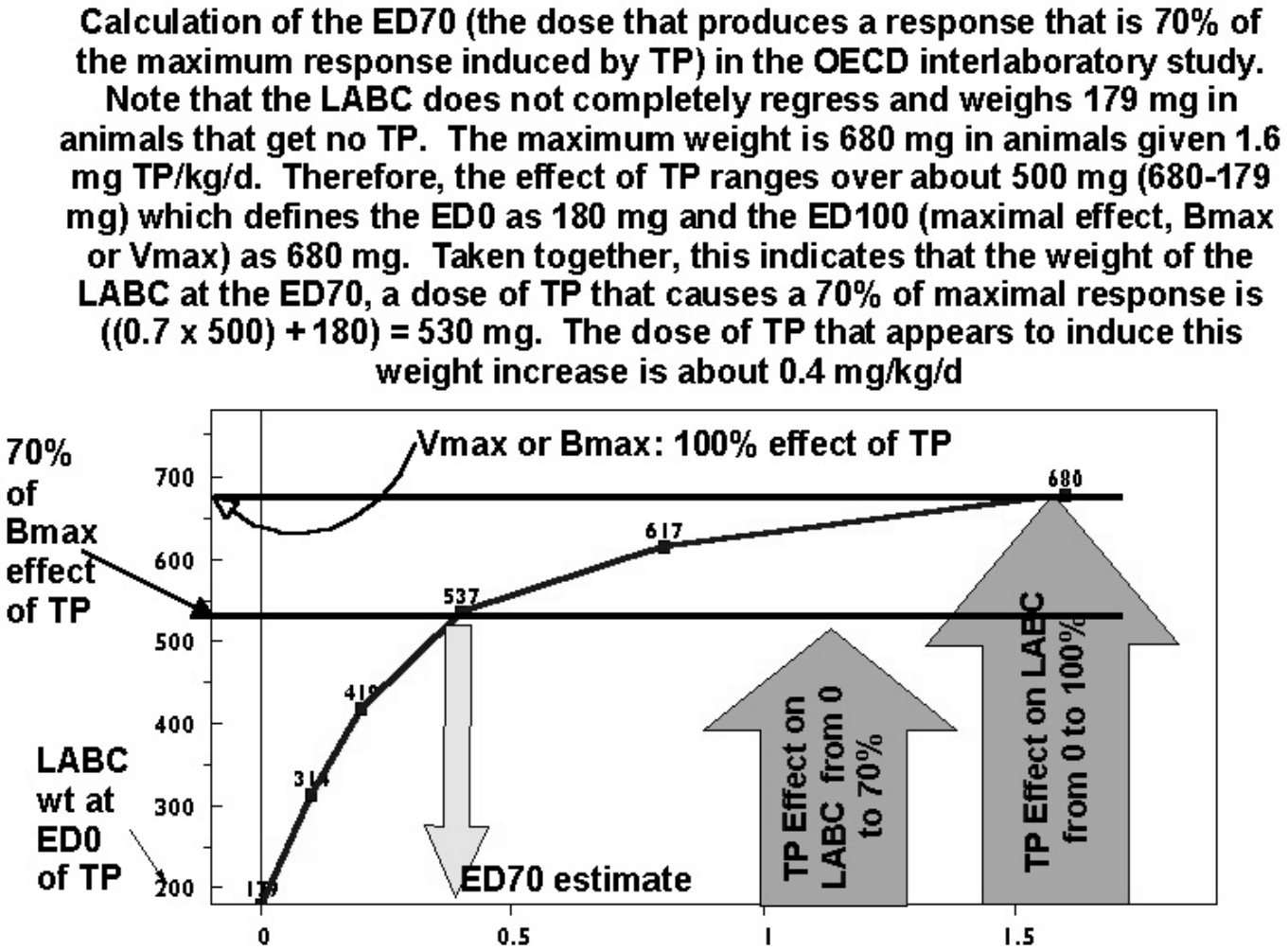


Figure 1. The x-axis is the testosterone propionate dose in mg/kg/d administered subcutaneously. The y-axis is the weight of the levator ani/bulbocavernosus muscle complex (LABC). This curve and markings demonstrate how an ED70 would be estimated from the data as 70% of the maximum effect as the dose response curve plateaus.

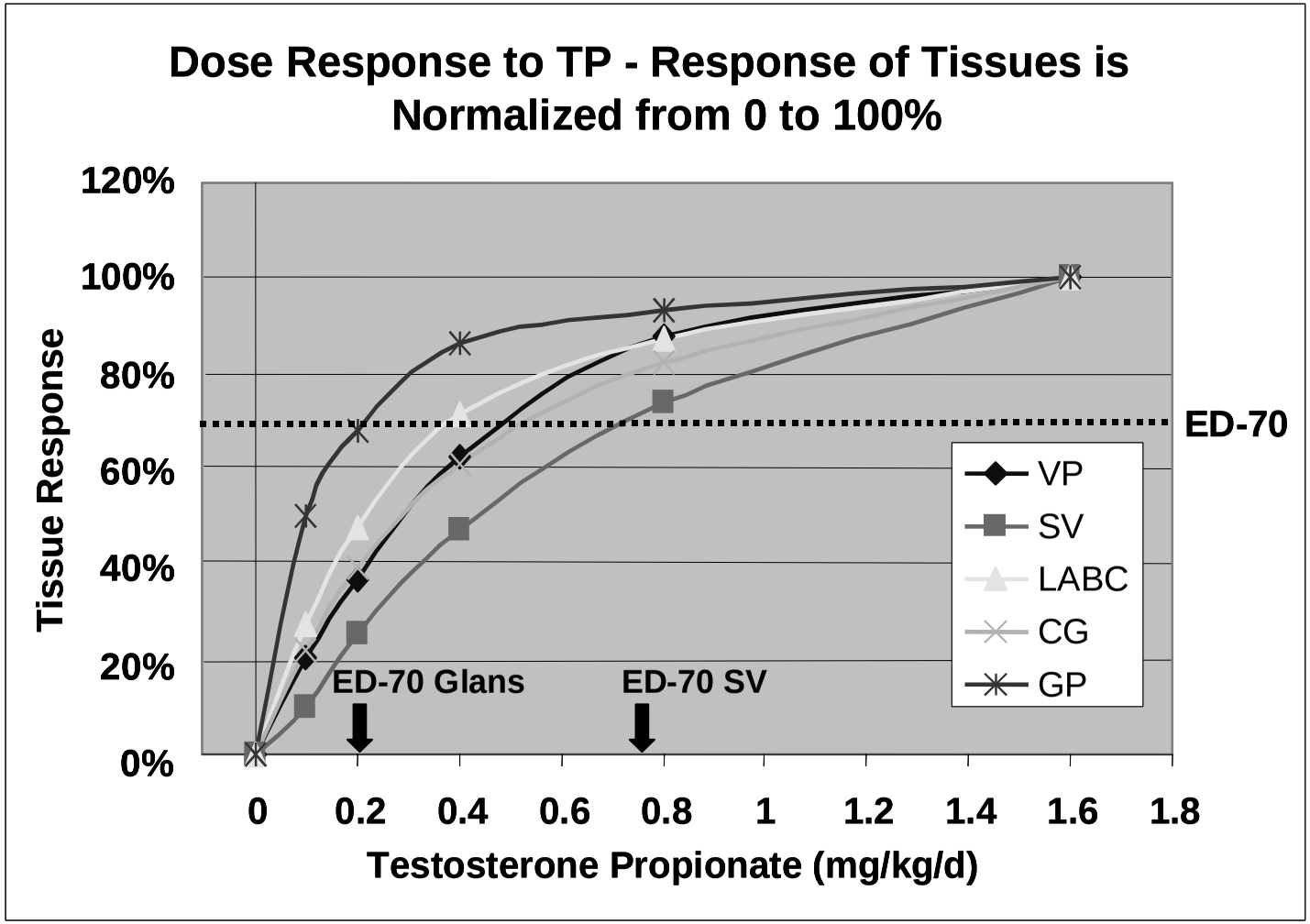


Figure 2. The pooled data across 16 laboratories for each tissue has been normalized so that the control value = 0% and the maximum reached at 1.6 mg/kg/d testosterone propionate (TP) = 100%. The differences between tissues in how rapidly a tissue reaches its maximum is then illustrated, e.g., glans penis first and seminal vesicles last. VP = ventral prostate; SV = seminal vesicles and coagulating glands; LABC = levator ani/bulbocavernosus muscle complex; CG = Cowper's glands; GP = glans penis.

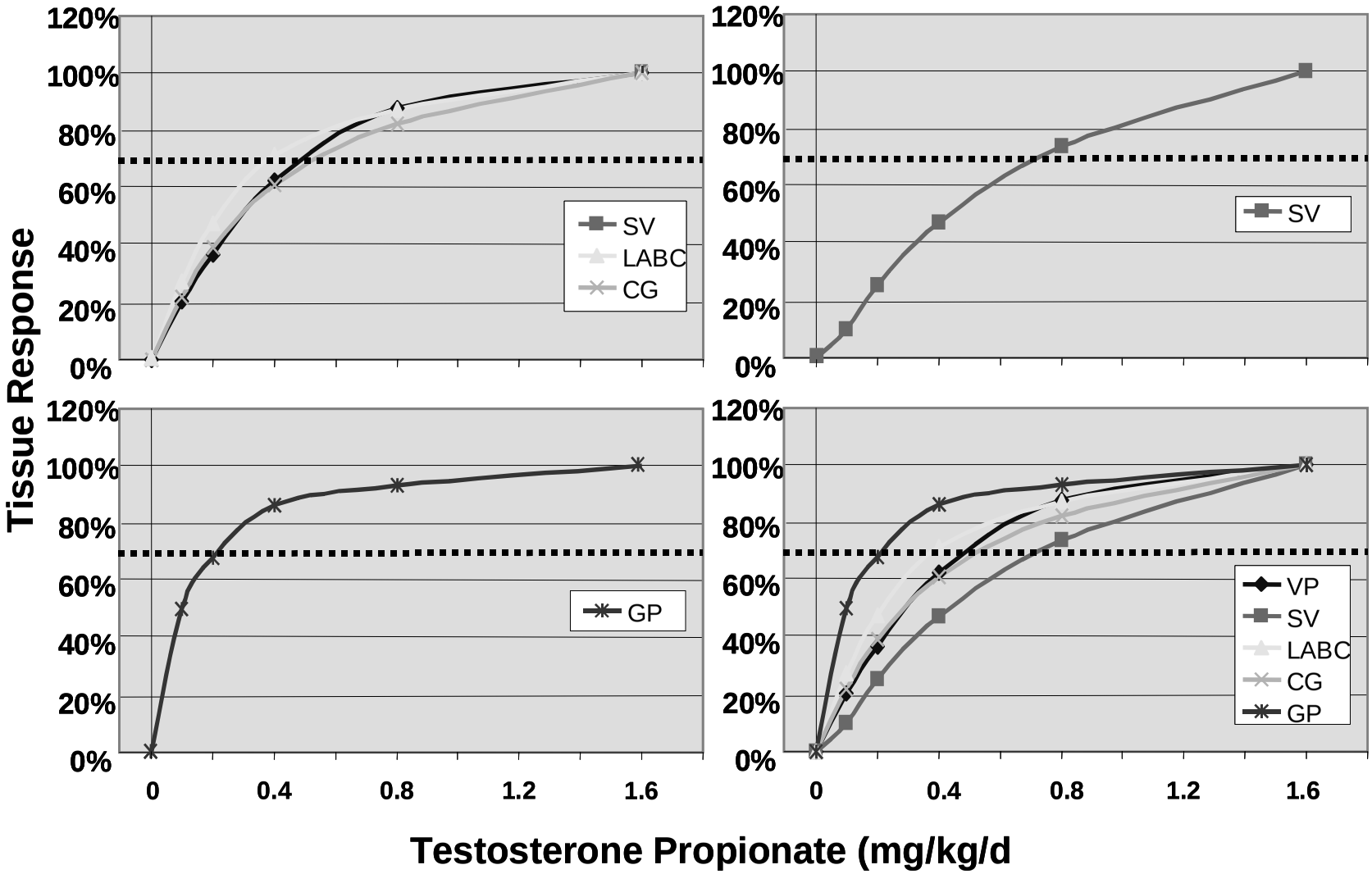
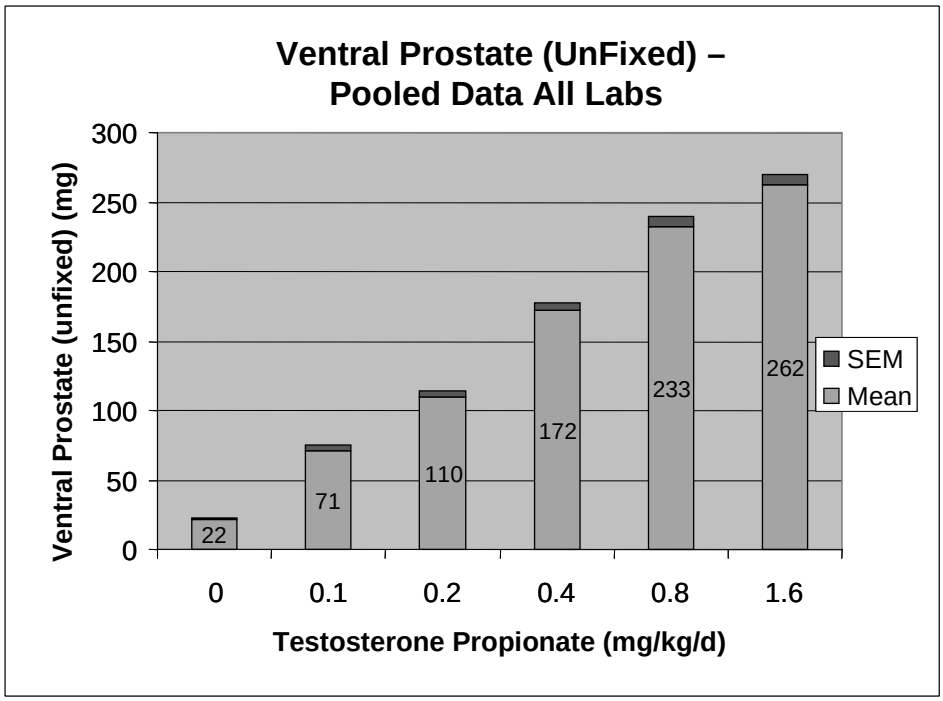


Figure 3. The three potential groups of male reproductive tissues are broken out using the normalized plot technique from Figure 2. Tissue abbreviations are also the same.



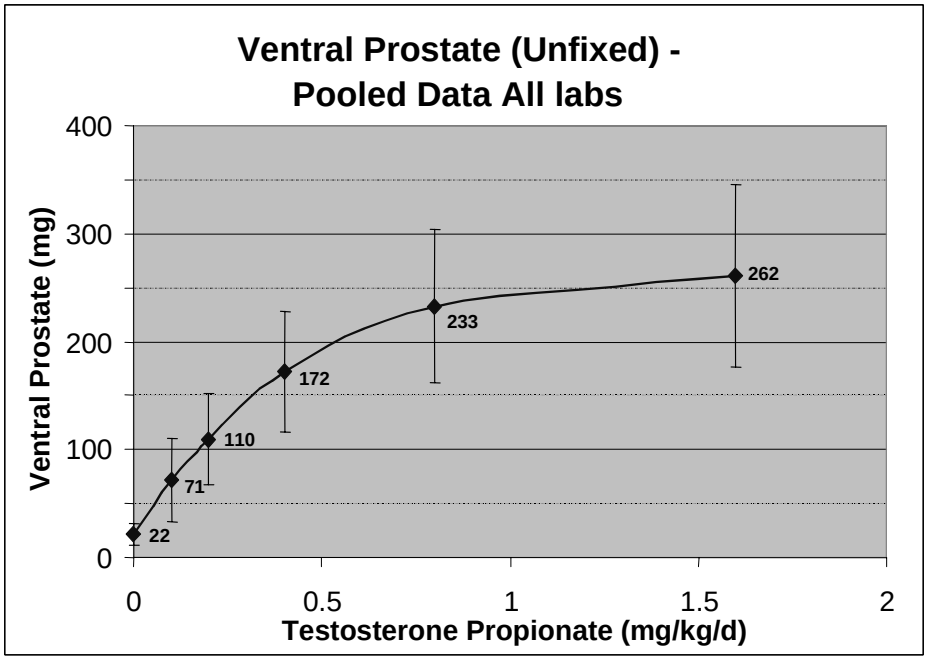


Figure 4. The response of the fresh, unfixed ventral prostate using the mean of pooled data from 16 laboratories in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (96 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the weight of the fresh, unfixed prostate.

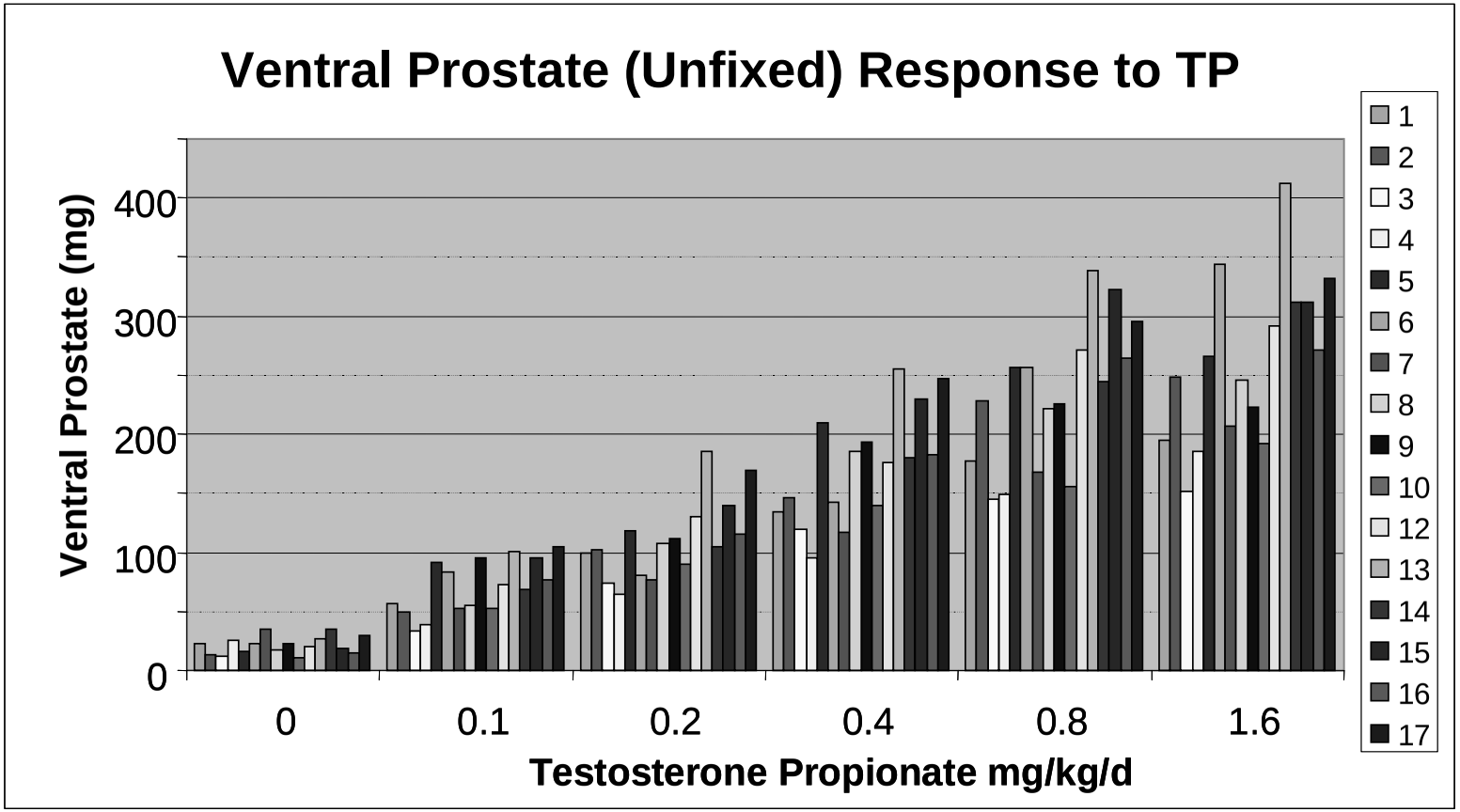
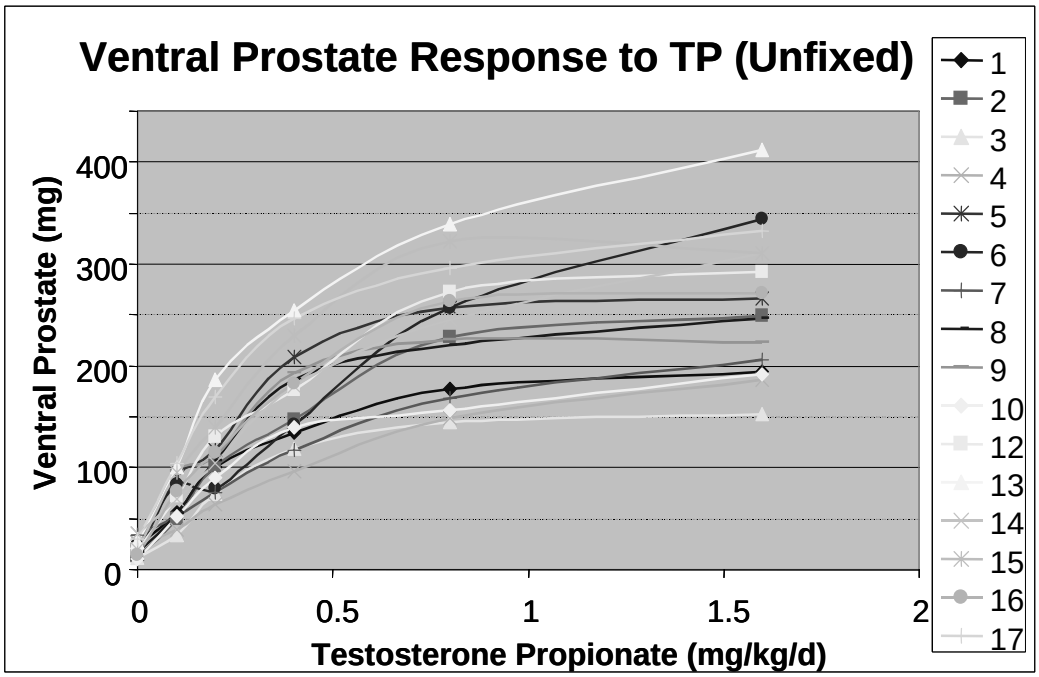


Figure 5. The individual responses of the mean of the fresh, unfixed ventral prostate weight from each of 16 laboratories (1 laboratory, #11, administered micrograms rather than milligrams of TP) plotted as a bar graph.



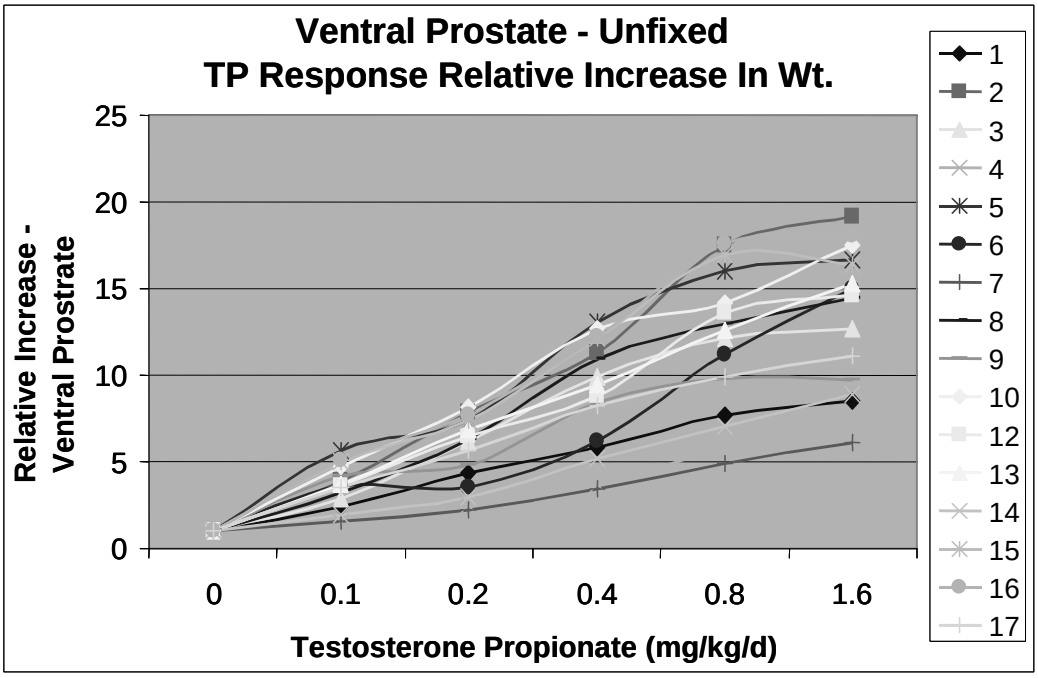
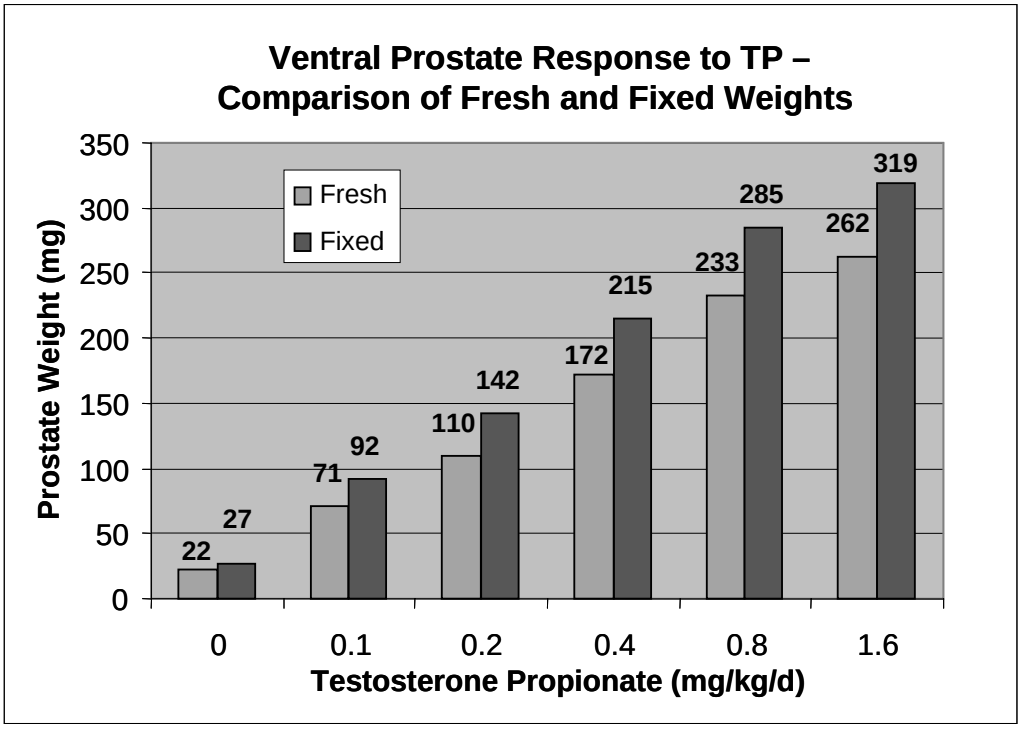


Figure 6. The individual responses of the mean of the fresh, unfixed ventral prostate weight from each of 16 laboratories. The top graph is plotted as the absolute weight for each laboratory. The bottom graph is plotted as the relative increase in weight versus the control with ventral prostate increasing between 6- and 19-fold, depending upon the laboratory.



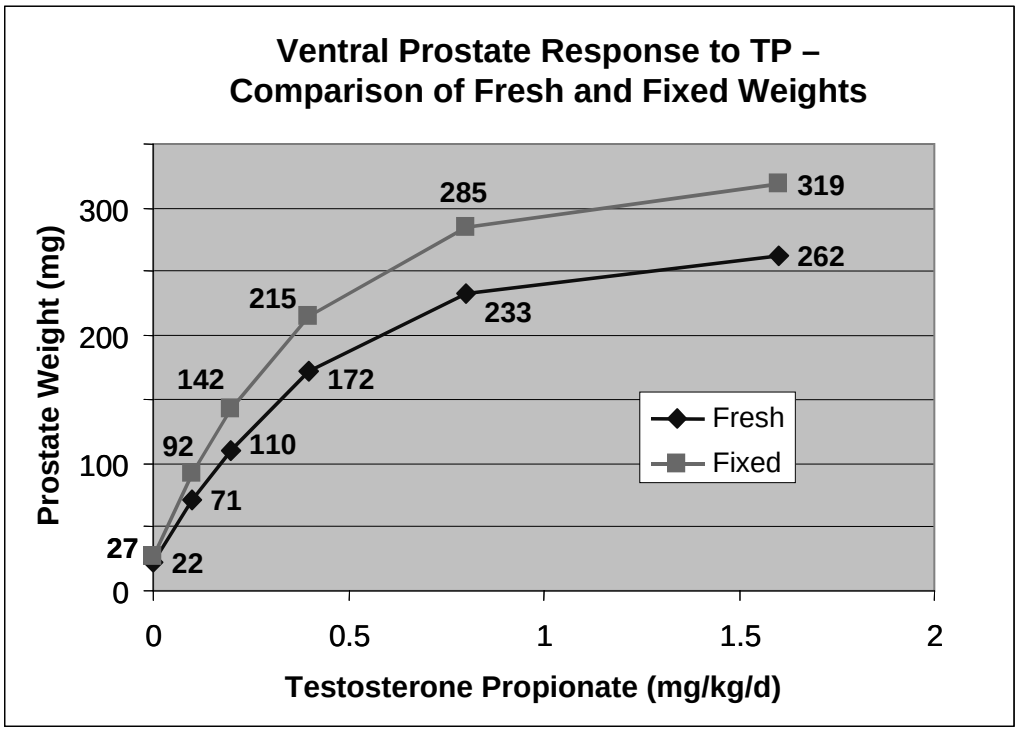
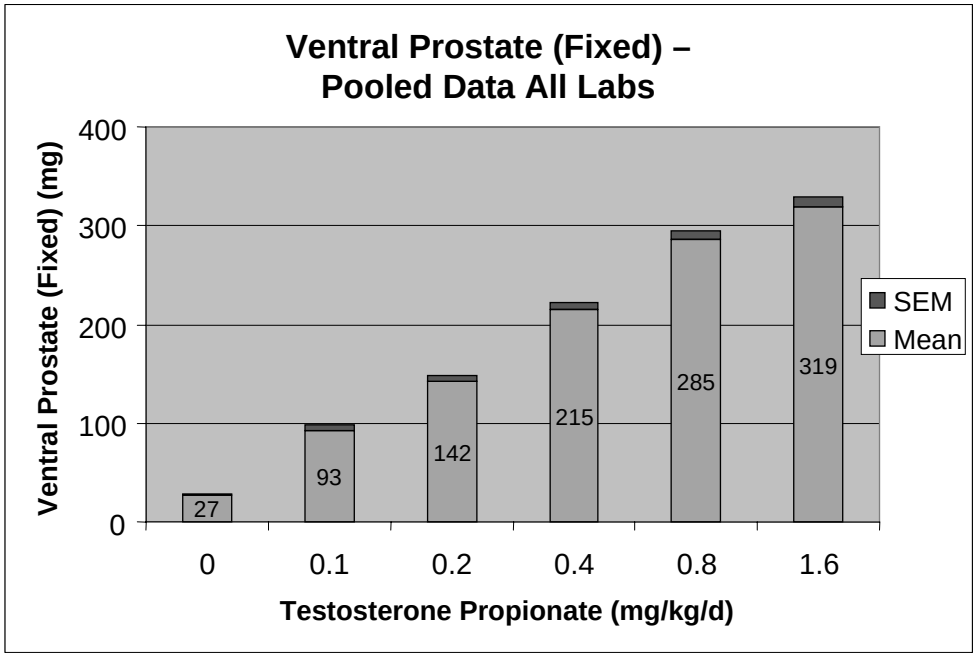


Figure 7. Comparison of pooled means of fresh prostates (16 labs) and fixed prostates weighed after 24 hours of fixation (15 labs). Upper graph is a bar graph comparison, and lower graph is a linear line graph comparison.



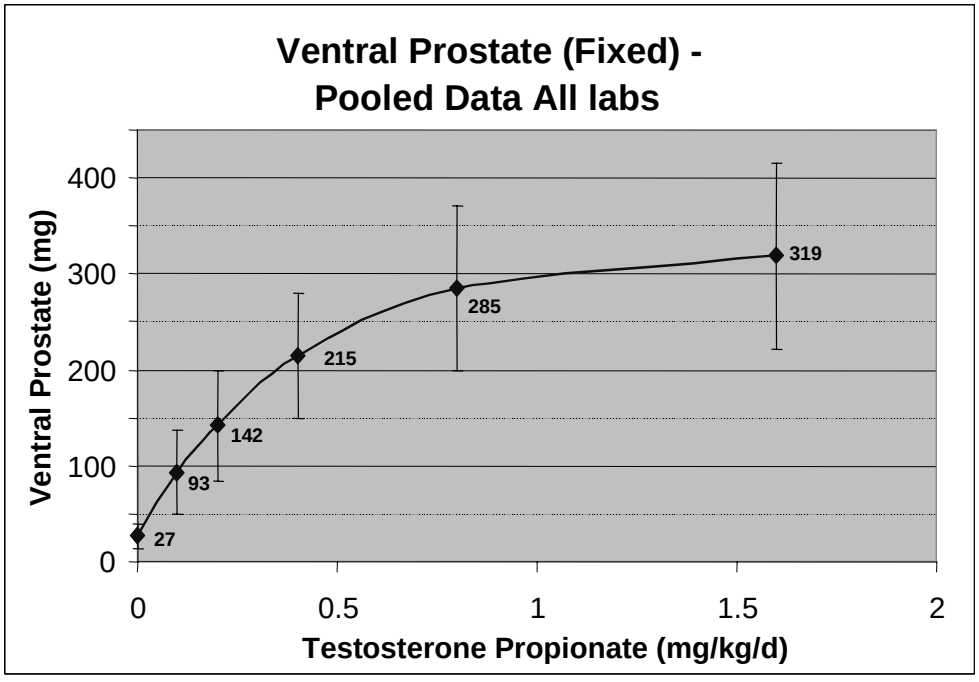
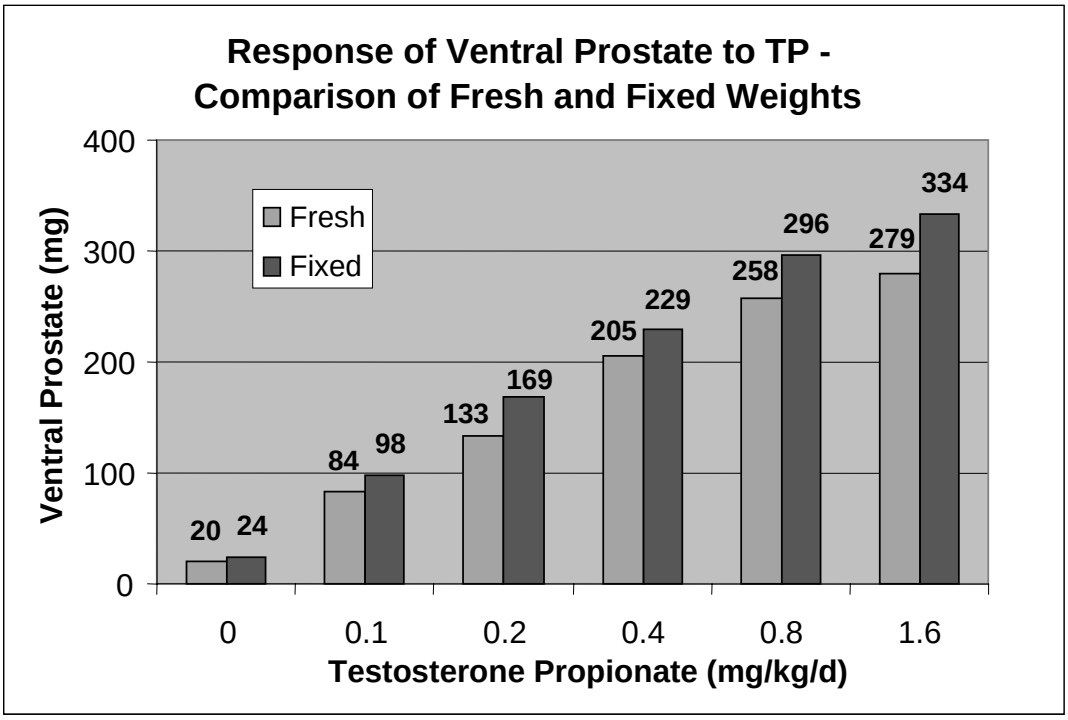


Figure 8. The response of the fixed ventral prostate (fixed after dissection and weighed 24 hours later) using the mean of pooled data from 15 laboratories (one laboratory did not fix and weight the prostate) in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (90 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the weight of fixed prostate.



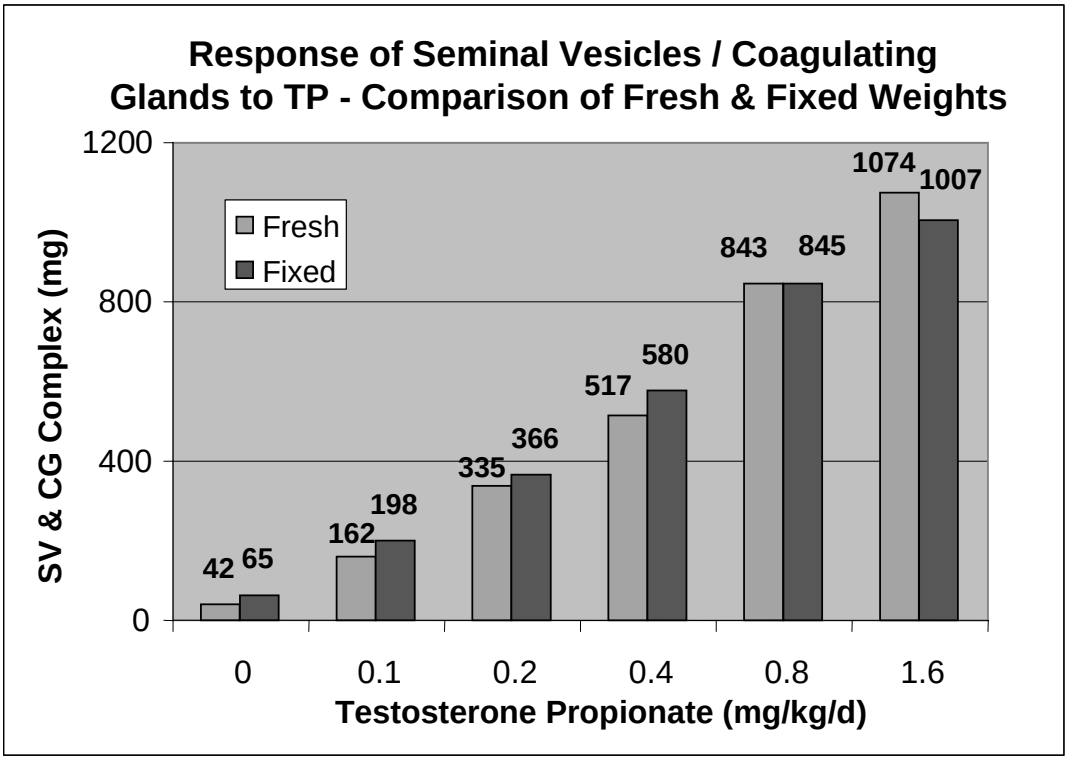


Figure 9 is continued on the next page.

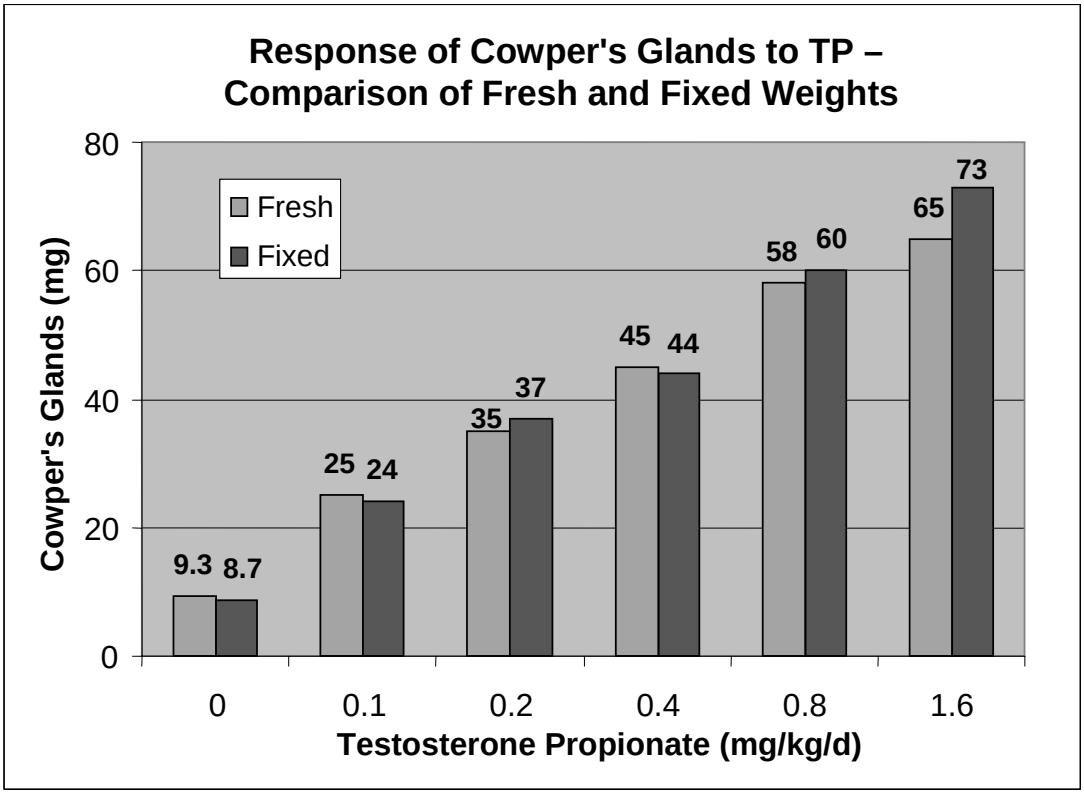
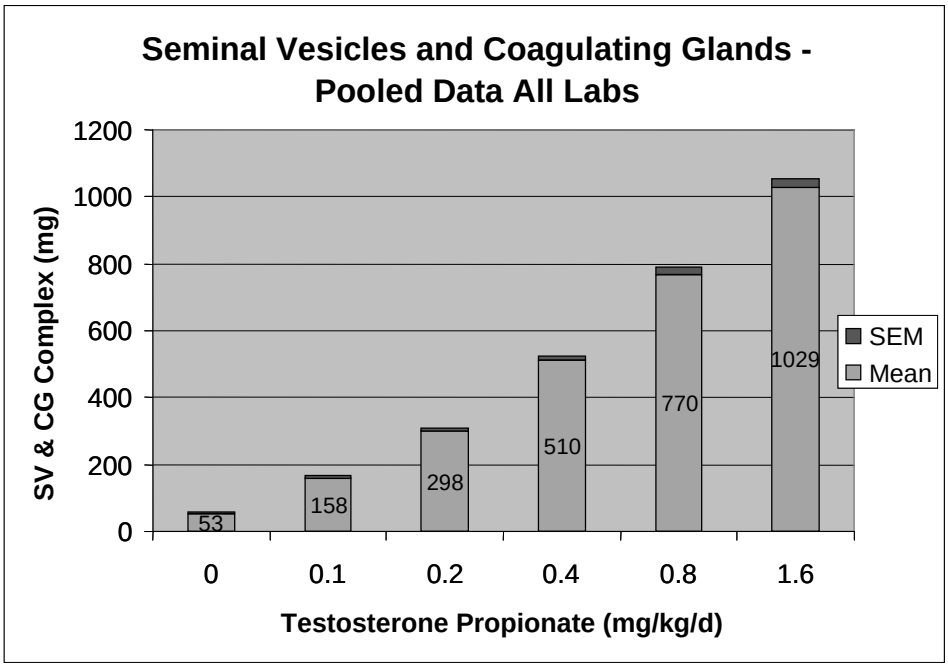


Figure 9. Graphs of data comparing fresh and unfixed weights of three tissues from three laboratories performing comparisons of two additional tissues (seminal vesicles/coagulating glands and Cowper's glands). Tissues were fixed for 24 hours after dissection and then weighed. Top graph – ventral prostate (data separated from other laboratories). Middle graph – seminal vesicles and coagulating glands. Bottom graph – Cowper's glands.



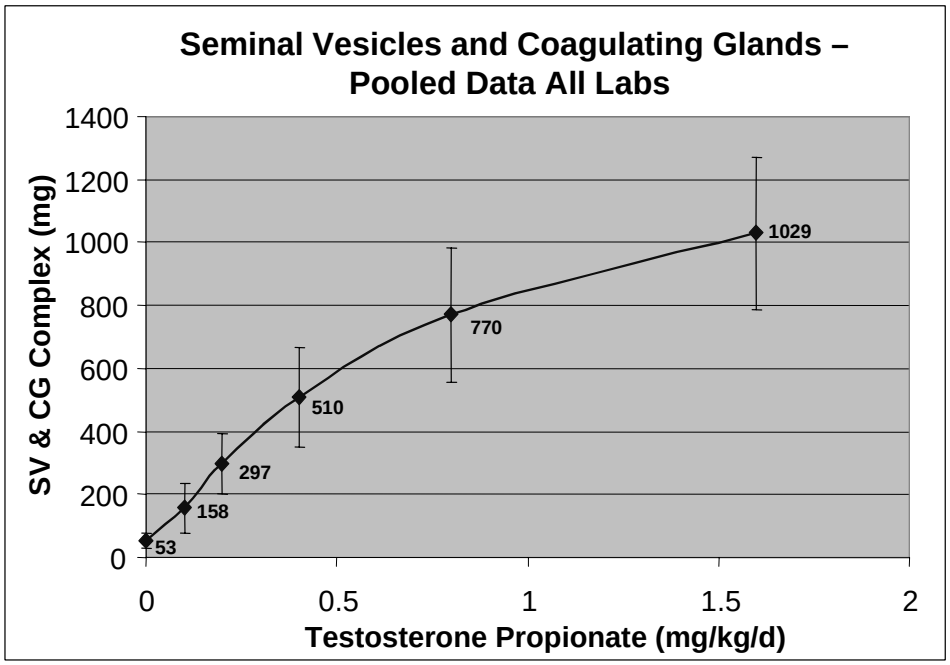


Figure 10. The response of the seminal vesicles and coagulating glands using the mean of pooled data from 16 laboratories in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (96 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the total weight of the pair of seminal vesicles and coagulating glands.

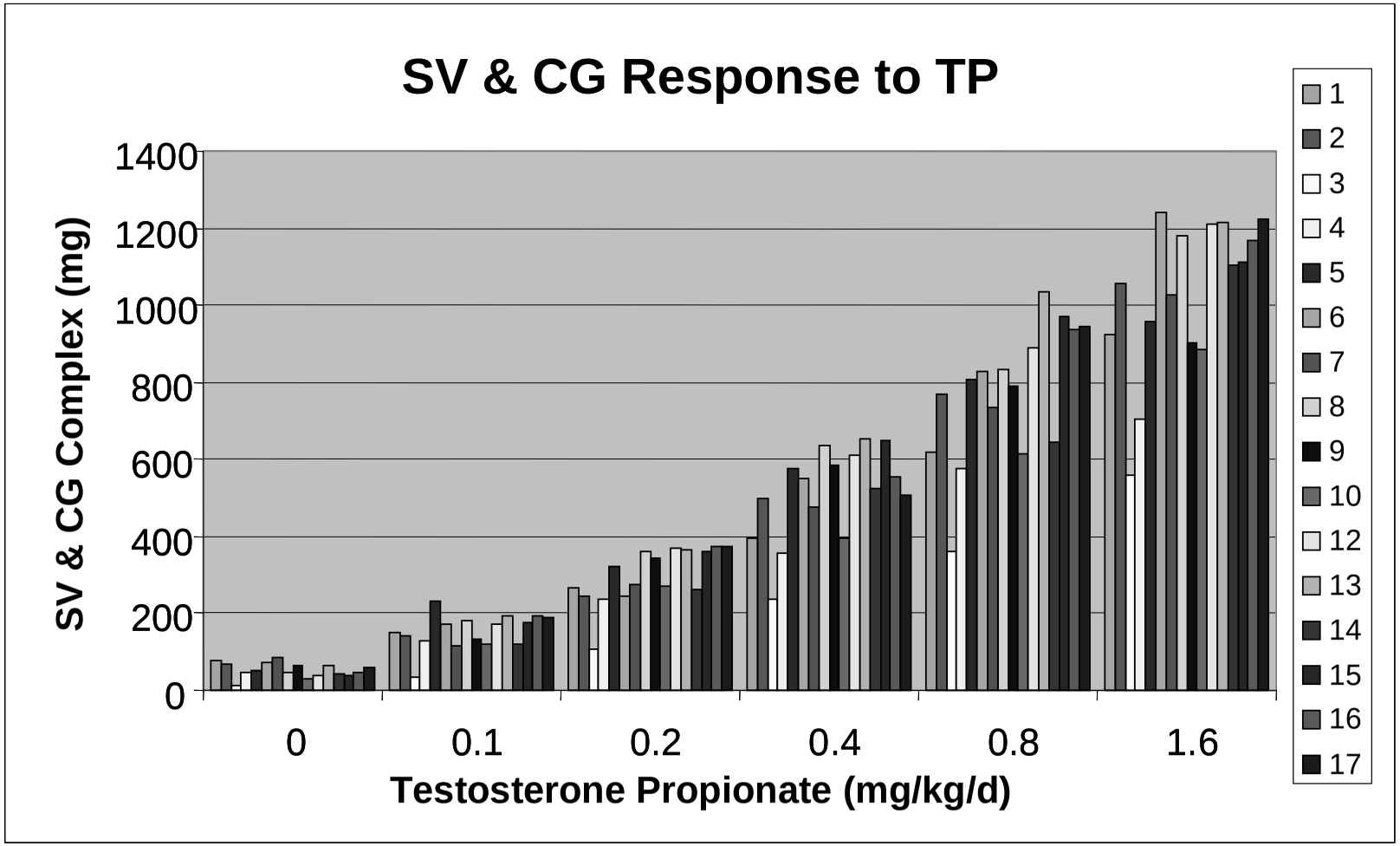


Figure 11. The individual responses of the mean of the seminal vesicle and coagulating gland weights from each of 16 laboratories (1 laboratory, #11, administered micrograms rather than milligrams of TP) plotted as a bar graph.

