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

28-Jul-2009

English - Or. English

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

ENV/JM/MONO(2009)13
Unclassified

SERIES ON TESTING AND ASSESSMENT
Number 108

REPORT OF THE VALIDATION OF THE HERSHBERGER BIOASSAY (WEANLING MODEL)

English - Or. English

JT03268291

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 108

REPORT OF THE VALIDATION OF THE HERSHBERGER BIOASSAY (WEANLING MODEL)

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNEP, UNIDO, UNITAR, WHO and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2009

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 (2004)*

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 (2007)*

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 (2005)*

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*

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, *Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment (2008)*

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 (Quantitative) Structure-Activity Relationships [(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)*

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 (2006)*

No. 63, *Guidance Document on the Definition of Residue (2006)*

No. 64, *Guidance Document on Overview of Residue Chemistry Studies (2006)*

No. 65, *OECD Report of the Initial Work Towards the Validation of the Rodent Utertrophic Assay - Phase 1 (2006)*

No. 66, *OECD Report of the Validation of the Rodent Uterotrophic Bioassay: Phase 2. Testing of Potent and Weak Oestrogen Agonists by Multiple Laboratories (2006)*

No. 67, *Additional data supporting the Test Guideline on the Uterotrophic Bioassay in rodents (2007)*

No. 68, *Summary Report of the Uterotrophic Bioassay Peer Review Panel, including Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the follow up of this report (2006)*

No. 69, *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models (2007)*

No. 70, *Report on the Preparation of GHS Implementation by the OECD Countries (2007)*

No. 71, *Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity (2007)*

No. 72, *Guidance Document on Pesticide Residue Analytical Methods (2007)*

No. 73, *Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen*

Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol (2007)

No. 74, *Detailed Review Paper for Avian Two-generation Toxicity Testing (2007)*

No. 75, *Guidance Document on the Honey Bee (Apis Mellifera L.) Brood test Under Semi-field Conditions (2007)*

No. 76, *Final Report of the Validation of the Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances: Phase 1 - Optimisation of the Test Protocol (2007)*

No. 77, *Final Report of the Validation of the Amphibian Metamorphosis Assay: Phase 2 - Multi-chemical Interlaboratory Study (2007)*

No. 78, *Final Report of the Validation of the 21-day Fish Screening Assay for the Detection of Endocrine Active Substances. Phase 2: Testing Negative Substances (2007)*

No. 79, *Validation Report of the Full Life-cycle Test with the Harpacticoid Copepods Nitocra Spinipes and Amphiascus Tenuiremis and the Calanoid Copepod Acartia Tonsa - Phase 1 (2007)*

No. 80, *Guidance on Grouping of Chemicals (2007)*

No. 81, *Summary Report of the Validation Peer Review for the Updated Test Guideline 407, and Agreement of the Working Group of National Coordinators of the Test Guidelines Programme on the follow-up of this report (2007)*

No. 82, *Guidance Document on Amphibian Thyroid Histology (2007)*

No. 83, *Summary Report of the Peer Review Panel on the Stably Transfected Transcriptional Activation Assay for Detecting Estrogenic Activity of Chemicals, and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 84, *Report on the Workshop on the Application of the GHS Classification Criteria to HPV Chemicals, 5-6 July Bern Switzerland (2007)*

No. 85, *Report of the Validation Peer Review for the Hershberger Bioassay, and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 86, *Report of the OECD Validation of the Rodent Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5 α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories (2008)*

No. 87, *Report of the Ring Test and Statistical Analysis of Performance of the Guidance on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (Transformation/ Dissolution Protocol) (2008)*

No.88, *Workshop on Integrated Approaches to Testing and Assessment (2008)*

No.89, *Retrospective Performance Assessment of the Test Guideline 426 on Developmental Neurotoxicity (2008)*

No.90, *Background Review Document on the Rodent Hershberger Bioassay (2008)*

No.91, *Report of the Validation of the Amphibian Metamorphosis Assay (Phase 3) (2008)*

No.92, *Report of the Validation Peer Review for the Amphibian Metamorphosis Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-Up of this Report (2008)*

No.93, *Report of the Validation of an Enhancement of OECD TG 211: Daphnia Magna Reproduction Test (2008)*

No.94, *Report of the Validation Peer Review for the 21-Day Fish Endocrine Screening Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2008)*

No.95, *Detailed Review Paper on Fish Life-Cycle Tests (2008)*

No.96, *Guidance Document on Magnitude of Pesticide Residues in Processed Commodities (2008)*

No.97, *Detailed Review Paper on the use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors (2008)*

No. 98, *Considerations Regarding Applicability of the Guidance on Transformation/Dissolution of Metals Compounds in Aqueous Media (Transformation/Dissolution Protocol) (2008)*

No. 99, *Comparison between OECD Test Guidelines and ISO Standards in the Areas of Ecotoxicology and Health Effects (2008)*

No.100, *Report of the Second Survey on Available Omics Tools (2009)*

No.101, *Report on the Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox (2009)*

No.102, *Guidance Document for using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping of Chemicals (2009)*

No.103, *Detailed Review Paper on Transgenic Rodent Mutation Assays (2009)*

No.104, *Performance Assessment: Conparsion of 403 and CxT Protocols via Simulation and for Selected Real Data Sets (2009)*

No. 105, *Report on Biostatistical Performance Assessment of the draft TG 436 Acute Toxic Class Testing Method for Acute Inhalation Toxicity (2009)*

No.106, Guidence Document for Histologic Evalution of Endorcine and Reproductive Test in Rodents (2009)

No.107, Preservative treated wood to the environment for wood held in storage after treatment and for wooden commodities that are not cover and are not in contact with ground. (2009)

No.108, Intact, Stimulated, Weanling Male Rat Version of the Hershberger Bioassay(2009)

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 developed in the IOMC context. The contents do not necessarily reflect the This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

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, UNEP, UNIDO, UNITAR, WHO and OECD. 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.

OECD Environment Directorate,
Environment, Health and Safety Division

2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80
E-mail: ehscont@oecd.org

FOREWORD

This validation report of the Hershberger Bioassay (weanling model) was developed by the Environmental Protection Agency of the United States in 2008. It was reviewed by the Validation Management Group for Mammalian Testing, and approved by the Working Group of National Coordinators of the Test Guidelines Programme, by written procedure, on 10 March 2009. Appendices A and B are attached to this report. Appendices C, D, E and F are available to governments at the OECD Secretariat, on request.

This document is published on the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology.

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

Tel : 33-1-45-24-16-74
E.mail : env.edcontact@oecd.org

TABLE OF CONTENTS

Page

Appendices:

LIST OF TABLES

Page

LIST OF TABLES (continued)

Page

LIST OF ABBREVIATIONS

The following abbreviations are used in this document:

accessory sex organ (ASO)
androgen receptor (AR)
body weight (BW)
coefficients of variation (CVs)
Cowper's glands (COW)
dihydrotestosterone (DHT)
dinitrophenol (DNP)
Endocrine Disruptors Screening and Testing Advisory Committee (EDSTAC)
Environmental Protection Agency (EPA)
estrogen receptor (ER)
finasteride (FIN)
flutamide (FLU)
follicle stimulating hormone (FSH)
glans penis (GP)
hypothalamus-pituitary (HP)
hypothalamic-pituitary-gonadal (HPG)
linuron (LIN)
luteinizing hormone (LH)
Organisation for Economic Cooperation and Development (OECD)
levator ani-bulbocavernosus (LABC) muscle complex
methyl testosterone (MT)
nonylphenol (NP)
p,p'-DDE (DDE)
postnatal day (PND)
preputial separation (PPS)
procymidone (PRO)
seminal vesicles (SV)

seminal vesicles with coagulating glands and their fluids (SVCG)
standard deviation (SD)
subcutaneous (SC)
testosterone (T)
testosterone propionate (TP)
Tier 1 screening (T1S)
Tier 2 testing (T2T)
trenbolone acetate (TREN)
ventral prostate (VP)
vinclozolin (VIN)

CHAPTER 1.0 SUMMARY

1.1 Introduction and Background

Recognizing that the environment contains compounds (both natural and man-made) that may interact with and subsequently affect the endocrine systems of humans and animals, the Organisation for Economic Cooperation and Development (OECD) in 1998 initiated revisions to existing test guidelines and development of new test guidelines for the screening and testing of potential endocrine disrupters (OECD, 1998). The Hershberger bioassay, a short-term *in vivo* screening test using accessory tissues of the male reproductive tract, originated in the 1930s and was subsequently modified to include androgen-responsive muscles in the male reproductive tract (Korenchevsky, 1932; Wainman and Shipounoff, 1941; Eisenberg et al., 1949; Eisenberg and Gordan, 1950; Hershberger et al., 1953; Dorfman, 1962; Hilgar and Hummel, 1964; Dorfman, 1969; OECD, 2000). It is a leading candidate for a Level 3 *in vivo* screening assay in the OECD Conceptual Framework (OECD, 2002). As a short-term *in vivo* screen, the information from the bioassay can be used to build on information from relevant *in vitro* screens, to reduce the list of chemicals that may need longer-term *in vivo* animal testing. The overall aim of the validation program is to demonstrate that the Hershberger bioassay is a robust, sensitive, reliable and reproducible bioassay that can be considered as the basis for an OECD Test Guideline. Once available, the test guideline is intended to be used as one element in an overall testing strategy for the detection and assessment of potential endocrine disruptors.

The focus of the Hershberger bioassay is on the detection of compounds that may mimic or interfere with the action of endogenous male sex hormones. The primary model for the Hershberger bioassay was the surgically castrated, peri-pubertal or adult male, a test protocol standardized by an official expert committee in 1962 for use as a screen for androgenic chemicals (Dorfman, 1962). The castrated peripubertal male rat is an animal model with low endogenous androgenic hormone levels (Ashby and Lefevre, 2000a). Based on the changes in weight of five androgen-dependent tissues, it evaluates the ability of a chemical to elicit biological activities consistent with androgen agonists, antagonists, or 5α-reductase inhibitors for a single endocrine class (i.e., [anti]androgens) and by a single mechanism, mediated only by the androgen receptor (AR). This model was the basis for the initial work in Phase 1, Phase 2, and Phase 3 of the Hershberger validation program. This included production of a Background Review Document, a detailed methods paper (Gray et al., 2005), a dissection guide, and intra- and interlaboratory studies to evaluate the reliability and reproducibility of the bioassay using the ventral prostate (VP), seminal vesicles with coagulating glands (SVCG), levator ani-bulbocavernosus (LABC) muscle, paired Cowper's glands (COW), and the glans penis (GP) as the five target tissues. These validation studies were conducted with a potent reference androgen (testosterone propionate [TP]), two potent synthetic androgens (trenbolone acetate [TREN] and methyl testosterone [MT]), a potent anti-androgenic pharmaceutical (flutamide [FLU]), a potent inhibitor of the synthesis of the natural androgen dihydrotestosterone (DHT) (i.e., finasteride [FIN]), several weakly anti-androgenic pesticides (linuron [LIN], vinclozolin [VIN], procymidone [PRO], p,p'-DDE [DDE]), a potent 5α-reductase inhibitor (FIN) and two known negative chemicals (dinitrophenol [DNP] and nonylphenol [NP]) (OECD, 2002, 2005; Owens et al., 2006; OECD, 2003, 2007, 2008). A draft test guideline resulted from this work.

Animal welfare concerns with the castration procedure prompted OECD to conduct a more limited validation program to assess the potential of the intact, stimulated weanling male rat as a reliable and relevant alternative model to the surgically castrated, adult male rat model. There is also the concern that the results from the adult castrate model might not be extrapolable to the gonad-intact animals, because the hypothalamic-pituitary-gonadal (HPG) axis is disrupted by the castration in the adult castrate model. The current stimulated weanling protocol specifics are based largely on the work performed at Lab A (Ashby and Lefevre, 2000b; Ashby et al., 2000a,b).

The OECD program on the Hershberger bioassay had previously demonstrated the response and the reliability of the surgical castrate model by measuring five male accessory sex glands and tissues (see above). In the case of the stimulated weanling, the mandatory target tissues were reduced to four: the VP, SVCG, LABC, and COWS. The list does not include the GP since the weanling male has not yet achieved preputial separation (PPS), and the prepuce cannot be reliably separated from the penile shaft in the male's sexually immature state. However, two tissues removed during castration are now available and optional for evaluation in this intact model, the paired testes and paired epididymides.

The validation process therefore proceeded as follows:

> The test substances were administered by oral gavage or subcutaneous (SC) injection to groups of six animals (per dose group) for ten consecutive days with daily body weights (beginning on postnatal day [PND] 21/22) in all studies in all stages. The animals were then necropsied approximately 24 hours later on the 11th day (24 hours after the last test substance administration, on PND 32-33). After dissection, the weights of the mandatory accessory sex tissues were measured, as well as the paired testes and paired epididymides weights and the optional organ weights (paired adrenal glands, paired kidneys, and liver) to provide supplementary information about the systemic toxicity, target organs, and other effects of the test substance, if employed.

The Hershberger bioassay with the stimulated intact weanling male provides data not only on (anti-)androgenicity mediated by the AR, as does the castrated male assay, but also by a number of additional possible mechanisms, such as metabolic inhibition and hypothalamic and/or pituitary regulation of the gonad, and/or indirect effects on the intact HPG axis, since in this model, the HPG axis is intact.

The rodent Hershberger bioassay has been based historically on changes in the weights of androgen-responsive male accessory sex tissues in peripubertal, castrated male rats. Accessory sex tissues and glands depend upon androgen stimulation to gain and maintain weight during and after puberty. When endogenous sources of androgen are low (the weanling before puberty, as well as the adult castrate), the biological activity of exogenous substances can be assayed by the increase (agonist response) in the weights of these accessory sex tissues or by blocking (antagonist response) the activity of administered androgens and by preventing an increase in the weights of these accessory sex tissues. The rodent Hershberger bioassay (modified to use the stimulated, intact weanling) therefore evaluates the ability of a chemical to show biological activities consistent with agonism or antagonism of potent endogenous androgens (e.g., testosterone [T] and DHT).