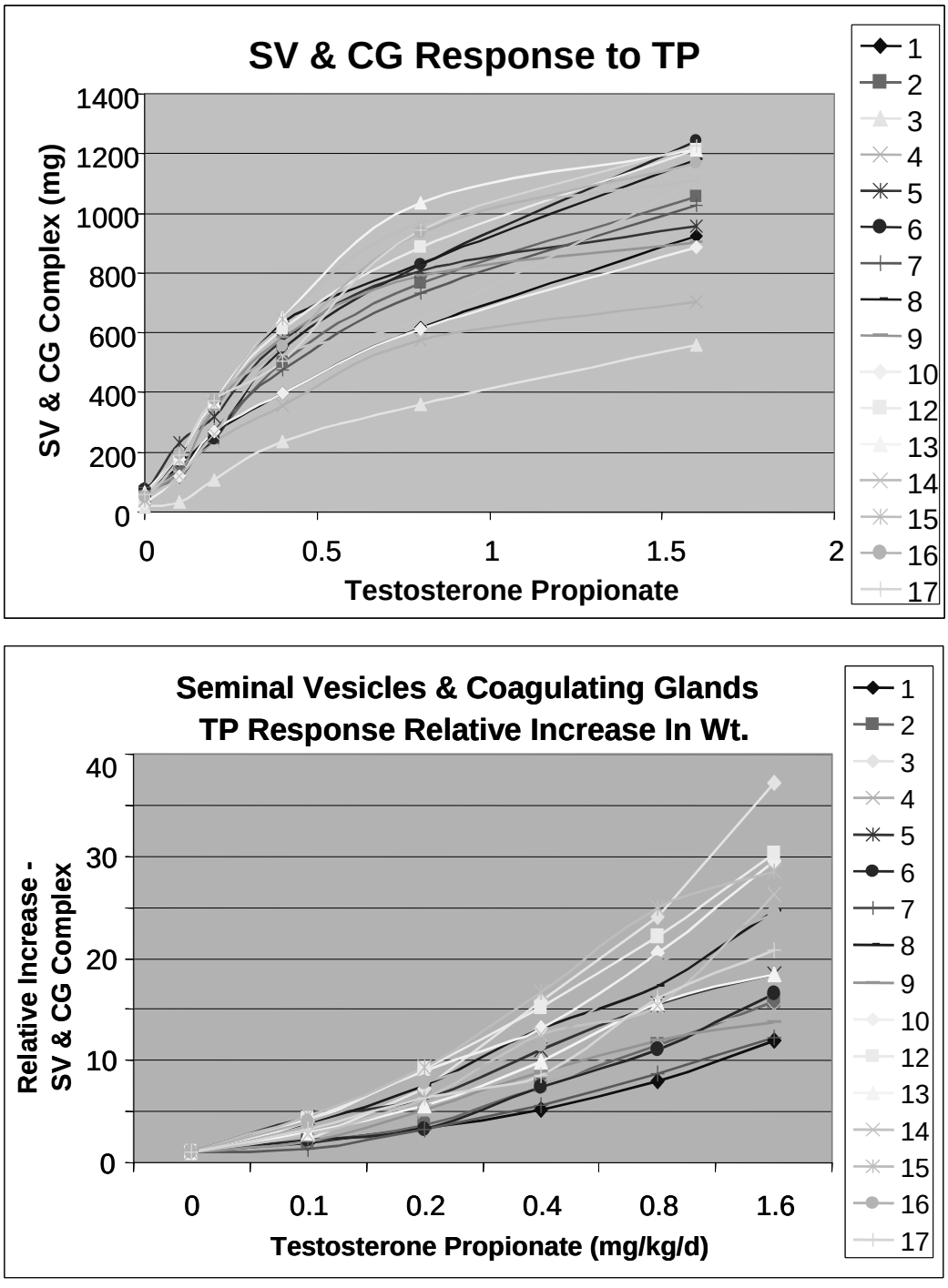
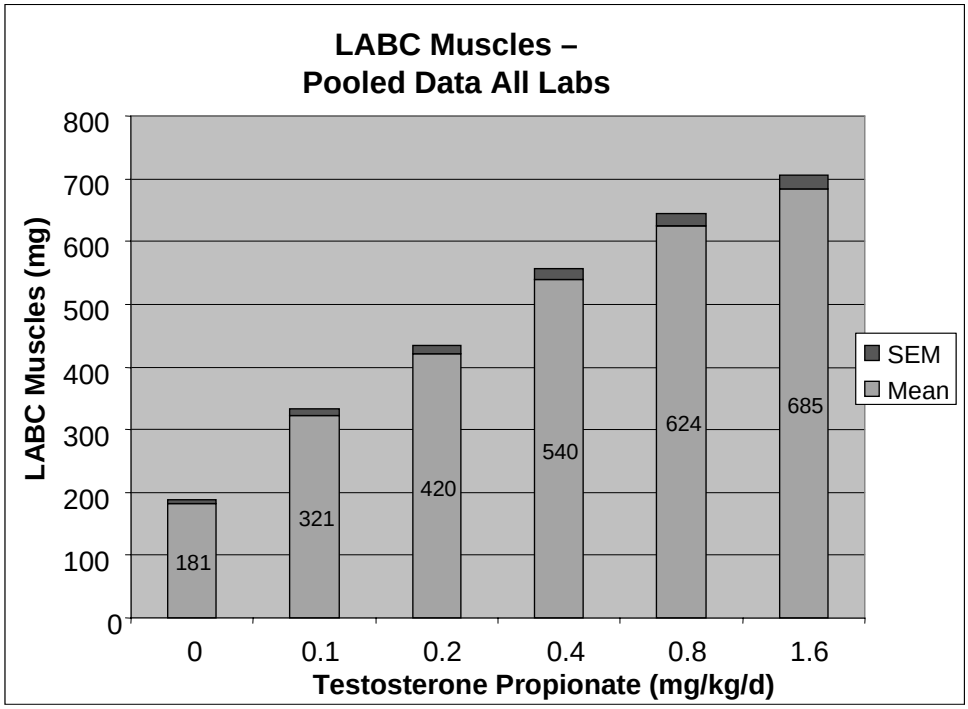


Figure 12. The individual responses of the mean of the seminal vesicles and coagulating gland weights from each of 16 laboratories. The top graph is plotted as the absolute weight for each laboratory. The bottom graph is plotted as the relative increase in weight versus the control with seminal vesicle-coagulating gland complex dramatically increasing between 12- and 40-fold, depending upon the laboratory.



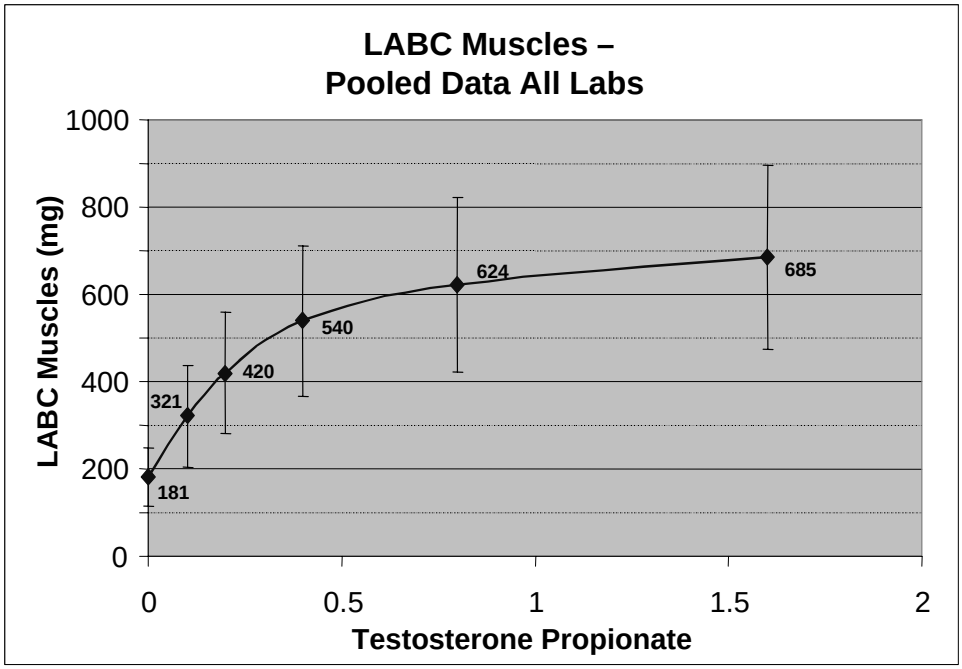


Figure 13. The response of the levator ani/bulbocavernosous muscles (LABC) using the mean of pooled data from 16 laboratories in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (96 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the total weight of the dissected LABC.

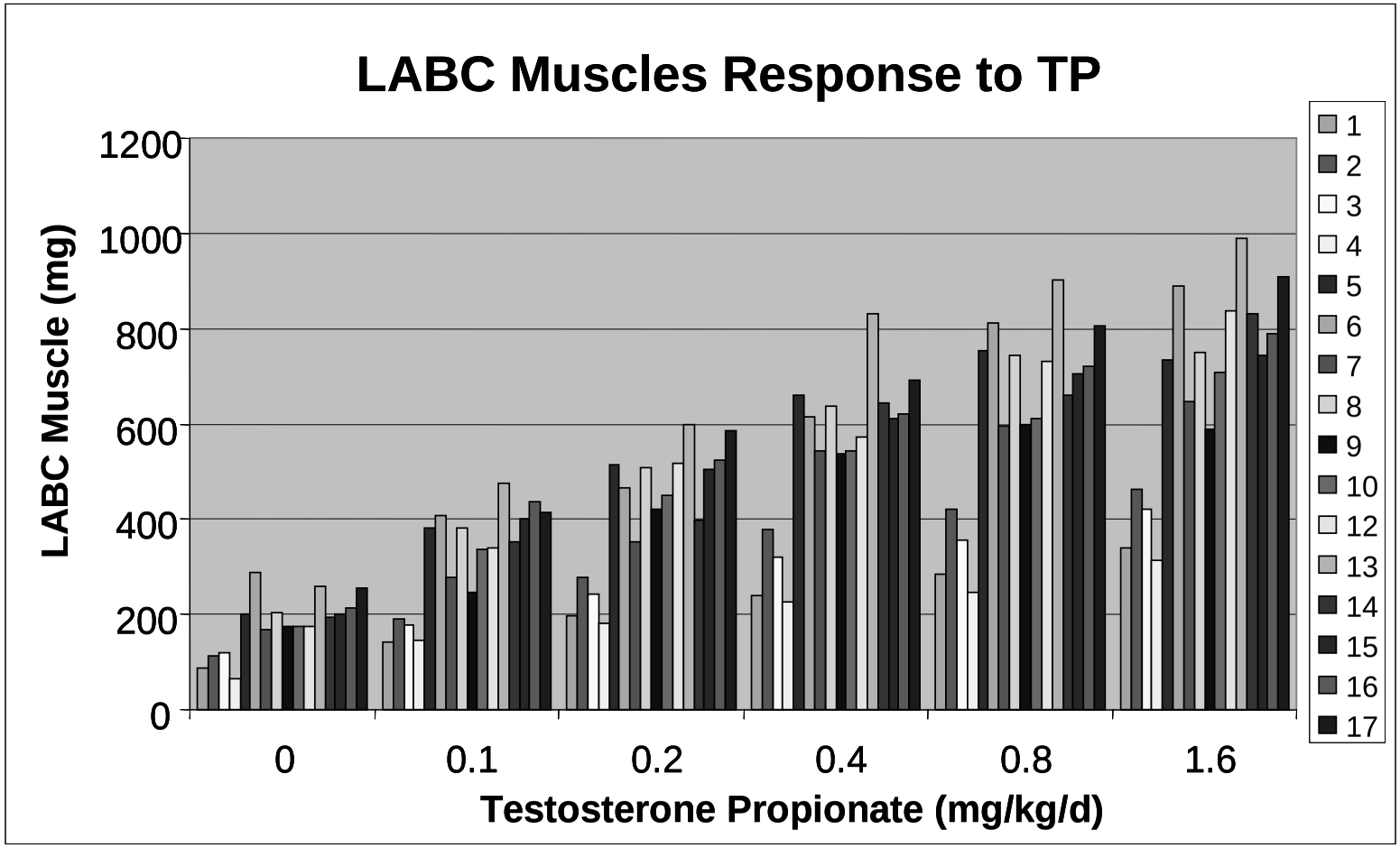
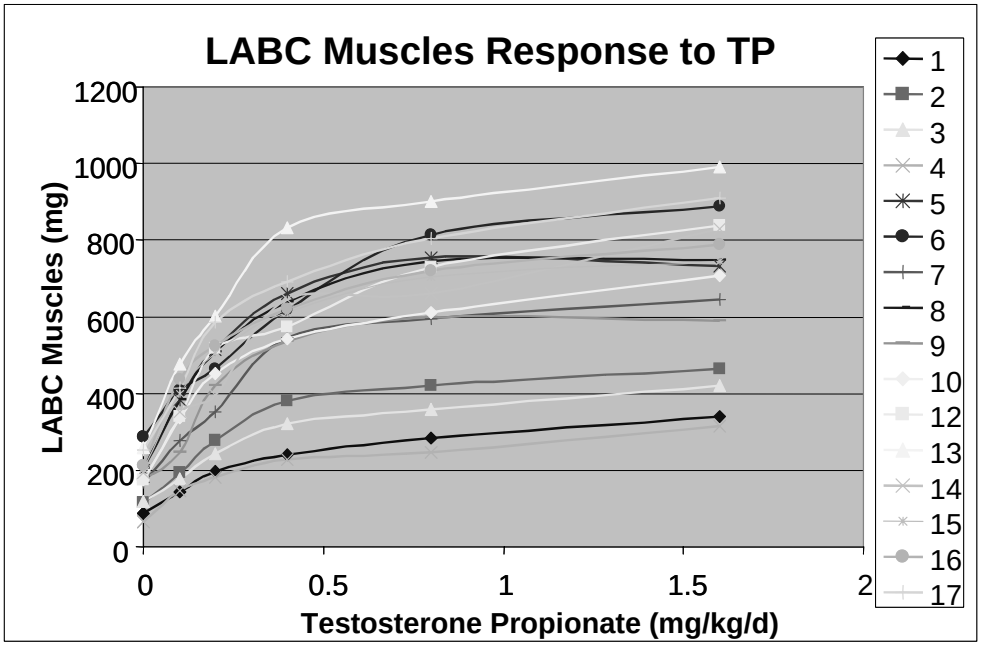


Figure 14. The individual responses of the mean of the LABC weights from each of 16 laboratories (1 laboratory, #11, administered micrograms rather than milligrams of TP) plotted as a bar graph.



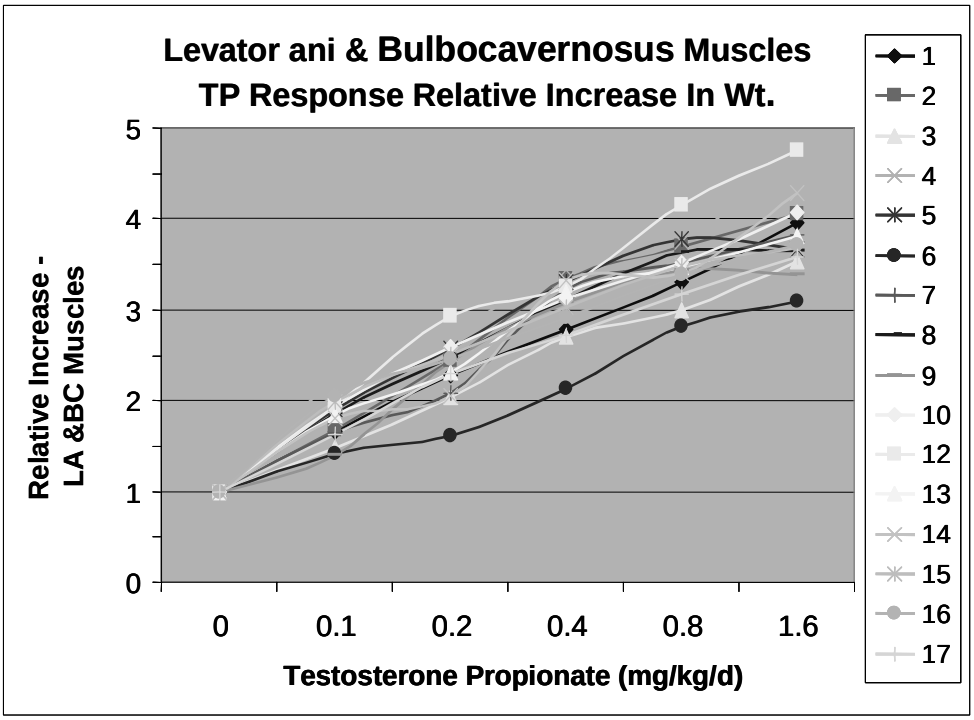
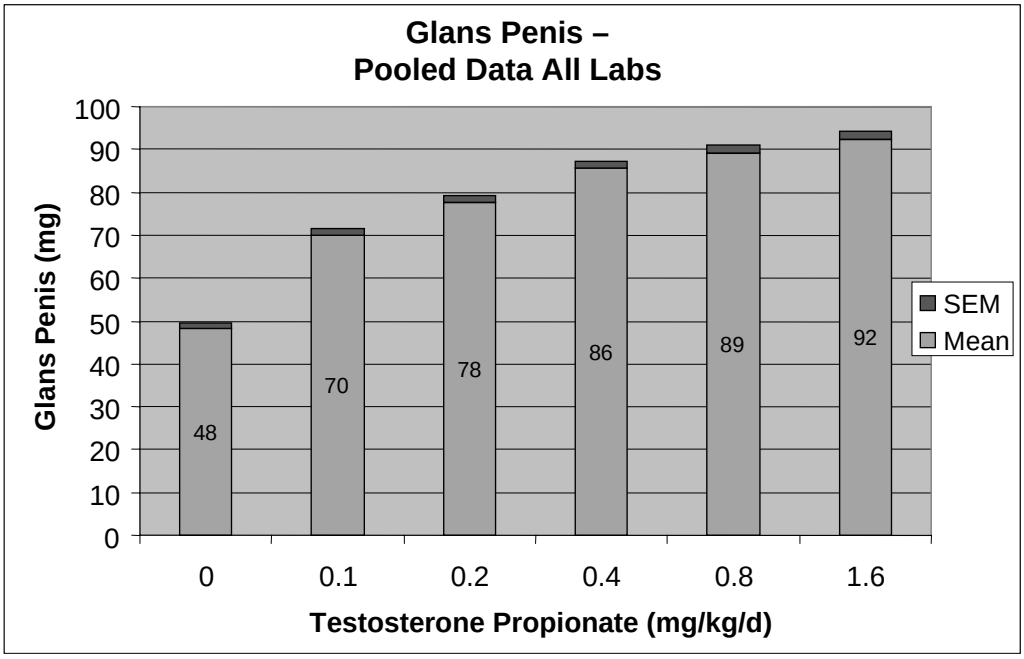


Figure 15. The individual responses of the mean of the LABC weights from each of 16 laboratories. The top graph is plotted as the absolute weight for each laboratory. The bottom graph is plotted as the relative increase in weight versus the control with the LABC increasing between 3- and 6-fold, depending upon the laboratory.



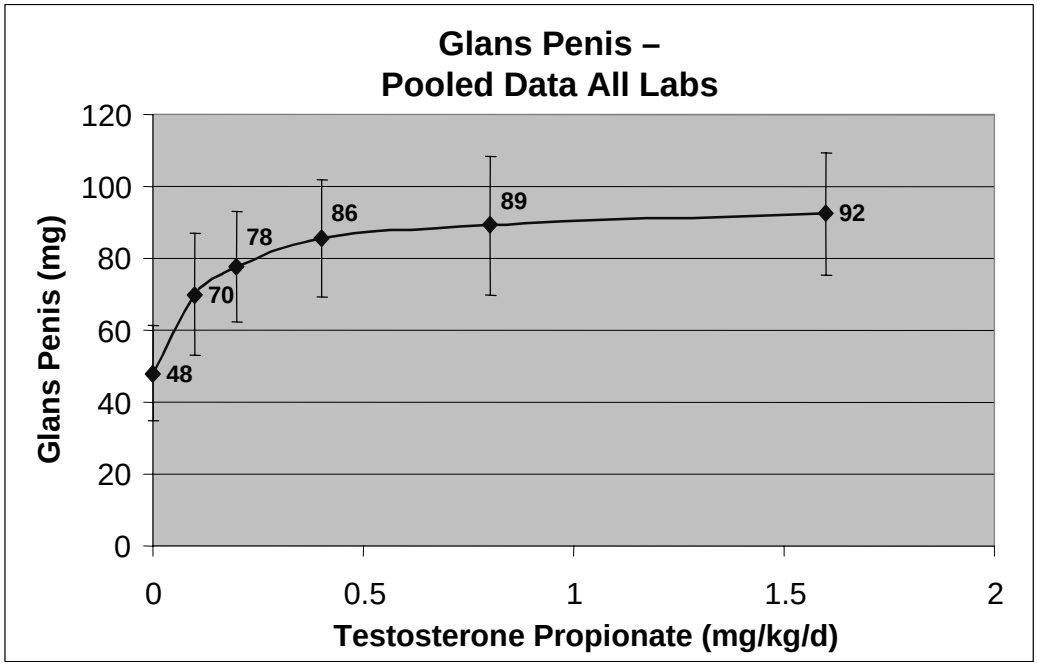


Figure 16. The response of glans penis using the mean of pooled data from 16 laboratories in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (96 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the total weight of the dissected glans penis.

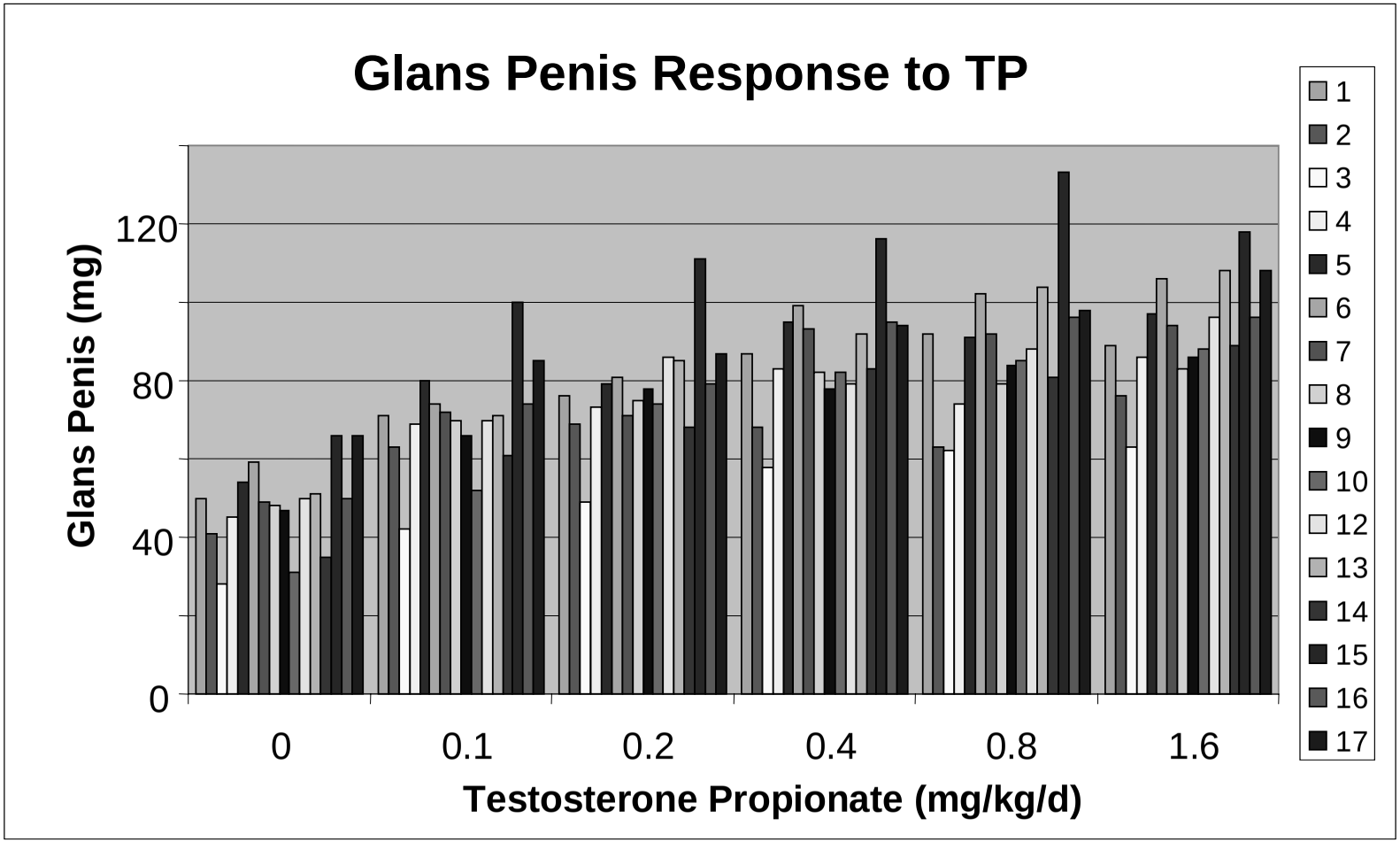


Figure 17. The individual responses of the mean of the glans penis weights from each of 16 laboratories (1 laboratory, #11, administered micrograms rather than milligrams of TP) plotted as a bar graph.

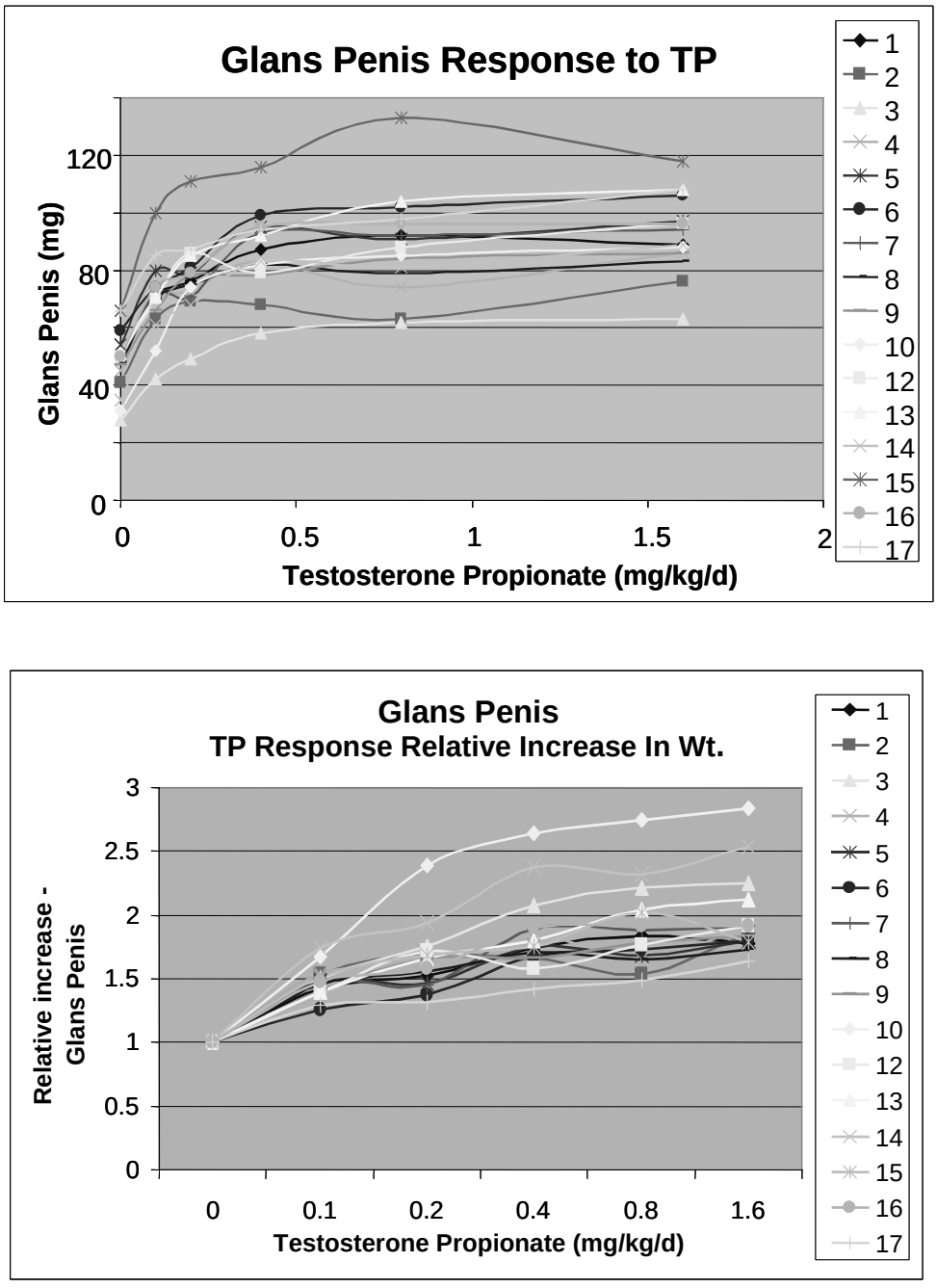


Figure 18. The individual responses of the mean of the glans penis weights from each of 16 laboratories. The top graph is plotted as the absolute weight for each laboratory. The bottom graph is plotted as the relative increase in weight versus the control with the glans penis increasing between 1.5- and 3-fold, depending upon the laboratory.

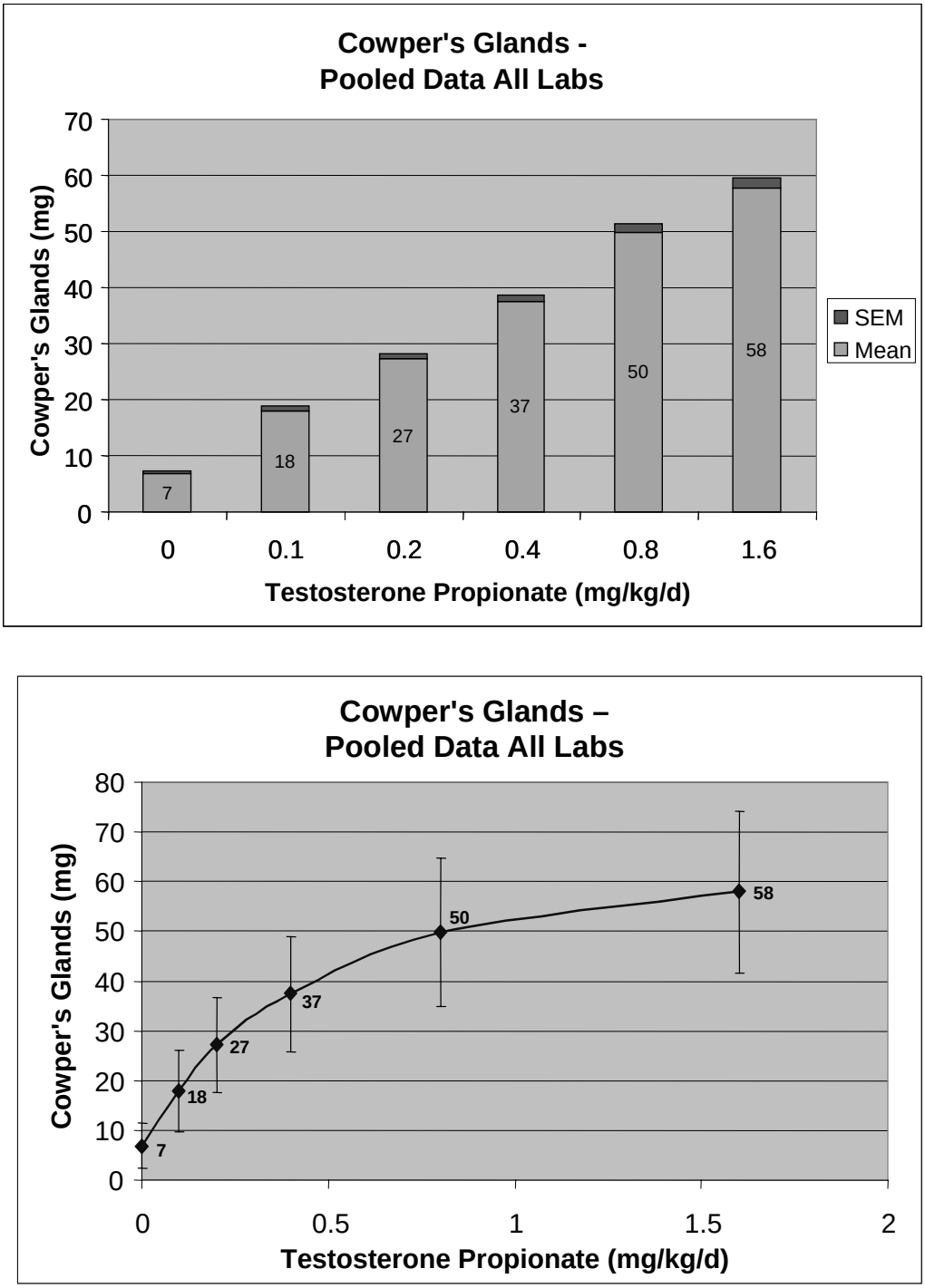


Figure 19. The response of Cowper's glands using the mean of pooled data from 16 laboratories in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (96 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the total weight of the dissected Cowper's Glands.

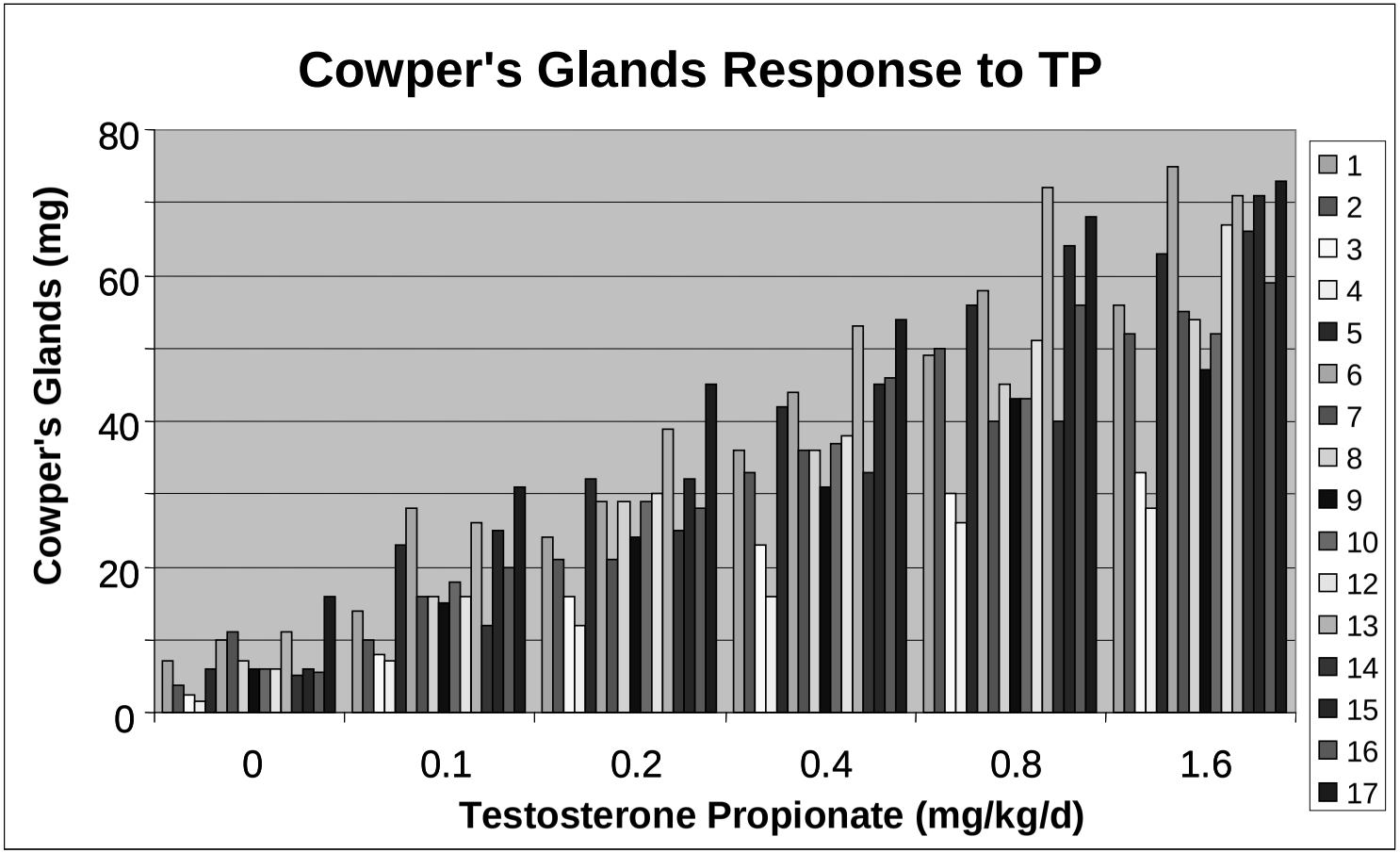


Figure 20. The individual responses of the mean of the Cowper's glands weights from each of 16 laboratories (1 laboratory, #11, administered micrograms rather than milligrams of TP) plotted as a bar graph.

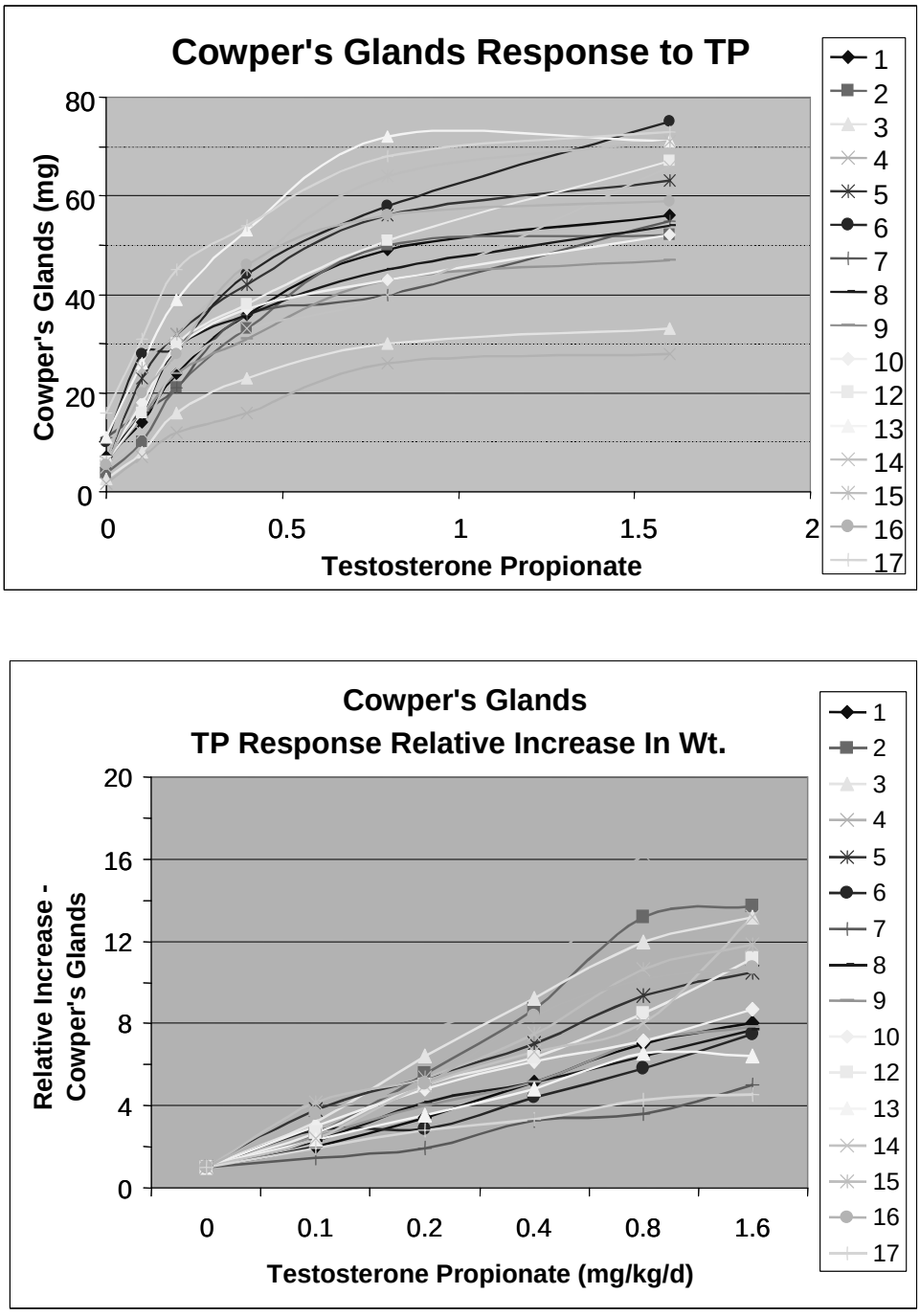


Figure 21. The individual responses of the mean of the Cowper's gland weights from each of 16 laboratories. The top graph is plotted as the absolute weight for each laboratory. The bottom graph is plotted as the relative increase in weight versus the control with the Cowper's glands increasing between 4- and 16-fold, depending upon the laboratory

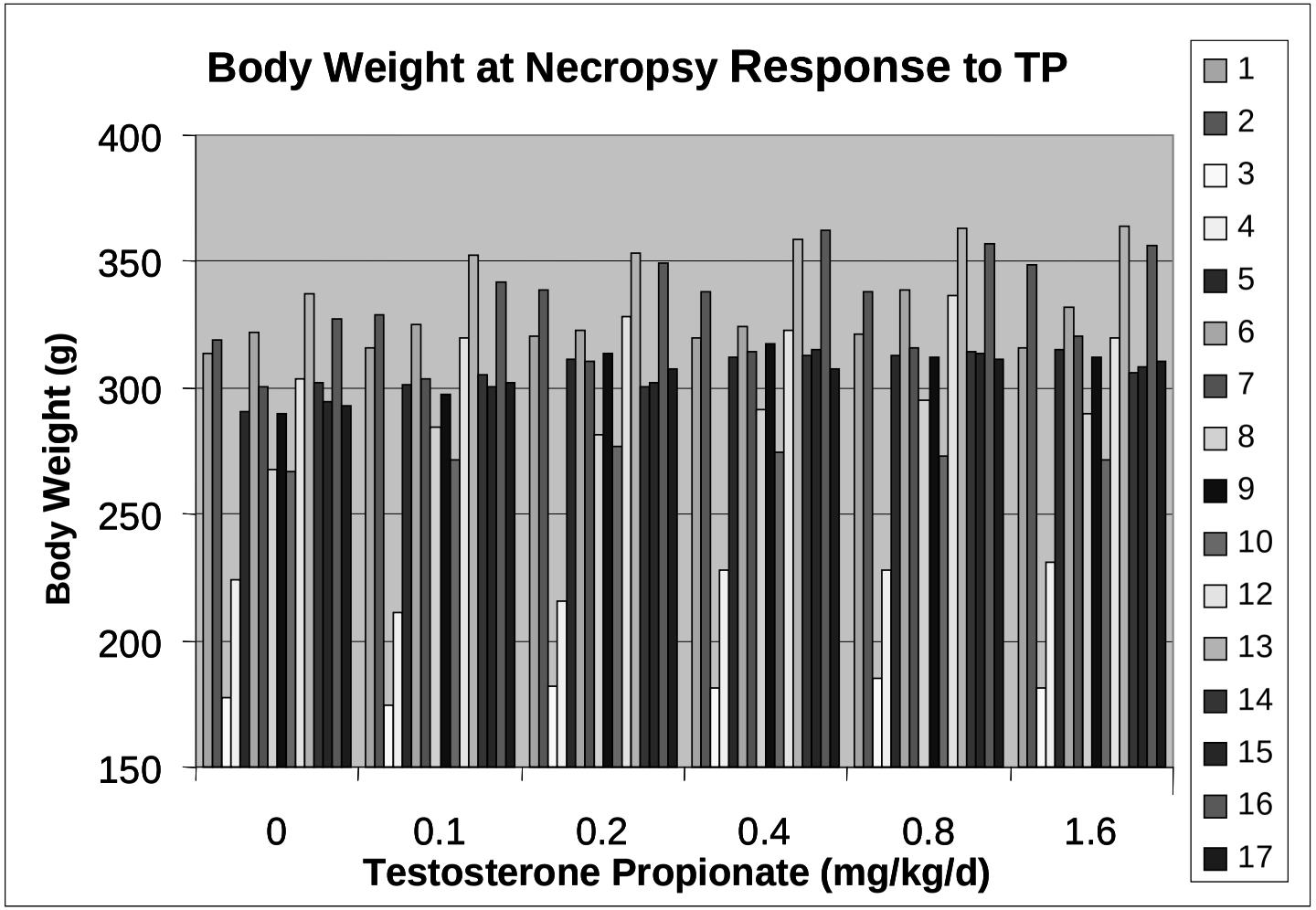


Figure 22. The individual responses of the mean of the body weights from each of 16 laboratories (1 laboratory, #11, administered micrograms rather than milligrams of TP) plotted as a bar graph.

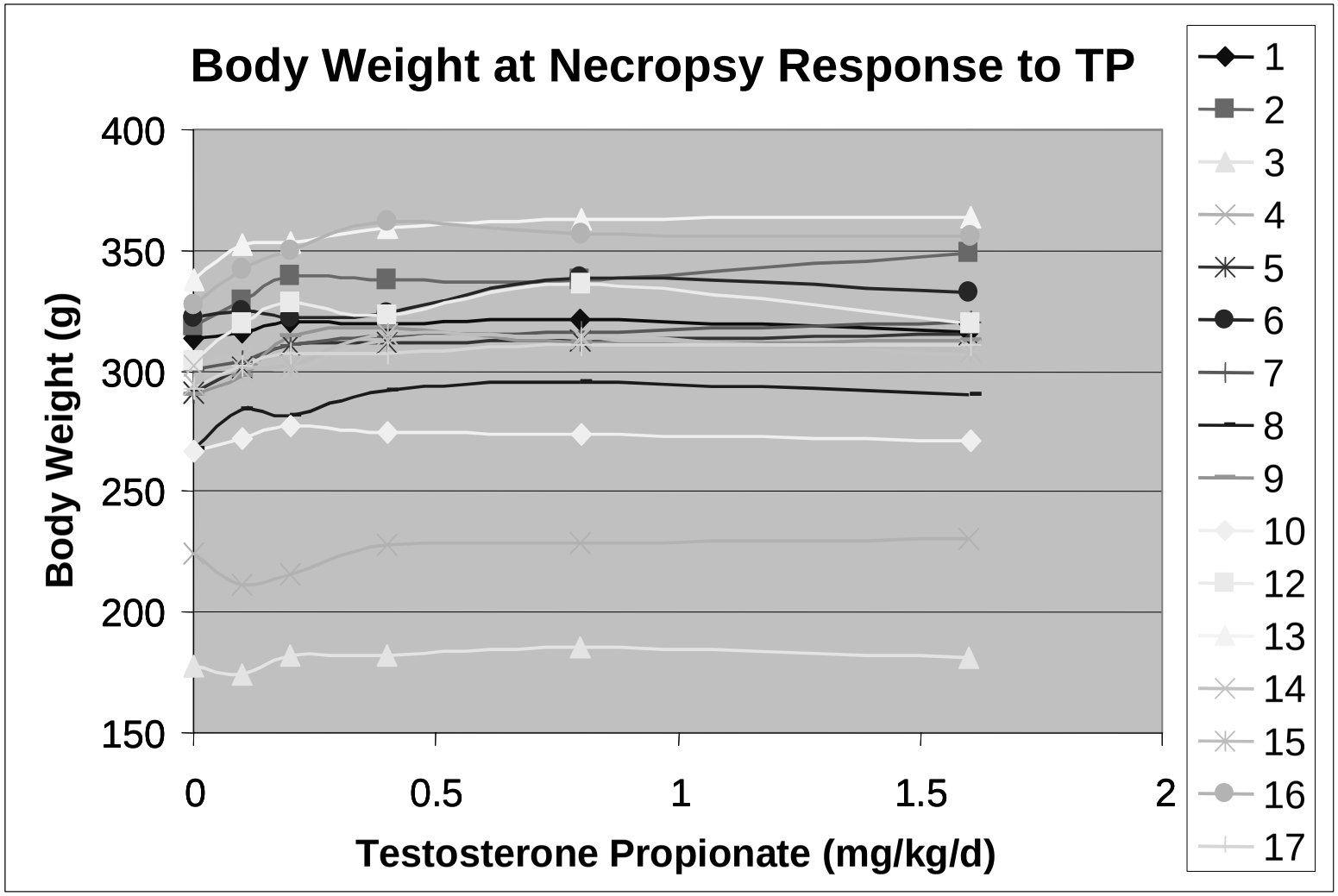
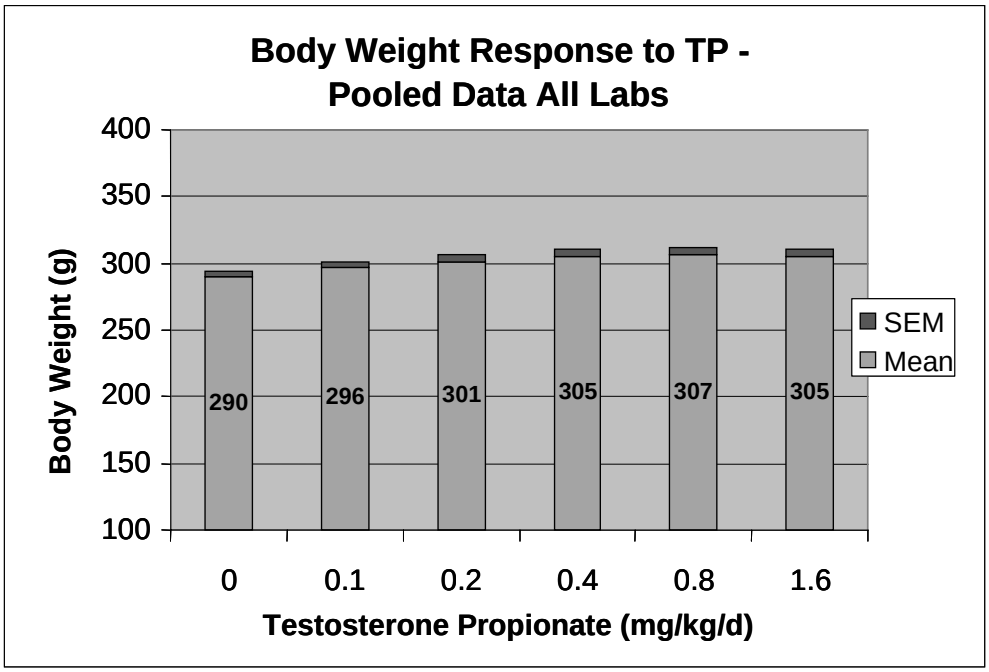


Figure 23. The individual responses of the mean of the necropsy body weights from each of 16 laboratories in response to increasing TP injections.