The available data indicate that the androgen-dependent accessory sex tissues of the intact weanling are also sensitive to exogenous, as well as endogenous, androgens since these animals also have ARs and appropriate steroidogenic enzymes necessary for agonist, antagonist, and 5α-reductase inhibitor responses. In addition to sensitivity, the weanling rodent accessory sex tissues have a small relative organ weight.

There are, however, differences between the stimulated weanling and the surgically castrated male. In the stimulated weanling, there are low prepubertal circulating levels of T produced by the paired testes, and there is the potential ability of the intact HPG axis to respond and compensate, to some degree, to chemical insult. The presence of the intact HPG axis is an asset in that it reflects the normal, intact animal; however, the presence of the HPG axis diminishes the responsiveness of the stimulated weanling version of the assay. The PND 20-35 time period in the male is not fully analogous to that in the female. There are no dramatic surges in the male in hormonal levels or tissue growth during this prepubertal time period. The change and rearrangement of the HPG axis is more gradual in males for circulating follicle stimulating hormone (FSH), luteinizing hormone (LH), T, and 5α-reduced steroids. At the same time, just as this is a period of rapid overall growth in the animal, the weights of the paired testes and the relevant tissues of the reproductive tract are also increasing in absolute terms, but the increase relative to body weight for these tissues is rather modest and constant, indicating no disproportionate surge in growth. Responses to hormone receptors during development, including the prepubertal period, may also

be different from those in adulthood. For example, the mode of action for thyroid hormones on the testis during development is not completely defined or understood, but is likely very different from that during adulthood.

1.2 Methods and Results

The program to assess the stimulated weanling as an alternative proceeded in the same fashion as that taken for the surgical castrate model. Due to the low circulating T levels and the presence of the HPG axis in the intact weanling, TP (CASRN 57-85-2) and FLU (CASRN 1311-84-7) standardization curves were necessary to ensure the proper TP dose was selected for the antagonist studies. Therefore, the first stage (designated Phase 1A and 1B), analogous to Phase 1A and Phase 1B with the surgical castrate included:

- Generation of a dose-response curve with TP as the first step (Phase 1A) to select optimum candidates for the TP co-administration dose with anti-androgens, to identify ED50-ED70 doses of TP that are applicable to all the responsive tissues. The doses employed were 0.4, 0.8, 1.0, 1.2, and 1.6 mg/kg/day in order to provide a complete comparison with the surgical castrate data set and adequately characterize the dose response for each of the four tissues (VP, SV, LABC, and COW), its relative maximum response, and provide insight on the CVs that labs can achieve with the smaller weanling tissues. With the vehicle control, this comprised a total of six dose groups per laboratory (36 animals based on n = 6 per dose group). The results were selection of two TP doses (0.8 and 1.0 mg/kg/day) to proceed into the second part of Stage 1, the FLU dose series experiments.
- Two selected doses of TP were used as the reference doses for co-administration to assess the response of stimulated weanling tissues to a dose series of FLU (Phase 1B). The doses of FLU employed were 0.3, 1.0, and 3.0 mg/kg/day. With the two TP reference dose groups, this comprised a total of eight dose groups per laboratory (48 animals based on n = 6 per dose group). Two additional groups were voluntary: a vehicle control and a fourth dose of FLU at 10 mg/kg/day. This series selected the optimum TP reference dose (1.0 mg/kg/day) and the optimum FLU reference dose (3.0 mg/kg/day) to proceed into the second stage, the dose-response experiments to known weak agonists and antagonists. In Stage 1A and 1B, all of the weights of the organs of interest exhibited statistically significant laboratory effects, i.e., one or more laboratories differed in the organ weights of interest from the organ weight values from the other laboratories.

The next stage (Phase 2) for the stimulated weanling comprised studies analogous to the Phase 2 dose response studies for the surgical castrate:

- Biological activity consistent with androgen agonists was tested by administering a test substance to intact, weanling male rats for ten consecutive days. The positive control for the tissue responses was TP, and the vehicle was the negative control. The weights of the accessory sex tissues of the test chemical groups were compared to the vehicle group for a statistically significant increase in weight.
- Biological activity, consistent with androgen antagonists and 5α-reductase inhibitors, was tested by administering the test substance to intact, weanling male rats for ten consecutive days together with a reference androgen agonist, TP. Administration of TP alone was the negative control. The weights of the accessory sex tissues after co-administration of the test chemical and the reference androgen TP together were then compared with the tissue weights of the reference androgen, TP, alone for a statistically significant decrease in weight versus the TP "negative control" group values. FLU was also co-administered with TP to another group as a positive control.
- The capability of the stimulated weanling was assessed to detect a weak androgen agonist, TREN,

and to detect two weak androgen antagonists, LIN and DDE. In order to make a comparison, the same dose series was used (as with the castrate model), with the lowest dose omitted since the weight(s) of one or more mandatory tissues in the castrate model in one or more laboratories did not reach statistically significant differences:

- TREN (CASRN 10161-33-8): 1.5, 8, and 40 mg/kg/day
- LIN (CASRN 330-55-2): 10, 30, 100 mg/kg/day
- DDE (CASRN 72-55-9): 16, 50, 160 mg DDE /kg/day

- For TP alone (the reference androgen), all three participating laboratories detected significant effects on all of the androgen-dependent organ weights.
- For TREN, only one of the three participating laboratories detected significant effects on VP and paired testes weights at 1.5 mg/kg/day. At 8 mg/kg/day, two of the participating laboratories detected effects on paired testes and on LABC, and one laboratory detected effects on the paired epididymides. At 40 mg/kg/day TREN, all laboratories detected significant effects on SV, LABC and paired testes; two laboratories detected effects on VP, and one laboratory detected effects on COW and paired epididymides. There were no significant effects on the weights of the optional systemic organs in any of the three participating laboratories.
- For LIN, none of the three participating laboratories detected any statistically significant effects on the androgen-dependent organ weights at 3, 10, or 30 mg/kg/day. At 100 mg/kg/day, all laboratories detected significant effects on all of the androgen-dependent organs. None of the laboratories detected any effects of TREN on any dose on the optional systemic organ weights.
- For DDE, none of the three participating laboratories detected significant effects on any of the androgen-dependent organ weights at 5 or 16 mg/kg/day. At 50 mg/kg/day DDE, two of the participating laboratories detected significant effects on SV only. At 160 mg/kg/day DDE, all three laboratories detected significant effects on the weights of the VP, SV, LABC and COW. Two laboratories detected significant effects on paired epididymide weights and only one laboratory detected significant effects on the paired testes weights. There were no significant effects detected on the weights of the optional systemic organs in any of the three participating laboratories.

  For all of the organs of interest, except the paired testes, in animals exposed to TREN or LIN, there was a significant laboratory effect. For DDE, all of the organs of interest, including the paired testes, exhibited a significant laboratory effect.

The third stage (designated Phase 3) was analogous to final Phase 3 of the surgical castrate validation. The reliability of the weanling version was assessed by its response to and ability to correctly identify coded positive and negative substances in both agonist and antagonist studies. In these studies, the doses of specific positive substances were identical to that in Phase 3 of the castrate model, so that the reproducibility of the bioassay over time was also assessed. In addition, the relative sensitivity and effectiveness of the different accessory sex tissues and glands in the assay continued to be assessed in addition to the weights of the paired kidneys, paired adrenal glands, and liver.