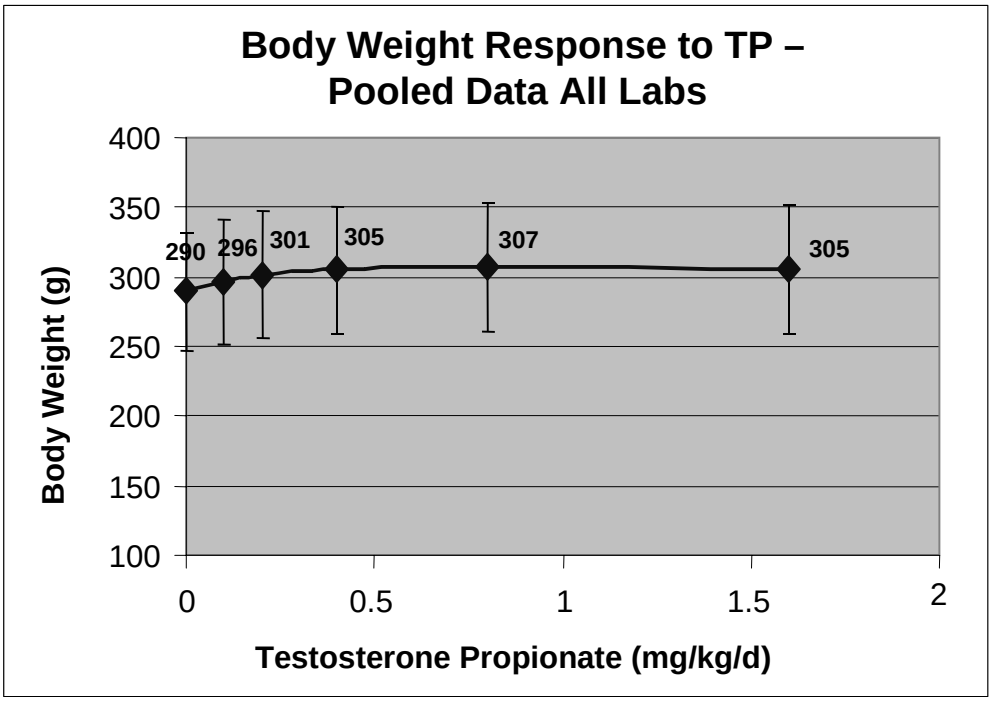


Figure 24. The response of body weight using the mean of pooled data from 16 laboratories in to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size (96 samples per dose in this case). The bottom graph uses the standard deviation from the pooled data. This incorporates significant laboratory variability and does not solely reflect the assay variability (see Figure 23). Both graphs are plotted as the total body weight at necropsy.

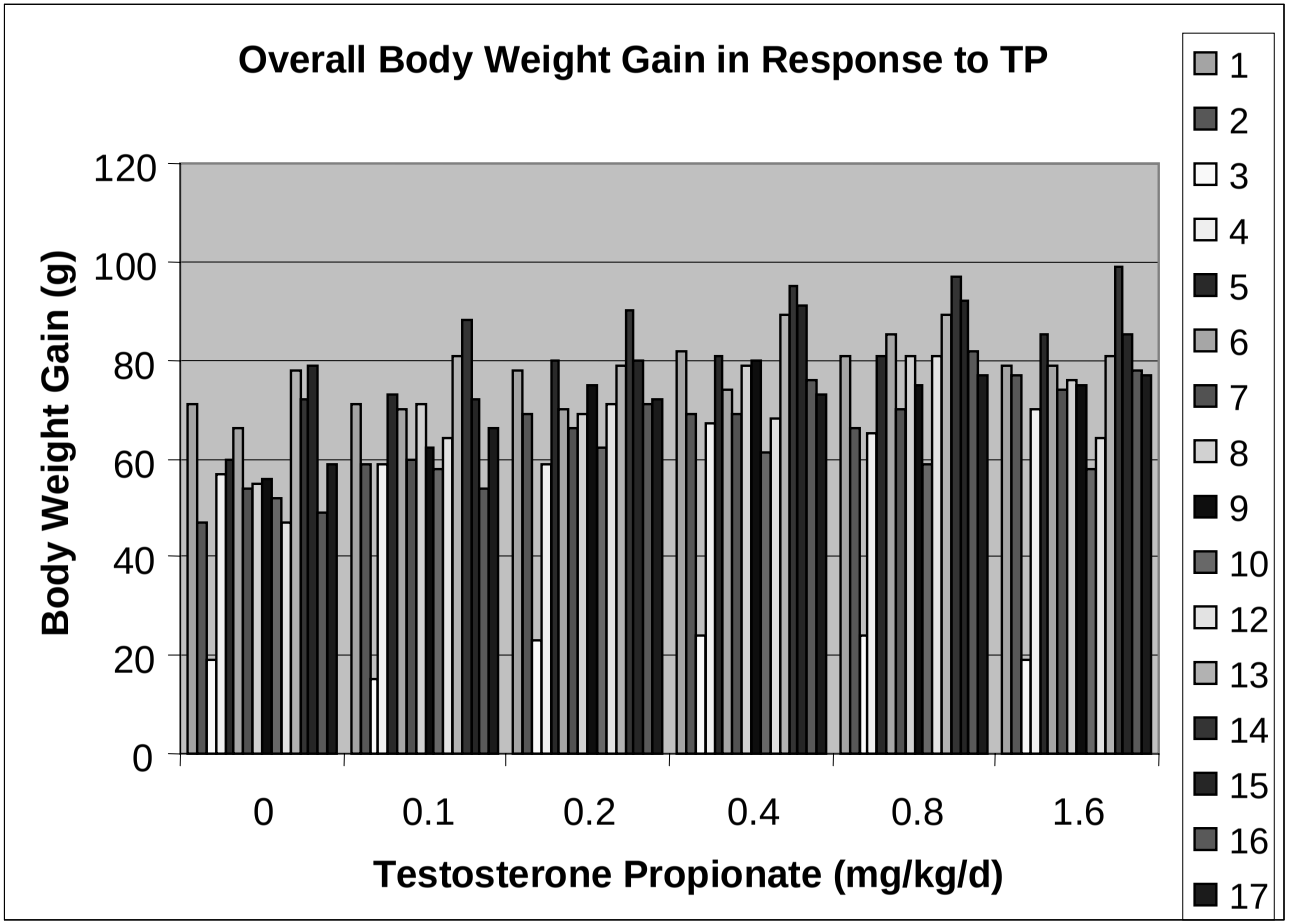
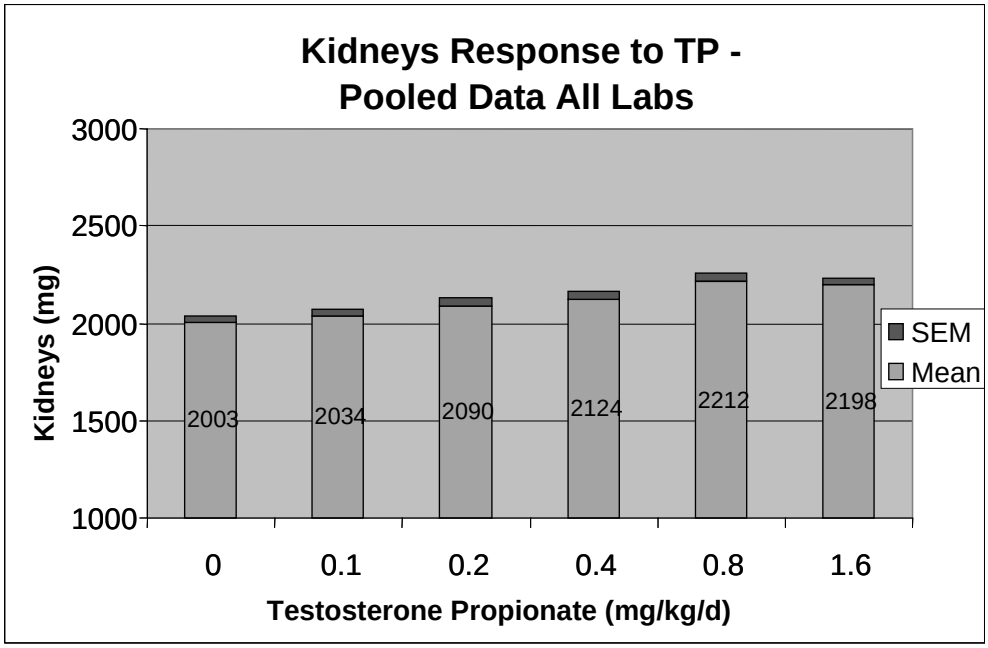


Figure 25. The individual responses of the mean gain in body weights from each of 16 laboratories for each dose group of TP. Note that laboratory 3 had the lowest starting body weights (see Figure 23) and here has the lowest body weight gains.



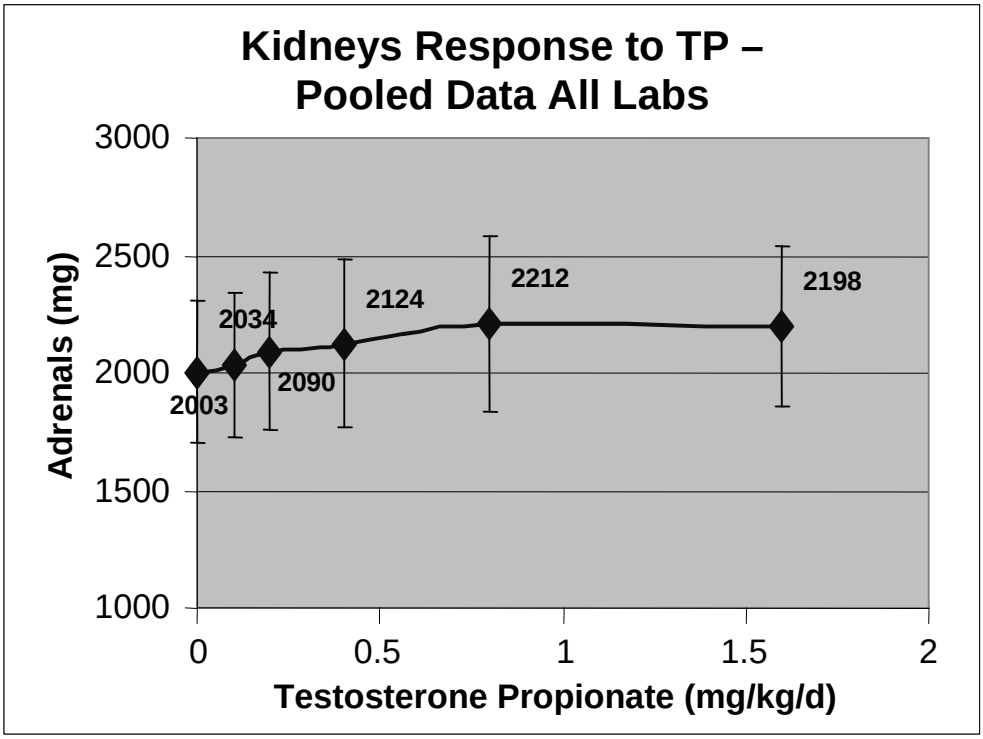
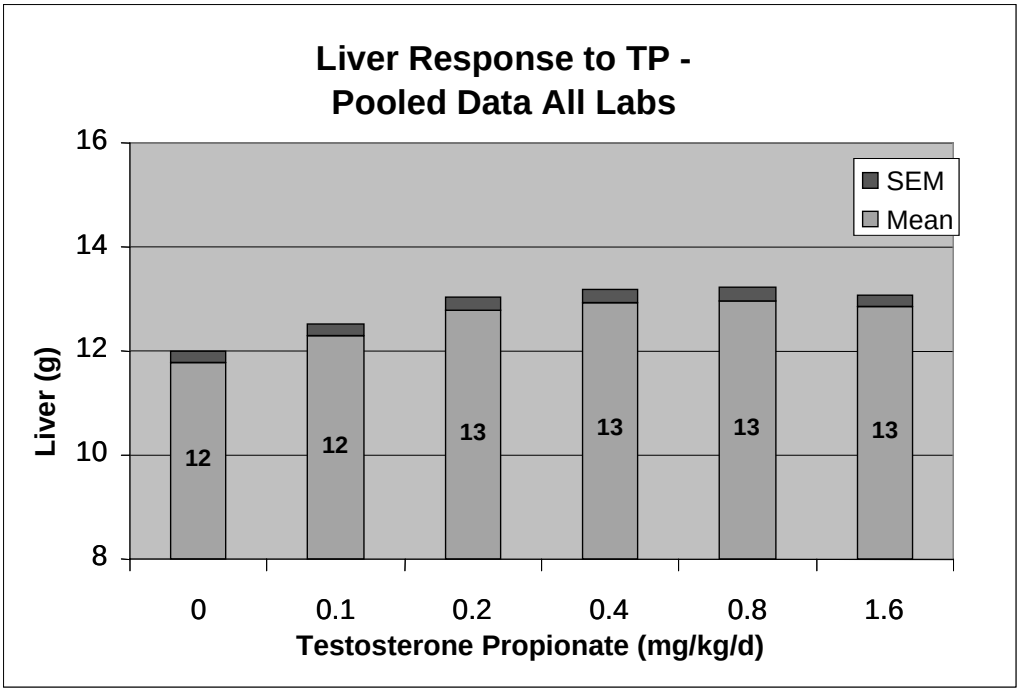


Figure 26. The response of kidney weights using the mean of pooled data to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size. The bottom graph uses the standard deviation from the pooled data. Note the reduced scale of the y-axis in both graphs. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the kidney weights at necropsy.



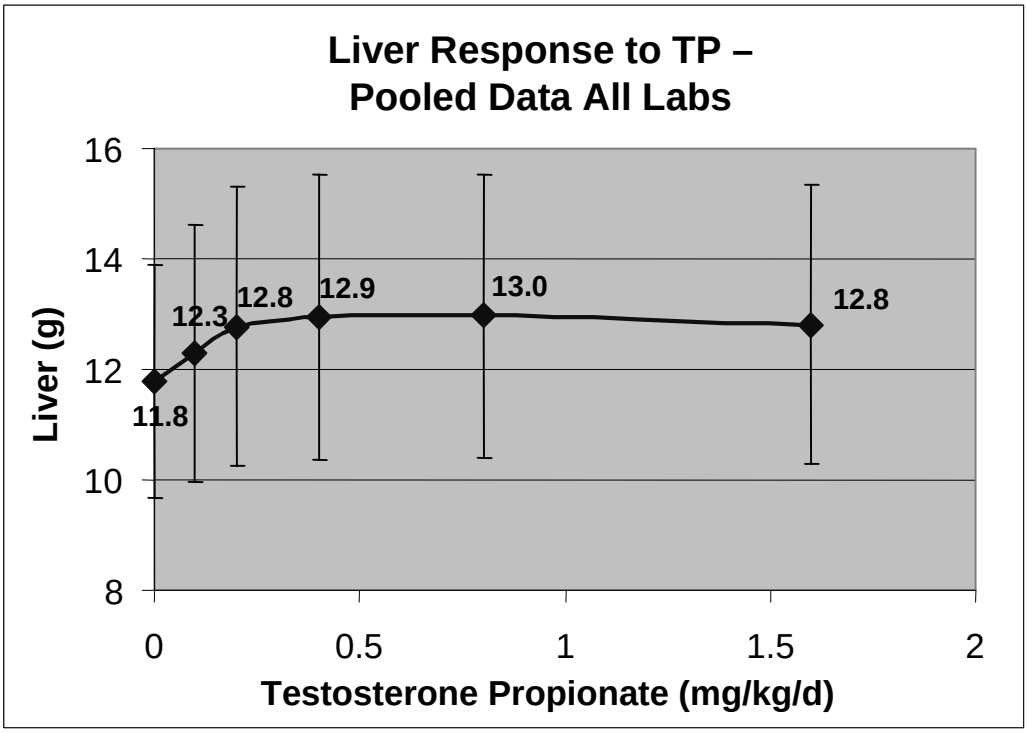
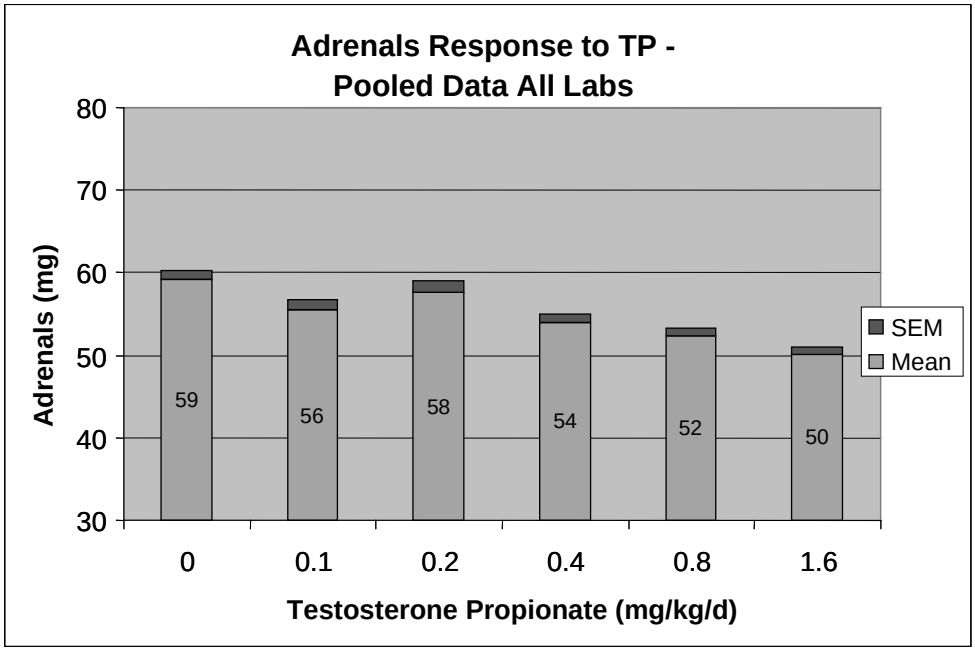


Figure 27. The response of liver weights using the mean of pooled data to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size. The bottom graph uses the standard deviation from the pooled data. Note the reduced scale of the y-axis in both graphs. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the liver weights at necropsy.



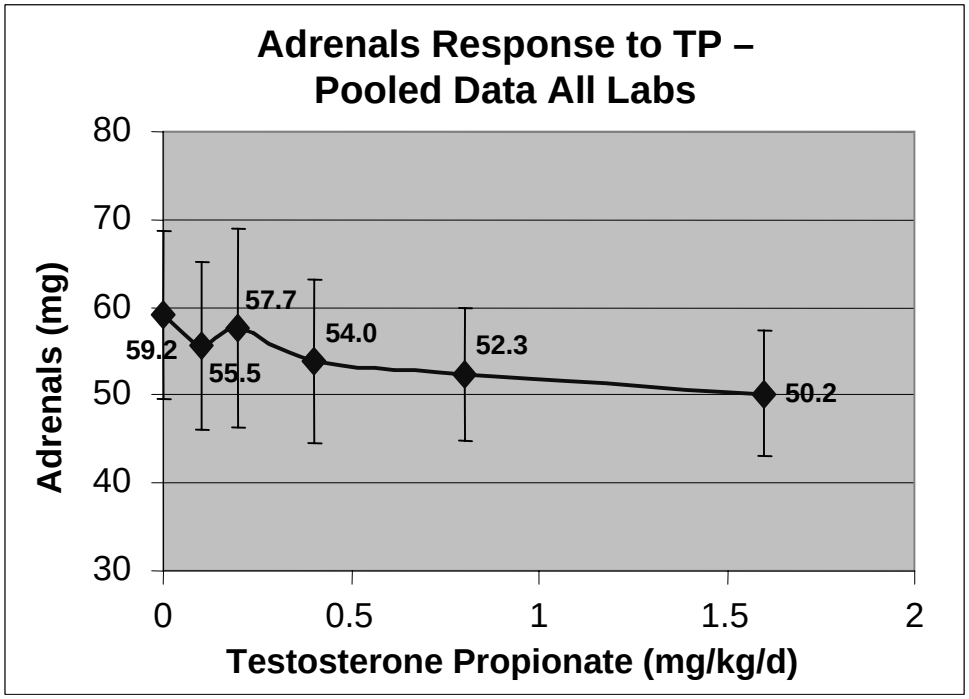


Figure 28. The response of adrenal weights using the mean of pooled data to increasing doses of testosterone propionate (TP) (sc). The top graph uses the standard error of the mean, which responses to sample size. The bottom graph uses the standard deviation from the pooled data. Note the reduced scale of the y-axis in both graphs. This incorporates significant laboratory variability and does not solely reflect the assay variability. Both graphs are plotted as the adrenal weights at necropsy.

Table 1. Description of Laboratory Animal, Treatment, and Diet Variables

LAB	AGE AT CASTRATION	POST-CASTRATION ACCLIMATION-DAYS	% CONTROL PPS +	AGE ON STUDY	RAT STRAIN	DIET
1	Final report not received in time					
2	33-47	10	100	50-65	SD	Pretreatment aliment type M20; lot 991125
3	38	12	0/6	50	Hsd/Cpb:WU	9433LL Meal, Eberle Nafag AG,CH - Lot 44/00
4	31	12	0/12	43	HSD/CPB-WU	Eberle Nafag AG, CH Gossau, NAFAG no.9349
5	43	9	100	49	Crj:CD(SD) IGS	Oriental Yeast Co., Ltd, Tokyo, Japan
6	42	10	100	52	Crl CD®(SD IGS BR	AO4 C pelleted maintenance diet, batch No: 00331
7	42	11-12	4/6	55-56	CD(SD)CR	PMI Certified 5002 Mash Lot # May 22 00 3A
8	40	7	100	47	Crj:CD(SD)IGS	CE-2; Clea Japan Co. Ltd.; Lot#, E2050-P8
9	42	11	100	53	CD Sprague Dawley	Special Diet Services RM1(E) SQC expanded pellet (lot 6706)
10	42	14	100	56	Jcl:Wistar	CE-2, Lot no. E2050-P8 (Clea Japan, Inc.)
11	35	14	0/36	49	Mol:WIST	Altromin 1324, Brogården
12	42	12	100	54	Crj CD(SD)IGS	Oriental Yeast Co., Ltd. Lot No. 000405
14	40	8	1/6	48	SD	PMI Lab Diet, 5014
13	45	7	100	52	Crj:CD(SD)IGS	Oriental Yeast, lot no. 000412A1
15	42	7	100	49	Crj:CD(SD)IGS	CRF-1, Lot No. 00.05.09
16	43	17	5/6	60	CD(SD)CR	PURINA 5002
17	44	8	100	52	Crj:CD(SD)IGS	Oriental Yeast Company, CRF-1, Lot # 000208, 000602

PPS + = Preputial Separation has occurred

Table 2. Ventral Prostate (fresh) Data and Statistical Analysis

Lab		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS						INITIAL	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.2	0.4	0.8	1.6	0		Log10	%
1	MEAN (mg)	23	56	100	135	177	195	241	114	91	94
	CV	34	21	15	6	18	21		20		
2	MEAN (mg)	13	50	102	147	228	249	271	132	65	91
	CV	22	36	60	14	18	15		27		
3	MEAN (mg)	12	34	74	119	145	152	160	90	120	95
	CV	35	23	24	18	9	16		18		
4	MEAN (mg)	26	39	64	96	149	186	163	93	50	89
	CV	43	25	18	25	29	15		26		
5	MEAN (mg)	16	91	118	209	257	266	231	160	143	96
	CV	26	28	20	11	11	23		20		
6	MEAN (mg)	23	83	80	142	257	344	254	155	60	91
	CV	33	39	25	18	32	29		36		
7	MEAN (mg)	34	52	76	117	168	207	311	109	12	67
	CV	54	39	61	40	30	33		42		
8	MEAN (mg)	17	55	107	186	221	246	213	138	146	96
	CV	33	22	21	9	13	13		16		
9	MEAN (mg)	23	96	111	193	226	223	237	145	27	82
	CV	47	119	26	34	17	11		40		
10	MEAN (mg)	11	52	90	140	156	192	214	107	145	96
	CV	25	16	27	18	19	18		22		
12	MEAN (mg)	20	72	130	176	272	292	256	160	154	96
	CV	29	30	9	10	16	15		18		
13	MEAN (mg)	27	101	186	255	339	412	265	220	125	95
	CV	33	16	6	24	17	24		24		
14	MEAN (mg)	35	69	105	180	245	311	224	157	141	96
	CV	7	10	24	15	10	13		16		
15	MEAN (mg)	19	95	140	230	322	311	223	186	210	97
	CV	23	25	10	15	16	13		17		
16	MEAN (mg)	15	77	115	183	264	271	349	154	95	**94**
	CV	35	36	26	17	13	18		21		
17	MEAN (mg)	30	105	169	247	296	332	234	196	140	96
	CV	21	9	26	14	14	20		20		
OVERALL MEAN		**22**	**71**	**110**	**172**	**233**	**262**	**301**			
CV		**48**	**54**	**38**	**32**	**30**	**32**	**5**			
Count		**96**	**96**	**96**	**96**	**96**	**96**				

Table 3. Ventral Prostate (fixed 24 hrs, then weighed) Data and Statistical Analysis

Lab		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS 0	0.1	0.2	0.4	0.8	1.6	INITIAL BODY WT	OVERALL	F VALUE	R-SQUARE
		Not Done									
2	MEAN (MG)	36	96	188	230	367	380	271	216	42	87
	CV	28	38	59	13	19	16		28		
3	MEAN (MG)	14	42	90	147	167	177	160	106	124	95
	CV	35	24	20	19	9	16		18		
4	MEAN (MG)	35	56	86	125	185	250	163	122	37	86
	CV	41	30	17	23	33	20		30		
5	MEAN (MG)	19	108	139	238	254	296	231	182	143	97
	CV	25	29	15	12	11	21		20		
6	MEAN (MG)	29	106	108	183	314	409	254	192	71	92
	CV	35	31	23	21	28	27		32		
7	MEAN (MG)	38	78	100	160	225	253	311	142	12	67
	CV	55	21	79	38	32	36		45		
8	MEAN (MG)	21	73	130	221	259	287	213	165	136	96
	CV	33	18	24	11	10	11		14		
9	MEAN (MG)	35	120	157	252	285	294	237	190	32	84
	CV	39	98	22	31	16	9		33		
10	MEAN (MG)	14	71	118	176	187	227	214	132	147	96
	CV	22	16	27	16	16	20		22		
12	MEAN (MG)	22	85	154	208	309	328	256	185	166	97
	CV	28	28	11	10	16	11		16		
13	MEAN (MG)	31	128	233	309	403	492	265	266	122	95
	CV	33	18	5	24	10	24		22		
14	MEAN (MG)	27	81	117	198	237	310	224	162	175	97
	CV	30	7	15	15	8	12		14		
15	MEAN (MG)	23	113	168	277	371	371	223	220	181	97
	CV	16	30	14	13	20	12		19		
16	MEAN (MG)	20	98	146	216	307	323	349	185	110	**95**
	CV	35	29	22	14	9	14		17		
17	MEAN (MG)	38	130	202	293	369	387	234	236	146	96
	CV	22	11	18	16	14	19		19		
OVERALL	MEAN (MG)	6.9	**27**	**92**	**142**	**215**	**285**	**319**			
	CV	68	**45**	**46**	**40**	**30**	**30**	**30**			
	Count	94	**90**	**90**	**90**	**90**	**90**	**90**			

Table 4. Seminal Vesicles and Coagulating Glands Data and Statistical Analysis

		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS						INITIAL			
Lab		0	0.1	0.2	0.4	0.8	1.6	BODY WT	OVERALL	F VALUE	R-SQUARE
1	MEAN (MG)	7	151	266	397	617	922	241	405	213	98
	CV	8	10	18	27	12	9		20		
2	MEAN (MG)	67	140	244	500	768	1058	271	462	65	91
	CV	23	29	47	45	21	12		29		
3	MEAN (MG)	15	33	109	238	360	559	160	219	287	98
	CV	15	30	23	24	15	10		19		
4	MEAN (MG)	47	130	235	356	577	704	163	342	220	97
	CV	19	11	13	26	12	13		18		
5I	MEAN (MG)	52	230	321	574	808	958	231	490	203	97
	CV	10	27	17	14	8	24		22		
6	MEAN (MG)	75	171	245	548	827	142	254	518	172	97
	CV	18	22	20	19	18	21		25		
7	MEAN (MG)	84	117	275	477	733	1027	311	452	69	92
	CV	43	38	44	24	10	11		20		
8	MEAN (MG)	48	181	361	634	832	1180	213	540	170	97
	CV	14	33	23	17	21	17		23		
9	MEAN (MG)	66	132	345	582	792	903	237	470	52	90
	CV	73	24	24	37	39	19		37		
10	MEAN (MG)	30	119	270	396	615	886	214	386	346	98
	CV	18	26	12	9	10	12		14		
11	MEAN (MG)	20	72	130	611	888	1212	256	547	231	97
	CV	29	30	9	20	13	15		19		
14	MEAN (MG)	42	119	261	524	646	1104	224	449	189	97
	CV	25	26	25	8	21	18		23		
13	MEAN (MG)	66	195	367	651	1034	1214	265	588	193	97
	CV	8	39	14	20	11	11		16		
15	MEAN (MG)	39	176	361	648	970	1111	223	551	375	98
	CV	25	14	16	18	12	9		15		
16	MEAN (MG)	48	194	375	553	936	1170	349	546	334	98
	CV	13	28	9	15	10	8		13		
17	MEAN (MG)	59	191	374	507	946	1226	234	550	325	98
	CV	13	19	17	10	7	16		17		
OVERALL	**MEAN (MG)**	**53**	**152**	**299**	**512**	**772**	**1029**	**300**			
	CV	**45**	**39**	**32**	**30**	**21**	**25**				
	Count	94	**90**	**90**	**90**	**90**	**90**	**90**			

Table 5. LABC Muscle Complex Data and Statistical Analysis

		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS						INITIAL			
Lab		**0**	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**	**BODY WT**	**OVERALL**	**F VALUE**	**R-SQUARE**
1	**MEAN (MG)**	86	143	196	240	284	340	241	215	110	95
	CV	15	13	11	9	7	13		12		
2	**MEAN (MG)**	114	192	278	380	421	463	271	308	36	86
	CV	15	21	34	23	6	10		19		
3	**MEAN (MG)**	119	177	243	321	357	420	160	273	67	97
	CV	11	22	14	10	13	10		14		
4	**MEAN (MG)**	66	146	182	227	247	314	163	197	33	84
	CV	33	25	20	21	29	13		23		
5	**MEAN (MG)**	200	382	514	660	755	734	231	541	186	97
	CV	9	14	715	8	4	6		5		
6	**MEAN (MG)**	288	407	465	616	812	889	254	579	69	92
	CV	17	14	10	9	10	13		12		
7	**MEAN (MG)**	169	278	352	477	594	646	311	430	19	76
	CV	48	15	21	24	17	22		21		
8	**MEAN (MG)**	205	382	509	638	744	749	213	538	199	97
	CV	10	6	8	8	9	10		9		
9	**MEAN (MG)**	174	245	421	536	597	590	237	427	17	74
	CV	24	40	7	5	13	11		14		
10	**MEAN (MG)**	174	337	451	543	612	709	214	471	171	97
	CV	14	10	8	6	10	6		9		
12	**MEAN (MG)**	176	340	516	574	730	838	256	529	137	96
	CV	18	20	10	9	12	5		11		
13	**MEAN (MG)**	260	477	600	832	901	991	265	677	133	96
	CV	12	15	8	11	11	5		10		
14	**MEAN (MG)**	194	351	397	644	661	830	224	513	111	95
	CV	18	20	12	3	8	8		10		
15	**MEAN (MG)**	202	400	506	611	705	744	223	528	140	96
	CV	13	10	11	8	9	7		10		
16	**MEAN (MG)**	213	437	523	621	720	790	349	551	132	96
	CV	12	12	7	9	11	8		10		
17	**MEAN (MG)**	254	413	585	691	946	910	234	610	111	95
	CV	14	16	8	7	7	12		11		
OVERALL	**MEAN (MG)**	**181**	**319**	**421**	**542**	**622**	**685**				
	CV	**37**	**37**	**33**	**32**	**32**	**31**				
	n	**96**	**96**	**96**	**96**	**96**	**96**				

Table 6. Glans Penis Data and Statistical Analysis

Lab		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS 0	0.1	0.2	0.4	0.8	1.6	INITIAL BODY WT	OVERALL	F VALUE	R-SQUARE
1	MEAN (MG)	50	71	76	87	92	89	241	76	46	89
	CV	7	7	9	8	12	6		9		
2	MEAN (MG)	41	63	69	68	63	76	271	59	4.6	44
	CV	22	22	24	24	25	26		24		
3	MEAN (MG)	28	42	49	58	62	63	160	56	16	72
	CV	21	29	22	9	6	10		15		
4	MEAN (MG)	45	69	73	73	74	86	163	76	2.5	29
	CV	43	44	13	18	16	11		26		
5	MEAN (MG)	54	80	79	95	91	97	231	82	34	85
	CV	12	4	6	13	8	7		9		
6	MEAN (MG)	59	74	81	99	102	106	254	84	20	77
	CV	19	8	15	9	3	12		11		
7	MEAN (MG)	49	72	71	93	92	94	311	80	9	61
	CV	30	16	9	15	24	16		18		
8	MEAN (MG)	48	70	75	82	79	83	213	74	80	93
	CV	5	4	6	3	7	7		6		
9	MEAN (MG)	47	66	78	78	84	86	237	73	21	78
	CV	14	12	17	12	6	10		12		
10	MEAN (MG)	31	52	74	82	85	88	214	70	62	91
	CV	12	21	9	9	10	10		11		
11	MEAN (MG)	50	70	86	79	88	96	256	76	23	79
	CV	6	15	11	9	16	14		13		
13	MEAN (MG)	51	71	85	92	104	108	265	82	32	84
	CV	6	14	10	14	10	15		13		
14	MEAN (MG)	35	61	68	83	81	89	224	70	38	87
	CV	7	11	4	16	14	16		14		
15	MEAN (MG)	66	100	111	116	133	118	223	108	28	82
	CV	13	8	14	11	9	12		11		
16	MEAN (MG)	50	74	79	95	96	96	349	82	31	85
	CV	18	10	8	9	7	8		9		
17	MEAN (MG)	66	85	87	94	98	108	234	89	16	74
	CV	13	13	8	5	9	10		10		
OVERALL	**MEAN (MG)**	**48**	**70**	**78**	**86**	**89**	**92**				
	CV	**27**	**24**	**19**	**19**	**22**	**18**				
	n	**92**	**96**	**96**	**96**	**96**	**94**				

Table 7. Cowper's Glands Data and Statistical Analysis

		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS						INITIAL BODY WT	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.2	0.4	0.8	1.6				
1	MEAN (MG)	7	14	24	36	49	56	241	31	172	97
	CV	24	8	15	13	16	7		14		
2	MEAN (MG)	3.8	10	21	33	50	52	271	28	27	84
	CV	30	52	26	27	12	27		27		
3	MEAN (MG)	2.5	8	16	23	30	33	160	19	167	97
	CV	22	31	19	9	13	11		15		
4	MEAN (MG)	1.6	7	12	16	26	28	163	15	39	87
	CV	95	28	21	23	11	10		18		
5	MEAN (MG)	6	23	32	42	56	63	231	37	48	89
	CV	46	20	29	14	17	26		25		
6	MEAN (MG)	10	28	29	44	58	75	254	41	77	93
	CV	23	27	16	10	23	12		17		
7	MEAN (MG)	11	16	21	36	40	55	311	30	9	60
	CV	59	43	46	25	48	23		38		
8	MEAN (MG)	7	16	29	36	45	54	213	31	66	92
	CV	31	19	24	15	23	26		26		
9	MEAN (MG)	6	15	24	31	43	47	237	28	33	84
	CV	64	15	25	24	15	12		20		
10	MEAN (MG)	6	18	29	37	43	52	214	31	198	97
	CV	15	13	14	16	12	7		13		
11	MEAN (MG)	6	16	30	38	51	67	256	35	63	91
	CV	45	29	14	10	26	21		25		
13	MEAN (MG)	11	26	39	53	72	71	265	45	76	93
	CV	34	10	20	16	20	15		19		
14	MEAN (MG)	5	12	25	33	40	66	224	30	143	96
	CV	23	14	24	20	15	17		21		
15	MEAN (MG)	6	25	32	45	64	71	223	41	97	94
	CV	32	38	15	19	14	20		22		
16	MEAN (MG)	5.5	20	28	46	56	59	349	36	115	95
	CV	14	29	11	20	17	20		22		
17	MEAN (MG)	16	31	45	54	68	73	234	48	36	86
	CV	56	20	20	12	15	5		16		
OVERALL	**MEAN (MG)**	**6.9**	**18**	**27**	**38**	**49**	**58**				
	CV	**68**	**45**	**35**	**30**	**31**	**28**				
	n	**94**	**95**	**96**	**96**	**95**	**95**				

Table 8. Body weight Gains in Response to Testosterone Propionate

Lab		DOSE OF TP (sc) – mg/kg/d - FOR 10 DAYS 0	0.1	0.2	0.4	0.8	1.6	p value ANOVA	OVERALL	F VALUE	R-SQUARE %
1	MEAN (G)	71	71	78	82	81	79	0.11	76.9g	2	25%
	CV	12	9	18	9	6	11		11%		
2	MEAN (G)	47	59	69	69	66	77	0.0001	64.4g	7.5	56%
	CV	15	13	20	9	18	8		14%		
3	MEAN (G)	19	15	23	24	24	19	0.04	20.6g	2.7	31%
	CV	32	25	19	24	30	33		28%		
4	MEAN (G)	57	59	59	67	65	70	0.02	62.7g	3.3	36%
	CV	13	8	9	18	9	6		11%		
5	MEAN (G)	60	73	80	81	81	85	0.0002	76.7g	7	54%
	CV	9	15	8	9	7	15		11%		
6	MEAN (G)	66	70	70	74	85	79	0.005	74.1g	4.4	42%
	CV	6	7	6	16	13	12		11%		
7	MEAN (G)	54	60	66	69	70	74	0.009	65.6	3.8	39%
	CV	19	16	12	14	14	5		13%		
8	MEAN (G)	55	71	69	79	81	76	0.0001	71.9g	9.3	61%
	CV	14	13	9	6	10	10		10%		
9	MEAN (G)	56	62	75	80	75	75	0.07	70.6	2.4	28%
	CV	18	17	26	17	18	24		21%		
10	MEAN (G)	52	58	62	61	59	58	0.22	58.4	1.5	20%
	CV	17	9	14	11	7	10		11%		
12	MEAN (G)	47	64	71	68	81	64	0.0001	66g	8.9	60%
	CV	18	10	14	13	14	14		14%		
13	MEAN (G)	72	88	90	95	97	99	0.0001	90g	7.3	55%
	CV	16	12	10	6	8	9		10%		
14	MEAN (G)	78	81	79	89	89	81	0.005	82.9g	4.4	42%
	CV	7	4	6	7	10	4		7%		
15	MEAN (G)	79	72	80	91	92	85	0.002	82g	5.1	46%
	CV	15	23	6	8	8	7		12%		
16	MEAN (G)	49	54	71	76	82	78	0.0006	68	6	50%
	CV	26	48	9	8	11	9		20		
17	MEAN (G)	59	66	72	73	77	77	0.004	71g	6.4	51%
	CV	14	4	5	11	12	8		9.5		
OVERALL	**MEAN (G)**	**57.2**	**63.8**	**69.7**	**73.8**	**75.4**	**75.4**		68.9g		
	CV	**27%**	**28**	**23**	**24**	**24**	**25**		12%		
	n	**96**	**96**	**96**	**96**	**96**	**96**				

ANNEX – 6

SUMMARY LABORATORY DATA TABLES

Table A. Weight of the fresh ventral prostate gland (VP) in castrated, immature rats administered testosterone propionate

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	23 [34]	56 [21]	100 [15]	135 [6]	177 [18]	195 [21]
2	13 [22]	50 [36]	102 [60]	147 [14]	228 [18]	249 [15]
3	12 [35]	34 [23]	74 [24]	119 [18]	145 [9]	152 [16]
4	26 [43]	39 [25]	64 [18]	96 [25]	149 [29]	186 [15]
5	16 [26]	91 [28]	118 [20]	209 [11]	257 [11]	266 [23]
6	23 [33]	83 [39]	80 [25]	142 [18]	257 [32]	344 [29]
7	34 [54]	52* [39]	76 [61]	117 [40]	168 [30]	207 [33]
8	17 [33]	55 [22]	107 [21]	186 [9]	221 [13]	246 [13]
9	23 [47]	96 [119]	111 [26]	193 [34]	226 [17]	223 [11]
10	11 [25]	52 [16]	90 [27]	140 [18]	156 [19]	192 [18]
11	20 [29]	72 [30]	130 [9]	176 [10]	272 [16]	292 [15]
12	35 [7]	69 [10]	105 [24]	180 [15]	245 [10]	311 [13]
13	19 [23]	95 [25]	140 [10]	230 [15]	322 [16]	311 [13]
14	27 [33]	101 [16]	186 [6]	255 [24]	339 [17]	412 [24]
15	15 [35]	77 [36]	115 [26]	183 [17]	264 [13]	271 [18]
16	30 [21]	105 [9]	169 [26]	247 [14]	296 [14]	332 [20]
All labs	***22 [48]***	***71 [54]***	***110 [38]***	***172 [32]***	***233 [30]***	***262 [32]***

mean weight in mg [CV]
* not significantly different from 0 mg/kg dose

Table B. Weight of the fixed ventral prostate gland (VP-fixed) in castrated, immature rats administered testosterone propionate

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	n.d.*	n.d.	n.d.	n.d.	n.d.	n.d.
2	36 [28]	96 [38]	188 [59]	230[13]	367 [19]	380 [16]
3	n.d.*	n.d.	n.d.	n.d.	n.d.	n.d.
4	14 [35]	42 [24]	90 [20]	147 [9]	167 [9]	177 [16]
5	35 [41]	56 [30]	86 [17]	125 [23]	185 [33]	250 [20]
6	19 [25]	108 [29]	139 [15]	238 [12]	254 [11]	296 [21]
7	29 [35]	106 [31]	108 [23]	183 [21]	314 [28]	409 [27]
8	38 [55]	78 [21]	100 [79]	160 [38]	225 [32]	253 [36]
9	21 [33]	73 [18]	130 [24]	221 [11]	259 [10]	287 [11]
10	35 [39]	120 [98]	157 [22]	252 [31]	285 [16]	294 [9]
11	14 [22]	71 [16]	118 [27]	176 [16]	187 [16]	227 [20]
12	22 [28]	85 [28]	154 [11]	208 [10]	309 [16]	328 [11]
13	31 [33]	128 [18]	233 [5]	309 [24]	403 [10]	492 [24]
14	27 [30]	81 [7]	117 [15]	198 [15]	237 [8]	310 [12]
15	23 [16]	113 [30]	168 [14]	277 [13]	371 [20]	371 [12]
16	20 [35]	98 [29]	146 [22]	216 [14]	307 [9]	323 [14]
17	38 [22]	130 [11]	202 [18]	293 [16]	369 [14]	387 [19]
All labs	***27 [45]***	***92 [46]***	***142 [40]***	***215 [30]***	***285 [30]***	***319 [30]***

mean weight in mg [CV]
* not done

Table C. Weight of the seminal vesicle plus coagulating glands (SV) in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	77 [8]	151 [10]	266 [18]	397 [27]	617 [12]	922 [9]
2	67 [23]	140 [29]	244 [47]	500 [45]	768 [21]	1050 [12]
3	15 [15]	33 [30]	109 [23]	238 [24]	360 [15]	559 [10]
4	47 [19]	130 [11]	235 [13]	356 [26]	577 [12]	704 [13]
5	52 [10]	230 [27]	321 [17]	574 [14]	808 [8]	958 [24]
6	75 [18]	171 [22]	245 [20]	548 [19]	827 [18]	1243 [21]
7	84 [13]	117* [38]	275 [44]	477 [24]	733 [10]	1027 [11]
8	48 [14]	181 [33]	361 [23]	634 [17]	832 [21]	1180 [17]
9	66 [73]	132 [24]	345 [24]	582 [37]	792 [39]	903 [19]
10	30 [18]	119 [26]	270 [12]	396 [9]	615 [10]	886 [12]
12	20 [29]	72 [30]	130 [9]	611 [20]	888 [13]	1212 [15]
13	66 [8]	195 [39]	367 [14]	651 [20]	1034 [11]	1214 [11]
14	42 [25]	119 [26]	261 [25]	524 [8]	646 [21]	1104 [18]
15	39 [25]	176 [14]	361 [16]	648 [18]	970 [12]	1111 [9]
16	48 [13]	194 [28]	375 [9]	553 [15]	936 [10]	1170 [8]
17	59 [13]	191 [19]	374 [17]	507 [10]	946 [7]	1226 [16]
All labs	***53 [45]***	***152 [39]***	***299 [32]***	***512 [30]***	***772 [21]***	***1029 [25]***

mean weight in mg [CV]

* not significantly different from 0 mg/kg dose

Table D. Weight of the levetor ani plus bulbocavernosus muscle (LABC) in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	86 [15]	143 [13]	196 [11]	240 [9]	284 [7]	340 [13]
2	114 [15]	192 [21]	278 [34]	380 [23]	421 [6	463 [10]
3	119 [11]	177 [22]	243 [14]	321 [10]	357 [13	420 [10]
4	66 [33]	146 [25]	182 [20]	227 [21]	247 [29	314 [13]
5	200 [9]	382 [14]	514 [715]	660 [8]	755 [4	734 [6]
6	288 [17]	407 [14]	465 [10]	616 [9]	812 [10	889 [13]
7	169 [48]	278 [15]	352 [21]	544 [24]	594 [17	646 [22]
8	205 [10]	382 [6]	509 [8]	638 [8]	744 [9	749 [10]
9	174 [24]	245* [40]	421 [7]	536 [5]	597 [13	590 [11]
10	174 [14]	337 [10]	451 [8]	543 [6]	612 [10	709 [6]
11	176 [18]	340 [20]	516 [10]	574 [9]	730 [12	838 [5]
12	194 [18]	351 [20]	397 [12]	644 [3]	661 [8	830 [8]
13	260 [12]	477 [15]	600 [8]	832 [11]	901 [11	991 [5]
15	202 [13]	400 [10]	506 [11]	611 [8]	705 [9	744 [7]
16	213 [12]	437 [12]	523 [7]	621 [9]	720 [11	790 [8]
17	254 [14]	413 [16]	585 [8]	691 [7]	946 [7	910 [12]
All labs	***181[37]***	***319 [37]***	***421 [33]***	***542 [32]***	***622 [32]***	***685 [31]***

mean weight in mg [CV]

* not significantly different from 0 mg/kg dose

Table E. Weight of the Cowper's glands (COWS) in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	7 [24]	14 [8]	24 [15]	36 [13]	49 [16]	56 [7]
2	3.8 [30]	10 [52]	21 [26]	33 [27]	50 [12]	52 [27]
3	2.5 [22]	8 [31]	16 [19]	23 [9]	30 [13]	33 [11]
4	1.6 [95]	7 [28]	12 [21]	16 [23]	26 [11]	28 [10]
5	6 [46]	23 [20]	32 [29]	42[14]	56 [17]	63 [26]
6	10 [23]	28 [27]	29 [16]	44 [10]	58 [23]	75 [12]
7	11 [59]	16* [43]	21 [46]	36 [25]	40 [48]	55 [23]
8	7 [31]	16 [19]	29 [24]	36 [15]	45 [23]	54 [26]
9	6 [64]	15 [15]	24 [25]	31 [24]	43 [15]	47 [12]
10	6 [15]	18 [13]	29 [14]	37 [16]	43 [12]	52 [7]
12	6 [45]	16 [29]	30 [14]	38 [10]	51 [26]	67 [21]
14	5 [23]	12 [14]	25 [24]	33 [20]	40 [15]	66 [17]
13	11 [34]	26 [10]	39 [20]	53 [16]	72 [20]	71 [15]
15	6 [32]	25 [38]	32 [15]	45 [19]	64 [14]	71 [20]
16	5.5 [14]	20 [29]	28 [11]	46 [20]	56 [17]	59 [20]
17	16 [56]	31 [20]	45 [20]	54 [12]	68 [15]	73 [15]
All labs	***6.9 [68]***	***18 [45]***	***27 [35]***	***38 [30]***	***49 [31]***	***58 [28]***

mean weight in mg [CV]
* not significantly different from 0 mg/kg dose

Table F. Weight of the glans penis (GLANS) in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	50 [7]	71 [7]	76 [9]	87 [8]	92 [12]	89 [6]
2	41 [22]	63 [22]	69 [24]	68 [24]	63 [25]	76 [26]
3	28 [21]	42 [29]	49 [22]	58 [9]	62 [6]	63 [10]
4	45 [43]	69* [44]	73 [13]	73 [18]	74 [16]	86 [11]
5	54 [12]	80 [4]	79 [6]	95 [13]	91 [8]	97 [7]
6	59 [19]	74 [8]	81 [15]	99 [9]	102 [3]	106 [12]
7	49 [30]	72 [16]	71 [9]	93 [15]	92 [24]	94 [16]
8	48 [5]	70 [4]	75 [6]	82 [3]	79 [7]	83 [7]
9	47 [14]	66 [12]	78 [17]	78 [12]	84 [6]	86 [10]
10	31 [12]	52 [21]	74 [9]	82 [9]	85 [10]	88 [10]
12	50 [6]	70 [15]	86 [11]	79 [9]	88 [16]	96 [14]
13	51 [6]	71 [14]	85 [10]	92 [14]	104 [10]	108 [15]
14	35 [7]	61 [11]	68 [4]	83 [16]	81 [14]	89 [16]
15	66 [13]	100 [8]	111 [14]	116 [11]	133 [9]	118 [12]
16	50 [18]	74 [10]	79 [8]	95 [9]	96 [7]	96 [8]
17	66 [13]	85 [13]	87 [8]	94 [5]	98 [9]	108 [10]
All labs	***48[27]***	***70 [24]***	***78 [19]***	***86 [19]***	***89 [22]***	***92 [18]***

mean weight in mg [CV]
* not significantly different from 0 mg/kg dose

Table G. Percent increases in body weight between Day 0 and Day 10 of s.c. administration of testosterone propionate

Lab.	Avg. wt. at day 0*	mg Testosterone Propionate/kg/day					
		0	0.1	0.2	0.4	0.8	1.6
1	241.0 (2.6)	25.6	25.9	27.3	30.0	29.3	29.1
2	270.7 (1.1)	15.1	18.9	21.9	21.6	22.5	25.0
3	159.8 (2.0)	28.5	25.0	31.9	33.8	32.3	28.2
4	163.2 (3.0)	33.7	36.0	36.6	41.4	40.4	42.4
5	230.7 (0.7)	24.2	29.3	31.6	33.3	32.9	33.8
6	256.0 (7.3)	18.0	26.3	29.4	29.6	33.1	32.4
7	245.5 (2.2)	19.9	21.6	24.1	24.8	25.8	26.7
8	213.2 (1.1)	24.4	30.5	30.7	34.4	34.9	32.7
9	236.7 (1.5)	24.4	25.4	30.5	32.5	31.2	30.8
10	214.2 (0.4)	22.9	25.2	26.5	26.6	26.6	24.8
12	256.0 (0.8)	17.1	23.0	24.1	24.7	29.0	25.0
13	265.2 (0.7)	23.7	29.8	30.7	32.1	34.6	34.3
14	224.3 (1.2)	31.3	31.6	32.0	34.4	35.6	31.4
15	223.7 (2.5)	29.7	27.1	32.0	36.2	37.4	35.0
16	280.7 (4.7)	17.2	20.8	22.6	25.2	26.9	25.9
17	234.3 (0.9)	23.2	24.6	27.7	27.8	31.2	29.9

* avg. weight in gms at day 0, prior to treatment, for all animals; mean (S.D.)