In Phase 3, the six participating laboratories evaluated coded samples of DDE and LIN (weak anti-androgens), and DNP and NP, negative chemicals, and FLU (the reference anti-androgen), and TP (the reference androgen, not coded). The single TP dose (1.0 mg/kg/day) and the single FLU (3.0 mg/kg/day) doses were selected from Stage 1A and 1B (Chapter 4.0). The two doses each of DDE (16 and 160 mg/kg/day) and LIN (10 and 100 mg/kg/day) were selected from Phase 3 (Chapter 5.0) and the single doses of DNP (10 mg/kg/day) and NP (160 mg/kg/day) were selected from the literature.

The results were as follows:

- For DDE, no laboratories detected significant organ weight changes for VP, LABC, or COW at 16 mg/kg/day; two of the six laboratories detected significant changes in SV weight, and one laboratory detected a significant change in paired testes weight. At 160 mg/kg/day DDE, all six laboratories detected significant effects on the weights of the VP, SV, LABC and COW, and five of the six detected significant effects on both paired testes and paired epididymide weights.
- For LIN, none of the six laboratories detected significant differences on the weights of any of the organs of interest at 10 mg/kg/day. At 100 mg/kg/day LIN, four of the six laboratories detected significant effects on SV weight, three of the six detected significant effects in LABC and on paired epididymides, and two of the six detected significant effects on paired testes.

None of the laboratories detected any significant effects on any of the optional systemic organ weights from exposure to TREN, LIN, or DDE.

The results for the coded negative chemicals, DNP and NP, are as follows. For 10 mg/kg/day DNP, there were no statistically significant effects observed in any of the participating laboratories on the weights of any of the organs of interest. For NP at 160 mg/kg/day, there were no statistically significant effects observed for VP, SV, COW, or paired epididymides in any laboratory. One laboratory detected a significant effect of NP for LABC, and a different laboratory detected a significant effect of NP on paired testes weight. For the inhibitory positive control group (TP + FLU), all participating laboratories detected significant effects on the weights of the VP, SV, LABC and paired epididymides, and three of the six laboratories also detected significant effects in paired testes weight.

There were significant laboratory effects for all of the chemicals used and the vehicle control group values, for the weights of all of the organs of interest (including the optional systemic organs) except for the VP.

1.3 Overall Discussion and Conclusions

The validation effort of the intact, stimulated weanling male version of the Hershberger bioassay, in addition to the castrated adult male version of this assay, is in response to animal welfare concerns for the surgical procedures.

The intact weanling is less sensitive than the castrate model in that it requires more TP to get sufficient stimulatory response (1 mg/kg/day) versus the castrate adult (0.2 or 0.4 mg/kg/day). At equivalent doses of the anti-androgen FLU (3 mg/kg/day), the intact weanling also appears less sensitive to weak anti-androgens such as LIN or DDE (in the presence of TP) than the castrate adult. The reduced sensitivity of the intact weanling versus the castrate adult is likely due to the differences in age and hormonal status of the two animal models. The intact weanling male has an intact HPG axis (testes are present), which offers both advantages and disadvantages in the Hershberger bioassay. The intact weanling is at a stage when the hypothalamus and pituitary are just beginning to produce the appropriate releasing, stimulatory, and inhibitory endocrine hormones, and the target end organs (gonads and accessory sex organs [ASO]) are just beginning to detect the endocrine signals and respond to them. There are also low levels of endogenous T (slowly increasing from weaning to puberty) from the intact gonads. This process of HPG axis maturation proceeds slowly and incrementally to puberty and adulthood in males. In contrast, the castrate adult has previously achieved maximum/optimum HPG signaling with, and gonad/ASO responsivity to, high endogenous levels of T. Castration (surgical removal of testes and epididymides) as an adult suddenly removes the source of T production and the site of receipt of the hypothalamus-pituitary (HP) signaling, but presumably retains the mature target organ maximum/optimal recognition and sensitivity (not yet developmentally achieved in the intact weanling).

It has been suggested that the intact, stimulated weanling version of the Hershberger bioassay has the potential to detect effects on androgen-producing and androgen-sensitive organs from insults to and mechanism(s) mode(s) of action involving the hypothalamus, and/or pituitary, and/or testes, and/or target-

end organs, as well as through AR binding, which would be an advantage over the castrate model; however, no formal validation data for that purpose are currently available.

CHAPTER 2.0 INTRODUCTION

Recognizing that the environment contains compounds (both natural and man-made) that may interact and subsequently affect the endocrine systems of humans and animals, the OECD initiated a high-priority activity in 1998 to revise existing test guidelines and to develop new test guidelines for the screening and testing of potential endocrine disrupters (OECD, 1998). The Hershberger bioassay, a short-term *in vivo* screening test using accessory tissues of the male reproductive tract that originated in the 1930s and subsequently modified to include androgen-responsive muscles in the male reproductive tract (Korenchevsky, 1932; Eisenberg et al., 1949; Eisenberg and Gordan, 1950; Hershberger et al., 1953; Dorfman, 1962; Hilgar and Hummer, 1964; Dorfman, 1969; OECD, 2000; Ashby and Lefevre, 2000a), is a leading candidate for a Level 3 *in vivo* screening assay of the OECD Conceptual Framework (OECD 2002). As a short-term screen, the information generated by the bioassay can be used to build on that information already available (e.g., from relevant *in vitro* screens) to narrow the field of chemicals that may need longer term animal testing. The overall aim of the validation program is to demonstrate that the Hershberger bioassay is a reliable and reproducible bioassay that can be considered as the basis for an OECD test guideline. Once available, the test guideline is intended to be used as one element in an overall testing strategy for the detection and assessment of potential endocrine disruptors. The OECD validation of the Hershberger bioassay followed a comparable validation program as that used for the rat uterotrophic bioassay (Kanno et al., 2001; 2003a,b; Owens and Ashby, 2002; Owens et al., 2003).

The primary model for the Hershberger bioassay was the surgically castrated pubertal male. After several decades of use by the pharmaceutical industry, a protocol using the castrated peripubertal male rat was standardized by an official expert committee in 1962 to be used as a screening tool for androgenic chemicals (Dorfman, 1962). The castrated peripubertal male rat is an animal model whose condition is designed to achieve low endogenous hormone levels and employ target tissues that are highly responsive to administration of exogenous (anti-)androgenic hormones. The focus of the Hershberger bioassay is on the detection of compounds that may mimic or interfere with the action of endogenous male sex hormones. Based on the changes in weight of five androgen-dependent tissues, it evaluates the ability of a chemical to elicit biological activities consistent with androgen agonists, antagonists or 5α-reductase inhibitors, providing data about a single endocrine mechanism (i.e., [anti-]androgenicity) mediated by the AR, with or without metabolic activation in the intact liver (Dorfman, 1969; Bunyan et al., 1972; Campbell et al., 1983; You et al., 1999). This model was the basis for work in Phase 1, Phase 2 and Phase 3 of the Hershberger validation program. From 2001 to 2007, the rat Hershberger bioassay has undergone an extensive validation program (e.g., Sloan et al., 2002; O'Connor et al., 2002; Yamasaki et al., 2003a,b; Kang et al., 2004; Kennel et al., 2004; Freyberger et al., 2005; Owens et al., 2006; Yamasaki et al., 2006; Shin et al., 2007) including generation of a Background Review Document, compilation of a detailed methods paper (Gray et al., 2005), development of a dissection guide, and conduct of intra- and interlaboratory studies to show the reliability and reproducibility of the bioassay, using the VP, SV (plus fluids and coagulating glands), LABC muscle, COW, and the GP as the five target tissues. These validation studies were conducted with a potent reference androgen (TP), two potent synthetic androgens (TREN and MT), a potent anti-androgenic pharmaceutical (FLU), a potent inhibitor of the synthesis of the natural androgen (DHT) (i.e., FIN), several weakly anti-androgenic pesticides (LIN, VIN, PROV, DDE), a potent 5α-reductase inhibitor (FIN) and two known negative chemicals (DNP and NP, the DDT