Table H. Weight of the livers in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	0.1	0.2	0.4	0.8	1.6
2	11.2 [12]	12.1 [10]	12.7 [10]	12.7 [6]	12.1 [10]	13.2 [8]
3	5.7 [6]	5.5 [6]	5.8 [4]	5.7 [9]	5.9 [10]	5.6 [12]
4	11.3 [11]	10.7 [4]	10.6 [8]	11.0 [6]	11.3 [8]	11.5 [7]
5	11.4 [6]	12.2 [9]	12.8 [9]	12.3 [6]	12.8 [7]	13.3 [14]
6	12.5 [6]	13.0 [5]	12.8 [3]	13.2 [15]	13.7 [11]	13.8 [7]
7	12.6 [6]	12.9 [10]	13.4 [10]	13.1 [8]	13.0 [9]	13.5 [4]
8	11.3 [7]	12.2 [5]	12.2 [8]	14.5 [33]	12.7 [8]	11.6 [7]
9	13.4 [12]	14.2 [5]	16.0 [10]	16.2 [6]	16.7 [10]	16.3 [9]
10	10.3 [9]	10.2 [6]	11.1 [9]	11.1 [9]	11.2 [5]	11.1 [5]
12	12.0 [8]	13.3 [8]	13.7 [7]	13.5 [6]	13.8 [7]	13.4 [8]
13	12.7 [12]	13.3 [7]	14.5 [6]	14.1 [10]	14.1 [7]	13.3 [10]
14	11.5 [4]	12.0 [4]	11.9 [4]	12.8 [1]	12.6 [7]	12.2 [4]
15	12.8 [11]	13.1 [4]	13.3 [5]	13.8 [7]	13.9 [6]	13.4 [8]
16	13.6 [7]	15.2 [10]	16.1 [11]	16.7 [7]	15.8 [9]	16.1 [6]
17	12.3 [8]	13.3 [8]	13.9 [5]	13.3 [4]	14.0 [6]	13.1 [8]
All labs	***11.6 [16]***	***12.2 [18]***	***12.7 [20]***	***12.9 [19]***	***12.9 [19]***	***12.8 [19]***

mean weight in mg [CV]

Table I. Weight of the kidneys in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
1	2089 [12]	2023 [9]	2208 [8]	2202 [9]	2142 [10]	2119 [9]
3	1468 [9]	1420 [6]	1483 [7]	1472 [9]	1563 [11]	1485 [9]
4	1627 [7]	1674 [5]	1677 [19]	1611 [7]	1658 [4]	1822 [4]
5	1960 [6]	2085 [8]	2125 [4]	2118 [9]	2175 [7]	2230 [8]
6	2057 [8]	2032 [7]	2131[8]	2143 [13]	2375 [7]	2337 [7]
7	2250 [7]	2360 [8]	2525 [9]	2485 [8]	2597 [5]	2527 [8]
8	2318 [9]	2138 [8]	2265 [12]	2210 [7]	2366 [9]	2360 [10]
10	1779 [7]	1878 [5]	1843 [9]	1899 [4]	2003 [5]	2075 [5]
12	2134 [8]	2275 [6]	2347 [10]	2317 [6]	2318 [5]	2362 [9]
13	2473 [6]	2480 [7]	2560 [6]	2742 [6]	2889 [7]	2790 [4]
14	2019 [4]	2124 [9]	1996 [4]	2165 [6]	2124 [7]	2093 [3]
15	1879 [12]	1924 [5]	1946 [4]	2071 [4]	2259 [11]	2132 [4]
16	1981 [7]	2024 [6]	2126 [5]	2146 [5]	2270 [8]	2237 [6]
All labs	***2003 [14]***	***2034 [14]***	***2095 [15]***	***2122 [16]***	***2211 [16]***	***2198 [15]***

mean weight in mg [CV]

Table J. Weight of the adrenal glands in castrated, immature rats administered testosterone propionate (fresh weight)

Lab	mg Testosterone Propionate/kg/day					
	0	**0.1**	**0.2**	**0.4**	**0.8**	**1.6**
3	71.8 [9]	67.7 [15]	71.3 [7]	63.3 [14]	60.5 [11]	61.0 [14]
4	55.2 [11]	48.2 [19]	52.2 [15]	48.0 [16]	47.2 [19]	44.7 [8]
5	48.7 [15]	50.6 [21]	55.1 [10]	46.5 [17]	47.5 [8]	47.2 [10]
6	62.0 [10]	56.4 [14]	54.5 [10]	57.6 [9]	49.4 [6]	48.3 [13]
7	62.4 [15]	65.2 [8]	71.4 [19]	55.3 [11]	56.8 [11]	52.4 [14]
10	65.8 [6]	61.5 [11]	56.2 [13]	62.3 [11]	57.1 [21]	52.8 [8]
12	54.8 [15]	52.9 [13]	49.5 [18]	48.1 [17]	47.7 [19]	45.0 [13]
13	60.6 [13]	62.2 [13]	60.5 [20]	56.5 [13]	55.1 [5]	54.4 [6]
14	53.8 [3]	47.7 [10]	47.5 [9]	49.2 [5]	47.9 [6]	49.4 [8]
16	58.1 [22]	49.6 [7]	52.1 [10]	49.9 [6]	50.6 [15]	44.9 [10]
17	51.7 [13]	51.0 [14]	52.0 [11]	48.2 [12]	51.5 [9]	47.0 [11]
All labs	***58.6 [11]***	***55.7 [13]***	***56.6 [14]***	***53.2 [11]***	***51.9 [9]***	***49.7 [10]***

mean weight in mg [CV]

Table K. Effects of testosterone propionate and flutamide administration on mean ventral prostate (VP) weights in each laboratory (fresh tissue)

A.		mg Flutamide/kg/day + 0.2 mg Testosterone Propionate/kg/day					
Lab	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	16.1 [23]	114 [9]	117 [9]	103 [18]	63.7* [24]	31.2* [17]	19.4* [12]
8	14.6 [37]	105 [32]	105 [33]	81 [24]	54* [36]	28* [34]	23* [42]
12	19 [43]	128 [11]	107 [22]	101* [19]	56* [20]	32* [17]	22* [15]
13	24.6 [24]	142 [20]	150 [20]	100* [17]	85* [32]	48* [16]	32* [18]
15	22.2 [19]	131 [13]	150 [28]	109 [17]	83* [18]	38* [12]	25* [17]
17	15.1 [31]	140 [18]	134 [15]	117 [30]	69* [17]	33* [30]	26* [16]
ALL LABS	*18.6 [39]*	*127 [20]*	*127 [26]*	*102* [23]*	*69* [30]*	*35* [27]*	*25* [25]*
B.		mg Flutamide/kg/day + 0.4 mg Testosterone Propionate/kg/day					
Lab	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	16.1 [23]	211 [17]	177 [20]	175 [13]	117* [19]	59* [38]	28* [16]
10	9.4 [26]	163 [18]	128*[19]	104* [16]	67* [26]	37* [14]	20* [15]
12	19 [43]	214 [10]	171 [27]	174 [14]	113* [29]	52* [31]	26* [36]
13	24.6 [24]	233 [20]	228 [20]	197 [12]	176* [15]	80* [17]	46* [28]
15	22.2 [19]	268 [22]	213* [16]	202* [10]	136* [10]	68* [26]	31* [8]
ALL LABS	*18.4 [39]*	*218 [24]*	*184* [27]*	*170* [24]*	*122* [34]*	*59* [35]*	*30* [37]*

mean weight in mg [CV]
* Significantly different from 0 mg/kg flutamide +TP response ($p<0.05$)

Table L. Effects of testosterone propionate and flutamide administration on mean seminal vesicle (SV) weights in each laboratory (fresh tissue)

A.		mg Flutamide/kg/day + 0.2 mg Testosterone Propionate/kg/day					
Lab	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	37 [25]	315 [13]	252* [11]	204* [23]	123* [18]	54* [22]	41* [15]
8	32 [20]	348 [22]	263 [33]	247* [33]	117* [32]	60* [35]	45* [16]
12	53 [28]	368 [24]	366 [20]	248* [27]	137* [37]	71* [12]	62* [22]
13	67 [13]	316 [21]	288 [28]	247 [29]	137* [25]	86* [20]	74* [11]
15	46 [18]	289 [28]	232 [32]	206* [18]	116* [28]	68* [26]	50* [9]
17	51 [22]	440 [22]	320* [15]	299* [28]	167* [40]	83* [21]	67* [10]
ALL LABS	*47.7 [34]*	*346 [25]*	*287* [27]*	*243* [29]*	*133* [33]*	*70* [27]*	*56* [25]*
B.		mg Flutamide/kg/day + 0.4 mg Testosterone Propionate/kg/day					
Lab	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	37 [25]	539 [12]	450 [17]	410* [13]	243* [22]	108* [26]	45* [17]
10	31 [17]	507 [7]	418* [23]	353*[14]	189* [25]	83* [16]	41* [19]
12	53 [28]	588 [20]	493 [21]	495 [28]	295* [22]	122* [32]	65* [12]
13	67 [13]	674 [23]	646 [113]	573 [21]	402* [24]	168* [35]	83* [17]
15	46 [18]	592 [11]	431* [9]	406* [16]	266* [14]	106* [26]	61* [16]
ALL LABS	*46.8 [34]*	*580 [18]*	*488* [23]*	*447* [24]*	*279* [33]*	*117* [38]*	*59* [30]*

mean weight in mg [CV]
* Significantly different from 0 mg/kg flutamide +TP response ($p<0.05$)

Table M. Effects of testosterone propionate and flutamide administration on mean levetor ani + bulbocavernosus muscle (LABC) weights in each laboratory (fresh tissue)

A.		**mg Flutamide/kg/day + 0.2 mg Testosterone Propionate/kg/day**					
Labs	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	178 [13]	444 [5]	413 [7]	387* [10]	306* [10]	239* [12]	189* [9]
8	175 [8]	488 [5]	493 [10]	422 [14]	359* [16]	243* [16]	214* [14]
12	187 [11]	482 [11]	461 [18]	412 [11]	273* [14]	223* [10]	217* [18]
13	268 [10]	515 [10]	516 [13]	441 [24]	414* [17]	316* [15]	268* [8]
15	208 [15]	442 [10]	454 [12]	415 [9]	332* [11]	259* [12]	220* [9]
17	249 [14]	565 [9]	511 [17]	514 [13]	419* [18]	280* [13]	268* [11]
ALL LABS	*210.8 [22]*	*490 [12]*	*474 [15]*	*433* [16]*	*351* [21]*	*260* [17]*	*227* [17]*
B.		**mg Flutamide/kg/day + 0.4 mg Testosterone Propionate/kg/day**					
Labs	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	178 [13]	578 [8]	555 [12]	547 [4]	419* [16]	314* [12]	202* [17]
10	167 [18]	603 [10]	596 [7]	497* [6]	435* [7]	328* [11]	207* [11]
12	187 [11]	606 [9]	585 [9]	549 [7]	406* [14]	265* [13]	235* [20]
13	268 [10]	784 [11]	710 [12]	659* [7]	577* [8]	391* [12]	321* [4]
15	208 [15]	654 [10]	588 [7]	512* [10]	445* [11]	337* [14]	252* [8]
ALL LABS	*201.6 [22]*	*645 [15]*	*607 [13]*	*553* [12]*	*456* [17]*	*327* [17]*	*243* [21]*

mean weight in mg [CV]

* Significantly different from 0 mg/kg flutamide +TP response ($p<0.05$)

Table N. Effects of testosterone propionate and flutamide administration on mean Cowper's glands (COWS) weights in each laboratory (fresh tissue)

A.		**mg Flutamide/kg/day + 0.2 mg Testosterone Propionate/kg/day**					
Labs	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	5 [30]	25.5 [28]	27.3 [22]	20.5 [18]	17.2* [27]	9.8* [30]	6.8* [11]
8	6.4 [43]	25.1 [15]	24.7 [18]	20.6 [15]	14.4* [17]	9.2* [22]	7* [26]
12	9.3 [12]	32.1 [15]	30.7 [10]	24.4* [19]	18.1* [17]	11.7* [29]	8.8* [13]
13	10.1 [27]	31.1 [10]	31.1 [15]	24.2* [28]	23.1* [21]	13.9* [35]	12.2* [20]
15	8.4 [14]	26.7 [13]	26.6 [30]	20* [24]	21.6 [15]	10.7* [29]	7.2* [22]
17	10 [34]	37.9 [30]	30.7 [10]	27.2 [24]	25.7 [25]	13.9* [31]	11* [32]
ALL LABS	*8.2 [36]*	*29.4 [24]*	*28.5 [19]*	*22.9* [24]*	*20* [28]*	*11.5* [33]*	*8.8* [32]*
B.		**mg Flutamide/kg/day + 0.4 mg Testosterone Propionate/kg/day**					
Labs	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	5 [30]	40 [14]	34.7 [16]	33.8 [13]	23.6* [16]	17.5* [41]	8.3* [21]
10	5.3 [18]	41.9 [5]	39.2 [10]	34.1 [12]	26* [9]	16.4* [27]	7.3* [28]
12	9.3 [12]	45.4 [9]	38.9 [21]	33.9* [13]	26* [15]	15.6* [13]	11.0* [20]
13	10.1 [27]	51.8 [19]	52.1 [15]	41.2* [9]	38.3* [19]	24.7* [18]	15.4* [18]
15	8.4 [14]	43.8 [13]	38.4 [19]	35.8 [13]	30.2* [13]	16.6* [24]	9.1* [26]
ALL LABS	*7.6 [36]*	*44.6 [16]*	*40.6* [21]*	*35.7* [14]*	*29* [24]*	*18.2* [30]*	*10.4* [35]*

mean weight in mg [CV]

* Significantly different from 0 mg/kg flutamide +TP response ($p<0.05$)

Table O. Effects of testosterone propionate and flutamide administration on mean glans penis (GLANS) weights in each laboratory (fresh tissue)

A.		mg Flutamide/kg/day + 0.2 mg Testosterone Propionate/kg/day					
Labs	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	48.6 [10]	77 [6]	78.4 [4]	75.8 [5]	67.6* [7]	57.7* [13]	51.3* [6]
8	40.1 [15]	68.7 [13]	70.8 [16]	66.6 [8]	66[12]	51.5* [11]	49.6* [11]
12	55.1 [5]	84.9 [8]	80.7 [8]	80 [7]	68.8* [4]	60.6* [11]	56.3* [6]
13	48.7 [8]	80.2 [7]	76.5 [3]	77.1 [13]	70.1* [9]	59.1* [11]	58.7* [18]
15	64.8 [13]	110 [12]	92.4* [12]	93* [7]	86.8* [12]	73.9* [13]	62.9* [10]
17	62.5 [21]	103 [9]	91.2 [8]	87.4* [20]	82.9* [9]	73.7* [15]	72.7* [14]
ALL LABS	*53.3 [25]*	*87.4 [19]*	*81.7* [13]*	*80.1* [15]*	*73.7* [14]*	*62.7* [18]*	*58.6* [18]*
B.		**mg Flutamide/kg/day + 0.4 mg Testosterone Propionate/kg/day**					
Labs	**no TP**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	48.6 [10]	82.7 [8]	87.9 [3]	81 [6]	77.4 [8]	70* [7]	86.6* [4]
10	30.8 [15]	80.9 [5]	84.7 [8]	75.9 [5]	64.6* [7]	51.9* [17]	39.5* [18]
12	55.1 [5]	90.8 [6]	87.9 [5]	88.9 [3]	81.5* [5]	69.2* [7]	59.9* [2]
13	48.7 [8]	95.8 [9]	93.4 [9]	88.7 [9]	82.9* [14]	67.4* [7]	56.8* [8]
15	64.8 [13]	114.7 [4]	118.1 [12]	106.9 [12]	102.9 [11]	87.2* [15]	77.4* [9]
ALL LABS	*49.6 [25]*	*93 [15]*	*94.4 [15]*	*88.3* [14]*	*81.8* [18]*	*69.1* [20]*	*58* [23]*

mean weight in mg [coefficient of variation]

* Significantly different from 0 mg/kg flutamide +TP response ($p<0.05$)

Table P. Weights of fresh and fixed VP in animals treated with TP and flutamide

			mg flutamide/kg/day						
Lab	**TP**	**type**	**0**	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
5	0	fresh	16 [23]						
		fixed	18 [21]						
	0.2	fresh		114 [9]	117 [8]	103 [18]	64 [24]	31 [17]	19 [12]
		fixed		120 [10]	123 [7]	111 [16]	70 [24]	34 [16]	21 [13]
	0.4	fresh		211 [17]	177 [20]	175 [13]	117 [19]	59 [38]	28 [16]
		fixed		222 [16]	190 [18]	187 [16]	122 [17]	66 [38]	31 [19]
8	0	fresh	15 [37]						
		fixed	17 [35]						
	0.2	fresh		105 [32]	106 [33]	80 [28]	54 [36]	28 [34]	23 [42]
		fixed		133 [28]	122 [16]	109 [26]	69 [36]	36 [32]	28 [50]
12	0	fresh	19 [43]						
		fixed	18 [46]						
	0.2	fresh		128 [11]	107 [22]	101 [19]	56 [20]	32 [17]	22 [15]
		fixed		138 [11]	117 [23]	111 [19]	61 [21]	32 [18]	21 [19]
	0.4	fresh		214 [10]	171 [27]	174 [14]	113 [29]	52 [31]	26 [36]
		fixed		227 [9]	179 [26]	183 [13]	120 [27]	57 [34]	26 [40]
15	0	fresh	22 [18]						
		fixed	25 [21]						
	0.2	fresh		131 [13]	150 [28]	109 [17]	83 [18]	38 [12]	25 [7]
		fixed		169 [19]	194 [24]	136 [17]	109 [20]	46 [14]	29 [7]

	0.4	fresh		268 [22]	213 [15]	202 [10]	136 [10]	68 [26]	31 [8]
		fixed		321 [20]	262 [16]	240 [9]	169 [12]	80 [31]	36 [13]

mean weight in mg [CV]

Table Q. Percent increase in body weight between Day 0 and Day 10 of administration of flutamide to castrated, immature rats receiving testosterone propionate.

Lab.	Avg. wt. at day 0*	Untreated control **	mg flutamide/kg/day + 0.2 mg TP/kg/day					
			0	**0.1**	**0.3**	**1.0**	**3.0**	**10**
5	235.0 (0.82)	29.1	32.9	32.2	35.3	32.3	31.1	29.7
8	216.3 (1.11)	24.8	28.6	28.7	29.0	29.0	28.2	27.2
12	241.7 (0.76)	17.4	21.2	21.1	19.0	20.7	18.7	19.3
13	258.9 (1.21)	25.9	31.1	33.0	29.6	29.1	28.3	28.1
15	217.9 (0.90)	26.6	29.8	29.4	30.1	29.8	28.9	29.2
17	237.1 (1.07)	16.5	22.0	25.6	22.6	20.3	21.5	19.4
			mg flutamide/kg/day + 0.4 mg TP/kg/day					
Lab			**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10**
5	235.0 (1.00)	29.1	38.1	36.9	35.9	34.6	33.6	32.2
10	229.4 (0.53)	18.8	24.8	27.3	21.7	21.0	21.8	22.2
12	241.7 (1.11)	17.4	24.1	23.3	25.1	19.8	21.5	18.9
13	265.0 (1.53)	23.1	31.8	31.3	31.1	28.0	27.4	27.7
15	217.9 (1.57)	26.6	34.4	33.9	32.9	32.1	31.3	29.0

* avg. weight in gms at day 0 for all untreated animals; mean (S.D.)

** Animals not receiving testosterone propionate or flutamide

Table R. Effects of 0.2 mg testosterone propionate and flutamide administration on mean liver, kidney, and adrenal weights

A.		**Control (no TP)**	**mg Flutamide/kg/day + 0.2 mg Testosterone Propionate /kg/day**					
		0	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
Liver (gms)								
Lab	**5**	12.9 [9]	1 3.9 [9]	14.4 [8]	1 3.8 [9]	1 3.8 [8]	1 4.1[6]	1 4.2 [8]
	8	12.2 [20]	1 2.7 [9]	12.2 [15]	1 2.1 [11]	1 2.7 [13]	1 2.6 [4]	1 2.9 [16]
	12	11.4 [6]	1 2.3 [8]	12.1 [8]	1 2.1 [10]	1 2.1 [11]	1 2.0 [11]	1 2.6 [7]
	13	13.4 [10]	1 4.8 [2]	14.2 [9]	1 4.0 [8]	1 4.2 [3]	1 3.3 [8]	1 4.6 [9]
	15	12.6 [4]	1 3.3 [6]	13.1 [5]	1 3.8 [9]	1 3.0 [8]	1 3.3 [8]	1 3.3 [8]
	17	12.1 [6]	1 2.7 [5]	13.8 [7]	1 2.0 [8]	1 3.6 [5]	1 2.8 [7]	1 2.8 [9]
Kidneys (mg)								
Lab	5	2035 [5]	2 069 [10]	2131 [9]	2 146 [6]	2 118 [6]	2 058 [11]	2 050 [5]
	8	1930 [12]	2 024 [6]	2000 [12]	1 953 [7]	2 003 [10]	1 974 [5]	2 026 [8]
	12	2010 [3]	2	1999 [7]	2	2	2	2

		115 [5]		035 [6]	061 [11]	073 [10]	073 [7]
15	1948 [5]	2064 [5]	1982 [6]	2003 [3]	2096 [4]	1888 [3]	1971 [11]
17	2008 [6]	1983 [5]	2121 [8]	2028 [7]	2150 [10]	1983 [9]	1952 [9]
Adrenals (mg)							
Lab 5	54.0 [10]	5 0.8 [13]	56.8 [17]	5 8.9 [14]	5 7.4 [14]	5 3.2 [10]	6 3.3 [11]
12	57.6 [16]	4 7.0 [14]	51.5 [14]	5 0.5 [13]	50.1 [8]	5 2.4 [22]	5 3.1 [12]
15	54.3 [9]	4 6.2 [6]	54.2 [15]	5 0.7 [11]	5 0.6 [10]	4 7.6 [11]	5 5.1 [12]
17	57.5 [22]	4 9.3 [14]	51.2 [11]	4 9.8 [7]	5 6.8 [9]	4 8.5 [20]	5 6.2 [16]

weight [CV]

Table S. Effects of 0.4 mg testosterone propionate and flutamide administration on mean liver, kidney, and adrenal weights

B.	Control (no TP)	mg Flutamide/kg/day + 0.4 mg Testosterone Propionate/kg/day					
	0	**0**	**0.1**	**0.3**	**1.0**	**3.0**	**10.0**
Liver (gms)							
Lab 5	13 [9]	14.8 [9]	1 4.9 [9]	1 4.6 [10]	1 4.0 [11]	1 4.2 [4]	1 3.8 [10]
12	11.2 [7]	1 2.6 [7]	1 2.4 [5]	1 2.5 [9]	1 1.6 [8]	1 2.3 [10]	1 2.2 [10]
13	13.7 [7]	1 5.7 [5]	1 5.1 [14]	1 4.9 [7]	1 4.3 [12]	1 4.7 [7]	1 4.6 [7]
15	12.6 [4]	1 3.6 [7]	1 3.5 [6]	1 4.0 [10]	1 3.5 [10]	1 3.3 [7]	1 3.6 [7]
Kidneys (mg)							
Lab 5	2035 [5]	2 192 [5]	2 294 [4]	2 203 [10]	2 131 [6]	2 108 [3]	2 161 [5]
12	2011 [3]	2 151 [8]	2 129 [8]	2 192 [8]	2 109 [10]	2 096 [5]	2 091 [7]
15	1948 [5]	2 183 [6]	2 092 [4]	2 015 [8]	1 998 [5]	2 039 [4]	2 067 [6]
Adrenals (mg)							
Lab 5	54.0 [14]	4 6.3 [14]	4 7.9 [15]	5 4.3 [14]	5 7.4 [13]	5 2.0 [14]	5 6.3 [19]
12	57.6 [16]	4 4.3 [31]	4 6.3 [19]	5 0.6 [5]	4 9.8 [16]	5 2.8 [14]	5 3.5 [19]
15	54.3 [9]	4 8.6 [14]	5 0.3 [9]	5 1.3 [9]	5 0.4 [10]	4 8.5 [11]	5 2.1 [8]

Mean tissue weights [coefficient of variation]

* significantly ($p<0.05$) decreased from untreated control (no TP)

** significantly increased ($P<0.05$) from 0 mg flutamide.

ANNEX – 7

LEAD LABORATORY SUMMARY REPORT OF INITIAL WORK TOWARDS THE VALIDATION OF THE RAT HERSHBERGER ASSAY: PHASE-1B, RESPONSE EFFECTS OF THE ANTI-ANDROGENIC SUBSTANCE FLUTAMIDE

Lead Laboratory SummaryReport on the OECD Interlaboratory Study on the Hershberger Assay: Phase IB - Dose Response Effects of the Antiandrogenic Drug Flutamide (oral)

Prepared by Leon Earl Gray Jr., Endocrinology Branch, Reproductive Toxicology Division, National Health and Environmental Effects Research Laboratory, US EPA, NC 27711

Reviewed by Mike Walker and William Owens
Their comments are incorporated

INTRODUCTION

1. In the summer of 2001, seven laboratories participated in an interlaboratory investigation of the Hershberger assay using a protocol developed by the OECD in 2000. This report presents the lead laboratory's summary of the results of the data analyses from these seven laboratories.

PRIMARY STUDY: THE OECD PHASE IB INTERLABORATORY STUDY: FLUTAMIDE DOSE RESPONSE (6 DOSAGE LEVELS) ANTAGONISM OF THE ANDROGENIC EFFECTS OF SC TESTOSTERONE PROPIONATE (TP) (2 DOSAGE LEVELS)

2. In this investigation, each laboratory examined the effects of graded doses of orally administered (gavage) flutamide (0, 0.1, 0.3, 1.0, 3.0 and 10 mg/kg/d) for ten consecutive days on androgen-dependent organ weights in the immature castrated-TP-treated male rat. Sample sizes were designed as six rats per group. Each laboratory included castrate-controls that did not receive TP. Flutamide is a potent antiandrogenic drug, which should inhibit the growth-promoting effects of TP on the seminal vesicles, ventral prostate, Cowper's glands, levator ani plus bulbocavernosus muscles and glans penis, which are all androgen-dependent.

3. Four of seven laboratories conducted two flutamide dose response studies using 0.2 and 0.4 mg TP/kg/d (labs 5, 12, 13, 15). The fifth and sixth labs competed flutamide against TP at 0.2 mg/kg/d (labs 8, 17) while the seventh lab used flutamide against 0.4 mg TP/kg/d (lab 10). As only one laboratory measured serum T and LH and only one laboratory conducted separate studies to evaluate the effects of tissue fixation on organ weights other than the ventral prostate, these data have not been analysed as there is no "interlaboratory component."

KEY OBSERVATIONS OR RESULTS

4. Flutamide was completely effective in blocking the stimulatory effects of TP in every one of the 11 studies analysed here.

5. One objective of the current study was to determine if either 0.2 or 0.4 mg TP/kg/d was more sensitive to the antagonistic effects of flutamide at low doses, as it is likely that most antiandrogenic toxicants will be much less potent on a mg/kg/d basis than flutamide. For three of five androgen-dependent tissues (VP, LABC and COWS), the effects of flutamide at 0.1 mg/kg/d were only statistically significant in the pooled analysis of the 0.4 mg TP dose group data. For the SV, the two TP dose groups were of equivalent utility. For the GLANS, the effect of flutamide at low doses was more evident in the 0.2 mg TP dose group. In all cases, the F-value for the flutamide effect in the pooled analysis was larger in the 0.4 mg TP/kg/d than the F-value in the 0.2 mg TP/kg/d group, indicating greater "significance" was achieved.

6. LOELs for the effect of flutamide varied among the seven labs by approximately:
- 3 fold for the LABC and Cowper's glands
- 10 fold for the VP and SV
- 30 fold for the GLANS

7. The order of sensitivity to flutamide, which is antagonising TP is SV > VP, LABC, and COWS > GLANS (sensitivity being defined here as the organ that shows the greatest change on a percent basis at a low dose). This is the opposite of the sensitivity to low doses of TP seen in the first OECD interlaboratory Hershberger study.

8. Very small, often non-significant lab-by-flutamide interactions were seen; the dose response curves being fairly parallel.

GENERAL METHODS

9. As in the first OECD interlaboratory study, a range in age at castration and the initiation of treatment were allowed and efforts were not made to standardise the strain of rat used or the diet among the laboratories. Each laboratory submitted their data in spreadsheet format, which was converted to text files and analysed.

10. All laboratories measured "fresh" weights of the ventral prostate, seminal vesicle (plus coagulating glands, presumably with their fluids), glans penis, Cowper's gland and levator ani plus bulbocavernosus muscle weights. Liver weights were also measured by all laboratories.

11. Several laboratories weighed the ventral prostate after being stored for 24h in fixative. "Optional" organ weights also were measured in some studies. Four laboratories weighed the adrenal glands (labs 5, 11, 17, 15) and five weighed the kidneys (labs 5, 8, 11, 15, 17).

AN ADDITIONAL COMPARATIVE STUDY

12. One laboratory executed a separate study in parallel to the primary study to determine how altering the method of tissue dissection and weighing affected the weights of the ventral prostate, seminal vesicles and Cowper's glands. In these studies, the sex accessory glands were dissected as a unit, not weighed fresh as above and the glands were preserved in fixative for 24 hours after which they were separated and weighed.

DATA ANALYSIS

13. Means, standard errors and coefficients of variation (the standard deviation divided by the mean, as an estimate of relative variability) were calculated for each endpoint using PROC MEANS in SAS (Version 6.08, available on the USEPA IBM Mainframe). Means and the coefficients of variation (CVs) for the ventral prostate (VP), seminal vesicles plus coagulating glands with fluids (SV), levator ani plus bulbocavernosus muscles (LABC), Cowper's glands (COWS), and glans penis (GLANS) data are presented in the attached Tables. These values were derived from the ANOVA tables of the untransformed data so the variation due to dose is not included in the variance used to calculate the COV. Examination of the COV among endpoints allows one to compare the statistical precision in the weight of a tissue among dose groups and laboratories.

14. The fact that CVs for each androgen-dependent organ weight are proportional to the means across the different dosage levels of flutamide indicates that heterogeneity of variance exists. Hence, the data were transformed using Log10 for statistical analysis. This transformation provides for a more valid

comparison of the effects of flutamide on organ weights at lower dosage levels.

15. For each dose of TP (0.2 and 0.4), two-way ANOVAS (dose of flutamide and laboratory as fixed main effects) were executed using PROC GLM for each tissue. Data also were analysed by one-way ANOVA on PROC GLM for each laboratory (flutamide dose as a main effect) with and without necropsy body weight as a covariate. The attached tables include the F and R2 values for each laboratory and pooled over all the laboratories, which provide indices of how strongly flutamide antagonised the action of TP. We expected flutamide to antagonise the stimulatory effect of TP on all five androgen-dependent tissues. The LSMEANS procedure was used (two-tailed t-tests, appropriate for a priori hypotheses) to determine lowest-observed-effect-levels (LOELs) and to compare effects at different dosage levels of flutamide to one another. Shaded values on the tables differ significantly from the castrate plus TP control by $p < 0.05$ using the log10-transformed data. In regards to the GLANS data, which display more uniform variances across the dose groups, analysis of the untransformed data provides exactly the same LOELS for each study, so the log transformed data are presented for consistency with the other endpoints.

16. Data for the five androgen-dependent sex accessory tissues (SV, VP, GLANS, COWS, and LABC) were "normalised" (see Figure 3 for an example of how this was done) in order to visually compare the shapes of the dose-response curves for each lab such that the data range from 0 to 100%. In this normalisation, the castrate-no TP value was used to establish a 0 % level, while the response seen in the castrate-TP without flutamide dose group was set as 100%. For comparative purposes, “normalised” TP dose response data from the previous OECD study also are presented.

17. One objective of the first investigation was to determine an approximate ED70-value for TP to be used in subsequent phases of the OECD Hershberger assay standardisation and validation exercise. As seen in Figure 1, the visually estimated ED70 values range from about 0.2 for the GLANS to 0.8 mg TP/kg/d for the SV. It appears from this figure that three distinct dose-response relationships exist for the five endpoints. The glans penis reaches a maximum response at a lower dosage level than the other four tissues, while, in contrast, the SV shows a more gradual and more linear response over the dose range used in the current studies. The LABC, COWS and VP appear to respond to the same degree to different doses of TP and are intermediate between the GLANS and SV curves with an ED70 of about 0.4 mg/kg/d. These normalised values were not analysed statistically.

RESULTS

"Normalised" dose-response curves for the five androgen-dependent tissues

18. Figures 1 (relative organ weight versus dose of TP from our previous study) and 2 (relative organ weights versus dose of flutamide) display the dose-response curves for the five tissues, pooled across all the labs and normalised such that the values for each organ range from 0 to 100%. Not surprisingly, the order of sensitivity to flutamide, which is antagonising TP (SV > VP, LABC, and COWS > GLANS; sensitivity being defined here as the organ that shows the greatest change on a percent basis at a low dose) is the opposite of the sensitivity to low doses of TP seen in the first OECD interlaboratory Hershberger study (GLANS > VP, LABC, and COWS> SV). Figure 3, shows how the relative values were “normalised” for the LABC as an example.

Effect of Flutamide on Ventral Prostate Weight in the OECD Phase IB Interlaboratory Study

19. The ventral prostate (fresh) weight data from each lab over the 6 flutamide dosage groups at each level of TP (0.2 and 0.4 mg/kg/d) are shown in Figures 4-6. Figure 4 displays the pooled mean values and standard errors (SE) (from PROC MEANS, with an SE, unadjusted for lab to lab variability). These pooled means, and the CVs also are shown at the bottom of Table 1. In the 0.2 mg and 0.4 TP/kg/d groups,

the overall CVs for VP weight was 24% and 21%, respectively, ranging from a low of 15% in one lab to a maximum of 36% in another.

20. In the pooled analysis of the VP data, the 0.1 flutamide dose differed from control only at the 0.4 mg TP/kg/d group. VP weight was significantly reduced in the 0.2 mg TP group at 0.3 mg flutamide/kg/d (Figure 4). When analysed by lab, flutamide inhibited the effect of TP at 0.2 and 0.4 mg/kg/d significantly ($p < 0.001$) in every experiment (11 individual experiments) (Figure 5). At 10 mg/kg/d, flutamide almost completely antagonised the action of TP. In every lab, the effects of flutamide on VP weight were significant at 1 mg/kg/d and above. At 0.2 mg TP/kg/d the LOELs for the six labs were 0.1, 0.3, 0.3, 1.0, 1.0 and 1.0 mg/kg/d. At 0.4 mg TP/kg/d the LOELS were similar being 0.1, 0.1, 1.0, 1.0, 1.0 mg flutamide/kg/d. Taken together, the results of the pooled analysis indicated that the low dose effect of flutamide only was apparent in the higher TP dose group. However, this advantage was not apparent on a lab-by-lab basis as the LOELS varied similarly in each TP dose group. As expected, the VPs are almost twice as large in the 0.4 mg/kg/d TP group (without flutamide) as compared to the 0.2 mg/kg/d group.

21. The lab-to-lab variability, termed a lab effect, is highly significant (Figure 5), but explained considerably less variance than did the flutamide effects (Table 1). The lab-by-dose interactions were not statistically significant, indicating that the dose response curves were relatively parallel (Figure 6).

22. Body weight at necropsy was not consistently a significant factor in any laboratory for any androgen-dependent tissue (being significant in only six of 55 ANOVAs). For this reason, it was not included further in the discussion of the data from the individual laboratories. The lack of relationship between VP and body size seen in this study results in part from the design of the study and this conclusion cannot be extended to some other protocol and should not be taken to indicate that controlling body weight is not important. Although the statistical analyses and effects here can be generalised to the SV, LABC, COWS and GLANS in this study, body weight often co-varies significantly with liver, kidney and adrenal weights. In the pooled analysis, body weight was ~~a~~ significantly related to VP weight (at 0.2 TP) and LABC weights (in both TP groups).

Effect of Flutamide on Seminal Vesicle (with coagulating glands and fluid) Weight in the OECD Phase IB Interlaboratory Study

23. The seminal vesicle (SV) (fresh) weight data from labs over the 6 flutamide dosage groups at each level of TP (0.2 and 0.4 mg/kg/d) are shown in Table 2 and Figures 7-9. Figure 7 displays the pooled mean values and standard errors (SE) (from PROC MEANS, with an SE, unadjusted for lab to lab variability). These means and the CVs are shown at the bottom of Table 2. In the 0.2 mg and 0.4 TP/kg/d groups, the overall CVs for VP weight was 29%, and 22%, respectively, ranging from a low of 14% in one lab to a maximum of 34% in another. These values are similar to those seen for VP weight, even though the SV is much larger than the VP. In general, as the CVs increased the F- and R2-values for the effect of TP declined.

24. In the overall analysis of the SV data, the lowest dose used, 0.1 mg flutamide/kg/d differed from control at each level of TP (Figure 7). For each lab, flutamide inhibited the effect of TP at 0.2 and 0.4 mg/kg/d significantly ($p < 0.0001$). At 10 mg/kg/d, flutamide almost completely antagonised the action of TP. In every lab the effects of flutamide on SV weight was significant at 1 mg flutamide/kg/d and above. The LOELs were not consistently lower in one TP dose group than the other. Taken together, these results indicate that one dose of TP does not have a significant advantage over the other for detecting low dose effects of an antiandrogen on this endpoint. As expected, the SVs are as much larger in the 0.4 mg/kg/d TP group (without flutamide) than they are at 0.2 mg TP /kg/d.

25. The lab-to-lab variability (Figure 8), termed a lab effect, is highly significant, but explained

considerably less variance than did the flutamide effects (Table 2). The lab-by-dose interactions were relatively small but still statistically significant. As seen in Figure 9, the dose-response curves for flutamide are quite similar from lab to lab.

Effect of Flutamide on Levator ani plus bulbocavernosus muscle (LABC) Weight in the OECD Phase IB Interlaboratory Study.

26. The LABC weight data over the 6 flutamide dosage groups at each level of TP (0.2 and 0.4 mg/kg/d) are shown in Table 3 and Figures 10-12. Figure 10 displays the mean values and standard errors (SE) (from PROC MEANS, with an SE, unadjusted for lab to lab variability). These pooled means, and the CVs are shown at the bottom of Table 3. In the 0.2 mg and 0.4 TP/kg/d groups, the pooled COV for LABC weight was 13%, and 10%, respectively, ranging from a low of 8% in one lab to a maximum of 16%. As is normally the case, the COV values for the LABC are lower than the CVs seen for the SV and VP weights.

27. In the pooled analysis of the LABC data, the 0.1 flutamide dose differed control in the 0.4 mg TP/kg/d dose group, while the LOEL in the 0.2 mg TP/kg/d group was 0.3 mg flutamide/kg/d (Figure 10). This resulted from the fact that the difference between the low dose flutamide means was more than twice as high in the 0.4 TP dose group versus the 0.2 TP dose group, but the CVs were equal. For each lab, flutamide inhibited the effect of TP at 0.2 and 0.4 mg/kg/d significantly ($p < 0.0001$). At 10 mg/kg/d, flutamide almost completely antagonised the action of TP. At 0.2 mg/kg/d the LOELs for the six labs were 0.1, 0.1, 0.3, 0.3, 0.3, and 1 mg/kg/d. At 0.4 mg TP/kg/d the LOELS were 0.1, 0.1, 0.3, 1.0, 1.0 mg flutamide/kg/d. Taken together, these results indicate that the high dose of TP had an advantage over the low dose for detecting low dose effects of flutamide on LABC weight. However, on a lab-by-lab basis the LOELS were not consistently lower in one TP group than the other. As expected, the LABCs are much larger in the 0.4 mg/kg/d TP group (without flutamide) than they are at 0.2 mg/kg/d.

28. The lab-to-lab variability, termed a lab effect, is highly significant (Figure 11), but explained considerably less variance than did the flutamide effects (Table 3). The lab-by-dose interactions were not statistically significant, indicating that the flutamide dose-response curves were parallel among the labs (Figure 12).

Effect of Flutamide on paired Cowper's Gland (COWS) Weights in the OECD Phase IB Interlaboratory Study

29. The COWS weight data over the 6 flutamide dosage groups at each level of TP (0.2 and 0.4 mg/kg/d) are shown in Table 4 and Figures 13-15. Figure 13 displays the mean values and standard errors (SE) (from PROC MEANS, with an SE, unadjusted for lab to lab variability). These means and the CVs are shown at the bottom of Table 4. In the 0.2 mg and 0.4 TP/kg/d groups, the overall CVs for LABC weight were 23%, and 17%, respectively, ranging from a low of 12% to a maximum of 26%. In some of the studies, the COV increased as the organ weights become smaller (Table 4), likely due to the difficulty of dissection of these glands when they are almost fully regressed.

30. In the lab-by-lab analysis, in 10 of 11 studies the 1.0 mg/kg/d flutamide dose differed from control at each level of TP. Flutamide inhibited the effect of TP at 0.2 and 0.4 mg/kg/d significantly ($p < 0.001$). At 10 mg/kg/d, flutamide completely antagonised the action of TP. In the overall analysis of the COWS weight data, the effect of flutamide pooled over all the labs was significant at 0.3 and 0.1 mg flutamide/kg/d in the 0.2 and 0.4 mg TP/kg/d groups, respectively (Table 4). However, on a lab-by-lab basis, the LOELs for the effects of flutamide at 0.2 and 0.4 mg TP were quite similar within each lab, but variable between labs (Figure 14). Taken together, these results indicate that the dose of 0.4 mg TP had a slight advantage over the lower dose of TP.

31. The lab-to-lab variability, termed a lab effect, is highly significant, but explained considerably less variance than did the flutamide effects (Table 4). The lab-by-dose interactions was not statistically significant in the 0.2 mg TP group but was statistically significant in the 0.4 mg TP group (Table 4). As shown in Figure 15 the dose-response curves are of similar shape in most cases.

Effect of Flutamide on Glans Penis (GLANS) Weight in the OECD Phase IB Interlaboratory Study

32. The GLANS weight data over the 6 flutamide dosage groups at each level of TP (0.2 and 0.4 mg/kg/d) are shown in Table 5 and Figures 16-18. Figure 16 displays the mean values and standard errors (SE) (from PROC MEANS, with an SE unadjusted for lab to lab variability). These means and the CVs are shown at the bottom of Table 5. In the 0.2 mg and 0.4 TP/kg/d groups, the pooled CVs for GLANS weight were 11%, and 9%, respectively, ranging from a low of 5% to a maximum of 13%. CVs for this organ and the LABC are smaller than the VP, SV and COWs.

33. In the pooled analysis of the GLANS data, the 3.0 mg flutamide/kg/d dose was always different from control at each level of TP (Figure 16). Flutamide inhibited the effect of TP at 0.2 and 0.4 mg/kg/d significantly ($p < 0.0001$) (Table 5). At 10 mg/kg/d, flutamide completely antagonised the action of TP. For this androgen-dependent organ, the effect of flutamide pooled over all the labs was significant at 0.1 and 0.3 mg flutamide/kg/d in the 0.2 and 0.4 mg TP/kg/d groups, respectively (the reverse of the VP, LABC and COWS data above). In addition 2/4 labs which ran both TP dose groups had lower LOELs at 0.2 mg TP than at 0.4 while the other 2 labs had equivalent LOELs for flutamide in the 0.2 and 0.4 mg TP groups. The LOELs were quite variable from lab to lab ranging from 0.1 to 3 mg flutamide/kg/d. Taken together, these results indicate that the dose of 0.2 mg TP had a slight advantage over the higher dose of TP.

34. The lab-to-lab variability, termed a lab effect, is highly significant, having an F value (relative to the F-value for flutamide) larger than that seen in the other androgen-dependent tissues (Figure 17). This suggests that although the dissections are fairly precise within each lab (the CVs being fairly low), the labs are executing the dissections slightly differently. The lab-by-dose interaction was not statistically significant in the 0.2 mg TP group but was statistically significant in the 0.4 mg TP group (Table 5, Figure 18).

Effect of preservation of the ventral prostate in fixative for 24 hours

35. Several of the labs weighed the ventral prostate after fixation as well as recording the fresh weight (4 labs at TP 0.2 and 3 labs at 0.4). Fixation of the VP significantly increased the weight of this tissue (weighed fresh prior to fixation) with larger tissues gaining more weight than smaller ones. Hence, the change in weight (vpdelta =vpfixed-vpfresh) was significantly affected by the dose of flutamide, e.g., at TP 0.2, the F-Flutamide (5,118 df) = 13.1, $p < 0.001$, while, for TP 0.4, the F-flutamide (5,189 df) = 17.3, $p < 0.001$. The lab and lab-by-flutamide effects also were significant in both TP groups). As the dose of flutamide increased, fresh VP weight decreased, resulting in smaller weight increases after fixation (Figure 19). With these effects in mind, however, it is important to note that statistical analyses of the fixed ventral prostate weight data yields the same results as analysis of the fresh tissue weights (individual lab data not shown. Overall data are in Figure 20.

2.

Effect of flutamide on body weight and body weight gain in the OECD Phase I Interlaboratory study

36. Body weights of the animals on study varied greatly from lab to lab. One source of variation in the size of the animals is obviously related to the age at which they assigned to treatments and necropsied. It must be emphasised that the fact that these rats appeared to differ in terms of weight for age, and growth

during the study, did not compromise the robust nature of the responses of the tissues to flutamide in these animals. Body weight was not significantly affected by flutamide treatment, but body weight gain appeared slightly reduced by flutamide treatment in both TP dose groups (Figure 21).

Effect of flutamide on nonreproductive organs measured by some or all laboratories in the OECD Phase I Interlaboratory study

37. All tissues contain androgen-receptors and, to some degree, respond to TP-treatment. Hence, many tissues may be affected by flutamide's antagonism of TP or from other effects of flutamide. It is known that flutamide treatment at high doses increases liver and adrenal weights. In the current study, flutamide treatment did not significantly affect kidney or liver weights (data not shown). Unlike the results of the statistical analysis of the relationship between body and reproductive organ weights, body and liver weights are highly correlated with one another.

38. Flutamide-treatment caused a dose-related increase in adrenal weights (Figure 22). Some of these effects were not significant in all labs when tested individually because they are much less robust than the effects seen in the sex accessory tissues, described above. In the TP 0.2 mg/kg/d dose group, flutamide significantly increased adrenal weights to castrate-no TP size, being significant at all dosage levels ($p < 0.05$) except 3 mg flutamide/kg/d. In the 0.4 mg TP dose group, flutamide also slightly increased adrenal weights, being statistically significant at dosage levels of 0.3 mg flutamide and above. (TP 0.2; F-flutamide (5,118 df) = 4.8, $p<0.005$) F-Lab (3,118 df) = 6.4, $p < 0.005$. While at TP 0.4; F-Flutamide (5,189) = 3.37, $p < 0.01$, F-Lab (2,893 df) = 1.8. $p > 0.15$. In both TP groups, the Lab*Flut interaction was not significant).

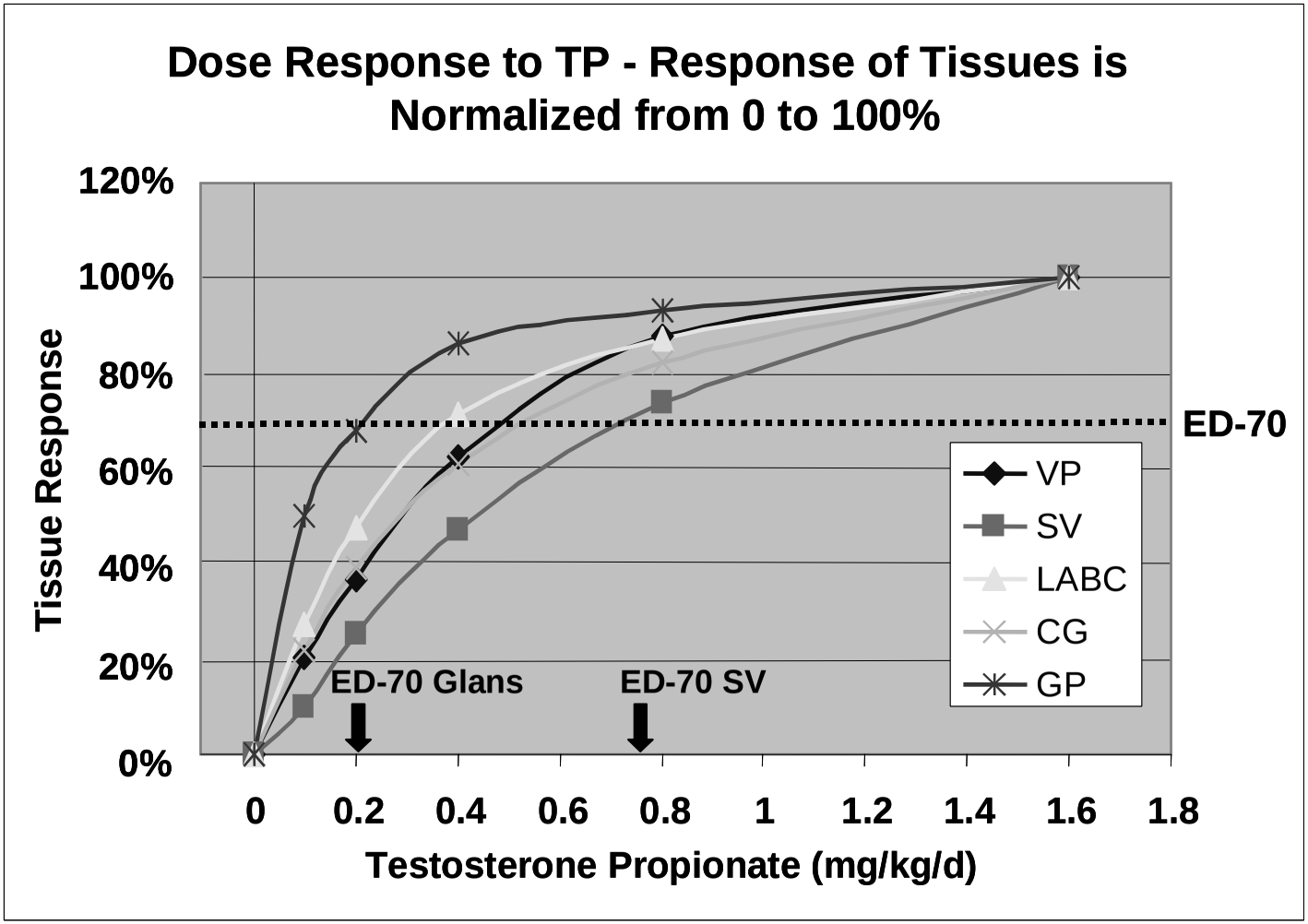
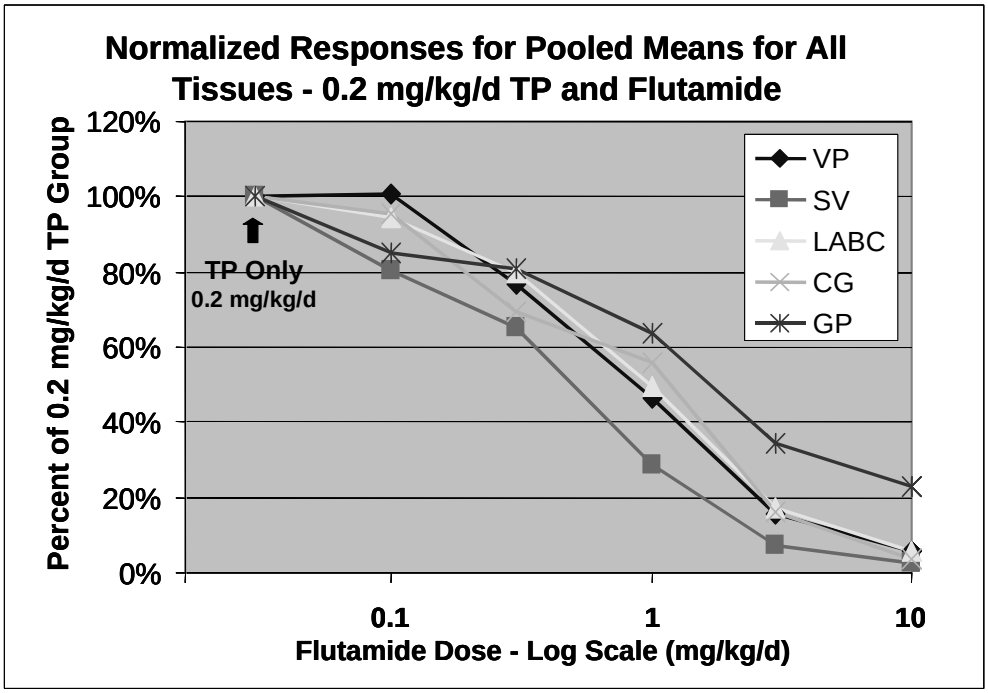


Figure 1. The pooled data across 16 laboratories for each tissue have been normalized so that the control value = 0% and the maximum reached at 1.6 mg/kg/d testosterone propionate (TP) = 100%. The differences between tissues in how rapidly a tissue reaches its maximum is then illustrated, e.g., glans penis first and seminal vesicles last. VP = ventral prostate; SV = seminal vesicles and coagulating glands; LABC = levator ani/bulbocavernosus muscle complex; CG = Cowper's glands; GP = glans penis.



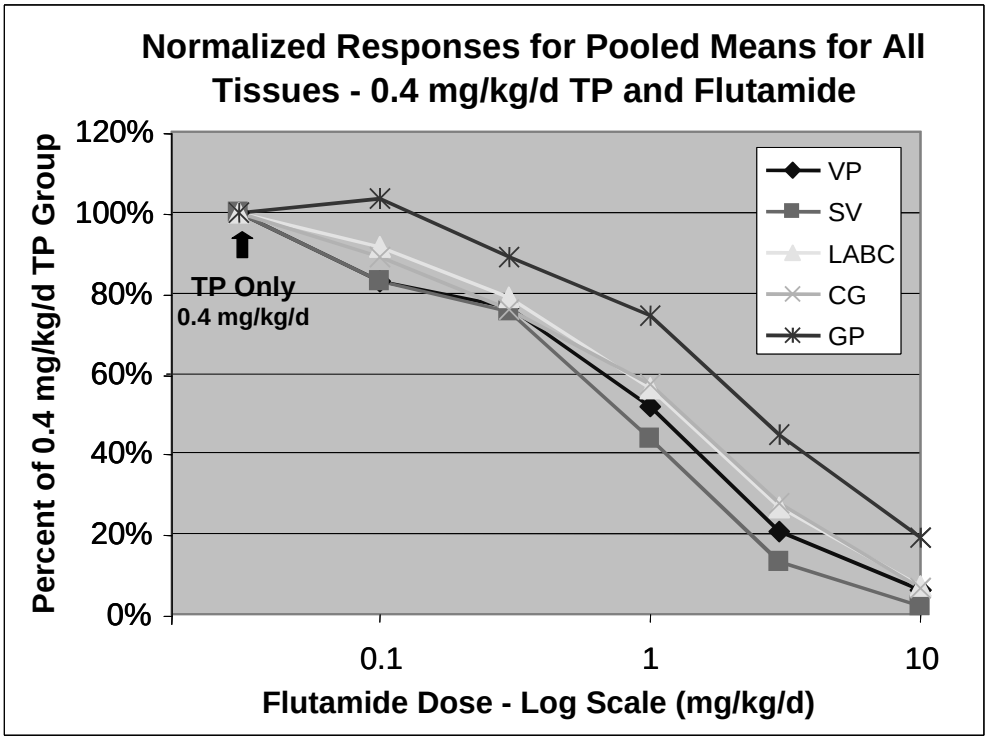


Figure 2. The pooled data across the laboratories for each tissue have been normalized so that the control value (or TP dose group) = 100%. The relative decrease in the tissue weights are then illustrated, e.g., seminal vesicles first and glans penis last. Note that a logarithmic scale is used for the x-axis. The top graph is the 0.2 mg/kg/d TP studies and the bottom graph is the 0.4 mg/kg/d studies. VP = ventral prostate; SV = seminal vesicles and coagulating glands; LABC = levator ani/bulbocavernosus muscle complex; CG = Cowper's glands; GP = glans penis.

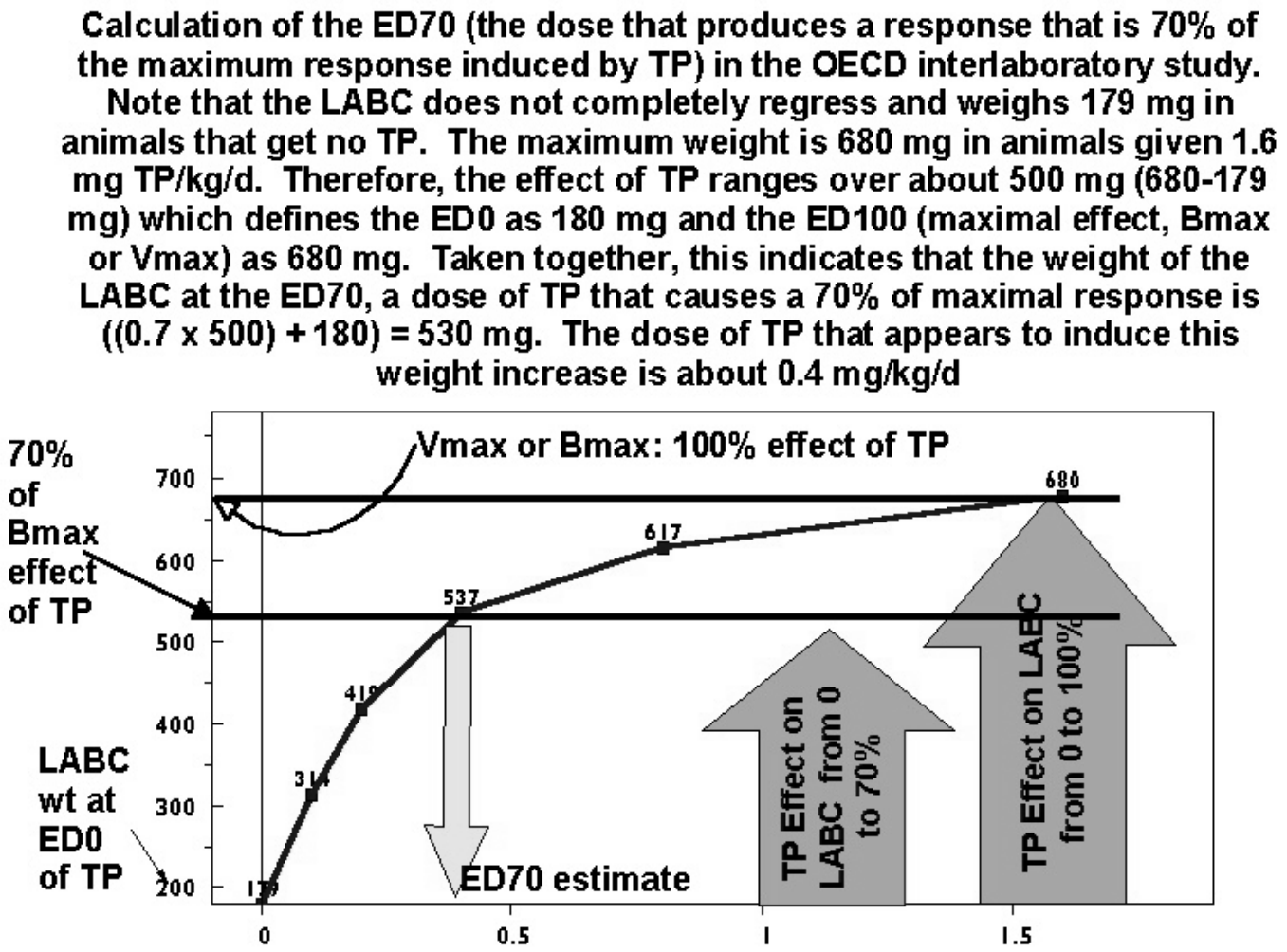


Figure 3. The x-axis is the testosterone propionate dose in mg/kg/d administered subcutaneously. The y-axis is the weight of the levator ani/bulbocavernosus muscle complex. This curve and markings demonstrate how an ED70 would be estimated from the data as 70% of the maximum effect as the dose response curve plateaus.

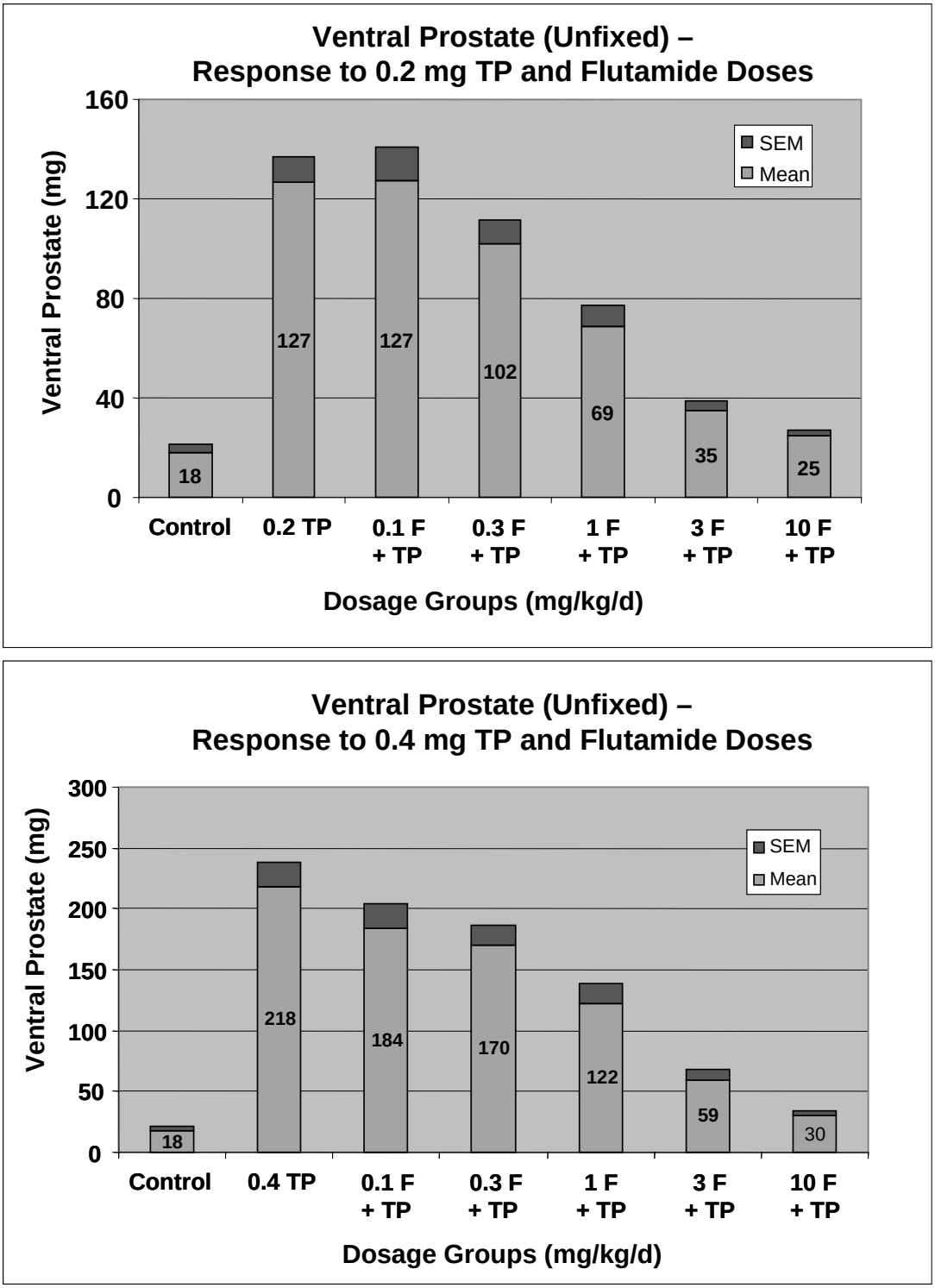
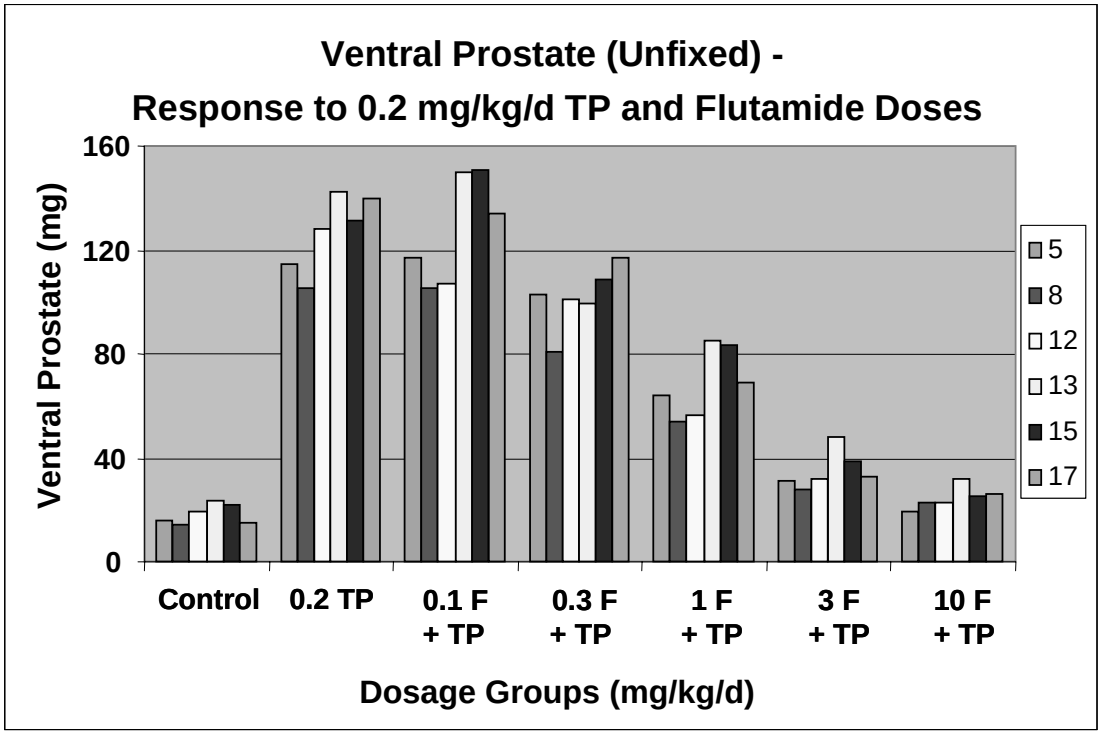


Figure 4. The results of TP and Flutamide administration on the ventral prostate. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.



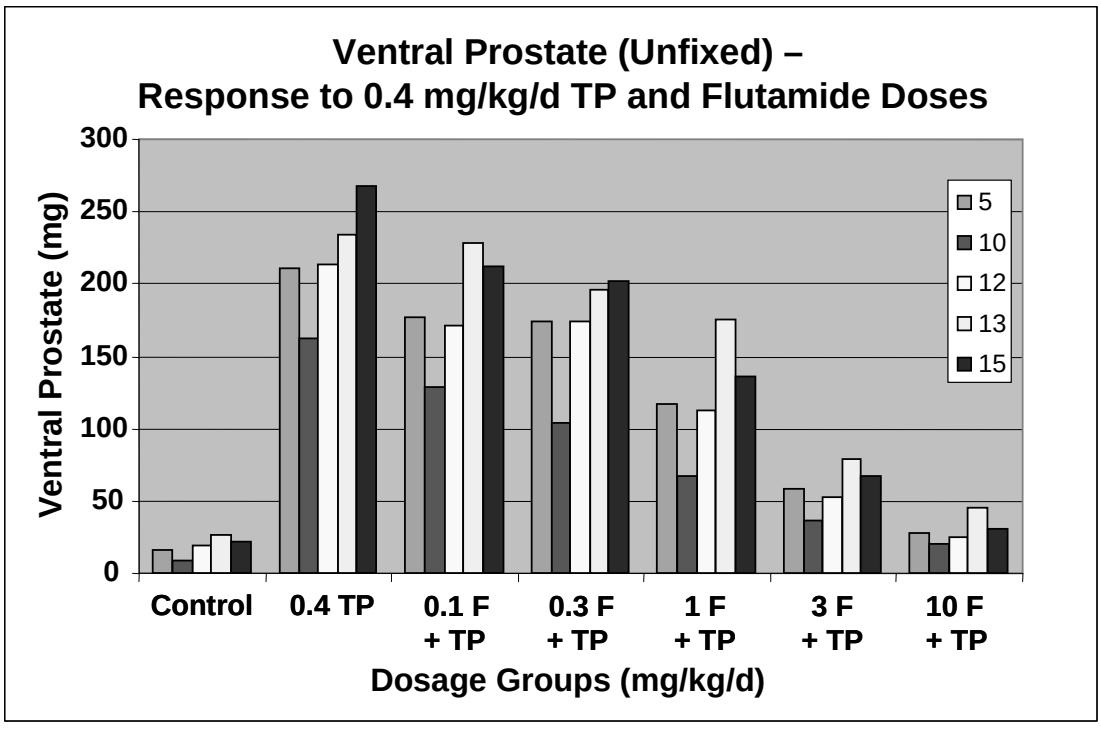
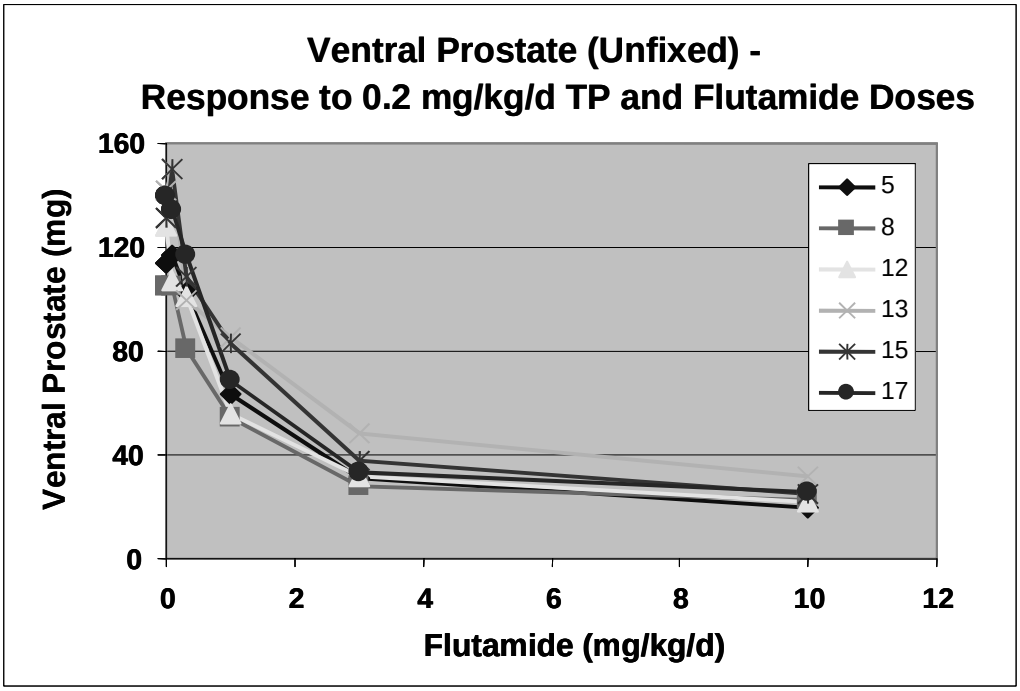


Figure 5. The results of TP and Flutamide administration on the ventral prostate. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are the individual means of the laboratories plotted by the control and substance administration groups.



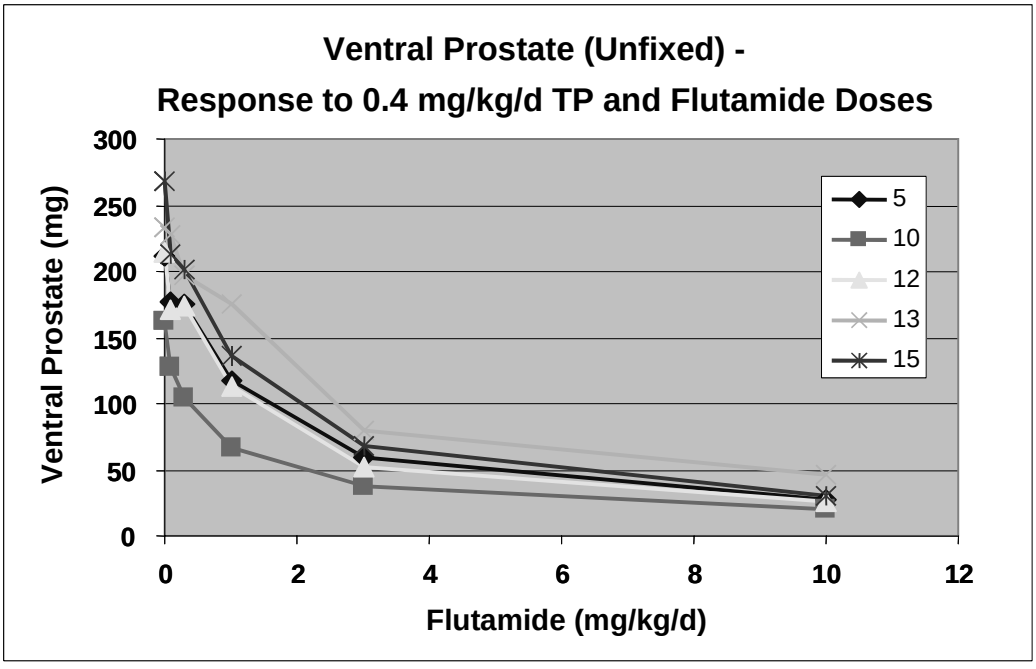
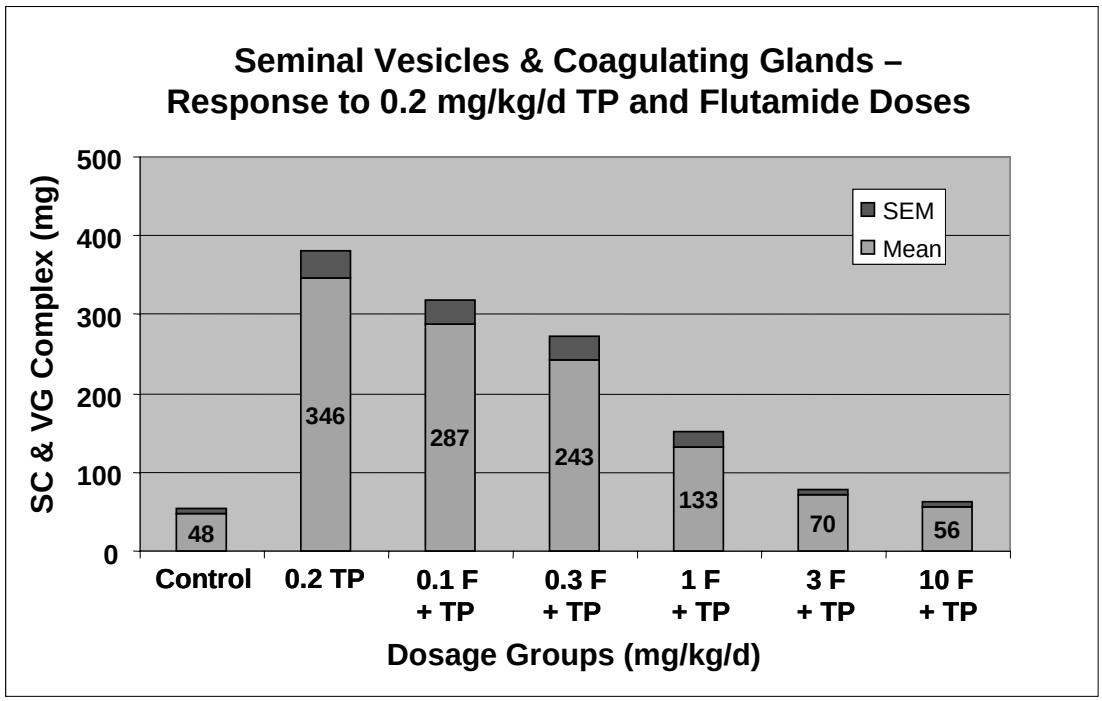


Figure 6. The results of TP and Flutamide administration on the ventral prostate. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are absolute weights of the pooled means tissue on the y-axis, and Flutamide dose on the x-axis.



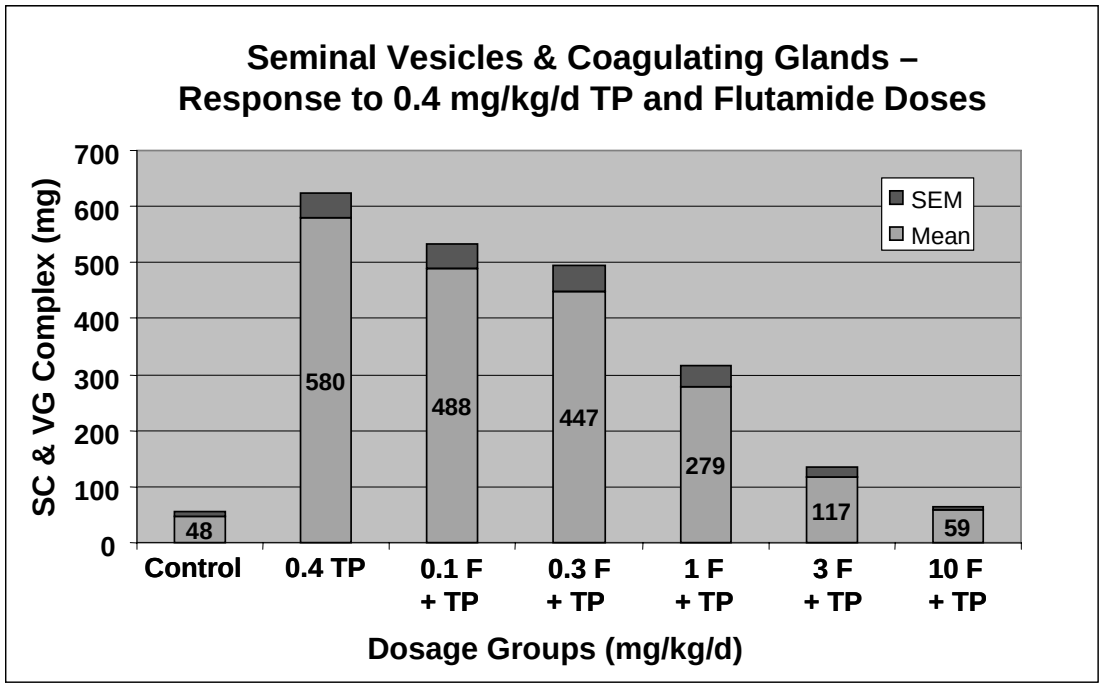
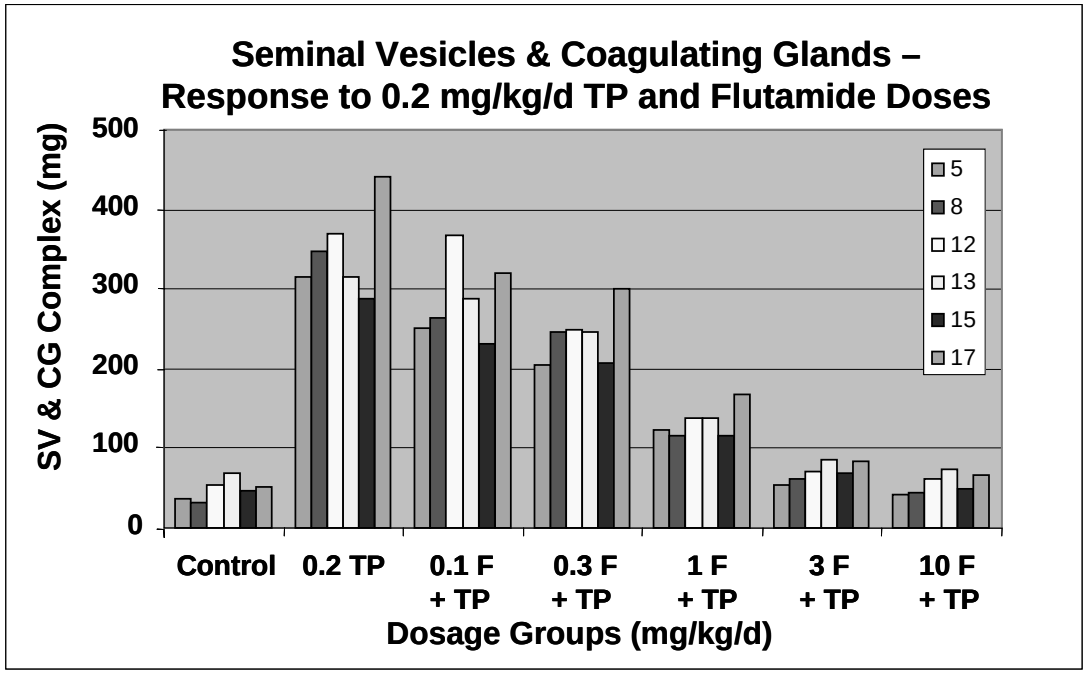


Figure 7. The results of TP and Flutamide administration on seminal vesicles and coagulating glands. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.



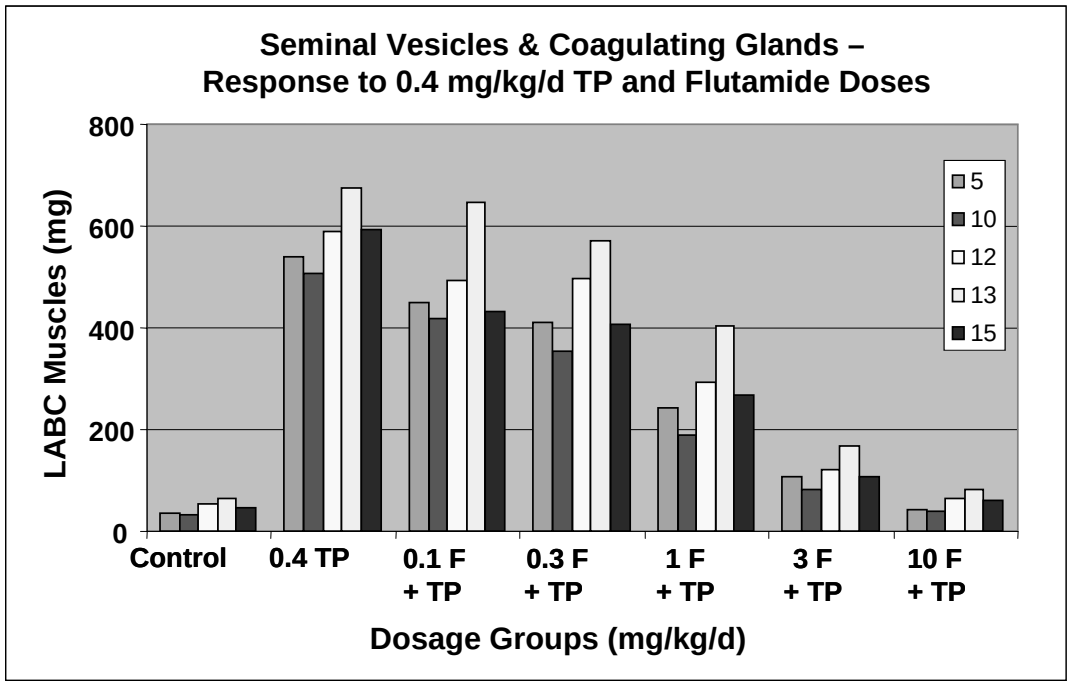
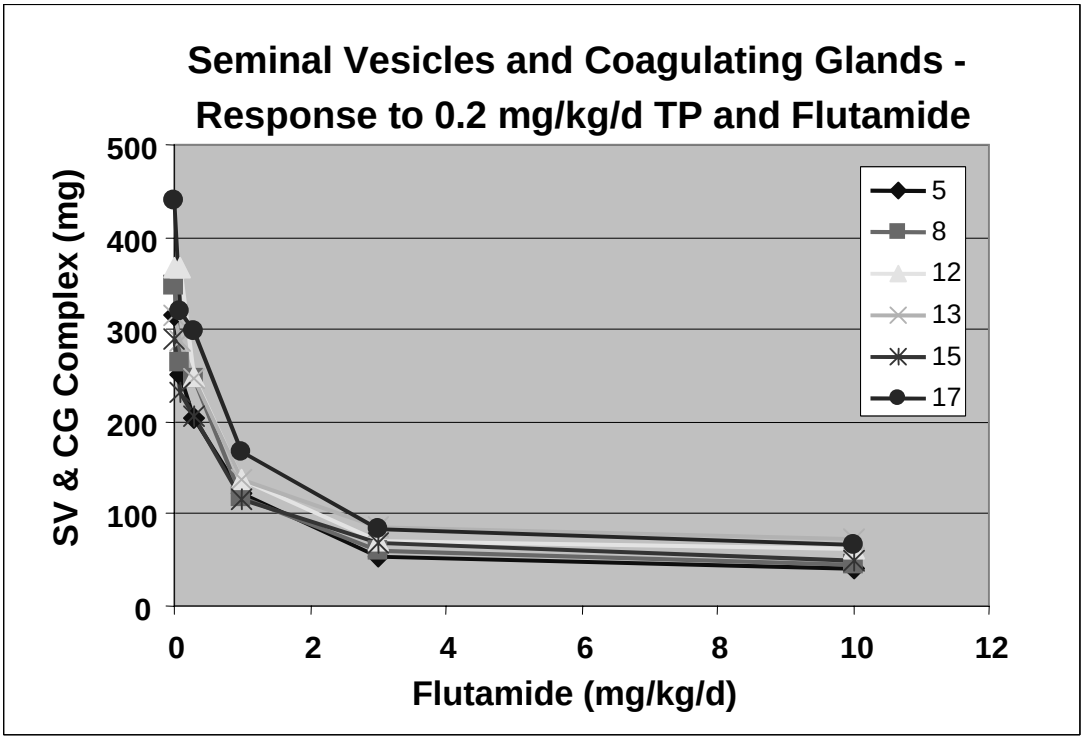


Figure 8. The results of TP and Flutamide administration on seminal vesicles and coagulating glands. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are the individual means of the laboratories plotted by the control and substance administration groups.



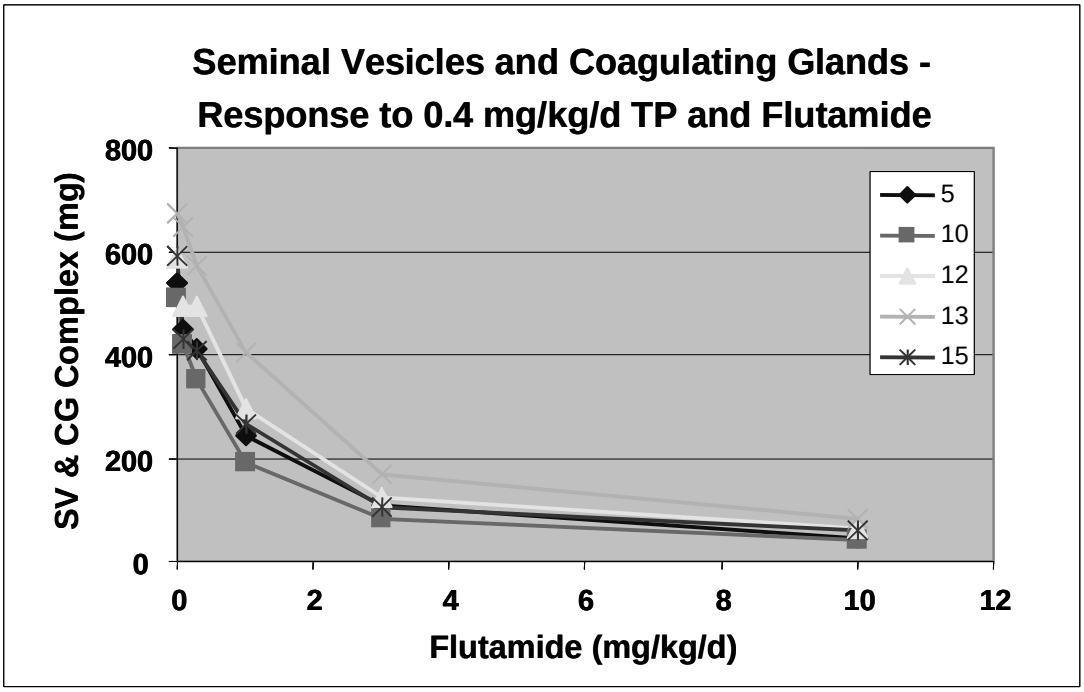
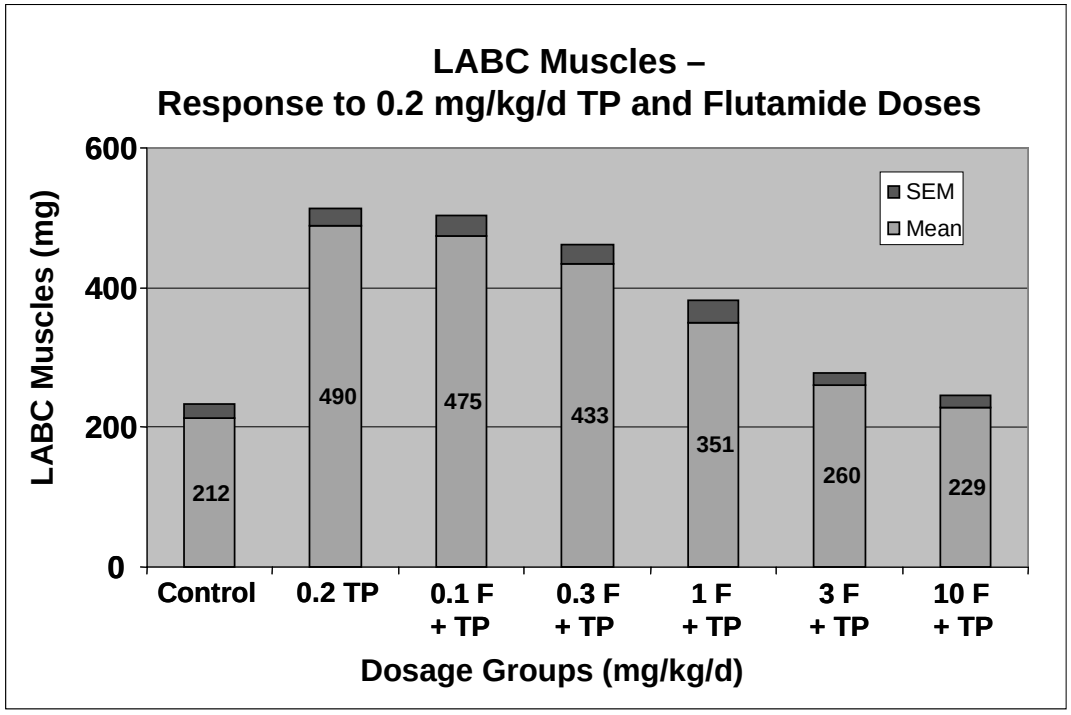


Figure 9. The results of TP and Flutamide administration on seminal vesicles and coagulating glands. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are absolute weights of the pooled means tissue on the y-axis, and Flutamide dose on the x-axis.



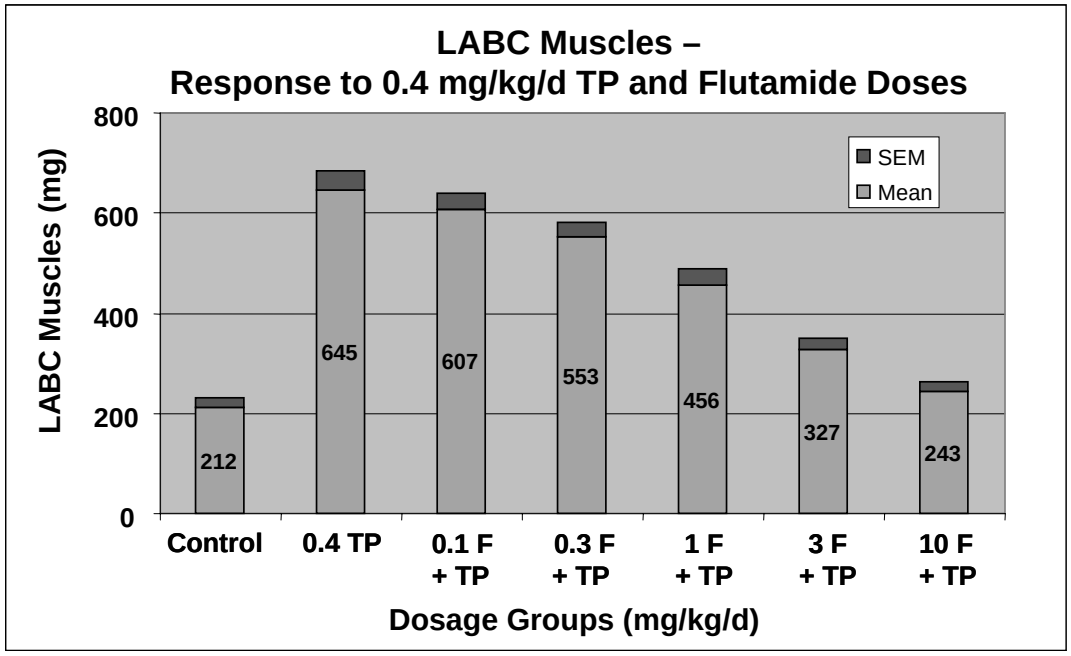
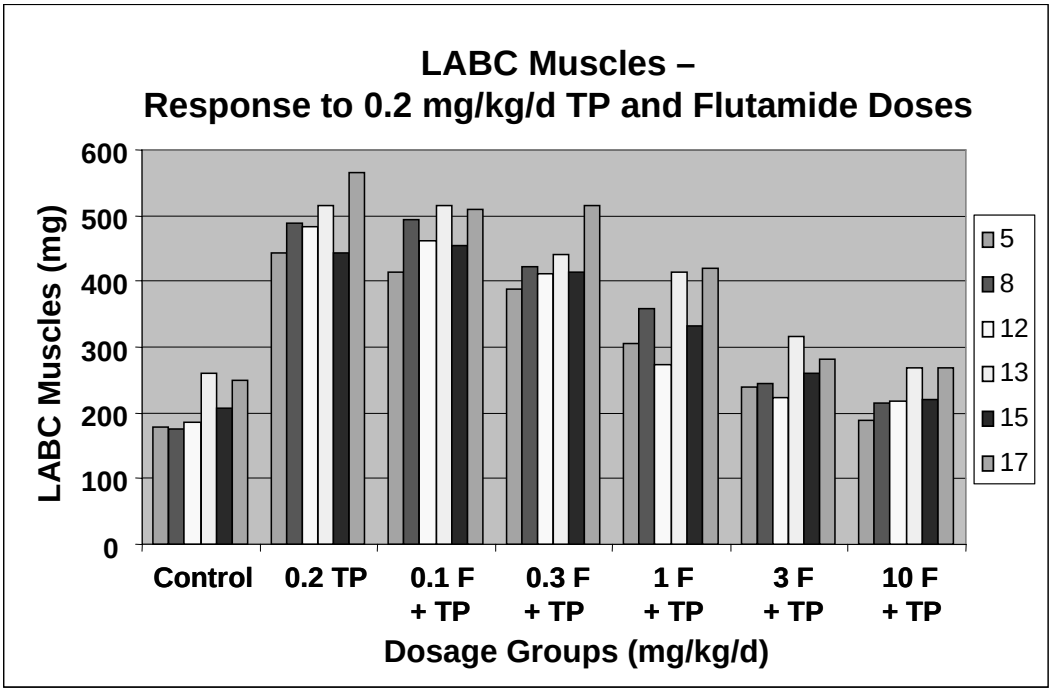


Figure 10. The results of TP and Flutamide administration on levator ani and bulbocavernosus muscles. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.