metabolites) (OECD, 2002, 2005; Owens et al., 2006; OECD, 2003, 2007, 2008). A draft test guideline was the culmination of the long historical experience with the bioassay, the experience gained during the validation test programme, and the results obtained therein.

Animal welfare concerns with the castration procedure prompted OECD to conduct a more limited validation program under ICCVAM (1997) and OECD (1996) in three laboratories using the intact (uncastrated) stimulated weanling male as an alternative model to the traditional Hershberger bioassay that avoided the castration step. The objective in Phase 3 of the validation with this model protocol is to assess the potential of the stimulated weanling as reliable and relevant alternative to the surgically castrated rat model. For the stimulated weanling, the program encompasses both previous and current validation work on the surgical castrate.[1] The current stimulated weanling protocol specifics are based largely on the standardization and optimization work performed in the Lab A (Ashby and Lefevre, 2000b; Ashby et al., 2002a,b).

The OECD program on the Hershberger bioassay has previously demonstrated the response and the reliability of the surgical castrate model by measuring male accessory sex glands and tissues, i.e., VP, SVCG, LABC, COW, and GP (Kim et al., 2002). In the case of the stimulated weanling, the mandatory target tissues were reduced to four: the VP, SVCG, LABC, and COW. The list does not include the GP since the weanling male has not yet achieved PPS, and the prepuce cannot be reliably separated from the penile shaft in the male's sexually immature state. However, two tissues removed during castration are available and optional organs for investigation in this intact model, the paired testes and paired epididymides.

The Hershberger bioassay with the stimulated intact weanling male provides data not only on (anti-)androgenicity mediated by the AR, as does the castrated peripubertal male assay, but also a number of additional possible mechanisms (inhibition of chemical metabolism, hypothalamic and/or pituitary regulation of the gonad, etc.), as the animal model goes through puberty (Ojeda and Urbanski, 1994; Monosson et al., 1999; Stoker et al., 2000; George et al., 2003; Schladt, 2008) since in this model, the HPG axis is intact.

Biological plausibility would suggest that the stimulated weanling and the surgically castrated versions of the Hershberger bioassay should respond in a qualitatively similar manner. In both versions, growth of the target mandatory tissues is controlled through the AR, and T is converted to DHT by the 5α-reductase enzyme. The same male reproductive tract tissues can be employed in both versions of the assay, with the exception of the GP, paired testes and paired epididymides, as noted above.

There are, however, differences to be noted between the stimulated weanling and the surgically castrated male. As a result, comparative work was needed to define the quantitative differences that may exist between the two versions. The key differences in the stimulated weanling are low prepubertal circulating levels of T produced by the testes and the potential ability of the intact HPG axis to respond and compensate to some degree to chemical insult. These may diminish the responsiveness of the stimulated weanling version of the assay. For further background, there are reviews and recent manuscripts that trace the development of the male rat for both hormonal levels and the growth of reproductive tract tissues prior to puberty (Owens et al., 2006; OECD, 2006; Joubert et al., 1994; Korenchevsky and Dennison, 1935; Korenchevsky, 1932; Korenchevsky et al., 1932). It should be noted that the PND 20-35 time period in the male is not fully analogous to the female. There are no dramatic surges in the male in hormonal levels or tissue growth in this prepubertal time period. The change and rearrangement of the HPG axis is more gradual in circulating FSH, LH, T, and 5α-reduced steroids. At the same time, just as this is a period of rapid overall growth in the animal, the weights of the testes and the relevant tissues of the reproductive tract are also increasing in absolute terms. The increase *relative to body weight* for these

[1] For the surgical castrate, Phase 3 refers only to work with coded test substances, including both positive and negative test substances.

tissues is rather modest and constant, indicating no disproportionate surge in growth.

The program to assess the stimulated weanling as an alternative proceeded in the same fashion as that taken by the surgical castrate model. Due to the low circulating T levels and the intact HPG axis in the weanling, TP (CASRN 57-85-2) and FLU (CASRN 1311-84-7) standardization curves with the stimulated weanling male were necessary to ensure the proper TP dose was selected for the antagonist studies. Therefore, the first stages, analogous to Phase 1A and Phase 1B with the surgical castrate (Eisenberg and Gordan, 1950; Hilgar and Hummel, 1964) were:

- A dose-response curve with TP was generated as the first step to select optimum candidates for the TP co-administration dose with anti-androgens. The basis for selection was intended to be the same as with the surgical castrate: to identify ED50-ED70 doses of TP that are applicable to all the responsive tissues. The doses employed were 0.4, 0.8, 1.0, 1.2, and 1.6 mg/kg/d in order to provide a complete comparison with the surgical castrate data set. This dose series was selected to adequately characterize the dose response for each of the four tissues (VP, SV, LABC, and COW), its relative maximum response, and provide insight on the CVs that labs can achieve with the smaller weanling tissues. With the vehicle control, this comprised a total of six dose groups per laboratory (36 animals based on n = 6 per dose group). The results allowed the selection of two TP doses to proceed into the second part of Stage 1, the FLU dose series experiments.
- Two selected doses of TP were used as the reference doses for co-administration to assess the response of stimulated weanling tissues to a dose series of FLU. The doses of FLU employed were 0.3, 1.0, and 3.0 mg/kg/d FLU. Higher doses were judged to be wasteful of animals, based on the available data. With the two TP reference dose groups, this comprised a total of eight dose groups per laboratory (48 animals based on n = 6 per dose group). Two additional groups were voluntary: a vehicle control and a fourth dose of FLU at 10 mg/kg/d. This allowed the selection of the optimum TP reference dose in order to proceed into the second stage, the dose response experiments to known weak agonists and antagonists.