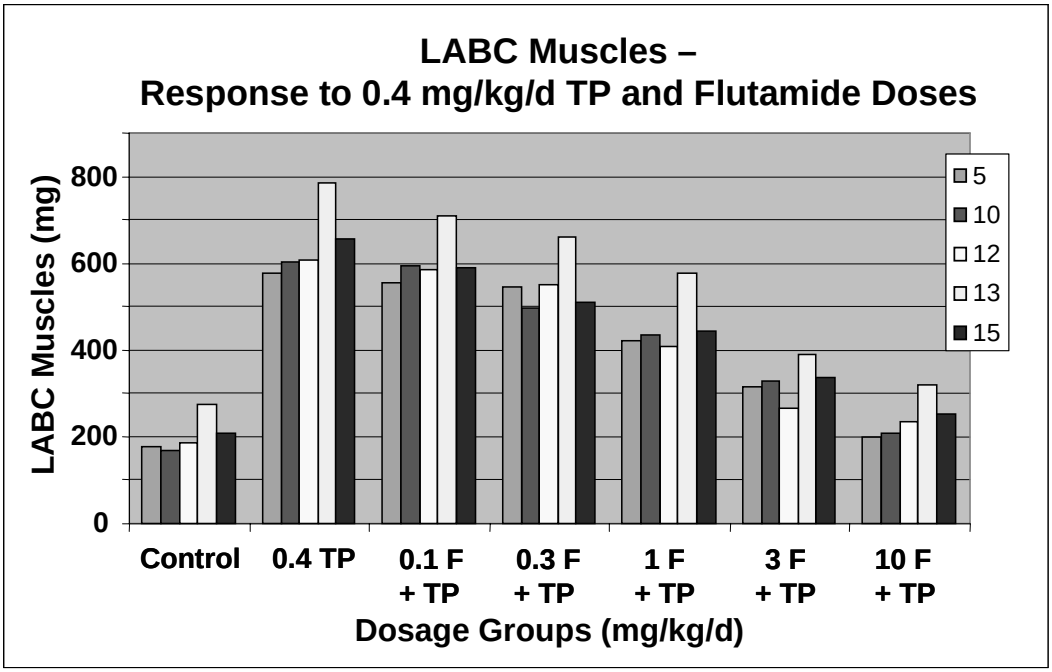
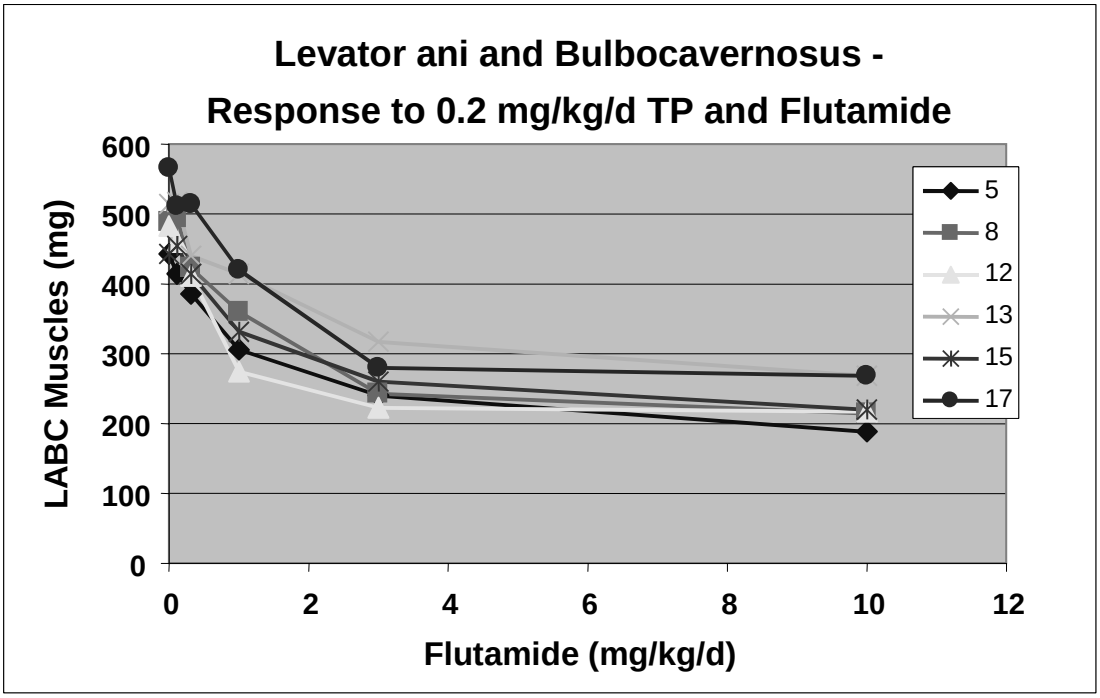


Figure 11. The results of TP and Flutamide administration on levator ani and bulbocavernosus muscles. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are the individual means of the laboratories plotted by the control and substance administration groups.



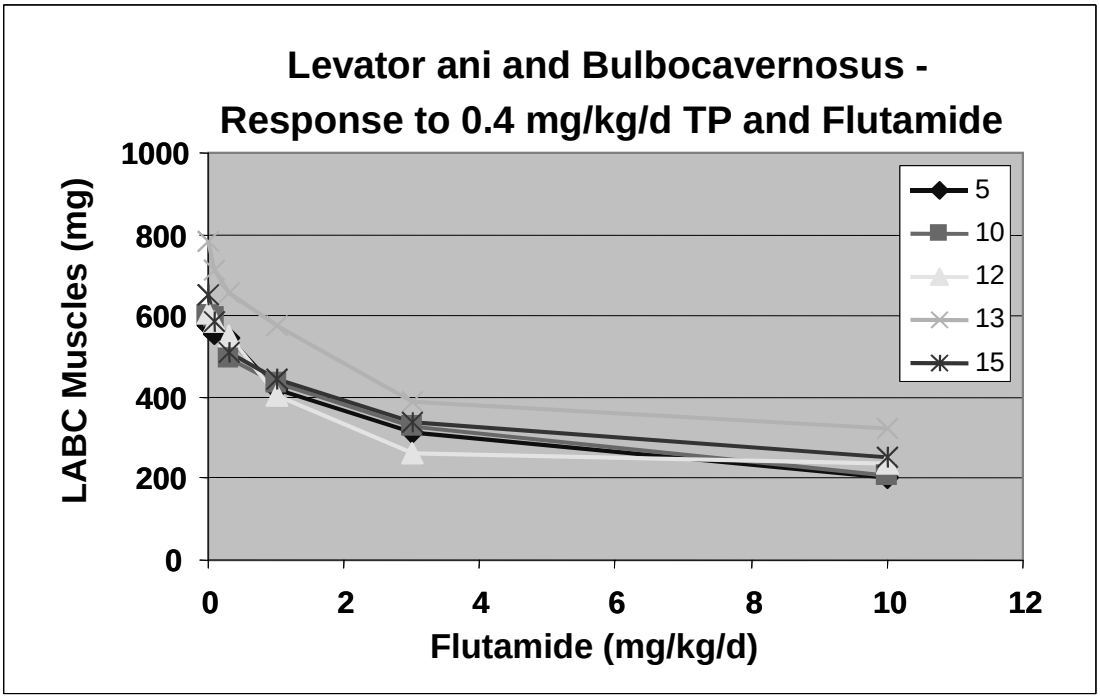
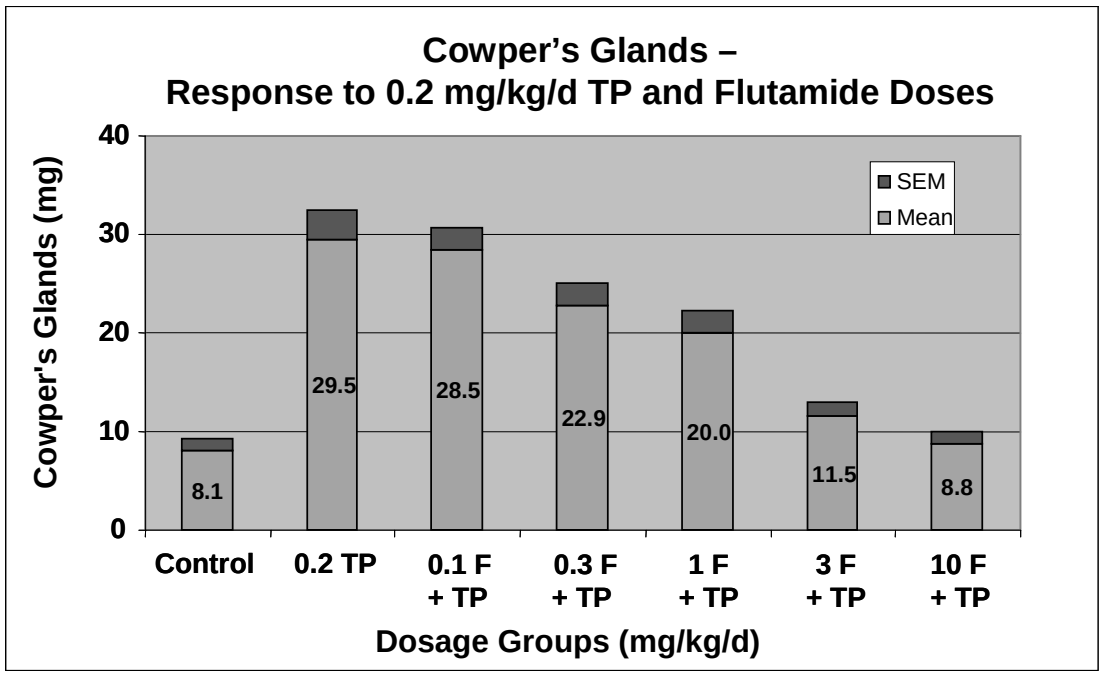


Figure 12. The results of TP and Flutamide administration on levator ani and bulbocavernosus muscles. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are absolute weights of the pooled means tissue on the y-axis, and Flutamide dose on the x-axis.



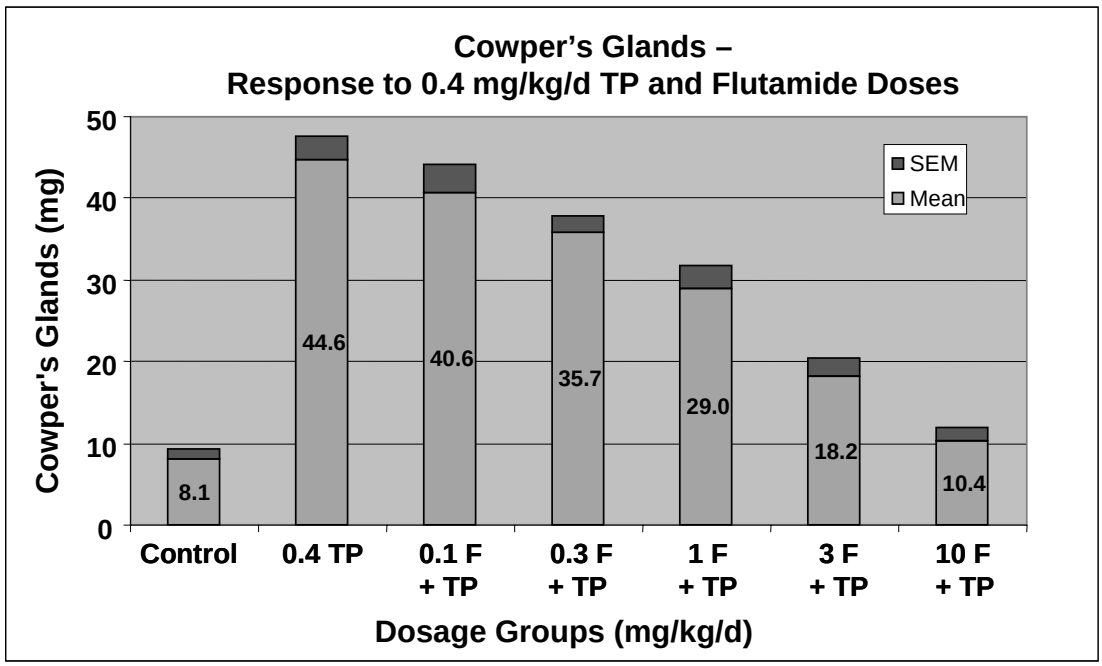
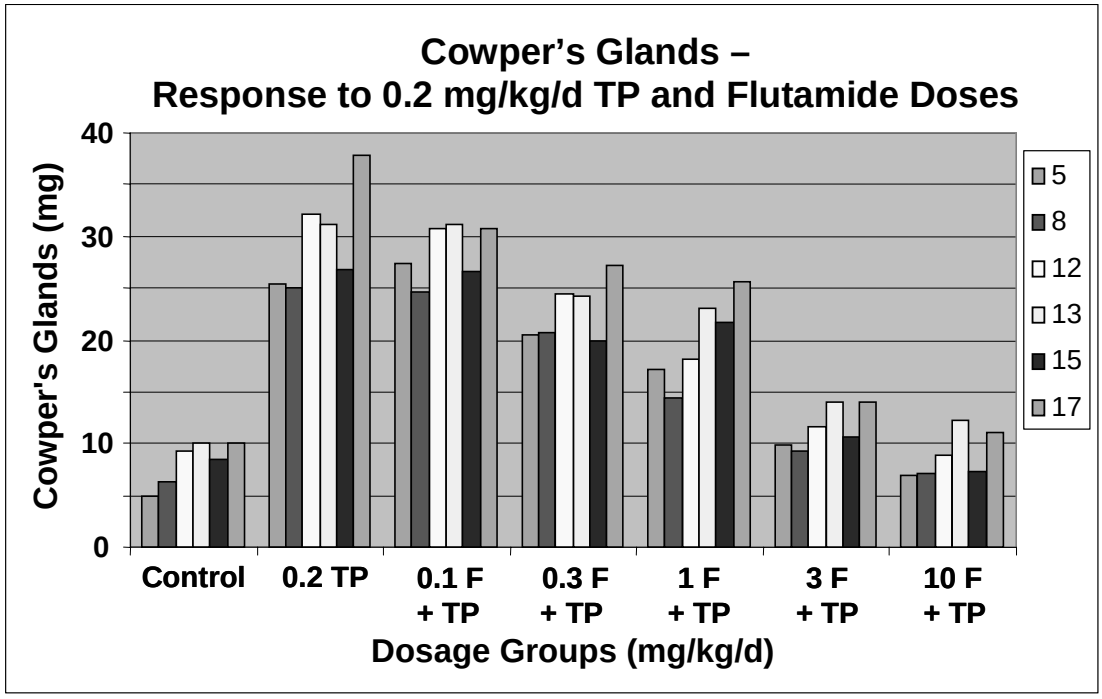


Figure 13. The results of TP and Flutamide administration on the Cowper's glands. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.



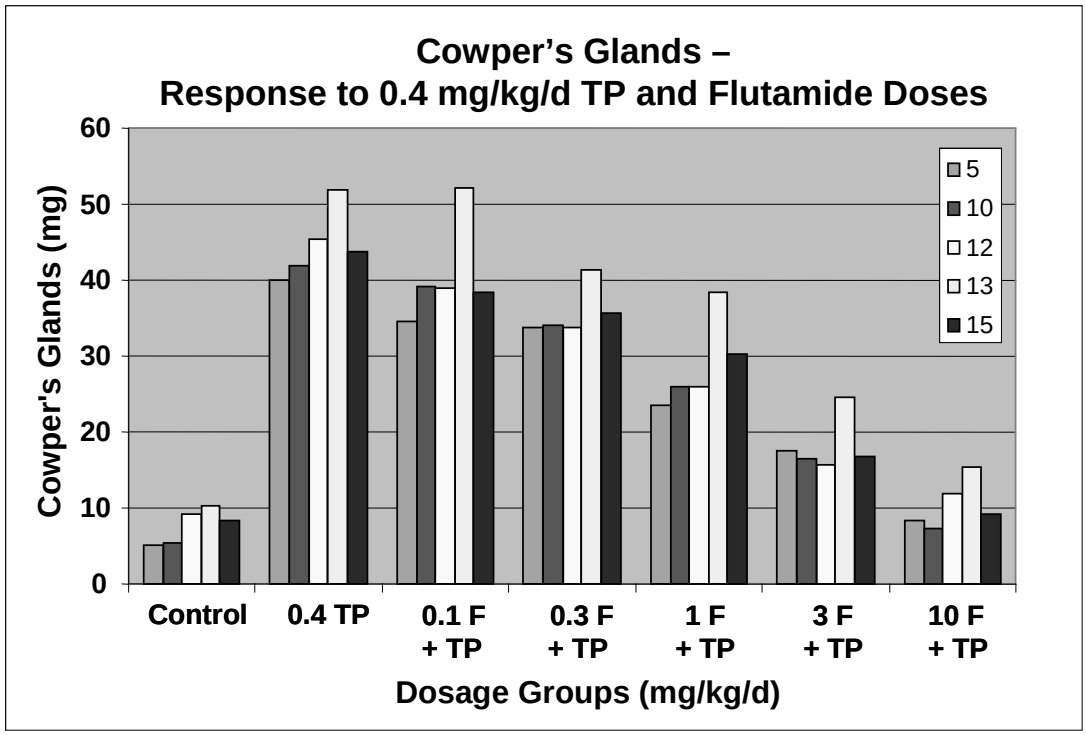
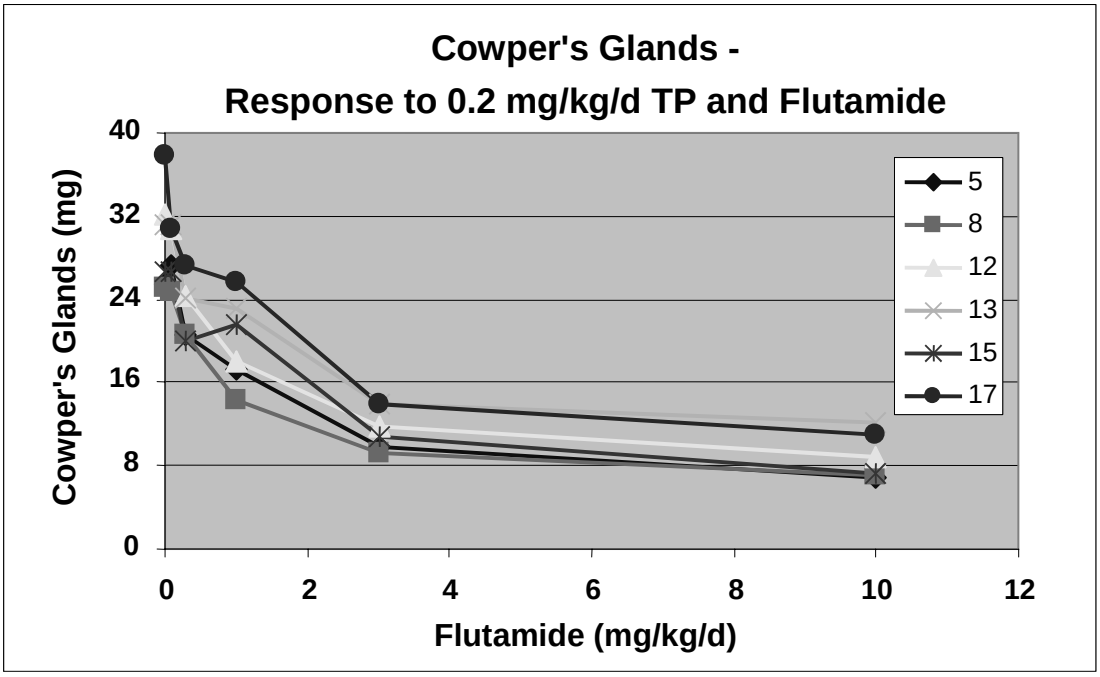


Figure 14. The results of TP and Flutamide administration on the Cowper's glands. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are the individual means of the laboratories plotted by the control and substance administration groups.



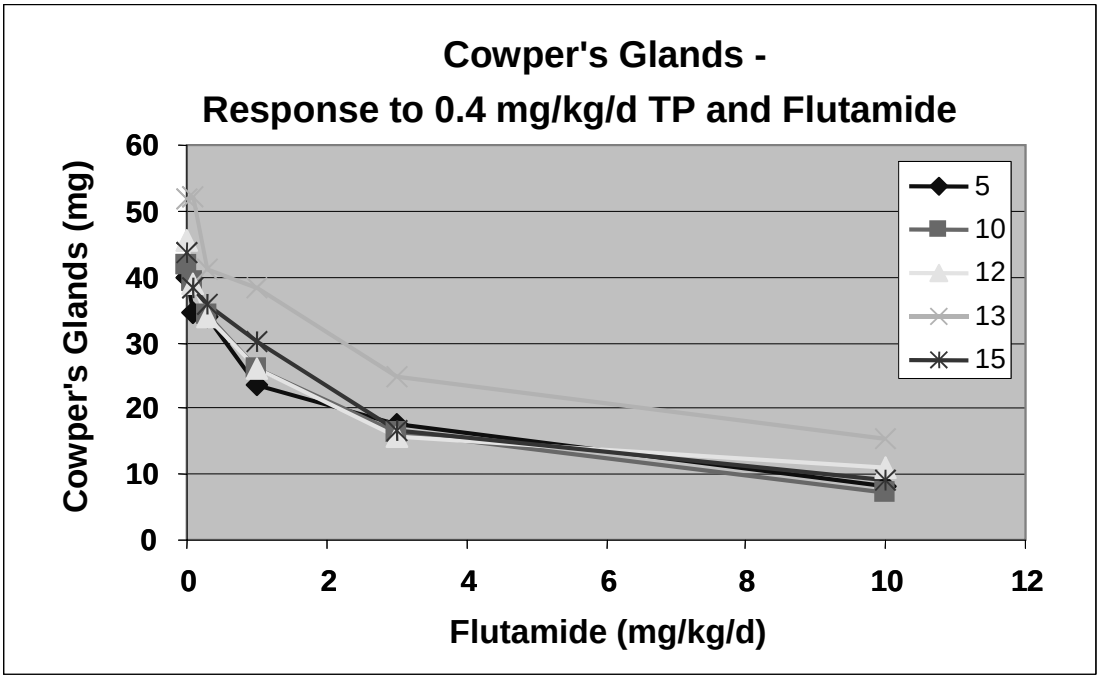
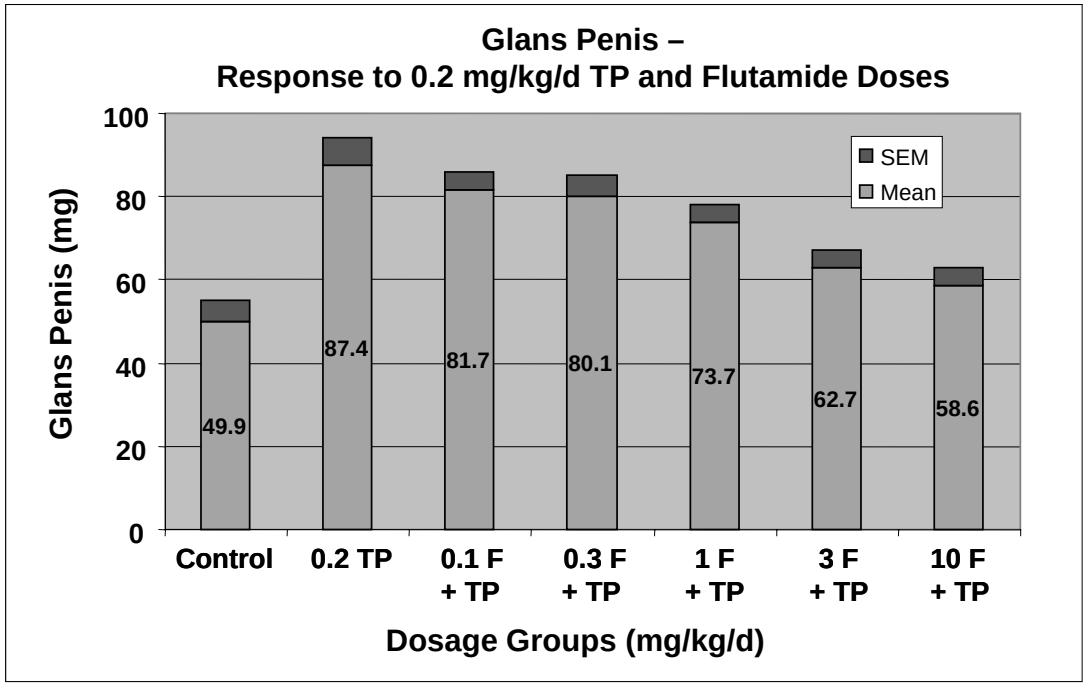


Figure 15. The results of TP and Flutamide administration on the Cowper's glands. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are absolute weights of the pooled means tissue on the y-axis, and Flutamide dose on the x-axis.



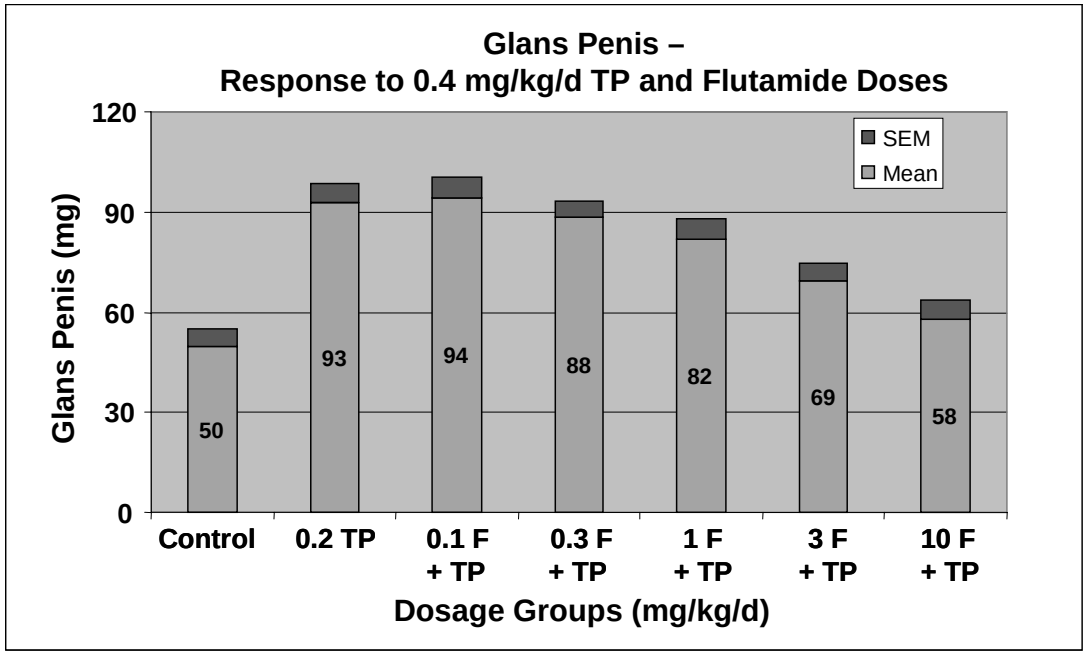
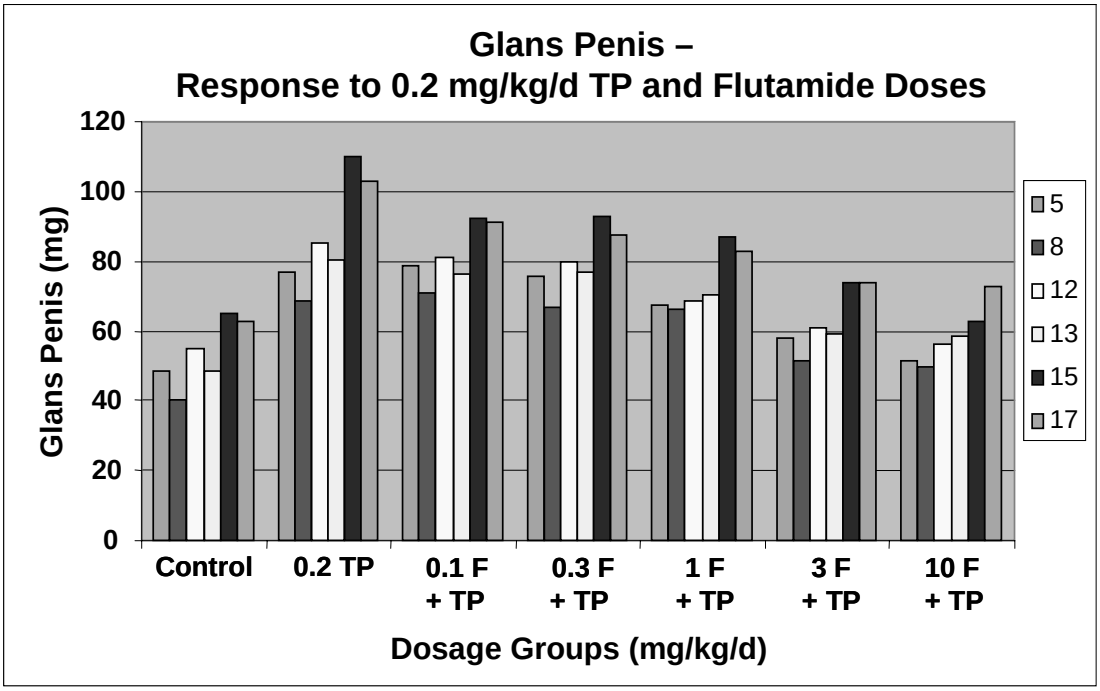


Figure 16. The results of TP and Flutamide administration on the glans penis. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.



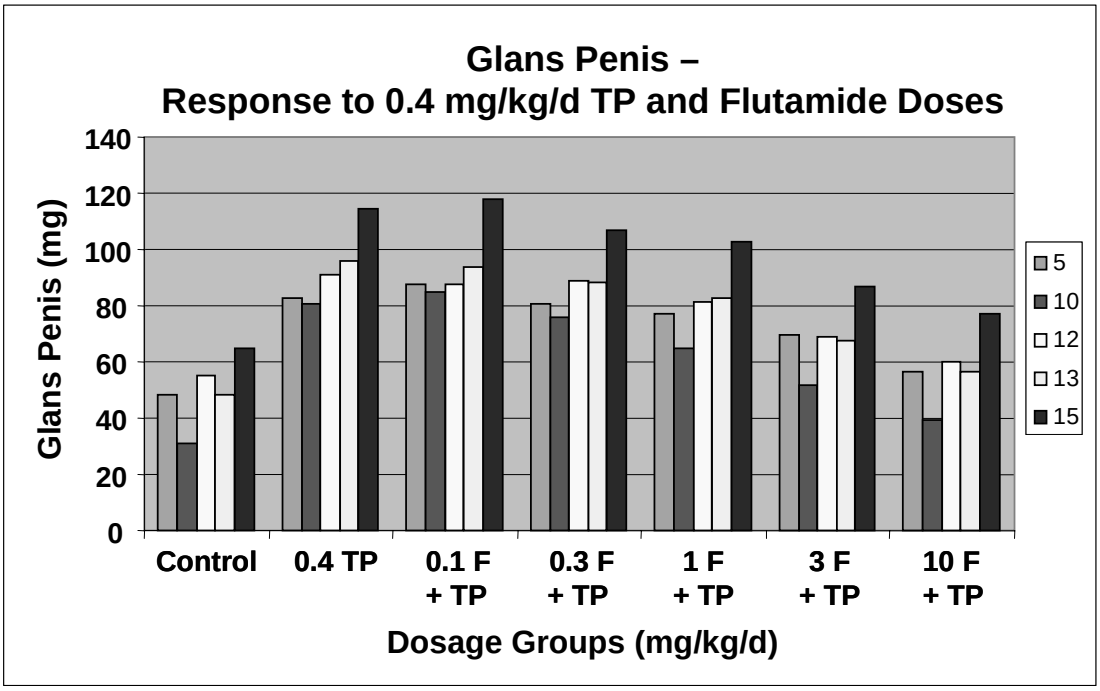


Figure 17. The results of TP and Flutamide administration on the glans penis. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are the individual means of the laboratories plotted by the control and substance administration groups.

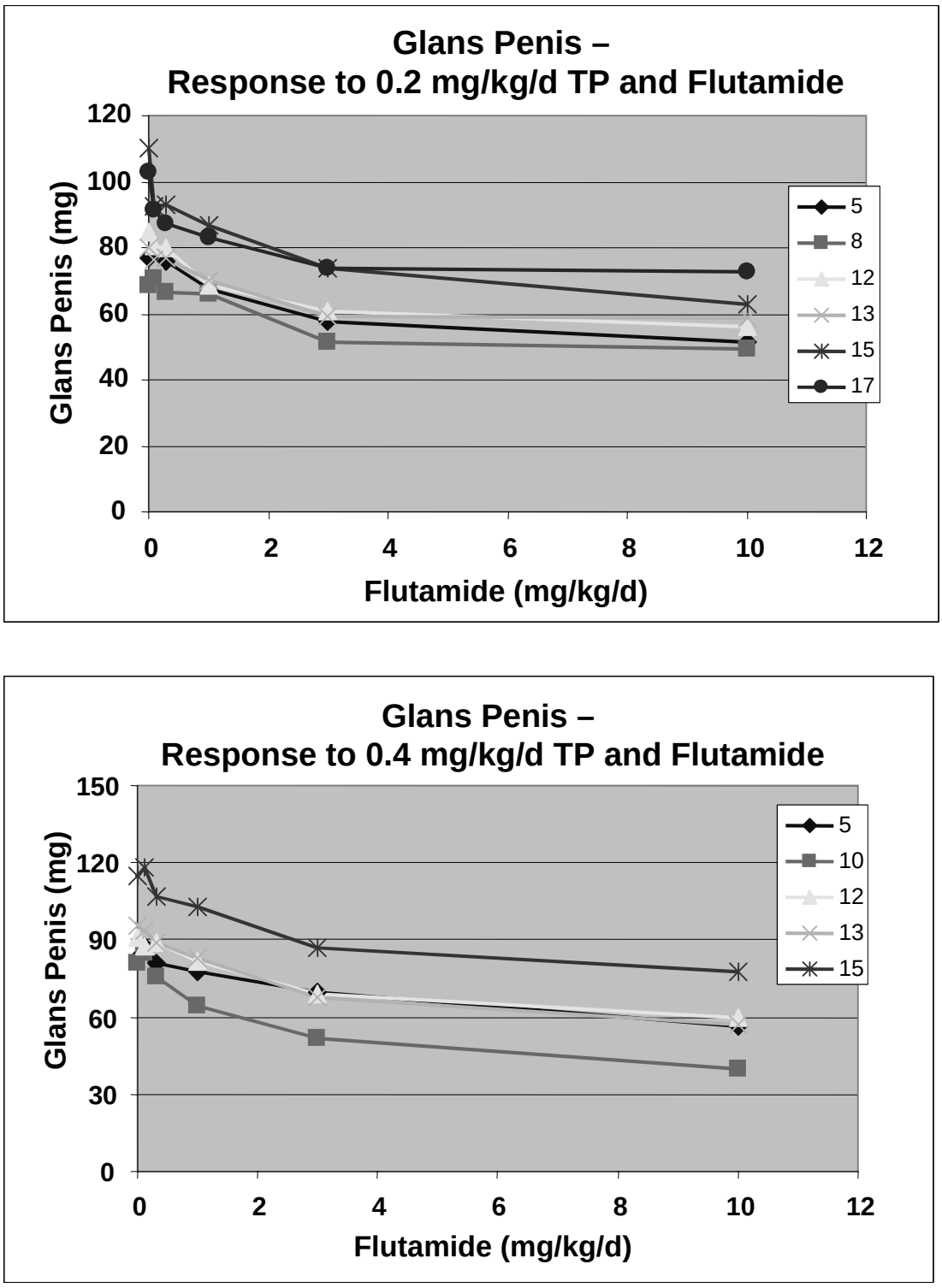
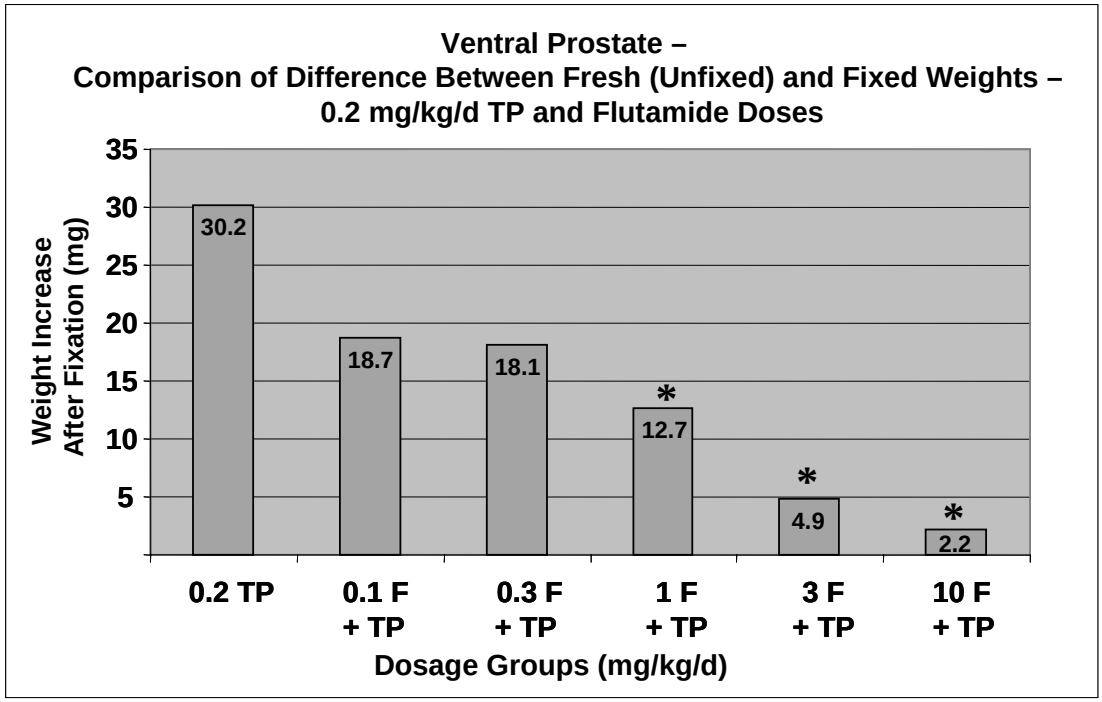


Figure 18. The results of TP and Flutamide administration on the glans penis. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data are absolute weights of the pooled means tissue on the y-axis, and Flutamide dose on the x-axis.



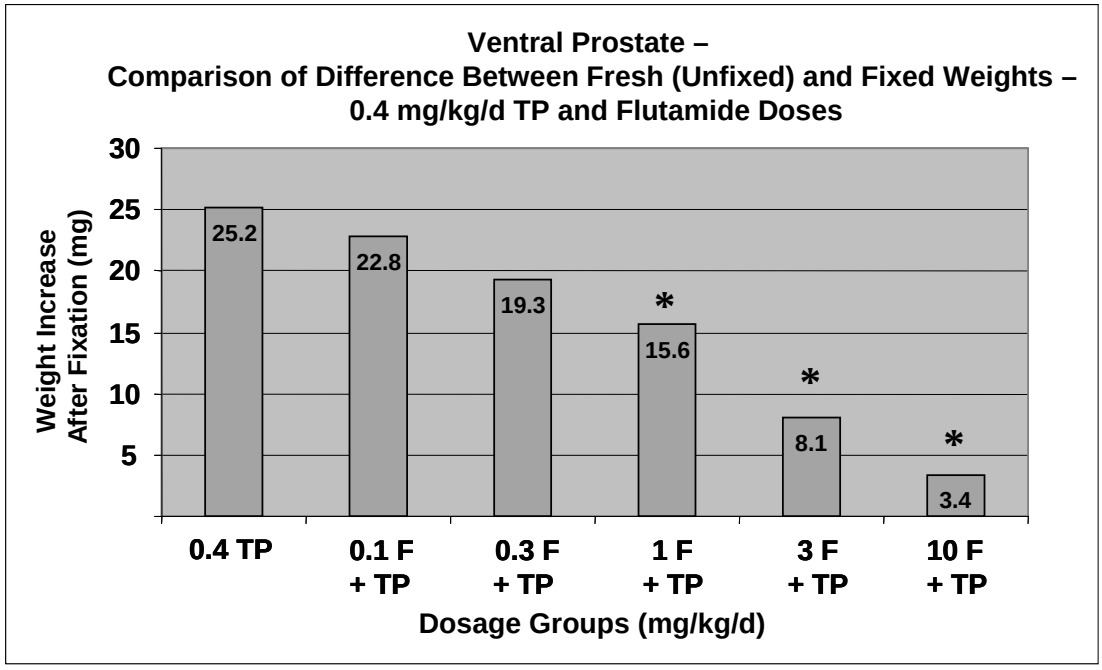
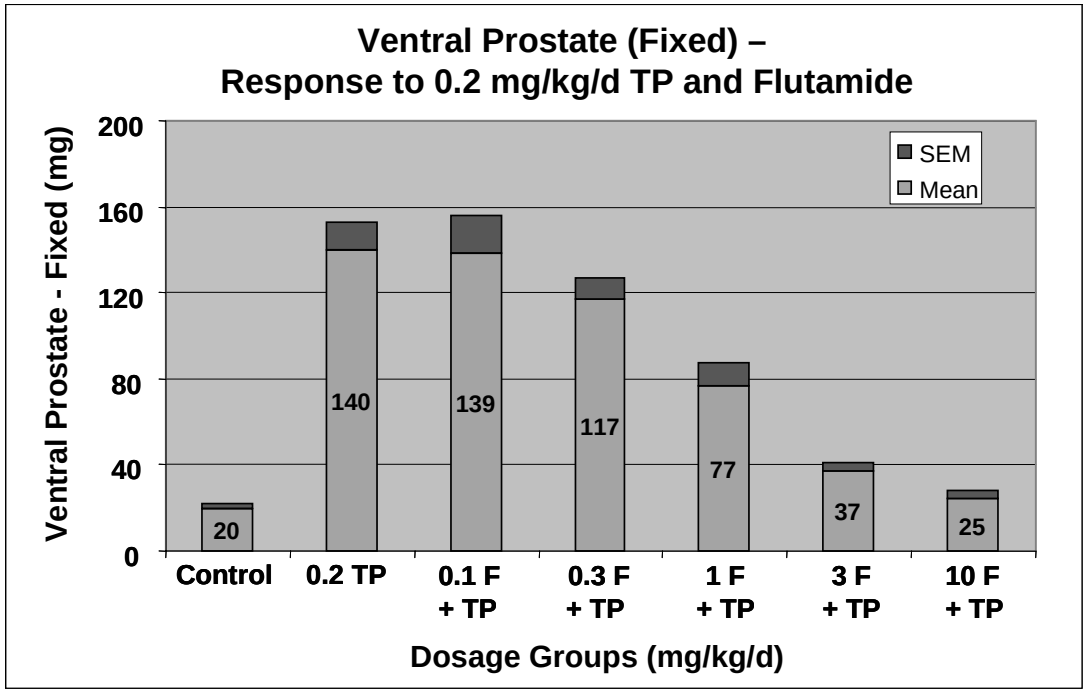


Figure 19. The impact of fixation on ventral prostate weights from a dosage range of TP and Flutamide administration. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean for the fresh data (vpfresh) and for the fixed data (vpfixed). The bars are labeled with the difference (vpfixed – vpfresh) or vpdelta values, and differences with $p < 0.05$ are labeled with an *.



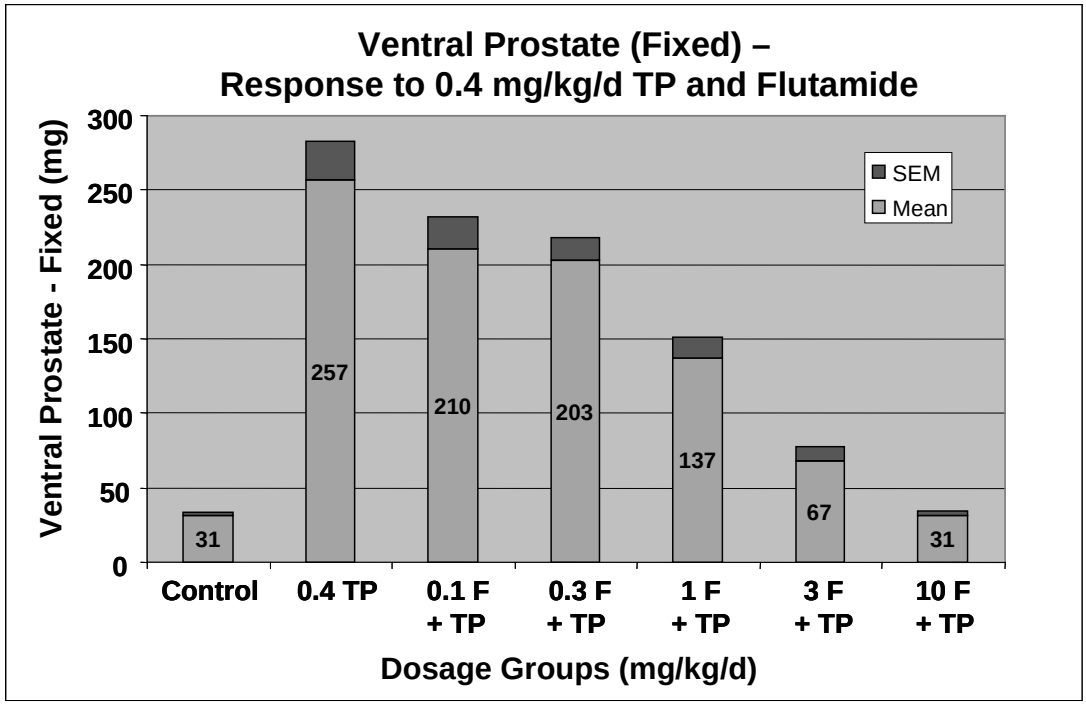
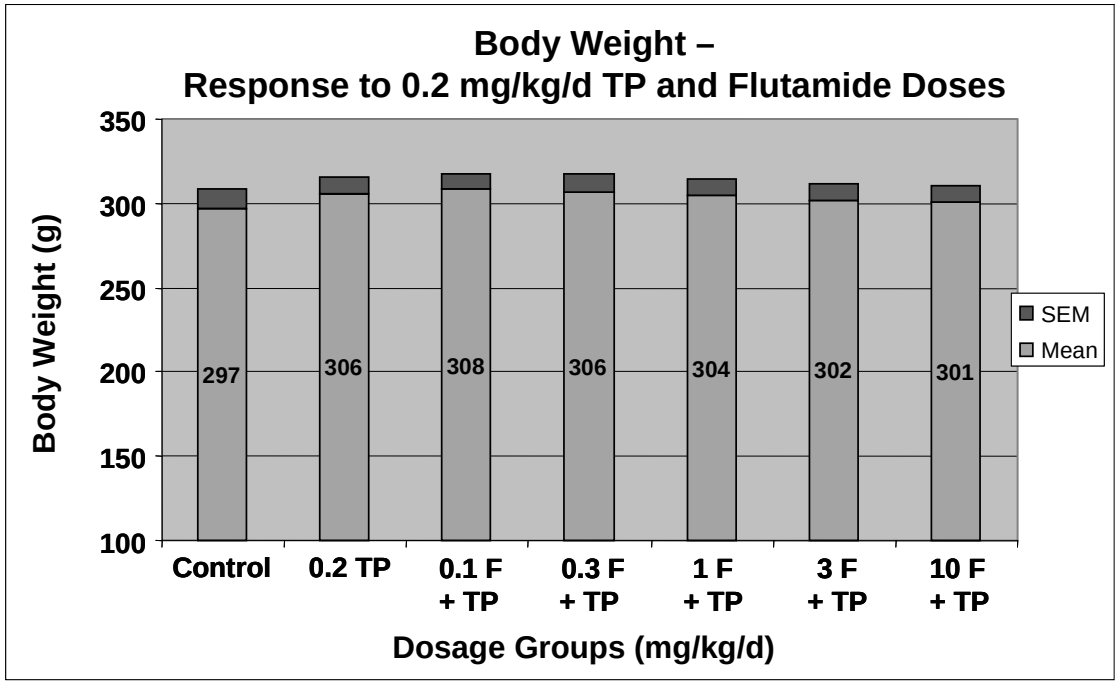


Figure 20. The results of TP and Flutamide administration on the fixed ventral prostate. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.



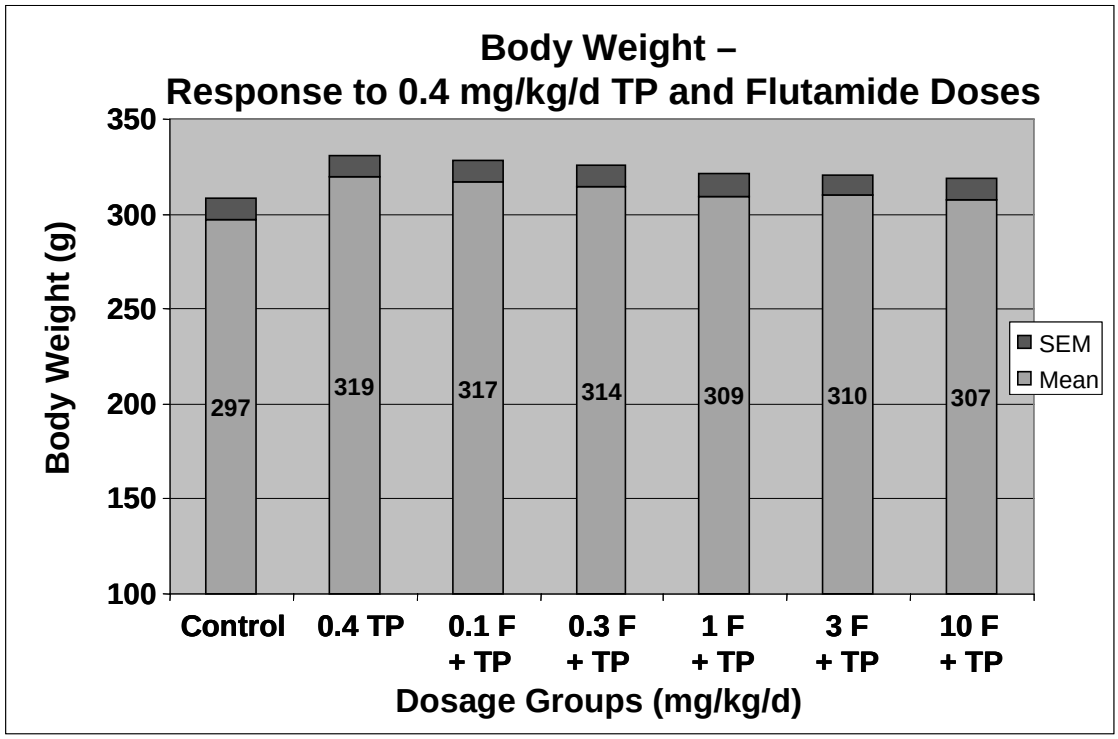
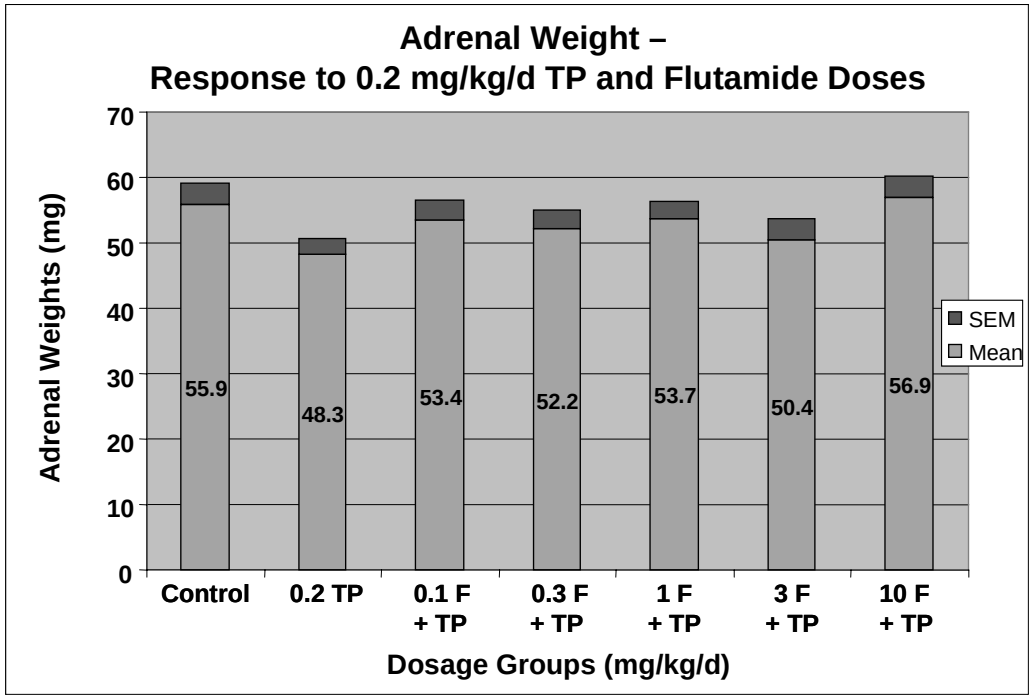


Figure 21. The impact of TP and Flutamide administration on necropsy body weights. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values. Note the reduced y-axis scale.



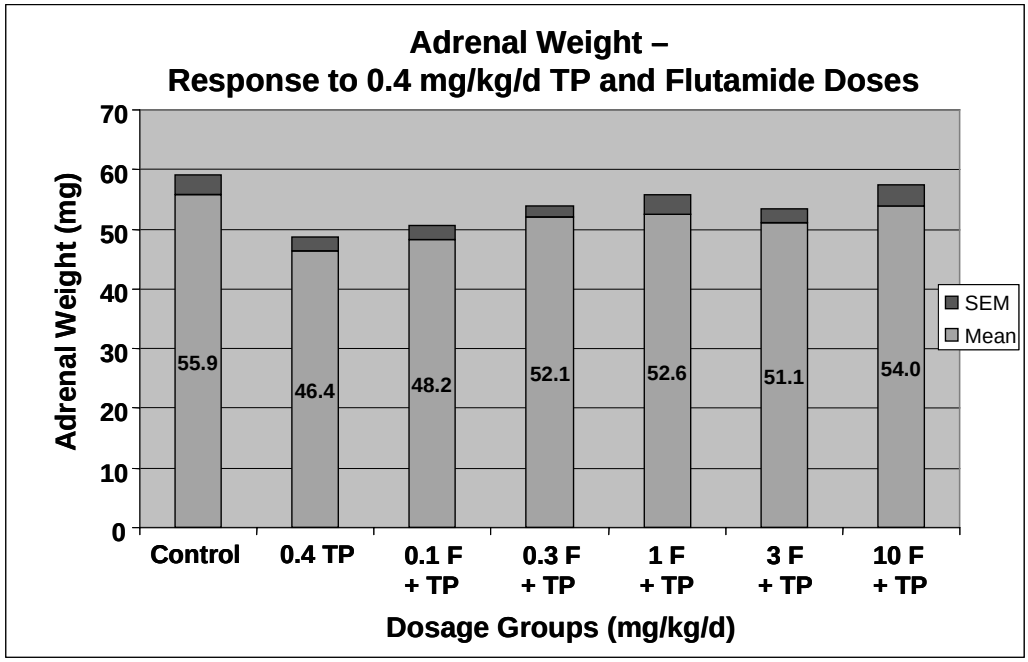


Figure 22. The impact of TP and Flutamide administration on necropsy adrenal weights. Data for 0.2 mg/kg/d TP is in the upper graph and 0.4 mg/kg/d TP is in the lower graph. The data from all labs has been pooled to calculate an overall mean and standard error (SEM). The bars are labeled with the mean values.

Table 1. Ventral Prostate Data and Statistical Analyses (VP in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 5	Mean	16.1						306		log 10 vp	log 10 vp
TP=0 mg/kg/d	CV	23									
	MEAN	114	117	103	63.7	59.2	19.4	319	74	139	96
TP=0.2 mg/kg/d	CV	9	9	18	24	17	12		15		
	MEAN	211	177	175	117	59	28	327	128	30	83
TP=0.4 mg/kg/d	CV	17	20	13	19	38	16		20		
Lab 8	MEAN	14.6						273			
TP=0 mg/kg/d	CV	37									
	MEAN	105	106	81	54	28	23	286	66	22	79
TP=0.2 mg/kg/d	CV	32	33	24	36	34	42		36		
Lab 10	MEAN	9.4						275			
TP=0 mg/kg/d	CV	26									
	MEAN	163	128	104	67	37	20	287	87	109	95
TP=0.4 mg/kg/d	CV	18	19	16	26	14	15		22		
Lab 12	MEAN	19						288			
TP=0 mg/kg/d	CV	43									
	MEAN	128	107	101	56	32	22	295	74	99	94
TP=0.2 mg/kg/d	CV	11	22	19	20	17	15		19		
	MEAN	214	171	174	113	52	26	300	125	63	91
TP=0.4 mg/kg/d	CV	10	27	14	29	31	36		22		
Lab 13	MEAN	24.6						337			
TP=0 mg/kg/d	CV	24									
	MEAN	142	150	100	85	48	32	343	93	44	88
TP=0.2 mg/kg/d	CV	20	20	17	32	16	18		23		
	MEAN	233	228	197	176	80	46	353	160	72	92
TP=0.4 mg/kg/d	CV	20	20	12	15	17	28		20		
Lab 15	MEAN	22.2						281			
TP=0 mg/kg/d	CV	19									
	MEAN	131	150	109	83	38	25	290	90	99	94
TP=0.2 mg/kg/d	CV	13	28	17	18	12	7		23		
	MEAN	268	213	202	136	68	31	296	153	154	96
TP=0.4 mg/kg/d	CV	22	16	10	10	26	8		20		

Table 1. Continued: Ventral Prostate Data and Statistical Analyses (VP in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 17	MEAN	15.1						281			
TP=0 mg/kg/d.0	CV	31									
	MEAN	140	134	117	69	33	26	296	86	58	91
TP=0.2 mg/kg/d	CV	18	15	30	17	30	16		24		
ALL LABS	MEAN	18.2									
TP=0 mg/kg/d	CV	39									
	MEAN	127	127	102	69	35	25		80.6		
TP=0.2 mg/kg/d	CV	20	26	23	30	27	25		24		
	MEAN	218	184	170	122	59	30		130		
TP=0.4 mg/kg/d	CV	24	27	24	34	35	37		21		

Table 2. Seminal Vesicle and Coagulating Gland Data and Statistical Analyses (SV in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF Flutamide (oral) mg/kg/d FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 5	Mean	37						306		log 10 SV	log 10 SV
TP=0 mg/kg/d	CV	25									
	MEAN	315	252	204	123	54	41	318	164	135	96
TP=0.2 mg/kg/d	CV	13	11	23	18	22	15		18		
	MEAN	539	450	410	243	108	45	327	299	166	97
TP=0.4 mg/kg/d	CV	12	17	13	22	26	17		18		
Lab 8	MEAN	32						273			
TP=0 mg/kg/d	CV	20									
	MEAN	348	263	247	117	60	45	286	180	52	90
TP=0.2 mg/kg/d	CV	22	33	33	32	35	16		34		
Lab 10	MEAN	31						275			
TP=0 mg/kg/d	CV	17									
	MEAN	507	418	353	189	83	41	287	265	204	97
TP=0.4 mg/kg/d	CV	7	23	14	25	16	15		19		
Lab 12	MEAN	53						288			
TP=0 mg/kg/d	CV	28									
	MEAN	368	366	248	137	71	62	295	209	69	92
TP=0.2 mg/kg/d	CV	24	20	27	37	12	22		28		
	MEAN	588	493	495	295	122	65	300	343	93	94
TP=0.4 mg/kg/d	CV	20	21	28	22	32	12		27		
Lab 13	MEAN	67						337			
TP=0 mg/kg/d	CV	13									
	MEAN	316	288	247	137	86	74	343	191	43	88
TP=0.2 mg/kg/d	CV	21	28	29	25	20	11		28		
	MEAN	674	646	573	402	168	83	353	424	71	92
TP=0.4 mg/kg/d	CV	23	13	21	24	35	17		23		
Lab 15	MEAN	46						281			
TP=0 mg/kg/d	CV	18									
	MEAN	289	232	206	116	68	50	290	160	49	89
TP=0.2 mg/kg/d	CV	28	32	18	28	26	9		31		
	MEAN	592	431	406	266	106	61	296	310	156	96
TP=0.4 mg/kg/d	CV	11	9	16	14	26	16		14		

Table 2. Continued: Seminal Vesicle and Coagulating Gland Data and Statistical Analyses (SV in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF Flutamide (oral) mg/kg/d FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 17	MEAN	51						281			
TP=0 mg/kg/d.0	CV	22									
	MEAN	440	320	299	167	83	67	296	229	60	91
TP=0.2 mg/kg/d	CV	22	15	28	40	21	10		28		
ALL LABS	MEAN	47.9									
TP=0 mg/kg/d	CV	34									
	MEAN	346	287	242	133	70	56		189		
TP=0.2 mg/kg/d	CV	25	27	29	33	27	25		29		
	MEAN	580	488	447	279	117	59		382		
TP=0.4 mg/kg/d	CV	18	23	24	33	38	30		22		

Table 3. Levator Ani & Bulbocavernosus and Statistical Analyses (LABC in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 5	Mean	178						306		log 10	log 10
TP=0 mg/kg/d	CV	13								LABC	LABC
	MEAN	444	413	387	306	239	189	318	328	87	94
TP=0.2 mg/kg/d	CV	5	7	10	10	12	9		8		
	MEAN	578	555	547	419	314	202	327	436	60	91
TP=0.4 mg/kg/d	CV	8	12	4	16	12	17		11		
Lab 8	MEAN	175						273			
TP=0 mg/kg/d	CV	8									
	MEAN	488	493	422	359	243	214	286	370	45	88
TP=0.2 mg/kg/d	CV	5	10	14	16	16	14		12		
Lab 10	MEAN	167						275			
TP=0 mg/kg/d	CV	18									
	MEAN	603	596	498	435	328	207	287	445	124	95
TP=0.4 mg/kg/d	CV	10	7	6	7	11	11		9		
Lab 12	MEAN	187						288			
TP=0 mg/kg/d	CV	11									
	MEAN	483	461	412	273	223	217	295	345	40	87
TP=0.2 mg/kg/d	CV	11	18	11	14	10	18		14		
	MEAN	606	585	549	406	265	235	300	441	62	91
TP=0.4 mg/kg/d	CV	9	9	7	14	13	20		11		
Lab 13	MEAN	268						337			
TP=0 mg/kg/d	CV	10									
	MEAN	515	516	441	414	316	268	343	411	18	75
TP=0.2 mg/kg/d	CV	10	13	24	17	15	8		16		
	MEAN	784	710	659	577	391	321	353	574	81	93
TP=0.4 mg/kg/d	CV	11	12	7	8	12	4		10		
Lab 15	MEAN	208						281			
TP=0 mg/kg/d	CV	15									
	MEAN	442	454	415	332	259	220	290	354	48	89
TP=0.2 mg/kg/d	CV	10	12	9	11	12	9		11		
	MEAN	654	588	512	445	337	252	296	465	74	93
TP=0.4 mg/kg/d	CV	10	7	10	11	14	8		10		

Table 3. Continued: Levator Ani & Bulbocavernosus and Statistical Analyses (LABC in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 17	MEAN	249						281			
TP=0 mg/kg/d.0	CV	14									
	MEAN	565	511	514	420	280	268	296	426	31	84
TP=0.2 mg/kg/d	CV	9	17	13	18	13	11		14		
ALL LABS	MEAN	212									
TP=0 mg/kg/d	CV	22									
	MEAN	490	475	432	351	260	229		373		
TP=0.2 mg/kg/d	CV	12	15	16	21	17	17		13		
	MEAN	645	607	553	456	327	243		472		
TP=0.4 mg/kg/d	CV	15	13	12	17	17	21		10		

Table 4. Cowper's Glands Data and Statistical Analyses (COWS in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 5	Mean	5						306		log 10	log 10
TP=0 mg/kg/d	CV	30								COWS	COWS
	MEAN	25.5	27.3	20.5	17.2	9.8	6.8	318	17.8	32	85
TP=0.2 mg/kg/d	CV	28	22	18	27	30	11		26		
	MEAN	40	34.7	33.8	23.6	17.5	8.3	327	26.3	46	88
TP=0.4 mg/kg/d	CV	14	16	13	16	41	21		19		
Lab 8	MEAN	6.4						273			
TP=0 mg/kg/d	CV	43									
	MEAN	25.1	24.7	20.6	14.5	9.2	7	286	16.8	44	88
TP=0.2 mg/kg/d	CV	15	18	15	17	22	26		19		
Lab 10	MEAN	5.3						275			
TP=0 mg/kg/d	CV	18									
	MEAN	41.9	39.2	34.1	26	16.4	7.3	287	27.5	80	94
TP=0.4 mg/kg/d	CV	5	10	12	10	27	28		12		
Lab 12	MEAN	9.3						288			
TP=0 mg/kg/d	CV	12									
	MEAN	32.1	30.7	24.4	18.1	11.7	8.8	295	20.9	54	87
TP=0.2 mg/kg/d	CV	15	10	19	17	29	13		17		
	MEAN	45.4	38.9	33.9	26	15.6	11.9	300	28.6	72	92
TP=0.4 mg/kg/d	CV	9	21	13	15	13	20		16		
Lab 13	MEAN	10.1						337			
TP=0 mg/kg/d	CV	27									
	MEAN	31.1	31.1	24.2	23.1	13.9	12.2	343	22.6	18	75
TP=0.2 mg/kg/d	CV	10	15	28	21	35	20		21		
	MEAN	51.8	52.1	41.2	38.3	24.7	15.4	353	37.3		
TP=0.4 mg/kg/d	CV	19	15	9	19	18	18		13		
Lab 15	MEAN	8.4						281			
TP=0 mg/kg/d	CV	14									
	MEAN	26.7	26.6	20	21.6	10.7	7.2	290	18.8	31.6	84
TP=0.2 mg/kg/d	CV	13	30	24	15	29	22		24		
	MEAN	43.8	38.4	35.8	30.2	16.6	9.1	296	28.9	64	91
TP=0.4 mg/kg/d	CV	13	19	13	13	24	26		17		

Table 4.Continued: Cowper's Glands Data and Statistical Analyses (COWS in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 17	MEAN	10						281			
TP=0 mg/kg/d.0	CV	34									
	MEAN	37.9	30.7	27.2	25.7	13.9	11	296	24	21	78
TP=0.2 mg/kg/d	CV	30	10	24	25	31	32		26		
ALL LABS	MEAN	8.1									
TP=0 mg/kg/d	CV	36									
	MEAN	29.5	28.5	22.9	20	11.5	8.8		20.2		
TP=0.2 mg/kg/d	CV	24	19	24	28	33	32		23		
	MEAN	44.6	40.6	35.7	29	18.2	10.4		29.5		
TP=0.4 mg/kg/d	CV	16	21	14	24	30	35		17		

Table 5. Glans Penis Data and Statistical Analyses (GP in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 5	Mean	48.6						306		log 10	log 10
TP=0 mg/kg/d	CV	10								GLANS	GLANS
	MEAN	77	78.4	75.8	67.6	57.7	51.3	318	67.7	39	87
TP=0.2 mg/kg/d	CV	6	4	5	7	13	6		6.6		
	MEAN	82.7	87.9	81	77.4	70	86.6	327	75.9	46	88
TP=0.4 mg/kg/d	CV	8	3	6	8	7	4		6.6		
Lab 8	MEAN	40.1						273			
TP=0 mg/kg/d	CV	15									
	MEAN	68.7	70.8	66.6	66	51.5	49.6	286	62.2	8.6	59
TP=0.2 mg/kg/d	CV	13	16	8	12	11	11		13		
Lab 10	MEAN	30.8						275			
TP=0 mg/kg/d	CV	15									
	MEAN	80.9	84.7	75.9	64.6	51.9	39.5	287	66.2	41.2	87
TP=0.4 mg/kg/d	CV	5	8	5	7	17	18		9.2		
Lab 12	MEAN	55.1						288			
TP=0 mg/kg/d	CV	5									
	MEAN	84.9	80.7	80	68.8	60.6	56.3	295	71.9	28.7	83
TP=0.2 mg/kg/d	CV	8	8	7	4	11	6		7.7		
	MEAN	90.8	87.9	88.9	81.5	69.2	59.9	300	79.7	70.6	92
TP=0.4 mg/kg/d	CV	6	5	3	5	7	2		5		
Lab 13	MEAN	48.7						337			
TP=0 mg/kg/d	CV	8									
	MEAN	80.2	76.5	77.1	70.1	59.1	58.7	343	70.3	9.5	61
TP=0.2 mg/kg/d	CV	7	3	13	9	11	18		11		
	MEAN	95.8	93.4	88.7	82.9	67.4	56.8	353	80.8	29.1	83
TP=0.4 mg/kg/d	CV	9	9	9	14	7	8		9.7		
Lab 15	MEAN	64.8						281			
TP=0 mg/kg/d	CV	13									
	MEAN	110	92.4	93	86.8	73.9	62.9	290	86.6	18.5	75
TP=0.2 mg/kg/d	CV	12	12	7	12	13	10		11.2		
	MEAN	114.7	118.1	106.9	102.9	87.2	77.4	296	101.1	12.8	68
TP=0.4 mg/kg/d	CV	4	12	12	11	15	9		10.9		

Table 5. Continued: Glans Penis Data and Statistical Analyses (GP in mg and Body Wt in g) – shaded cells statistically significant

		DOSE OF FLUTAMIDE (ORAL) MG/KG/D FOR 10 DAYS						Necropsy	OVERALL	F VALUE	R-SQUARE
		0	0.1	0.3	1	3	10	BODY WT		Flutamide	Flutamide
Lab 17	MEAN	62.5						281			
TP=0 mg/kg/d.0	CV	21									
	MEAN	103	91.2	87.4	82.9	73.7	72.7	296	89.2	5.4	47
TP=0.2 mg/kg/d	CV	9	8	20	9	15	14		13		
ALL LABS	MEAN	49.9									
TP=0 mg/kg/d	CV	25									
	MEAN	87.4	81.7	80.1	73.7	62.7	58.6		74		
TP=0.2 mg/kg/d	CV	19	13	15	14	18	18		11		
	MEAN	93	94.4	88.3	81.8	69.1	58		80.8		
TP=0.4 mg/kg/d	CV	15	15	14	18	20	23		9		

ANNEX – 8

STATISTICAL REVIEW AND ANALYSIS OF THE PHASE 1 STUDIES - TESTOSTERONE PROPIONATE AND FLUTAMIDE

SUMMARY

1. The purpose of this work was to review and validate the statistical analysis of steps one and two of the first phase of the Testosterone propionate study, and to perform any additional analysis that was felt needed.

2. As a first step, I verified that correct statistical methods were used and that the results were interpreted correctly. Secondly, I repeated the analyses to further validate the results. Finally, I undertook four additional analyses for both studies which I felt might further strengthen the conclusions of the report: rigorous validation of ANOVA model assumptions, comparison of LOELs across endpoints and laboratories, comparison of benchmark doses (ED_{05}s) across endpoints and laboratories and treating the LAB effect as random.

3. In my opinion, appropriate analyses were performed and the interpretation of the results was correct for both studies we examined. The additional analysis I undertook could be used to strengthen the analysis, but none of the new analyses substantially changed the interpretation of the data.

STATISTICAL REVIEW

Phase 1a Study (TP)

4. The purpose of the phase 1a study was to assess the level of standardization of a common OECD protocol for the Hershberger assay. In the standardize assay, castrated male rats are exposed to five dose levels of the androgen Testosterone propionate (TP). The experiment was conducted by 17 laboratories with the intent of confirming that results can be duplicated across laboratories. The data generated by the 17 laboratories were subjected to statistical analysis in order to:

- Demonstrate the reproducibility and sensitivity of various measured endpoints within and among participating laboratories;
- Demonstrate the responses of five sex accessory tissues to the action of TP; and
- Enable a standard reference dose of TP to be calculated.

5. The analysis consisted primarily of summary statistics and ANOVA models fit to either the raw data or log10 transformed data, within each laboratory and with all laboratories combined. Results reported included coefficients of variation and R^2 values for terms in the model.

6. I believe that the correct analyses were performed. It is always possible to fit more complex models to such data, but given the objectives of these study, it is almost certainly not necessary. For all of the models, the effect of laboratory was highly significant, in addition to that of dose. This is not unexpected and not of concern given that each laboratory will have its own differences including the initial body weights of animals on test. Of potentially greater concern is the LABxDOSE interaction, which was significant for all endpoints. This may indicate that different laboratories would produce different dose-response curves even though a similar LOEL was detected. This may or may not be an issue depending on

the goal of the assay.

Phase 1b Study (Flutamide)

7. The phase 1b study was designed to see if the assay could detect the effects of the anti-androgen flutamide given on top of doses of 0.2 and 0.4 mg/kg-bw/day TP. It was also intended to determine which dose of TP was more sensitive to low dose antagonistic effects of flutamide.

8. As with the phase 1a study, the analysis consisted of summary statistics and ANOVA models fit to either the raw or log10 transformed data, within each laboratory and with all laboratories combined. Results reported included coefficients of variation and R2 values for terms in the model.

9. The results of the analysis indicate that the assay can indeed detect flutamide effects at both levels of TP. The LOELs were variable between labs and endpoints, falling between 0.1 and 3 mg/kg-bw/day. The most sensitive endpoint appeared to be SV at the higher dose of TP and the least sensitive endpoint was GLANS at the highest dose of TP. Discounting GLANS, which was more variable and tended to be less sensitive in general, the dose of TP that provided the best sensitivity overall was 0.4 mg/kg-bw/day.

10. I believe that the correct analyses were performed and that these analyses were interpreted correctly. In my opinion, the conclusions reached by the authors are supported by the data.

ADDITIONAL ANALYSIS

Methods

11. I thought it would be helpful to rigorously validate the ANOVA model assumptions - particularly that the error terms are independently normal with constant variance. If these assumptions are not satisfied, then the test statistics can yield misleading results. The authors noted increasing variance in most cases and applied a log10 transformation to correct it, but no formal diagnostics were performed. I examined normal probability plots of the residuals and applied normality tests (Wilk-Shapiro) to assess whether the transformation satisfied the model assumptions. I also examined the square root transformation to see if it properly normalized the data. The best transformation was the one which gave the largest (non-significant) p-value for the normality test statistic. If no transformation was available, the best transformation was listed as “no obvious transformation”.

12. In cases where a transformation better than log10 was available, I compared the LOELs from the log10 transformed data to the correctly transformed data to determine the impact of applying the ANOVA model to data that does not satisfy the assumptions.

13. We next examined the LOELs for each endpoint and laboratory, both using the log10 transformation applied in the given analysis, and using the “correct” transformation from table 1. This would be similar to the kind of analysis performed by Peddada and Haseman (2001) for the uterotrophic data.

14. We also compared benchmark doses across endpoints and laboratories, using log10 transformation and “correct” transformations. The benchmark dose method results in the BMD, which is defined here as the dose which causes the mean response to increase (or decrease) by twice the standard deviation of the control group. This definition of the BMD allows for better comparisons among endpoints with inherently different variability. The mean model fit was the Hill model, which is of the form

$$Y(dose) = intercept + v*dose^n/(k^n + dose^n).$$

15. Here *intercept*, *v*, *k* and *n* are parameters to be estimated. The US EPA=s benchmark dose software was used to fit the models. In cases where the Hill model did not converge, the highest dose group was iteratively dropped until convergence was obtained. This method was used since it was determined that most of the convergence problems were due to a flat dose-response relationship in the high dose region. The benchmark dose method depends on a good fit in the range of the benchmark dose, and not necessarily in the high dose region. Moderate lack of fit was not deemed a concern as long as the model visually fit the data in the range of the benchmark dose. When eliminating dose groups did not eliminate gross lack of fit or convergence problems, the power model was used. In three cases, the lower bound on the BMD (BMDL) was left as NA (not available) when the "correct" model otherwise fit, but the shape of the curve in the low-dose region would not allow the lower bound to converge. Ideally, more effort would be spent finding a model which would allow for computation of a BMDL, but for the sake of simplicity and consistency, this was not done here. The benchmark dose for all laboratories combined was calculated by adjusting the each animal's response by the corresponding LAB and LABxDOSE effects.

16. In our final analysis, we examined the implication of allowing the LAB term to be random in the overall ANOVA models for each endpoint. This allows us to generalise our conclusions to any laboratory as opposed to just the participating laboratories. Applying this type of analysis does not change the R^2 associated with dose, only the F-statistic.

Results: Phase 1a (TP) study

17. No single transformation adequately normalized the data across all labs and endpoints. However, the log10 transformation often stabilized the variance and normalized the residuals, even if a better transformation was available. Table 1 indicates which transformation was most appropriate for which each laboratory and endpoint. For some labs, no obvious transformation was available to transform the data to normality.

18. Unfortunately, the choice of transformation can have an impact on the resulting LOEL, as seen in tables 2 to 6. In five cases the better transformation yields a higher LOEL than does the log10 transformation. Despite this, the LOELs are still reasonably consistent across laboratories and endpoints, supporting the claim that the Hershberger assay is robust.

19. This analysis leads us to conclude that, for the purpose of this assay, applying one transformation to all laboratories and endpoints is acceptable for the sake of consistency and simplicity. However, more care will be required to ensure that some proper transformation is applied when individual laboratories run this assay for their own purposes in future.

20. Benchmark doses were computed for all endpoints and laboratories using the Hill model. As shown in table 7, the most sensitive chemical was VP and the least sensitive was GLANS. Benchmark doses were reasonably consistent across transformations and endpoints (tables 8 to 12). In some cases though, the choice of transformation can have a large impact on the resulting benchmark dose, highlighting the importance of using the correct variance-stabilizing transformation.

21. Allowing LAB to be random caused the F-statistic for DOSE to be between 35% and 62% of the F-statistic when LAB was fixed. In other words, treating LAB as a random effect caused DOSE to be less significant. However, DOSE was still highly significant for all endpoints.

Results: Phase 1b (Flutamide) study

22. As with the phase 1a study, no single transformation adequately normalised the data across all labs and endpoints. Table 13 indicates which transformation was most appropriate for which each laboratory and endpoint. For two models, no obvious transformation was available to transform the data to normality.

23. The resulting LOELs are displayed in tables 14 to 18. In four cases the better transformation yields a lower LOEL than does the log10 transformation. Despite this, the LOELs are still reasonably consistent across laboratories and endpoints, ranging from 0.1 to 1 for VP, SV, LABC and COWS. GLANS was the most variable endpoint, and LOELs ranged from 0.1 to 3 for that endpoint.

24. Based on benchmark dose estimates, SV was the most sensitive endpoint and GLANS was the least sensitive (table 19). As with the phase 1 study, the benchmark doses were reasonably consistent across endpoints and transformations (tables 20 to 24). In some cases though, the choice of transformation can have a large impact on the resulting benchmark dose, highlighting the importance of using the correct variance-stabilising transformation.

25. Allowing LAB to be random did not have as large an impact on the resulting F-statistic for dose as it did in the phase 1a study. F-statistics when LAB was random were occasionally larger than when LAB was fixed. This is because the LABxDOSE interactions were not as significant for the flutamide data. F-statistics wen LAB was random were between 50% and 125% of those when LAB was fixed. Again, this change was not enough to prevent DOSE from being a highly significant predictor for all endpoints.

Table 1: Transformation needed to normalize the phase 1a data

Variable	Overall model (all labs)	Individual labs
VP	Untransformed	1,4,14,15: LOG10 2,3,7,8,17: Untransformed 5,6,10,12,16: Square Root 9,13: no obvious transformation
SV	Square Root	All labs except 2, 6, 12, 13, 16, 17: LOG10 2, 6, 12, 13, 16, 17: Square Root
LABC	Square Root	1, 3, 4, 10, 12, 13: LOG10 6, 7, 16, 17: Untransformed 8, 15: Square Root 2, 5, 9, 14: no obvious transformation
GLANS	Untransformed	1, 9, 13, 15: LOG10 2, 3, 5, 6, 7, 14, 16, 17: Untransformed 8, 10, 12: Square Root 4: no obvious transformation
COWS	Untransformed	1, 8, 12, 13, 16: LOG10 2, 5, 7, 9, 15, 17:Untransformed 3, 4, 6, 10, 14: Square Root

Table 2: LOELs (mg/kg-bw/day) for VP from phase 1a study

Variable	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.1	Untransformed	0.1
Lab 1	0.1	Log10	0.1
Lab 2	0.1	Untransformed	0.2
Lab 3	0.1	Untransformed	0.1
Lab 4	0.2	Log10	0.2
Lab 5	0.1	Square Root	0.1
Lab 6	0.1	Square Root	0.1
Lab 7	0.2	Untransformed	0.4
Lab 8	0.1	Untransformed	0.1
Lab 9	0.1	No obvious transformation	-
Lab 10	0.1	Square Root	0.1
Lab 12	0.1	Square Root	0.1
Lab 14	0.1	Log10	0.1
Lab 13	0.1	No obvious transformation	-
Lab 15	0.1	Log10	0.1
Lab 16	0.1	Square Root	0.1
Lab 17	0.1	Untransformed	0.1

Table 3: LOELs (mg/kg-bw/day) for SV from phase 1a study

Variable	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.1	Square Root	0.1
Lab 1	0.1	Log10	0.1
Lab 2	0.1	Square Root	0.1
Lab 3	0.1	Log10	0.1
Lab 4	0.1	Log10	0.1
Lab 5	0.1	Log10	0.1
Lab 6	0.1	Square Root	0.1
Lab 7	0.2	Log10	0.2
Lab 8	0.1	Log10	0.1
Lab 9	0.1	Log10	0.1
Lab 10	0.1	Log10	0.1
Lab 12	0.1	Square Root	0.1
Lab 14	0.1	Log10	0.1
Lab 13	0.1	Square Root	0.1
Lab 15	0.1	Log10	0.1
Lab 16	0.1	Square Root	0.1
Lab 17	0.1	Square Root	0.1

Table 4: LOELs (mg/kg-bw/day) for LABC from phase 1a study

Variable	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.1	Untransformed	0.1
Lab 1	0.1	Log10	0.1
Lab 2	0.1	No obvious transformation	-
Lab 3	0.1	Log10	0.1
Lab 4	0.1	Log10	0.1
Lab 5	0.1	No obvious transformation	-
Lab 6	0.1	Untransformed	0.1
Lab 7	0.1	Untransformed	0.1
Lab 8	0.1	Square Root	0.1
Lab 9	0.2	No obvious transformation	-
Lab 10	0.1	Log10	0.1
Lab 12	0.1	Log10	0.1
Lab 14	0.1	No obvious transformation	-
Lab 13	0.1	Log10	0.1
Lab 15	0.1	Square Root	0.1
Lab 16	0.1	Untransformed	0.1
Lab 17	0.1	Untransformed	0.1

Table 5: LOELs (mg/kg-bw/day) for GLANS from phase 1a study

Variable	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.1	Untransformed	0.1
Lab 1	0.1	Log10	0.1
Lab 2	0.1	Untransformed	0.1
Lab 3	0.1	Untransformed	0.1
Lab 4	1.6	No obvious transformation	-
Lab 5	0.1	Untransformed	0.1
Lab 6	0.1	Untransformed	0.1
Lab 7	0.1	Untransformed	0.1
Lab 8	0.1	Square Root	0.1
Lab 9	0.1	Log10	0.1
Lab 10	0.1	Square Root	0.1
Lab 12	0.1	Square Root	0.1
Lab 14	0.1	Untransformed	0.1
Lab 13	0.1	Log10	0.1
Lab 15	0.1	Log10	0.1
Lab 16	0.1	Untransformed	0.1
Lab 17	0.1	Untransformed	0.1

Table 6: LOELs (mg/kg-bw/day) for COWS from phase 1a study

Variable	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.1	Square Root	0.1
Lab 1	0.1	Log10	0.1
Lab 2	0.1	Untransformed	0.2
Lab 3	0.1	Square Root	0.1
Lab 4	0.1	Square Root	0.1
Lab 5	0.1	Untransformed	0.1
Lab 6	0.1	Square Root	0.1
Lab 7	0.2	Untransformed	0.4
Lab 8	0.1	Log10	0.1
Lab 9	0.1	Untransformed	0.1
Lab 10	0.1	Square Root	0.1
Lab 12	0.1	Log10	0.1
Lab 14	0.1	Square Root	0.1
Lab 13	0.1	Log10	0.1
Lab 15	0.1	Untransformed	0.1
Lab 16	0.1	Log10	0.1
Lab 17	0.1	Untransformed	0.1

Table 7: Benchmark Doses (BMDs) (mg/kg-bw/day) for phase 1a study combined across all laboratories and listed in order of decreasing potency

Variable	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
VP	0.066 (0.061)	Untransformed	0.089 (0.080)
SV	0.083 (0.077)	Square Root	0.120 (0.107)
COWS	0.079 (0.0676)	Square Root	0.129 (0.112)
LABC	0.184 (0.152)	Untransformed	0.310 (0.256)
GLANS	0.226 (0.196)	Untransformed	0.256 (0.189)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 8: Benchmark Doses (BMDs) (mg/kg-bw/day) for VP from phase 1a study

Variable	BMD, Log10 transformed (BMDL)	"Correct" transformation	BMD, correct transformed (BMDL*)
Overall	0.066 (0.061)	Untransformed	0.089 (0.080)
Lab 1	0.042 (0.024)	Log10	0.042 (0.024)
Lab 2	0.041 (0.019)	Untransformed	0.166 (0.109)
Lab 3	0.050 (0.036)	Untransformed	0.119 (0.092)
Lab 4	0.116 (0.077)	Log10	0.116 (0.077)
Lab 5	0.013 (0.004)	Square Root	0.041 (0.023)
Lab 6	0.033 (0.016)	Square Root	0.158 (0.112)
Lab 7	0.255 (0.140)	Untransformed	0.401 (0.239)
Lab 8	0.038 (0.026)	Untransformed	0.108 (0.083)
Lab 9	0.071 (0.032)	No obvious transformation	-
Lab 10	0.018 (0.007)	Square Root	0.046 (0.025)
Lab 12	0.022 (0.011)	Square Root	0.044 (0.028)
Lab 14	0.043 (0.028)	Log10	0.043 (0.028)
Lab 13	0.021 (0.008)	No obvious transformation	-
Lab 15	0.014 (0.007)	Log10	0.014 (0.007)
Lab 16	0.019 (0.055)	Square Root	0.049 (0.028)
Lab 17	0.020 (0.009)	Untransformed	0.097 (0.061)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 9: Benchmark Doses (BMDs) (mg/kg-bw/day) for SV from phase 1a study

Variable	BMD, Log10 transformed (BMDL)	"Correct" transformation	BMD, correct transformed (BMDL*)
Overall	0.083 (0.077)	Square Root	0.120 (0.107)
Lab 1	0.032 (0.020)	Log10	0.032 (0.020)
Lab 2	0.085 (0.054)	Square Root	0.148 (0.099)
Lab 3	0.070 (0.058)	Log10	0.070 (0.058)
Lab 4	0.042 (0.032)	Log10	0.042 (0.032)
Lab 5	0.012 (0.005)	Log10	0.012 (0.005)
Lab 6	0.048 (0.032)	Square Root	0.109 (0.073)
Lab 7	0.115 (0.085)	Log10	0.115 (0.085)
Lab 8	0.027 (0.016)	Log10	0.027 (0.016)
Lab 9	0.080 (0.055)	Log10	0.080 (0.055)
Lab 10	0.039 (0.024)	Log10	0.039 (0.024)
Lab 12	0.023 (0.014)	Square Root	0.048 (0.033)
Lab 14	0.050 (0.036)	Log10	0.050 (0.036)
Lab 13	0.035 (0.024)	Square Root	0.068 (0.050)
Lab 15	0.017 (0.012)	Log10	0.017 (0.012)
Lab 16	0.013 (0.007)	Square Root	0.034 (0.023)
Lab 17	0.043 (0.033)	Square Root	0.046 (0.029)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 10: Benchmark Doses (BMDs) (mg/kg-bw/day) for LABC from phase 1a study

Variable	BMD, Log10 transformed (BMDL)	"Correct" transformation	BMD, correct transformed (BMDL*)
Overall	0.184 (0.152)	Untransformed	0.310 (0.256)
Lab 1	0.047 (0.026)	Log10	0.047 (0.026)
Lab 2	0.079 (0.045)	No obvious transformation	-
Lab 3	0.069 (0.043)	Log10	0.069 (0.043)
Lab 4	0.042 (0.028)	Log10	0.042 (0.028)
Lab 5	0.031 (0.021)	No obvious transformation	-
Lab 6	0.075 (0.054)	Untransformed	0.149 (0.096)
Lab 7	0.098 (0.043)	Untransformed	0.170 (0.110)
Lab 8	0.048 (0.037)	Square Root	0.038 (0.026)
Lab 9	0.257 (0.186)	No obvious transformation	-
Lab 10	0.015 (0.006)	Log10	0.015 (0.006)
Lab 12	0.052 (0.040)	Log10	0.052 (0.040)
Lab 14	0.054 (0.039)	No obvious transformation	-
Lab 3	0.031 (0.020)	Log10	0.031 (0.020)
Lab 15	0.066 (0.050)	Square Root	0.025 (0.012)
Lab 16	0.007 (0.001)	Untransformed	0.028 (0.010)
Lab 17	0.046 (0.027)	Untransformed	0.063 (0.038)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 11: Benchmark Doses (BMDs) (mg/kg-bw/day) for GLANS from phase 1a study

Variable	BMD, Log10 transformed (BMDL)	"Correct" transformation	BMD, correct transformed (BMDL*)
Overall	0.226 (0.196)	Untransformed	0.256 (0.189)
Lab 1	0.090 (0.066)	Log10	0.090 (0.066)
Lab 2	0.161 (1.3e-9)	Untransformed	0.848 (2.4e-20)
Lab 3	0.095 (0.034)	Untransformed	0.111 (0.051)
Lab 4	0.705 (0.384)	No obvious transformation	-
Lab 5	0.029 (0.017)	Untransformed	0.037 (0.002)
Lab 6	0.111 (0.056)	Untransformed	0.140 (0.082)
Lab 7	0.093 (0.005)	Untransformed	0.189 (0.073)
Lab 8	0.074 (0.055)	Square Root	0.026 (0.003)
Lab 9	0.051 (0.002)	Log10	0.051 (0.002)
Lab 10	0.059 (0.038)	Square Root	0.066 (0.045)
Lab 12	0.047 (0.026)	Square Root	0.063 (0.034)
Lab 14	0.032 (0.016)	Untransformed	0.066 (0.038)
Lab 13	0.057 (0.036)	Log10	0.057 (0.036)
Lab 15	0.038 (0.001)	Log10	0.038 (0.001)
Lab 16	0.205 (0.152)	Untransformed	0.061 (0.025)
Lab 17	0.057 (0.007)	Untransformed	0.121 (0.063)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 12: Benchmark Doses (BMDs) (mg/kg-bw/day) for COWS from phase 1a study

Variable	BMD, Log10 transformed (BMDL)	"Correct" transformation	BMD, correct transformed (BMDL*)
Overall	0.079 (0.0676)	Square Root	0.129 (0.112)
Lab 1	0.036 (0.024)	Log10	0.036 (0.024)
Lab 2	0.095 (0.059)	Untransformed	0.105 (0.119)
Lab 3	0.031 (0.020)	Square Root	0.049 (0.035)
Lab 4	0.022 (0.001)	Square Root	0.042 (0.021)
Lab 5	0.017 (0.001)	Untransformed	0.107 (0.050)
Lab 6	0.036 (0.025)	Square Root	0.077 (0.058)
Lab 7	0.308 (0.144)	Untransformed	0.408 (0.225)
Lab 8	0.042 (0.019)	Log10	0.042 (0.019)
Lab 9	0.037 (0.004)	Untransformed	0.108 (0.065)
Lab 10	0.015 (0.006)	Square Root	0.026 (0.012)
Lab 12	0.035 (0.011)	Log10	0.035 (0.011)
Lab 14	0.029 (0.024)	Square Root	0.075 (0.058)
Lab 13	0.038 (0.019)	Log10	0.038 (0.019)
Lab 15	0.007 (0.0004)	Untransformed	0.098 (0.048)
Lab 16	0.025 (0.011)	Log10	0.025 (0.011)
Lab 17	0.045 (0.012)	Untransformed	0.086 (0.049)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 13: Transformation needed to normalize the phase 1b data

Variable	Overall model (all labs)	Individual labs
VP	0.2TP: Square root 0.4TP: Log10	0.2TP: 11, 12, 17: Square root 5, 15: Log10 13: Untransformed 0.4TP: 5, 10, 12: Square root 13, 15: Log10
SV	0.2TP: Log10 0.4TP: Square root	0.2TP: 5: Untransformed 11, 12, 13, 15: Log10 17: Square root 0.4TP: 5: Untransformed 10, 13, 15: Square root 12: Log10
LABC	0.2TP: Square root 0.4TP: Untransformed	0.2TP: 5, 13: Square root 11, 13: Log10 12, 15: Untransformed 0.4TP: 5, 15: Untransformed 10: Square root 12, 13: Log10
GLANS	0.2TP: No obvious transformation 0.4TP: Square root	0.2TP: 5: No obvious transformation 11, 12, 17: Untransformed 13: Log10 15: Square root 0.4TP: 5: Square root 10, 13, 15: Untransformed 12: Log10
COWS	0.2TP: Log10 0.4TP: Log10	0.2TP: all except 13: Log10 13: Untransformed 0.4TP: 5: Square root 10, 13: Untransformed 12, 15: Log10

Table 14: LOELs (mg/kg-bw/day) for VP from phase 1b study

Lab	TP Level	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
All Labs	0.2	0.3	Square root	0.3
	0.4	0.3	Log10	0.3
5	0.2	1	Log10	1
	0.4	1	Square root	1
11	0.2	1	Square root	1
10	0.4	0.1	Square root	0.1
12	0.2	0.1	Square root	0.1
	0.4	1	Square root	0.3
13	0.2	0.3	Untransformed	0.3
	0.4	1	Log10	1
15	0.2	1	Log10	1
	0.4	0.1	Log10	0.1
17	0.2	1	Square root	1

Table 15: LOELs (mg/kg-bw/day) for SV from phase 1b study

Lab	TP Level	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
All Labs	0.2	0.1	Log10	0.1
	0.4	0.1	Square root	0.1
5	0.2	0.1	Untransformed	0.1
	0.4	0.3	Untransformed	0.1
11	0.2	0.3	Log10	0.3
10	0.4	0.1	Square root	0.1
12	0.2	0.3	Log10	0.3
	0.4	1	Log10	1
13	0.2	1	Log10	1
	0.4	1	Square root	1
15	0.2	0.3	Log10	0.3
	0.4	0.1	Square root	0.1
17	0.2	0.1	Square root	0.1

Table 16: LOELs (mg/kg-bw/day) for LABC from phase 1b study

Lab	TP Level	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.2	0.3	Square root	0.3
	0.4	0.1	Untransformed	0.1
5	0.2	0.3	Square root	0.3
	0.4	1	Untransformed	1
11	0.2	1	Log10	1
10	0.4	0.3	Square root	0.3
12	0.2	1	Untransformed	0.3
	0.4	1	Log10	1
13	0.2	1	Square root	1
	0.4	0.3	Log10	0.3
15	0.2	1	Untransformed	1
	0.4	0.3	Untransformed	0.1
17	0.2	1	Log10	1

Table 17: LOELs (mg/kg-bw/day) for GLANS from phase 1b study

Lab	TP Level	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
Overall	0.2	0.1	No obvious transformation	-
	0.4	0.3	Square root	0.3
5	0.2	1	No obvious transformation	-
	0.4	3	Square root	3
11	0.2	3	Untransformed	3
10	0.4	1	Untransformed	1
12	0.2	1	Untransformed	1
	0.4	1	Log10	1
13	0.2	1	Log10	1
	0.4	1	Untransformed	1
15	0.2	0.1	Square root	0.1
	0.4	3	Untransformed	3
17	0.2	0.3	Untransformed	0.3

Table 18: LOELs (mg/kg-bw/day) for COWS from phase 1b study

Lab	TP Level	LOEL, Log10 transformed	"Correct" transformation	LOEL, correct transformation
All Labs	0.2	0.3	Log10	0.3
	0.4	0.1	Log10	0.1
5	0.2	1	Log10	1
	0.4	1	Square root	1
11	0.2	1	Log10	1
10	0.4	1	Untransformed	0.3
12	0.2	0.3	Log10	0.3
	0.4	0.3	Log10	0.3
13	0.2	0.3	Untransformed	0.3
	0.4	0.3	Untransformed	0.3
15	0.2	0.3	Log10	0.3
	0.4	1	Log10	1
17	0.2	1	Log10	1

Table 19: Benchmark Doses (BMDs) (mg/kg-bw/day) for phase 1b study combined across all laboratories and listed in order of decreasing potency

Variable	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
GLANS	0.2TP: 0.502 (NA**) 0.4TP: 1.308 (NA)	0.2TP: Untransformed 0.4TP: Square root	0.2TP: 0.332 (0.218) 0.4TP: 1.067 (0.825)
SV	0.2TP: 0.542 (0.477) 0.4TP: 0.510 (NA)	0.2TP: Log10 0.4TP: Square root	0.2TP: 0.542 (0.477) 0.4TP: 0.311 (0.271)
VP	0.2TP: 0.603 (0.512) 0.4TP: 0.609 (0.525)	0.2TP: Square root 0.4TP: Log10	0.2TP: 0.499 (0.418) 0.4TP: 0.609 (0.525)
LABC	0.2TP: 1.115 (1.007) 0.4TP: 0.501 (NA)	0.2TP: Square root 0.4TP: Untransformed	0.2TP: 0.917 (0.790) 0.4TP: 0.293 (0.240)
COWS	0.2TP: 1.333 (NA) 0.4TP: 0.948 (0.737)	0.2TP: Log10 0.4TP: Log10	0.2TP: 1.333 (NA) 0.4TP: 0.948 (0.737)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

Table 20: Benchmark Doses (BMDs) (mg/kg-bw/day) for VP from phase 1b study

Lab	TP Level	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
Overall	0.2	0.603 (0.512)	Square root	0.499 (0.418)
	0.4	0.609 (0.525)	Log10	0.609 (0.525)
5	0.2	0.365 (0.254)	Log10	0.365 (0.254)
	0.4	0.436 (0.250)	Square root	0.250 (0.120)
11	0.2	0.253 (0.086)	Square root	0.176 (0.059)
10	0.4	0.161 (0.098)	Square root	0.065 (0.033)
12	0.2	0.225 (0.128)	Square root	0.159 (0.078)
	0.4	0.418 (0.224)	Square root	0.231 (0.097)
13	0.2	0.198 (0.091)	Untransformed	0.096 (0.037)
	0.4	0.787 (0.473)	Log10	0.787 (0.473)
15	0.2	0.397 (0.241)	Log10	0.397 (0.241)
	0.4	0.374 (0.255)	Log10	0.374 (0.255)
17	0.2	0.353 (0.186)	Square root	0.244 (0.124)

*BMDL = 95% Lower Confidence Limit on the BMD

Table 21: Benchmark Doses (BMDs) (mg/kg-bw/day) for SV from phase 1b study

Lab	TP Level	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
Overall	0.2	0.542 (0.477)	Log10	0.542 (0.477)
	0.4	0.510 (NA**)	Square root	0.311 (0.271)
5	0.2	0.190 (0.107)	Untransformed	0.046 (0.023)
	0.4	0.350 (0.246)	Untransformed	0.093 (0.048)
11	0.2	0.183 (0.076)	Log10	0.183 (0.076)
10	0.4	0.232 (0.163)	Square root	0.118 (0.072)
12	0.2	0.169 (0.096)	Log10	0.169 (0.096)
	0.4	0.435 (0.264)	Log10	0.435 (0.264)
13	0.2	0.193 (0.088)	Log10	0.193 (0.088)
	0.4	0.526 (0.312)	Square root	0.349 (0.183)
15	0.2	0.168 (0.073)	Log10	0.168 (0.073)
	0.4	0.316 (0.176)	Square root	0.098 (0.055)
17	0.2	0.158 (0.067)	Square root	0.072 (0.025)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

Table 22: Benchmark Doses (BMDs) (mg/kg-bw/day) for LABC from phase 1b study

Lab	TP Level	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
Overall	0.2	1.115 (1.007)	Square root	0.917 (0.790)
	0.4	0.501 (NA**)	Untransformed	0.293 (0.240)
5	0.2	0.184 (0.105)	Square root	0.144 (0.080)
	0.4	0.452 (0.260)	Untransformed	0.259 (0.140)
11	0.2	0.281 (0.125)	Log10	0.281 (0.125)
10	0.4	0.542 (NA**)	Square root	0.250 (0.145)
12	0.2	0.203 (0.099)	Untransformed	0.164 (0.078)
	0.4	0.356 (NA)	Log10	0.356 (NA)
13	0.2	0.191 (0.053)	Square root	0.162 (0.042)
	0.4	0.306 (0.141)	Log10	0.306 (0.141)
15	0.2	0.278 (0.147)	Untransformed	0.228 (0.115)
	0.4	0.167 (0.083)	Untransformed	0.057 (0.028)
17	0.2	0.545 (0.218)	Log10	0.545 (0.218)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

Table 23: Benchmark Doses (BMDs) (mg/kg-bw/day) for GLANS from phase 1b study

Lab	TP Level	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
Overall	0.2	0.502 (NA**)	Untransformed	0.332 (0.218)
	0.4	1.308 (NA)	Square root	1.067 (0.825)
5	0.2	0.390 (NA)	No obvious transformation	-
	0.4	0.902 (NA)	Square root	0.738 (0.328)
11	0.2	0.947 (NA)	Untransformed	0.893 (0.234)
10	0.4	0.350 (0.179)	Untransformed	0.239 (0.127)
12	0.2	0.219 (0.073)	Untransformed	0.177 (0.053)
	0.4	0.557 (NA)	Log10	0.557 (NA)
13	0.2	0.539 (0.066)	Log10	0.539 (0.066)
	0.4	0.374 (0.150)	Untransformed	0.274 (0.091)
15	0.2	2.612 (NA)	Square root	9.194 (0.776)
	0.4	0.310 (0.074)	Untransformed	0.256 (0.061)
17	0.2	0.012 (NA)	Untransformed	0.008 (0.053)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

Table 24: Benchmark Doses (BMDs) (mg/kg-bw/day) for COWS from phase 1b study

Lab	TP Level	BMD, Log10 transformed (BMDL*)	"Correct" transformation	BMD, correct transformation (BMDL*)
Overall	0.2	1.333 (NA**)	Log10	1.333 (NA)
	0.4	0.948 (0.737)	Log10	0.948 (0.737)
5	0.2	0.331 (0.138)	Log10	0.331 (0.138)
	0.4	0.736 (0.237)	Square root	0.736 (0.237)
11	0.2	0.229 (0.109)	Log10	0.229 (0.109)
10	0.4	0.721 (0.345)	Untransformed	0.136 (0.077)
12	0.2	0.171 (0.084)	Log10	0.171 (0.084)
	0.4	0.169 (0.084)	Log10	0.169 (0.084)
13	0.2	0.243 (0.052)	Untransformed	0.120 (0.030)
	0.4	0.423 (0.213)	Untransformed	0.160 (0.060)
15	0.2	0.962 (0.285)	Log10	0.962 (0.285)
	0.4	0.558 (0.313)	Log10	0.558 (0.313)
17	0.2	0.569 (0.058)	Log10	0.569 (0.058)

*BMDL = 95% Lower Confidence Limit on the BMD
**NA = Lower bound computation did not converge

ANNEX - 9

AGREED RESEARCH PROPOSAL FOR PHASE-2 OF THE OECD HERSHBERGER ASSAY INTERLABORATORY STUDY

INTRODUCTION

1. The following is the research proposal for Phase-2 of the OECD Hershberger Assay interlaboratory validation study drafted by L. Earl Gray Jr. based on discussions and recommendations of the 3rd VMG-mammalian and agreed by the VMG-mammalian in written procedure following a review of the draft Phase 1 report in April – May 2002.