The next stage for the stimulated weanling comprised studies analogous to the Phase 2 dose response studies for the surgical castrate:

- The stimulated weanling was used to assess its capability to detect a weaker androgen agonist, TREN, and to detect two weak androgen antagonists, LIN and DDE. All three compounds challenged the surgical castrate in one or more laboratories where one or more of the five castrate mandatory tissues did not achieve statistical significance. In order to make a comparison, the same dose series was used, but the lowest dose was omitted in the interests of animal welfare and conserving resources:
  - TREN (CASRN 10161-33-8):
    1.5, 8, and 40 mg/kg/d (rationale: no laboratory reached statistical significance with the surgical castrate at 1.5 mg TREN/kg/d; therefore, the 0.3 mg/kg/d dose is not an apparent sound use of animals or resources)
  - LIN (CASRN 330-55-2):
    10, 30, 100 mg/kg/d (rationale: no laboratory reached statistical significance with the surgical castrate at 10 mg LIN/kg/d; therefore, the 3 mg/kg/d dose is not an apparent sound use of animals or resources)
  - DDE (CASRN 72-55-9):
    16, 50, 160 mg DDE /kg/d (rationale: no laboratory reached statistical significance with the surgical castrate at 10 or 16 mg/kg/d – doses inside and outside of Japan, respectively; therefore, the 5 mg/kg/d or similar dose is not an apparent sound use of animals or resources)

The third stage was analogous to final Phase 3 of the surgical castrate validation. The weanling version's reliability was assessed by its response to and ability to correctly identify coded positive and negative substances in both agonist and antagonist studies. In these studies, the doses of certain positive substances were identical to that in Stage 3, so that the reproducibility of the bioassay over time was also assessed (e.g., Yamasaki et al., 2006). In addition, the relative sensitivity and effectiveness of the different accessory sex tissues and glands in the assay continued to be assessed. The model protocol employed is described in more detail in Chapter 3.0. The six participating testing laboratories, and the stages in which they participated, are presented in Table 2.1.

The rodent Hershberger bioassay has been based historically on changes in the weights of androgen-responsive male accessory sex tissues largely in peripubertal, castrated male rats. Accessory sex tissues and glands depend upon androgen stimulation to gain and maintain weight during and after puberty. When endogenous sources of androgen are low (the weanling before puberty), the biological activity of exogenous substances can be assayed by the increase (agonist response) in the weights of these accessory sex tissues or by blocking (antagonist response) the activity of administered androgens and by preventing an increase in the weights of these accessory sex tissues. The rodent Hershberger bioassay modified to use the stimulated weanling therefore evaluates the ability of a chemical to show biological activities consistent with the agonism or antagonism of potent endogenous androgens (e.g., T and DHT).

The available data indicate that the androgen-dependent accessory sex tissues of the weanling are sensitive to androgens. This is plausible as these animals have both ARs and appropriate steroidogenic enzymes necessary for agonist, antagonist, and 5α-reductase inhibitor responses. In addition to sensitivity, the weanling rodent accessory sex tissues have a small relative weight.

Table 2.1. The Six Participating Testing Laboratories

Stages Performed	Laboratories[1]
1, 2, 3	Lab A
1, 2, 3	Lab B
1, 2, 3	Lab C
2	Lab D
3	Lab E
3	Lab F

[1] The identity of the participating laboratories is available to Government Representatives of OECD Member countries.

Primary objectives of Phase 3 of the validation program were to demonstrate:

- In Stage 1, (a) the dose response of the stimulated weanling version of the Hershberger bioassay to a series of TP doses for comparisons with the surgical castrate and as the basis for selecting two doses for co-administration studies with FLU, and, (b) the dose response of the stimulated weanling version of the Hershberger bioassay to a series of FLU doses using two selected TP co-administration doses for comparisons with the surgical castrate and as the basis for selecting a single TP dose for co-administration with antagonists in dose response studies
- In Stage 2, the dose response of the stimulated weanling version of the Hershberger bioassay to a series of doses with a weak agonist, TREN, and two weak antagonists, LIN and DDE
- In Stage 3, to assess the capability of the stimulated weanling version of the Hershberger bioassay to identify coded positive and negative substances
- The relative effectiveness of the different accessory sex tissues and glands in the assay
- The reproducibility of the stimulated weanling version of the bioassay over time by comparing appropriate data from Stage 3 to that generated in Stages 1 and 2
- Continue the investigation of the value of the different accessory tissues and glands

The test substances were coded and administered to groups of six animals (n = 6 per dose group) for ten consecutive days in all studies in all stages. The animals were then necropsied approximately 24 hours later on the 11th day (24 hours after the last test substance administration). After dissection, the weights of the mandatory accessory sex tissues were measured as well as the paired testes and paired epididymides weights.

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

Biological activity consistent with androgen agonists was tested by administering a test substance to

intact, weanling male rats for ten consecutive days. The positive control for the tissue responses was TP. The vehicle was the negative control. The weights of the accessory sex tissues of the test chemical groups were compared to the vehicle group for a statistically significant increase in weight.

Biological activity consistent with androgen antagonists and 5α-reductase inhibitors was tested by administering the test substance to intact, weanling male rats for ten consecutive days together with a reference androgen agonist, TP. Administration of TP alone was the negative control. The weights of the accessory sex tissues after co-administration of the test chemical and the reference androgen TP together were then compared with the weights of tissues of the reference androgen TP alone for a statistically significant decrease in weight. FLU may be co-administered with TP to another group as a positive control.

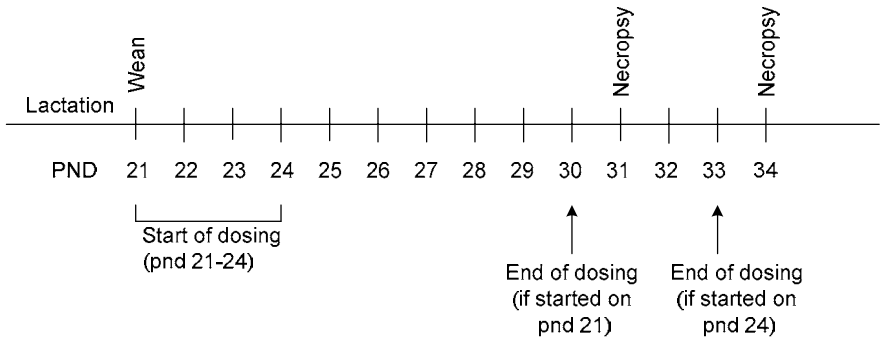
This document provides the comparison and interpretation (intralaboratory and interlaboratory) of the statistically analyzed data generated by the six testing laboratories participating in the validation effort for the intact, stimulated weanling version. It also provides interpretation and discussion of the statistical comparison of within and across laboratory variability for the endpoints of interest.

Chapter 3.0 will present the model protocol. Chapter 4.0 will present the results of Stage 1: studies with TP and FLU. Chapter 5.0 will present the results of Stage 2: studies with weak agonists and antagonists. Chapter 6.0 will present the result of Stage 3: studies with coded chemicals. Chapter 7.0 will present the comparison of the intact stimulated weanling male version of the Hershberger bioassay with the adult castrated male version of the Hershberger bioassay. Chapter 8.0 will provide overall discussion and conclusions. Chapter 9.0 will be acknowledgements. Chapter 10.0 will consist of references for the entire document.

CHAPTER 3.0 THE MODEL PROTOCOL

The model protocol for the OECD intact stimulated weanling male version of the Hershberger bioassay is taken directly from the Validation Program Phase 3 Draft (dated 17 October 2003, paged 7 through 15), since this protocol was followed by the 6 testing laboratories during this validation effort. The list of the six testing laboratories and the stages in which each participated are presented in Table 2.1, Chapter 2.0; see Figure 1 for a graphical representation of the intact, stimulated weanling male version of the Hershberger bioassay.

Figure 1. Study Design for the Intact Stimulated Weanling Version of the Hershberger Bioassay



In-Life

- Body weight: 50 ± 10 g at study start
- Ten consecutive days of dosing
- Route: Oral gavage or SC injection
- Numbers: Six males (3/cage and 2 cages)/group; minimum 2 doses plus negative control (positive control group optional)
- Data: Body weights and clinical observations daily, feed consumption by cage for entire interval (optional)

Necropsy

- 24 hours after last dose
- Body weight, euthanasia, exsanguination

Mandatory Accessory Sex Tissues

- VP
- SVCG
- LABC
- COW

Optional Tissues	**Optional Assessments in Blood**
Reproductive: Paired testes	Serum LH
Paired epididymides	Serum T
Systemic: Liver	Serum T_4 (tetraiodothyronine; thyroxine)
Paired kidneys	Serum T_3 (triiodothyronine)
Paired adrenal glands	

Animal Species and Strain

Laboratories are allowed to select the strain of rat to be used in the validation of the assay with this protocol. The strain used should be the strain historically used by the participating laboratory (where the screening assay may be preliminary to a repeated dose oral study, a reproductive and developmental study, or a long-term study). Preferably, animals from the same strain and source should be used in all studies. If a laboratory is planning to use an unusual rat strain or one unique to their own facility, the strain selected should be based on the sexual development criteria noted under "Initial Considerations and Principle of the Assay" in Chapter 2.0.

Age and Acclimatisation

Young weanling animals should be employed in a relatively small time window between weaning and before puberty (i.e., PND 20 to 34). Animals should be observed daily, and any animals with evidence of disease or physical abnormalities should be removed. The treatment with initiation of dosing (on study) may commence as early as pnd 21 days of age, but preferable not later than pnd 24. The laboratory is allowed some flexibility to schedule the experimental work efficiently.

Housing and Feeding Conditions

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, with the photoperiod 12 hours light, 12 hours dark.

Laboratories participating in the validation should use the laboratory diet normally used in their chemical testing work. In previous phases, no effects or variability were observed that were attributable to the diet. The diet used should be recorded, and a sample of the laboratory diet retained for possible future analysis. Both diet and drinking water should be supplied *ad libitum*.

Weanling animals should be caged in groups of no more than 6 similarly treated rats per cage, with a minimum of 1 cage of 6 rats/cage per treatment group and a maximum of 2 cages of 3 rats/cage per treatment group. When cages are properly sized (~2000 square centimeters), 6 animals or less per cage avoids crowding. Cages should be thoroughly cleaned to remove possible contaminants and arranged in such a way to minimize possible effects due to cage placement.

Each animal should be identified individually (e.g., ear mark or tag, or tattoo) and the method of identification recorded.

Body Weight and the Selection of Animals for the Study

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

Experimental control of BW is accomplished in 2 steps. The first step involves selection of animals with relatively small variation in BW for the study cohort from the larger population of animals that have been

supplied. Unusually small or large animals should be avoided and not be placed in the study cohort. A reasonable level of BW variation within the study cohort should be tolerated. Here, ± 20% of the mean BW for the cohort population is judged to be reasonable (e.g., 50 g ± 10 g). The second step involves the assignment of animals to different treatment groups (n= 6) by a randomized complete block approach. Under this approach, animals are randomly assigned to treatment groups so that each group has the same mean and standard deviation (SD) in weight at the beginning of the study. The procedure used for block randomization should be recorded.

Non-routine Health and Safety Requirements

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

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

PROCEDURE - VALIDATION OF THE STIMULATED WEANLING

Administration of Doses

TP should be administered by sc injection in all stages of the validation program. All other test substances should be administered by oral gavage.

Sc injections should be on the dorsal surface of the animal after shaving or trimming of fur. Multiple injections sites may be used. The maximum limit on the volume administered per animal is approximately 1.0 ml/kg BW per day.

Oral gavage should be the delivery of the test substance in vehicle by means such as intubation with an oral gavage syringe. The maximum limit on the volume administered per animal should be 5 ml/kg/day. Since the weanling animals are small, the technical staff conducting the gavage should be experienced in order to avoid gavage errors that might lead to morbidity or mortality.

The animals should be dosed in the same manner and time sequence for 10 consecutive days at approximately 24-hour intervals. The dosage level should be adjusted daily based on the concurrent daily measures of BW. The volume of dose and time administered should be recorded on each day of exposure.

Good Laboratory Practice

Work should be conducted according to the principles of Good Laboratory Practice (OECD Good Laboratory Practice and Compliance Monitoring [12]). In particular, data should have a full audit trail and be retained on file. Data should be collected in a manner that will allow independent peer review, and all written records should be maintained.

Vehicle

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

The following procedures should be conducted in 4 consecutive stages.

STAGE 1 (analogous to Phase-IA with the surgical castrate)

Test Substance and Doses in Stage 1 of the Stimulated Weanling Validation

The test substance for Stage 1 of the stimulated weanling validation is TP.

The reference androgen agonist, TP, should be administered in a series of doses comprising 0.4, 0.6, 1.0, 1.2, and 1.6 mg/kg/day.

Test Groups in Stage 1 of the Stimulated Weanling Validation

Six animals of the same age and cohort should be used for the vehicle, the TP doses, and any other control group or treatment used by the laboratory.

The response of the accessory sex tissues and glands to a reference agonist, TP, should be studied. This work involves 5 test groups with increasing doses of TP and 1 vehicle control group. The test groups are illustrated in Table 3.1.

Table 3.1. Test Groups for Stage 1 of the Stimulated Weanling Validation: TP Dose Series

	Agonist Response
Group A Vehicle Control	Vehicle only
Group B	0.4 mg/kg/day TP
Group C	0.8 mg/kg/day TP
Group D	1.0 mg/kg/day TP
Group E	1.2 mg/kg/day TP
Group F	1.6 mg/kg/day TP

STAGE 2 (analogous to Phase-IB with the surgical castrate)

Test Substance and Doses in Stage 2 of the Stimulated Weanling Validation

The test substances for Stage 2 of the stimulated weanling validation are TP and FLU.

Two doses of the reference androgen agonist, TP, will be selected based upon Stage 1 data, and these selected doses will be co-administered with a series of FLU doses. The FLU doses are 0.3, 1.0, and 3.0 mg/kg/day. Laboratories may voluntarily extend the FLU dose series by adding test groups with the selected TP dose plus 10 mg/kg/day.

Test Groups in Stage 2 of the Stimulated Weanling Validation

Six animals of the same age and cohort should be used for the vehicle, the 2 selected TP doses, and 2 series of 3 groups each where a selected TP dose is co-administered with an increasing series of FLU doses.

The response of the accessory sex tissues and glands to selected doses of a reference agonist (TP), co-administered to a series of FLU doses, should be studied. This work involves TP control groups for each selected dose and 3 test groups for each selected dose, where the TP dose is co-administered with FLU, having a prescribed dose of the test substance and 1 vehicle control group. The required and voluntary test groups are illustrated in Table 3.2.