2. Phase-1 of the OECD validation study of the rodent Hershberger Assay has been completed [see report ENV/JM/TG/EDTA(2002)1/REV2]. The 17 laboratories were successful in demonstrating the androgenic activity of testosterone propionate (TP) (Phase-1a), and all 7 laboratories that participated in Phase-1b successfully demonstrated the anti-androgenic activity of flutamide.

EXPERIMENTAL DESIGN OF PHASE-2 OF THE HERSHBERGER VALIDATION STUDY

3.

3. Phase-2 of the validation study is designed to demonstrate the ability of the OECD protocols to reproducibly detect the activity of weak androgen agonists, and antagonists that act through different molecular mechanisms.

Protocol to be used

4. The protocol that was used in the Phase-1a agonism and Phase-1b antagonism studies will also be used for the Phase-2 agonism and antagonism studies (see attachment to this Annex). The sex accessory tissues to be examined for weight gain (fresh weight) will be the ventral prostate, seminal vesicles plus coagulating glands, levator ani plus bulbocavernosus muscles, Cowper's glands, and glans penis. The participating laboratories will have the option to also weigh the ventral prostate after fixation. Other mandatory measurements will be body weight and adrenal weight. Liver and kidney weights are optional.

5. No changes are recommended in the Phase-1 protocols with regards to duration, age of the animals, sample size, and endpoints (other than the optional weight of liver and kidney), because the Phase-1 protocol performed extremely well in the TP and flutamide dose-response studies. Shortening the treatment duration or using older animals would reduce the sensitivity of the assay.

6. All of the androgen-dependent endpoints evaluated in Phase-1 should be retained, because most of them have a unique characteristic in terms of responsiveness to different chemicals, or sensitivity at different ranges of the androgen dose-response curve. Inclusion of the paired adrenal gland weights as a required endpoint will broaden the types of endocrine-disrupting chemicals that can be detected. Adrenal weight and function is affected by androgen receptor antagonists like vinclozolin, procymidone, flutamide, and linuron, agonists like trenbolone, and inhibitors of steroidogenic P450 enzymes, like ketoconazole.

7. Based upon the results of the Phase-1b TP and flutamide dose-response studies, it is recommended that the Phase-2 antagonist protocol use 0.4 mg TP/kg/day. This was the approximate ED_{70} for three of the five androgenic tissues, and was as sensitive to flutamide as was the 0.2 mg/kg/day dose regimen, if not more sensitive. In addition, the tissues from the 0.4 mg TP animals are larger and easier to weigh.

8. The Phase-1 studies showed that the liver and kidney weights, and the serum testosterone and lutenizing hormone levels, were not sensitive indicators of androgenic or anti-androgenic effects. These measurements will not be mandatory in Phase-2.

Laboratories

9. It is estimated that 14 laboratories will participate in Phase-2 of the Hershberger assay validation study; each will test one or two chemicals. The laboratories may include 6 from Japan, 6 from Europe, 1 from Korea, and 1 from the US. However, it is possible that the number of laboratories will be different from those participating in Phase 1.

10. In order to allow adequate interlaboratory comparison, statistical analysis and scientific interpretation of results, Phase-2 of the validation will be organised in a way as to preferably meet the following:

- Each chemical will ideally be tested in at least four laboratories, preferably geographically balanced;
- One laboratory each from Europe, Japan, Korea and the US are requested to test trenbolone as an agonist using the oral route, using at least the following treatment groups:
  - Oral route: 0, 0.3, 1.5, 8, 40 mg/kg/day;
- Four laboratories from Europe are requested to evaluate methyl testosterone as an agonist, and linuron as antagonist against testosterone propionate;
- Four laboratories from Japan will evaluate *p,p'* –DDE and finasteride as antagonist against testosterone propionate.

Blind Testing and Testing of Negative Chemicals

11. The testing of negative chemicals as well as "bling testing"will be addressed in a third phase of the validation after the successful completion of Phase 2.

Test Chemicals

12. The criteria for selecting chemicals for Phase-2 of the OECD Hershberger interlaboratory study were:

- The chemical should display a mechanism of action that is detectable by the OECD Hershberger assay protocol, i.e., it is either an androgen receptor agonist (androgen) or antagonist (anti-androgen).
- The chemical's androgenic or anti-androgenic effects *in vivo* should be well documented. There should be sufficient information to assure that effects can be produced at dose levels that do not induce severe systemic toxicity (i.e., below the MTD). An MTD can be operationally defined as not causing death, a reduction in body weight at necropsy of greater than 10 %, or any signs of overt toxicity.
- The chemical is known or suspected to alter reproductive development or pregnancy via an androgen receptor-mediated mechanism. This criterion is included because screening assays will be used to detect chemicals that have the potential to be developmental reproductive toxicants.
- The chemical must be commercially available at "reasonable" cost

13. The following chemicals have been selected for testing in Phase-2:

- Androgens: testosterone propionate, trenbolone, and methyl testosterone
- Anti-androgens: vinclozolin, procymidone, linuron, *p,p'*-DDE, and finasteride.

14. Concerns have been expressed about including vinclozolin and procymidone in the Phase-2

studies because they share a common mechanism of toxicity with flutamide. However, although vinclozolin and procymidone do share a common mechanism of toxicity with flutamide they are about 1-2 orders magnitude less potent in both the Hershberger assay and as developmental toxicants. Therefore, the ability of the OECD protocol to detect these weaker chemicals is an important component of this validation study.

15. Another concern expressed was that "neutral" substances should be preferred over commercially important substances, because isolated selections will generally lead to unpredictable advantages or disadvantages for the individual producer. Because the androgenic and anti-androgenic activities of these chemicals are well-documented, and there are very few well-studied androgens and anti-androgens, their elimination from the study would leave too few well-defined substances (e.g., *p,p*'-DDE) to support a validation study.

16. There was a lack of consensus within the VMG regarding the need to include in Phase 2 negative chemicals, i.e. chemicals known to have no androgenic or anti-androgenic activity. A number of experts suggested that at least two negative chemicals should be considered, and suggested to use metabolic toxicants that would give body loss. Other experts strongly believed that vehicle controls, together with weak and strong actors provide adequate data to assess the validity of the test. Besides, the VMG was unable to suggest negative chemicals with otherwise sufficient toxicity. From an animal welfare point of view and because of the lack of consensus at the VMG, negative chemcials are not formally included in Phase 2. Although participating laboratories are free to add negative chemicals to the core set of androgens and antiandrogens that are formally included in this phase, a follow-up third phase will address this issue, as well as blind testing, specifically.

17. The following provides scientific justification for the selection of the recommended chemicals.

DHT synthesis inhibitors

18. Finasteride is a 5-α reductase inhibitor. It should be administered orally at 1, 3, and 10 mg/kg/day (as per Di Salle *et al*., *J Steroid Biochem Mol Biol* 48, 241-248, 1994). This chemical will alter ventral prostate and seminal vesicle weights to a much greater degree than the levator ani plus bulbocavernosus muscle (LABC) weight because they rely on 5-α reductase to convert testosterone to 5-alpha dihydrotestosterone (DHT). An alternative 5-α reductase inhibitor would be turosteride (Di Salle *et al*., 1994) at 1, 3, and 10 mg/kg/day.

Androgens

19. Testosterone propionate was the androgen used in Phase-1. For this reason, it is recommended for use in Phase-2 as the control against which the anti-androgen responses will be measured.

20. Trenbolone (17-β) is an anabolic steroid used in the US to promote muscle growth in cattle, and is excreted into the environment, in part, in an active form. Like several other anabolic steroids, trenbolone's "anabolic effects" on muscles, including the LABC, far exceeds it's "androgenic effects" on the ventral prostate and seminal vesicles. This pattern of responses demonstrates the importance of measuring LABC weights in the Hershberger assay. When administered s.c., trenbolone is as potent as is testosterone in stimulating LABC growth. It is recommended that trenbolone be administered s.c. at 0.2, 0.4, and 0.8 mg/kg/day.

21. Like many steroids, trenbolone is less effective when take orally than via the s.c. route. For oral administration in oil, trenbolone should be given at 1, 5 and 10 mg/kg/day. If administered in 2.5 ml corn oil/kg, the 10 mg/kg/day dose remains as a suspension and requires a few days stirring prior to use to

obtain a uniform, stable suspension. Use of 5 ml corn oil/kg may result in the formation of a solution, but this has not been determined.

22. Methyl testosterone differs from testosterone and trenbolone in that it is relatively potent when administered orally. Dose levels similar to trenbolone are recommended.

Androgen-receptor antagonists (Anti-androgens)

23. Vinclozolin and procymidone are dicarboximide fungicides that display AR antagonist activities *in vitro* and *in vivo*. The are effective in the Hershberger assay at doses as low as 25 mg/kg/day (Price *et al.*, 1999; Gray *et al.*, 2001). Although they act via the same mechanism of action as does flutamide, these fungicides are at least ten fold less potent *in vitro* and *in vivo*. One or the other could be used, and, if so, doses of 50, 100, and 200 mg/kg/day orally are recommended. This dose range has been used for both chemicals several times in the Lead Laboratory with consistent success.

24. Linuron is a urea-based herbicide that displays weak AR agonist activity. It inhibits growth of all androgen-dependent tissues at 100 mg/kg/day in the Hershberger assay using the castrate, TP-treated immature (but not adult) male rat. Higher dosage levels should not be used because of the induction of overt neurotoxicity (e.g., salivation, lacrimation, urination), as reported by Cook *et al.*, 1993. Dosage levels of 25, 50, and 100 mg/kg/day are recommended; this should produce coverage of the majority of the dose-effect curve, from little effect to a rather large reduction in organ weight gains. The value of this chemical, like *p,p'*-DDE, is that the AR-mediated effects in the Hershberger assay occur at dose levels just below those that induce overt toxicity (body weight reductions of 10%, which is beyond the MTD). The sensitivity of the OECD Phase-1 Hershberger protocol (a ten-day assay with immature, castrate male rat) with linuron has been demonstrated. When linuron was administered to immature male rats at 100 mg/kg/day for ten days, it produced a significant reduction in the ventral prostate, seminal vesicle, and LABC weight gains. Linuron had no effect on the weights of these tissues in a 7-day assay using adult (rather than immature) castrate males or a 5-day (rather than 10-day) assay with castrate-immature males (Lambright *et al.*, 2000). In contrast to its effects in the OECD Hershberger protocol, linuron does not produce reproducible anti-androgenic effects in adult, intact male rats, even at an overtly toxic dose of 200 mg/kg/day (Cook *et al.*, 1993; Gray *et al.*, 1999). In the pubertal male, linuron produces a delay in the onset of puberty, as indicated by a delay in preputial separation. For these reasons, it is important to include linuron in the Phase-2 tests because the results will illustrate the sensitivity and tissue-specific selectivity of the OECD protocol.

25. *p,p'*-DDE is an AR antagonist that produces it's most dramatic and reproducible Hershberger assay effects in the immature animals following 10-day dosing. Several different laboratories that performed the Hershberger assay using castrate immature rats have found robust changes in androgen-dependent tissue weight gains when *p,p'*-DDE is administered at 100 mg/kg/day. A similar exposure to adult castrate rats produced much smaller changes in organ weight gains (Yamada *et al.*, 2001). In contrast, administration of *p,p'*-DDE was completely without anti-androgenic effect in the intact male rat, even at overtly toxic doses of 200-300 mg/kg/day (O'Conner *et al.*, 2000). In a pubertal male assay, *p,p'*-DDE significantly delayed puberty by a few days in male rats, but sex accessory glands and serum hormone levels were not altered by 30 days of treatment at 100 mg/kg/day. For these reasons, it is important to include *p,p'*-DDE in Phase-2 of the validation study because the results will illustrate the sensitivity of the OECD protocol.

JUSTIFICATIONS FOR CHEMICAL DOSING REGIMENS

26. The following is a brief description of the literature about the anti-androgenic and developmental reproductive effects, and mechanisms of action, of several of the chemicals selected for inclusion in Phase-

2 of the OECD validation of the Hershberger assay. If the Hershberger assay fails to detect any of these chemicals it is an indication that the assay is producing serious false negatives, and that chemicals that cause reproductive malformation are being missed.

ANDROGEN-RECEPTOR ANTAGONISTS

Mechanisms of Action: vinclozolin, procymidone, *p,p'*-DDE, and linuron

27. The pesticides vinclozolin, procymidone, *p,p'*-DDE, DDE, and linuron are AR antagonists. The vinclozolin metabolites, M1 and M2, and procymidone, *p,p'*-DDE (and its metabolites), and linuron competitively inhibit the binding of endogenous androgens to the human androgen receptor (hAR), and therefore inhibit subsequent androgen-induced gene expression. It also has been demonstrated that vinclozolin, *p,p'*-DDE (Kelce et al., 1997), and linuron (Lambright et al., 2000) alter androgen-dependent ventral prostatic gene expression *in vivo*. None of these pesticides appear to display significant affinity for the estrogen receptor, or inhibit 5α-reductase *in vitro* (Kelce *et al.*, 1995; Waller *et al.*, 1996a), although M1 binds the rat progesterone receptor, albeit with relatively low affinity (Laws *et al.*, 1996).

Dose-Response Effects of Vinclozolin on Rat Reproductive Development

28. Vinclozolin-treated male offspring display female-like anogenital distance (AGD) at birth, retained nipples, cleft phallus with hypospadias, undescended testes, vaginal pouch, epididymal granulomas, and small-to-absent sex accessory glands. An examination of mating behaviour in these males indicated that vinclozolin-treatment did not alter mounting behaviour, based upon the percentage of male mounting or latencies to mount, but malformed treated males were incapable of attaining intromission (Gray and Ostby 1998, Gray *et al.*, 1994). Lactational transfer of vinclozolin to the neonatal rat does not appear to provide sufficient levels of the active metabolites to affect the organization of rough-and-tumble play behaviour, because pups directly exposed to vinclozolin show female-like rough and tumble play levels when measured during peripubertal life (Hotchkiss *et al.*, 2001). Dose-response curves for different effects of vinclozolin vary in shape and ED_{50} values for different androgen-dependent tissues. Some of these dose response curves failed to display an obvious threshold, i.e. anogenital distance, induction of areolas, and ventral prostate weight (Gray *et al.*, 1999b), and appear linear in the low-dose range. Wolf *et al.* (2000) found that the most sensitive period of fetal development to the antiandrogenic effects of vinclozolin was on gestational days (GD) 16-17, but some malformations and other effects also were seen in male rat offspring dosed with vinclozolin on GD 14-15, and GD 18-19.

Effect of Vinclozolin Treatment on Pubertal Development of the Male Rat

29. Androgens play a key role in pubertal maturation in young males (Korenbrot *et al.*, 1977) and anti-androgens delay this process. Peripubertal treatment with vinclozolin (Monosson *et al.*, 1999), *p,p'*-DDE (Kelce *et al.*, 1997), methoxychlor (Gray *et al.*, 1989), linuron, or di-*n*-butyl phthalate (Gray *et al.*, 1999a) delay the onset of androgen-dependent preputial separation (PPS). This model appears to have potential as an assay to screen for endocrine-disrupting chemicals (EDC), the assay being more sensitive to AR antagonists than the Adult Male Assay (O'Conner *et al.*, 1999), but slightly less sensitive than the Hershberger Assay.

30. Monosson *et al.* (1999) conducted a study to examine the effects of peripubertal oral administration of vinclozolin on morphological landmarks of puberty, hormone levels, and sex accessory gland development in male rats. They also examined the effects of vinclozolin on AR distribution in the target cells and measured serum levels of vinclozolin, M1, and M2. Vinclozolin treatment delayed pubertal maturation, and retarded sex accessory gland and epididymal growth. Serum lutenizing hormone,

testosterone, and 5α-androstane-3α,17β-diol levels were increased. These effects were concurrent with subtle, but statistically significant, alterations in the subcellular distribution of AR. In control animals, most of the AR was in the high salt cell fraction, apparently bound to the natural ligand and DNA, while the AR distribution was altered in treated males.

31. M1 and M2 concentrations in the serum of affected animals were below their Ki values for AR. These results suggest that when the vinclozolin metabolites occupy a modest percentage of the AR they prevent maximal AR-DNA binding and alter *in vivo* androgen-dependent gene expression and protein synthesis. This, in turn, alters morphological development and serum hormone levels. Although vinclozolin treatment has been shown to alter both adrenal and liver functions, the mechanism(s) of action for these effects have not been elucidated, and the role of AR, if any, is unknown.

PROCYMIDONE, CHLOZOLINATE, AND IPRODIONE

32. When administered by gavage from GD 14 to post-natal day (PND) 3 after birth, at doses ranging from 25 to 200 mg/kg/day, effects were noted in all dosage groups (Ostby *et al.*, 1999). Procymidone reduced anogenital distance in male pups and induced retained nipples, hypospadias, cleft phallus, vaginal pouch, and reduced sex accessory gland size in rat offspring. Procymidone also had marked effects on the histology of the dorsolateral and ventral prostatic, and seminal vesicular, tissues (at 50 mg/kg/day, and above). The effects consisted of fibrosis, cellular infiltration, and epithelial hyperplasia (Ostby *et al.*, 1999).

33. Chlozolinate and iprodione are dicarboximide fungicides, similar in structure to the anti-androgens vinclozolin and procymidone. However, when chlozolinate and iprodione were administered at 100 mg/kg/day from GD 14 to PND 3, the male rat offspring were not demasculinized or feminized (Gray *et al.*, 1999a).

EFFECTS OF MIXTURES OF ANTI-ANDROGENS: CUMULATIVE RISK

34. The U.S. Food Quality Protection Act of 1996 mandated that the risk assessment process consider combinations of chemicals that act via the same mechanism, rather than evaluate the potential risk on a individual basis. People are not exposed to one chemical at a time, but rather are exposed to mixtures, e.g., pesticides and toxic substances, from many different sources. Hence, several studies have been initiated to examine if mixtures of anti-androgens act in an additive or synergistic manner. In one of the studies, two "anti-androgenic" chemicals that altered foetal development via different mechanisms of action were combined to see if this combination also produced cumulative responses.

Cumulative Effects of Vinclozolin plus Procymidone

35. In the first study, the effects of graded doses of the AR antagonists procymidone and vinclozolin, ranging from 25 to 100 mg/kg/day, individually or together, were evaluated in the castrate-immature-testosterone-treated male rat Hershberger assay for seven days (Price *et al.*, 2000). At low doses the mixtures of vinclozolin plus procymidone reduced ventral prostate and levator ani weights in an additive fashion. When the higher doses were combined the effects were less than additive, because each chemical by itself nearly completely inhibited the effects of testosterone. These results provide scientific support for the concept that risk assessments for pesticides that act via a common mechanism of action should consider "cumulative" risk as opposed to examining risk on an individual, chemical-by-chemical, basis.

DDT - BACKGROUND

36. Although use of DDT has been banned in some countries, as a result of decades of former use

and the persistent nature of this pesticide, some wildlife populations still display high total DDT residue levels (Elliot *et al.*, 1994; Williams, 1999; Guillette *et al.*, 1999b). In the orchards and fields sampled by Elliot *et al.* (1994), birds had tissue levels of *p,p'*-DDE of up to 103 ppm in fat, and fat samples in Lake Apopka birds had even higher values. The most widely know endocrine effect of *p,p'*-DDE, the induction of eggshell thinning in avian and reptilian oviparous vertebrates, is hypothesized to result via an inhibition of prostaglandin synthesis in the shell membrane (Lundholm *et al.*, 1987; 1994; Lundholm and Bartonek, 1992a;b; Lundholm, 1994; Guillette *et al.*, 2000).

In utero effects of p,p'-*DDE in the Rat*

37. When *p,p'*-DDE is administered at 100 mg/kg/day (days 14-18 of gestation) to Long Evans Hooded (LE) and Sprague-Dawley (SD) rats it reduced AGD and induced hypospadias, retained nipples, and weights of androgen-dependent tissues in treated male rat offspring (Gray *et al.*, 1999a). While the alterations were evident in both rat strains, only the SD strain displayed hypospadias, and the other effects of DDE were of a greater magnitude in this strain. You *et al.* (1998) also found that *p,p'*-DDE induced anti-androgenic effects on AGD and areola development in both LE and SD rat strains, and prostate glands in the DDE-treated group displayed chronic suppurative prostatitis (You *et al.*, 1999b), which is not an uncommon observation for males exposed *in utero* to an endocrine disrupting chemical (i.e., PCBs; procymidone). Foetal rat tissue *p,p'*-DDE levels ranged from 1 to 2 μg/g on GD 21, and 10-20 μg/g on GD 20 (You et al., 1999a), following oral maternal treatment with *p,p'*-DDE at 100 mg/kg/day, as above.

Effects of *p,p'*-DDE administered to Pubertal and Adult Male Rats

38. When *p,p'*-DDE is administered at 0, 30, or 100 mg/kg/day from weaning until about 50 days of age, PPS was delayed about five days in male rats treated with 100 mg/kg/day, but sex accessory weights and serum hormone levels were not significantly altered (Kelce *et al.*, 1995). *p,p'*-DDE produces marked reductions in androgen-dependent, tissue weight gains in the Hershberger assay (100 mg/kg/day for seven days) but, in contrast to the positive responses seen in the "Pubertal Male" and Hershberger assays, it is negative in the adult male rat assay (O'Conner *et al.*, 1999).

39. Brien *et al.* (2000) reported that *p,p'*-DDE markedly interferes with erectile function, an androgen-dependent process, in an established rat model of apomorphine-induced erections. A single dose of *p,p'*-DDE (500 mg./kg., i.p.) decreased apomorphine-induced erections for at least two weeks. Testosterone supplementation restored function in castrated rats to pre-castrated levels, but the *p,p'*-DDE treated rats required 4 times as much testosterone to recover erections as compared to control males.

LINURON

Mechanistic Studies

40. The urea-based herbicide, linuron, binds rat prostatic AR and hAR, and inhibits DHT-hAR induced gene expression *in vitro* (Lambright *et al.*, 2000; McIntyre *et al.*, 2000; Cook *et al.*,1993; Waller *et al.*, 1996b). The anti-androgenicity of linuron is quite apparent when administered during gestation (McIntyre *et al.*, 2000; Lambright *et al.*, 2000; Gray *et al.*, 1999a) or in a Hershberger assay (Lambright *et al.*, 2000). Linuron-treatment (100 mg/kg/day) produces robust reductions in testosterone- and DHT-dependent tissue weight gains in castrate-immature-testosterone propionate-treated male rats in the Hershberger assay (Hershberger, 1953; Lambright *et al.*, 2000). In contrast, the anti-androgenic effects of linuron on androgen-dependent tissue weights are difficult, if not impossible, to detect in the intact adult male rat, except at overtly toxic dosage levels (i.e. 200 mg/kg/day; Cook *et al.*, 1993). The finding that the effects of linuron and *p,p'*-DDE (O'Connor *et al.*, 1999) on these organ weights, are seen only at overtly toxic dosage levels (after 14 days of treatment in the intact adult male rat) negates the use if the intact adult

rat as an animal model as a screening assay for anti-androgens.

Developmental Effects of Linuron in the Rat

41. In a multi-generational study, the linuron-treated (40 mg/kg/day) F1 offspring (Gray *et al.*, 1999a) sired 40% fewer pups, and treated F1 males had reduced testicular and epididymal weights, and lower testicular spermatid numbers. These effects were unexpected because it had been reported in an earlier multi-generational study that linuron did not produce reproductive malformations (Hodge *et al.*, 1968), and Khera *et al.* (1978) reported that linuron was not teratogenic at dosage levels up to 100 mg/kg/day. To resolve this discrepancy, linuron was administered at 100 mg/kg/day from days 14-18 of gestation (Gray *et al.*, 1999a). The AGD was reduced in male offspring and the incidence of areolas (with and without nipples) was increased in linuron treated, infant males. Linuron treatment induced epispadias and reduced the weight gains of the androgen-dependent tissues, including the seminal vesicles, ventral prostate, levator ani/bulbocavernosus muscles, and epididymides, and caused agenesis of the caput and/or corpus epididymides. In addition, some testes were atrophic, fluid filled, and flaccid (Gray *et al.*, 1999a; Lambright *et al.*, 2000; McIntyre *et al.*, 2000). Although linuron is an AR antagonist, it produces a profile of effects that resembles the effects seen with dibutylphthalate or di(2-ethylhexyl)phthalate treatment (i.e., relatively high incidences of testis and epididymal malformations) (Table 1). It is possible that that linuron can alter sexual differentiation by more than one mechanism of action.

42. A comparison of the sensitivity of the Hershberger assay to other assays that assess AR-mediated effects is shown in Table 1. It demonstrates the sensitivity of this assay as compared to the pubertal male and adult, intact male assays. Developmental toxicants are detected in the Hershberger and PPS assays, and intact animals are least affected. Sex accessory gland size, serum testosterone and lutinizing hormone levels are not always altered. When affected, seminal vesicle weights are more reduced in immature than adult animals.

Table 1. A comparison of the responses of male assays to anti-androgens.

	Hershberger assay	Develop. malformation	PPS delay	Pubertal male (other)	Intact adult
Fenitrothion	+++	?????	no	no	no??
p,p'-DDE	+++	+	4-day	no	no
Vinclozolin	+++	+++	7-day	+++	+
Linuron	+++	+++	2-day	?????	overt toxicity
Dibutylphthalate	+	++	5-day	+++	??
Flutamide	+++	+++	long	++++	++

REFERENCES

Brien et al. 2000. *****

Cook JC, Mullin LS, Frame SR, Biegel LB. 1993. Investigation of a mechanism for Leydig cell tumorigenesis by linuron in rats. Toxicol. Appl. Pharmacol. 119:195-204 (1993).

Di Salle E, Briatico G, Giudici D, Ornati G, Panzeri A. 1994. Endocrine properties of the testosterone 5 alpha-reductase inhibitor turosteride (FCE 26073). J. Steroid Biochem. Mol. Biol. 48(2-3):241-248.

Elliot et al. 1994. *****

Gray and Ostby, 1998. *****

Gray LE, Ostby J, Furr J, Wolf CJ, Lambright C, Parks L, Veeramachaneni DN, Wilson V, Price M, Hotchkiss A, Orlando E, Guillette L. 2001. Effects of environmental antiandrogens on reproductive development in experimental animals. Hum. Reprod. Update. 7(3):248-264. [Review article]

Gray LE Jr, Wolf C, Lambright C, Mann P, Price M, Cooper RL, Ostby J. 1999. Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p'-DDE, and ketoconazole) and toxic substances (dibutyl- and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differentiation produces diverse profiles of reproductive malformations in the male rat. Toxicol. Ind. Health. 15(1-2):94-118.

Gray et al. 1989. *****

Gray et al. 1994. *****

Gray et al. 1999a. *****

Gray et al. 1999b. *****

Guillette et al. 1999b. *****

Guillette et al. 2000. *****

Hershberger, 1953. *****

Hodge et al. 1968. *****

Hotchkiss et al. 2001. *****

Kelce WR, Stone CR, Laws SC, Gray LE, Kemppainen JA, Wilson EM. 1995. Persistent DDT metabolite p,p'-DDE is a potent androgen receptor antagonist. Nature 375(6532):581-585.

Kelce et al. 1997. *****

Khera et al. 1978. *****

Korenbrot et al. 1977. *****

Lambright C, Ostby J, Bobseine K, Wilson V, Hotchkiss AK, Mann PC, Gray LE Jr. 2000. Cellular and molecular mechanisms of action of linuron: an antiandrogenic herbicide that produces reproductive malformations in male rats. Toxicol. Sci. 56(2):389-399.

Laws et al. 1996. *****

Lundholm, 1994. *****

Lundholm et al. 1987. *****

Lundholm et al. 1994. *****

Lundholm and Bartonek, 1992a. *****

Lundholm and Bartonek, 1992b. *****

McIntyre et al. 2000. *****

Monosson et al. 1999. *****

O'Connor JC, Frame SR, Davis LG, Cook JC. 1999. Detection of the environmental antiandrogen p,p-DDE in CD and long-evans rats using a tier I screening battery and a Hershberger assay. Toxicol. Sci. 51(1):44-53.

O'Connor et al. 2000. *****

Ostby et al. 1999. *****

Price et al. 1999. *****

Price et al. 2000. *****

Sunami O, Kunimatsu T, Yamada T, Yabushita S, Sukata T, Miyata K, Kamita Y, Okuno Y, Seki T, Nakatsuka I, Matsuo M. 2000. Evaluation of a 5-day Hershberger assay using young mature male rats: methyltestosterone and p,p'-DDE, but not fenitrothion, exhibited androgenic or antiandrogenic activity in vivo. J Toxicol. Sci. 25(5):403-415.

Waller et al. 1996a. *****

Waller et al. 1996b. *****

Williams, 1999. *****

Wolf et al. 2000. *****

Yamada T, Sunami O, Kunimatsu T, Kamita Y, Okuno Y, Seki T, Nakatsuka I, Matsuo M. 2000. Dissection and weighing of accessory sex glands after formalin fixation, and a 5-day assay using young mature rats are reliable and feasible in the Hershberger assay. Toxicology. 162(2):103-119.

Yamada et al. 2001. *****

You et al. 1998. *****

You et al. 1999a. *****

You et al. 1999b. *****

***** For the complete citation, contact Dr. L. Earl Gray Jr., Endocrinology Branch, National Health and Environmental Effect Research Laboratory, U.S. EPA, Research Triangle Park, NC 27711, USA.
Tel: 1-919-541-7750; Fax: 1-919-541-4017; Email: gray.earl@epamail.epa.gov.

ATTACHMENT TO ANNEX 9

OECD PROTOCOL AND GUIDANCE FOR THE CONDUCT OF THE RODENT HERSHBERGER ASSAY

PHASE 2 OF THE VALIDATION OF THE RODENT HERSHBERGER ASSAY

PHASE 2 OF THE VALIDATION OF THE RODENT HERSHBERGER ASSAY[1]

INTRODUCTION

1. The overall aim of the validation work is to develop a robust assay that can be considered as the basis for an OECD Test Guideline. This document provides the essential elements of the study protocol to be used for Phase 2 of the validation of the rodent Hershberger assay. More detailed practical laboratory protocols for the OECD Validation work may be built on the essential requirements contained in this document.

2. The rodent Hershberger assay was first described in 1953 (Hershberger *et. al.*, 1953). Since that time it has been used primarily in the pharmaceutical industry. A standardised and validated protocol has not been available for consideration internationally. This protocol provides all essential study elements for a standard protocol for the assay. The protocol was also used in the Phase 1 of the validation study with minor specific modifications in treatment regimen only as these are test substance specific. The protocol was first agreed at the Second meeting of the OECD Validation Management Group (VMG) for the Screening and Testing of Endocrine Disrupters (20-21 January 2000) and subsequently revised further at the teleconference of the VMG on 6 March 2000. The protocol appeared to perform extremely well in Phase 1 of the validation work and, therefore, no changes were considered necessary.

3. The Hershberger assay is an *in vivo* short-term assay for chemicals that have the potential to act like endogenous sex hormones. The rodent Hershberger assay is similar in concept to the rodent uterotrophic assay - both measure as endpoints changes in specific tissues that normally respond to endogenous hormones. The focus of the Hershberger assay is on male sex hormone interactions while the uterotrophic assay's focus is on female sex hormone interactions.

4. The Hershberger and uterotrophic assays are both being considered by OECD as potential short term screening assays. The information generated by use of the assay 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.

PHASE 2 VALIDATION WORK

5. The aims of the Phase 2 validation of the Hershberger assay are to:

- Demonstrate the ability of the protocol to reproducibly detect the activity of weak androgen agonists and antagonists that act through different mechanisms;
- Demonstrate the reliability of the protocol by testing a series of chemicals in a number of laboratories in Europe, North America and Southeast Asia;
- Confirm the effectiveness of the standard reference dose of TP, as used in Phase 1 of the validation for routine use as positive control when detecting androgen agonists and as the negative control in case of detection of antagonists.

INITIAL CONSIDERATIONS

[1] As agreed at the Second meeting of the OECD Validation Management Group (VMG) for the Screening and testing of Endocrine Disrupters (20-21 January 2000) and subsequently revised further at the teleconference of the VMG on 6 March 2000.

6. The rodent Hershberger assay evaluates the ability of a chemical to show biological activities consistent with the agonism or antagonism of natural hormones, that have masculinising effects. These hormones are known as androgens (e.g., testosterone).

7. The rodent Hershberger assay is based on changes in weight of male sex accessory tissues in sexually immature castrated male rats.

8. Test substances may stimulate or, in the presence of a reference androgen, inhibit the stimulated development of sex accessory tissues.

9. Accessory sex glands and accessory sex tissues are dependent upon androgen stimulation to gain and maintain weight during and after puberty. If endogenous sources of androgen are removed, exogenous sources of androgen are necessary to increase or maintain the weights of these sex accessory tissues.

10. The sex accessory glands and tissues for this protocol are the:

- Ventral Prostrate (VP);
- Seminal Vesicles (SV);
- Coagulating Glands (CG);
- Levator ani plus Bulbocavernosus muscles (LABC);
- Glans Penis (GP) and
- Cowpers (or bulbourethral) Glands (CP).

11. This protocol uses sexually immature male rats, castrated at peripuberty by removal of testes and epididymides (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. TP will initiate prepuce separation so that the Glans Penis (GP) can be 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 rodent is that the sex accessory tissues have a high sensitivity and small relative weight which both help to minimise variation in responses between individual animals.

12. As part of the development of this protocol, study variables have been standardised as far as possible based on historical experience and current research. Results from Phase 1 of the validation of this assay have shown that standardisation of dissection techniques of the selected sex accessory glands and tissues was sufficiently adequate to reduce the variability of their weights to acceptable limits in all participating laboratories. The key variables not standardised in this protocol are the strain of rodent, diet, and housing conditions.

PRINCIPLE OF TEST

13. The test substance is administered in graduated doses to several groups of male rodents for a number of consecutive days. Measurement of the weight of sex accessory tissues provides information on the androgenic nature of a chemical, however it can also provide additional information on whether effects are due to the effects on the androgen hormone receptor *in vivo* or on other relevant biochemical mechanisms, e.g., effect on other enzymes involved in the production of sex hormones such as 5-alphareductase.

14. In addition to the sex accessory tissues, body weight gain is a mandatory measurement to provide

information on the general health and wellbeing of the animals. Liver and kidney weight at necropsy are not mandatory endpoints, but liver weight is highly recommended, as some test substances may appear to be anti-androgenic by inducing an increased metabolism of TP by the liver. This may be indicated by an increase in liver size. Because of the wider range of chemicals studied in Phase 2, kidney weight is recommended because this may serve as a sensitive indicator of systemic toxicity. Necropsy of the adrenals and kidneys may provide supplementary information about the effects of the test substance on other related biochemical pathways and are therefore optional supplementary endpoints. Measurement of serum testosterone and leutinising hormone may also be investigated in this context.

Androgen agonists

15. To test for androgen agonists the test substance is administered to immature castrated rats for ten consecutive days. TP is administered by daily sub-cutaneous injection. TP provides the positive control in studies with substances of unknown androgenic activity. The vehicle provides the negative control.

Androgen antagonists

16. To test for androgen antagonists, the test substance is administered to immature castrated rats for ten consecutive days together with a reference androgen agonist (TP). Administration of TP alone is used as the negative control which treatments are compared to for antiandrogenic activity. The weights of the sex accessory tissues after co-administration of the test chemical and reference androgen are compared with the weights of tissues from this control group.

DESCRIPTION OF METHOD/PREPARATIONS FOR THE TEST

Animal Species and Strain

17. 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, which has a different schedule of sexual development compared to other more commonly used strains such as Sprague Dawley, Long Evans or Wistar strains. 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 mentioned in paragraph 11 are met. The strain of rat used should be recorded in the report.

Acclimatisation

18. Healthy young animals that have been acclimatised to the laboratory conditions for 1-2 weeks following castration will be used. Animals will be observed daily, and any animals with evidence of disease or physical abnormalities will be removed from the study. If castrated animals are purchased from an animal supplier the age of animals and stage of sexual maturity should be assured by the supplier and the time between castration and initiation of dosing will be counted as part of the acclimatisation period. In such cases the animals will be no more than 8 weeks of age at the initiation of dosing. A period of between one and two weeks acclimatisation has been chosen to allow sufficient period of acclimatisation while also allowing a laboratory to schedule the experimental work efficiently.

Housing and feeding conditions

19. Temperature in the experimental animal room should be 22 °C (±3°). The relative humidity should preferably be 50 to 60%, but not exceed limits of 30 to 70% except during room cleaning. Lighting should be

artificial, the photoperiod being 12 hours light, 12 hours dark.

20. There is currently insufficient information showing any influence of laboratory diets on the identification of androgenic substances *in vivo*. Laboratories participating in the validation should use the laboratory diet normally used in their chemical testing work. The diet used should be recorded and a sample of the laboratory diet should be retained for possible future analysis. Both diet and drinking water should be supplied *ad libitum.*

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

22. Each animal will be identified individually and uniquely (e.g., ear mark or tag).

23. Six animals of the same age and cohort will be used per treatment and control group.

Body Weight and the selection of animals for the study

24. Variations in body weight may be a source of variation in the weight of tissues of interest (especially the liver). This variation, if present, will increase variability within a group or among groups of animals. This may interfere with assay sensitivity, and possibly lead to false positives or false negatives.

25. Body weights will vary from study to study and different rodent strains. Each participating laboratory should establish its own procedure for limiting the variability in body weight. These procedures will be recorded in the report and should ensure that all groups of animals reflect normal variations expected for healthy animals.

26. As a precautionary measure, any effect of body weight on sex accessory tissue weight will be controlled in both the experimental design and data analysis phases of the study.

27. Within the experimental design the variation in body weight will be both experimentally and statistically controlled. Within the data analysis phase, body weight will be used as a covariate in the overall analysis.

28. Experimental control is accomplished in two steps. The first step involves selection of animals with relatively small variation in body weight from the larger population. Avoiding unusually small or large animals achieves this. A reasonable level of body weight variation within a study should be tolerated to ±20% of the mean body weight (e.g. 175g ± 35g). While this degree of variability may seem large it is not expected to alter the outcome of the study, as long as the animals are healthy, and will reduce the numbers of animals that would be rejected.

29. The second part of "experimental" control of body weight involves the assignment of animals to different treatment groups by a randomised complete block approach rather than by completely randomisation. 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. This variable is then included in the data analysis to adjust for differences in body weight.

Non-routine health and safety requirements

30. The test substances are known as possible reproductive and developmental toxicants and therefore appropriate precautions should be taken to protect personnel during the study, e.g. necessary training, labelling and storage procedures, and protective handling procedures during dose preparation and dose administration. 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.

PROCEDURE

Reference substance and vehicle

31. The reference androgen will be Testosterone Propionate (TP), CAS No 57-85-2. TP will be administered in a specified laboratory grade stripped corn oil. All participating laboratories will use stripped corn oil to eliminate potential differences in absorption as a source of variation. Participating laboratories will be supplied with TP from the central chemical repository.

Test Chemicals

32. The following chemicals have been selected by the VMG for testing in Phase 2:

Androgens
- Testosterone propionate (reference chemical)
- Trenbolone
- Methyl testosterone

Anti-androgens
- Vinclozolin
- Procymidone
- Linuron
- *P,p'* –DDE
- Finasteride / Turosteride

33. Phase 2 will be organised in a way that each chemical is tested in at least four laboratories. Laboratories that select finasteride as anti-androgen should not test turosteride and vice versa, as both chemicals have the same mechanism of action.

The number of test groups

34. All tests will be conducted with at least three dose levels and one (vehicle) control group. In testing anti-androgens, animals at each dose level will also be dosed with testosterone propionate (0.2 and/or 0.4 mg/kg/day), and a second, negative, control group will be added to these tests receiving only TP.

Doses

35. The following recommended dose levels are based on suggestions made by the lead laboratory, subsequent comments by the VMG and discussions of a number of leading experts (John Ashby, Alexius

Freyberger, Earl Grey, Elard Jacob and William Owens:

Chemical	Dose Level (mg/kg/day)	Route of Administration
Testosterone propionate	(0.2)*, 0.4	Subcutaneous
Trenbolone	0.3, 1.5, 8, 40	Oral
Methyl testosterone	0.5, 2, 10, 40	Oral
Vinclozolin	3, 10, 30, 100	Oral
Procymidone	3, 10, 30, 100	Oral
Linuron	3, 10, 30, 100	Oral
P,p' –DDE	5, 16, 50, 160	Oral
Finasteride / Turosteride	1, 5, 25	Oral

* The 0.4 mg/kg/day dose level is preferred, although some laboratories may prefer 0.2 mg/kg/day.

36. Laboratories are advised to perform sighting studies, as appropriate, to confirm the appropriateness of the suggested dose levels for the various test chemicals.

37. Laboratories may consider including one or more chemicals expected to have no (anti) androgenic activity as a demonstration that the stress of being dosed with a potentially toxic substance will not initiate a positive response in the Hershberger assay. However, following successful completion of Phase 2, an additional Phase will specifically focus on the (lack of) effects of negative chemicals and on blind testing.

Administration of doses

38. For subcutaneous administration, all treatments are administered by s.c. injections on the dorsal surface of the animal. The maximum limit on the volume administered per animal is approximately 0.5 ml/kg body weight per day.

39. For oral administration, all treatments are administered by gavage. The maximum limit on the volume administered per animal will be 5 ml/kg/day.

40. 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 for changes in body weight. The volume of dose and time that it is administered will be recorded on each day of exposure.

Good Laboratory Practice

41. Work will be conducted according to the principles of Good Laboratory Practice (OECD Good Laboratory Practice and Compliance Monitoring (OECD, 1998). In particular data will have a full audit trail and be retained on file. Data will be collected in a manner that will allow independent peer review. Calibration data for all balances used should be determined a part of the study and written records maintained.

OBSERVATIONS

Clinical observations

42. Animals will be evaluated 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.

43. 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 the reasons.

Body weight and food consumption

44. Individual body weights will be recorded prior to start of treatment (to the nearest 0.1g), on each day of administration period and prior to necropsy. Group means and standard deviations will be calculated.

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

Necropsy

46. 50. Approximately 24 hours after the last administration of the test substance, the rats will be euthanized and exsanquinated 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.

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

48. The sex accessory tissue and liver weights are mandatory measurements. Adrenal and kidney weights are optional additional measurements.

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

50. The sex accessory tissues will be excised and their weights determined, for comparison with the weights of sex accessory tissues from the vehicle control group, or reference TP group (in the case of antagonist response). If serum hormones are to be measured as an option, the rodents will be anaesthetised prior to necropsy and blood taken by cardiac puncture. If serum hormones are to be measured, the method of anaesthesia should be chosen with care so that it does not affect hormone measurement.

51. It is important that persons carrying out the dissection of the sex accessory tissues are familiar with standard dissection procedures for these tissues. This will minimise a potential source of variation in the study. Ideally the same prosector should be responsible for the weighing a given tissue to eliminate inter-individual differences in tissue processing. If this is not possible, the necropsy should be designed such that each prosector weighs a given tissue from all treatment groups as opposed to one individual weighing all tissues from a control group, while someone else is responsible for the treated groups.

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

Measurement of sex accessory tissues

53. After necropsy, the sex accessory tissues will be removed and weighed without blotting (to the nearest 0.1mg). The excised tissues will be trimmed of any fat. Participating laboratories should ensure that the excision procedures used are reproducible over time and pay particular care to prevent variations in fluid losses from tissues during processing. A standard operating procedure will be followed for the

excision of sex accessory tissue. This procedure will be provided by the Lead Laboratory.

54. After excision and weighing of the ventral prostate it will be fixed for 24 hours in 10% neutral buffered formalin (4% formaldehyde) and weighed again.

55. The following weight of the following sex accessory tissues will be measured:

- Ventral Prostate (VP) – fresh and fixed tissue weight (24 hours)
- Seminal vesicles together with coagulating gland (SV and CG) – fresh tissue weight
- Levator ani and bulbocavernous muscles (LABC) – fresh tissue weight
- Glans penis (GP) - fresh tissue weight
- Cowpers (or bulbourethral) Glands (CG) – fresh tissue weight

56. The weight of the adrenal glands and the kidneys and levels of serum leutinising hormone and testosterone may be measured as optional endpoints.

REPORTING

Data

57. Data will be reported individually and for each group of animals (i.e. body weights, liver weight, accessory sex tissue weights, optional measurements and other responses and observations). 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. To assist data reporting and compilation a standardised electronic spreadsheet will be used by participating laboratories to report data during the initial validation work.

Test report

58. The test report must include the following information:

Laboratory identification

Test substance:

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

Vehicle:

Test animals:

- Species/strain used
- Number, age and sex of animals
- Source, housing conditions, diet, and bedding
- Individual weights of animals at the start of the study (to nearest 0.1 g)

Test conditions:

- Housing conditions
- Number of animals per cage
- Necropsy procedures
- Diet

Results:

- Daily observations

- 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
  - Animal age
  - Order of animal in the necropsy

  - TP treatment (Yes or No and dosage level)

  - Body weight at start of dosing (to nearest 0.1g)
  - Body weight at necropsy (to nearest 0.1g)

- Weights of sex accessory tissues[2] (to the nearest 0.1g)
  - Ventral prostate (fresh weight and weight after fixation)
  - Seminal vesicle plus coagulating gland, including fluid (fresh weight)
  - Levator ani plus bulbocavernosus muscle (fresh weight)
  - Glans penis (fresh weight)
  - Cowpers Gland (fresh weight)
  - Liver (fresh weight)
  - Kidney weight (optional)
  - Adrenal weight (optional)
  - Serum LH (optional)
  - Serum T (optional)

- General remarks and comments

Discussion

Conclusions

[2] In a parallel protocol, identical in all aspects to this, some laboratories may generate data by fixing the sex accessory tissues before separation and weighing. This is an optional additional protocol for comparative purposes.

Interpretation of results

59. Statistical comparisons in individual laboratories will be made for the different sex accessory by analysis of variance. For androgen agonism, the test substance groups will be compared to the vehicle control. A statistically significant increase in tissue weight will be considered 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. A statistically significant decrease in tissue weight will be considered a positive antagonist result. If more than one set of comparisons is required, all comparisons will be conducted separately for each test group against its control.