Table 3.2. Test Groups for Stage 2 of the Stimulated Weanling Validation: Two Selected TP Doses Co-administered With FLU Dose Series

	Antagonist Response Selected TP Dose #1	**Antagonist Response Selected TP Dose #2**
Group A Vehicle Control **Voluntary**	Vehicle **Voluntary**	Vehicle **Voluntary**
Group B Negative Control	Selected Dose of TP #1	Selected Dose of TP #2
Group C	Selected Dose of TP #1 + 0.3 mg/kg/day FLU	Selected Dose of TP #2 + 0.3 mg/kg/day FLU
Group D	Selected Dose of TP# 1 + 1.0 mg/kg/day FLU	Selected Dose of TP #2 + 1.0 mg/kg/day FLU
Group E	Selected Dose of TP #1 + 3.0 mg/kg/day FLU	Selected Dose of TP #2 + 3.0 mg/kg/day FLU
Group F **Voluntary**	Selected Dose of TP #1 + 10 mg/kg/day FLU **Voluntary**	Selected Dose of TP #2 + 3.0 mg/kg/day FLU **Voluntary**

STAGE 3 (analogous to Phase-2 with the surgical castrate)

Test Substance in Stage 3 of the Stimulated Weanling Validation

The reference androgen agonist is TP, and the reference antagonist is FLU. The appropriate doses of these substances are selected based upon Stage 2 data.

The test substances for Phase 3, Stage 3 dose-response studies for the stimulated weanling are:

TREN CAS No. 10161-33-8
LIN CAS No. 330-55-2
DDE CAS No. 72-55-9

The agonist test substance doses for Phase 3, Stage 3 dose-response studies for the stimulated weanling is:

TREN 1.5, 8, and 40 mg/kg/day

The 2 weak antagonists are administered along with a dose of TP selected from the Stage 2 studies with each of the following doses:

LIN 10, 30, 100 mg/kg/day
DDE 16, 50, 160 mg/kg/day

Test Groups in Stage 3 of the Stimulated Weanling Validation

Six animals of the same age and cohort should be used for the vehicle, TP, any other control group, and for each treatment or test substance group.

The response of the accessory sex tissues and glands to a weak agonist, TREN, should be studied. This work involves 3 test groups for each agonist having a prescribed dose of the test substance and 1 vehicle control group used as the negative control. The required test groups are illustrated in Table 3.3.

The response of the accessory sex tissues and glands to 2 weak antagonists, LIN and *p,p* -DDE, should be studied. This work involves 3 test groups for each antagonist and the co-administration of a dose of the reference agonist, TP, to each group. Each test substance should have a prescribed dose for each of its groups. The positive TP control group for the antagonist series and the dose of co-administered TP (1.0 mg/kg/day) should be the same and selected based upon Stage 2 results. The required test groups are illustrated in Table 3.

Table 3.3. Test Groups for Stage 3 of the Stimulated Weanling Validation: Agonist and Antagonist Dose Responses

	TREN Response	**LIN Response**	**DDE Response**
Group A Vehicle Control	Vehicle **Mandatory** for agonist	Vehicle **Voluntary** for antagonist	Vehicle **Voluntary** for antagonist
Group B Negative Control	Provided by vehicle control (no additional group needed for agonist)	Selected TP dose from Stage 2	Selected TP dose from Stage 2
Group C	1.5 mg/kg/d TREN	Selected TP dose from Stage 2 + 10 mg/kg/d LIN	Selected TP dose from Stage 2 + 16 mg/kg/d DDE
Group D	8 mg/kg/d trenbolone	Selected TP dose from Stage 2 + 30 mg/kg/d LIN	Selected TP dose from Stage 2 + 50 mg/kg/d DDE
Group E	40 mg/kg/d trenbolone	Selected TP dose from Stage 2 + 100 mg/kg/d LIN	Selected TP dose from Stage 2 + 160 mg/kg/d DDE
Group F Positive Treatment Group **Voluntary**	TP dose selected from Stage 2 **Voluntary**	TP dose selected from Stage 2 + FLU dose selected from Stage 2 **Voluntary**	TP dose selected from Stage 2 + FLU dose selected from Stage 2 **Voluntary**

STAGE 4 (analogous to Phase 3 with the surgical castrate)

Coded Samples in Stage 4 of the Stimulated Weanling Validation

Six animals of the same age and cohort should be used for the vehicle, TP, and any other control group, and for each treatment or test substance group.

All participating laboratories should use the coded test substances supplied and follow the instructions for the preparation of the proper doses, in order to achieve specified dosages that can be compared to data generated in Stage 3. There are 2 coded series. One series is for androgen agonists, and this series requires a positive vehicle control. A second series is for androgen antagonists and requires a reference TP dose group using a 1.0 TP mg/kg/d dose to be selected based upon the results from Stage 2. In the latter antagonist series, a vehicle control group and a positive control group of FLU, plus the selected TP dose, are voluntary. The required test groups are illustrated in Table 3.4.

Table 3.4. Test Groups for Stage 3 of the Stimulated Weanling Validation: Coded Dose Studies

	Agonist Coded Samples	**Antagonist Coded Samples**
Group A Vehicle Control	Vehicle **Mandatory** in agonist series	Vehicle **Voluntary** in antagonist series
Group B Negative Treatment Control	Not applicable - provided by vehicle control (no additional group needed for agonist series)	Selected TP from Stage 2
Group C	Coded Agonist Test Substance #1[a]	Selected TP from Stage 2 + Coded Antagonist Test Substance #5[a]
Group D	Coded Agonist Test Substance #2[a]	Selected TP from Stage 2 + Coded Antagonist Test substance #6[a]
Group E	Coded Agonist Test Substance #3[a]	Selected TP from Stage 2 + Coded Antagonist Test Substance #7[a]
Group F	Coded Agonist Test Substance #4[a]	Selected TP from Stage 2 + Coded Antagonist Test Substance #8[a]
Group G	Not applicable	Selected TP from Stage 2 + Coded Antagonist Test Substance #9"
Group H	Not applicable	Selected TP from Stage 2 + Coded Antagonist Test Substance #10[a]
Group F Positive Treatment Group	TP dose selected from Stage 2 **Voluntary**	TP dose selected from Stage 2 + FLU dose selected from Stage 2 **Voluntary**

[a] The doses of each test substance should be prescribed in order to have comparable data for the analysis of variability among labs and for comparison to previous data. The test substances in the agonist and antagonist coded samples may or may not be identical. Therefore, sequential numbering is used across the series to avoid any suggestion that they might be the same.

Clinical Observations

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

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

Body Weight and Food Consumption

Individual BWs should be recorded prior to start of treatment (to the nearest 1 g or to the nearest 0.1 g if that is normal practice of the laboratory with smaller animals), on each day of the administration period, and prior to necropsy. Group means and SD should be calculated.

Food consumption should generally be observed and any significant changes recorded. It may be voluntarily recorded per cage, and an average value per animal calculated based upon the number of animals per cage.

Necropsy

Approximately 24 hours after the last administration of the test substance, the rats should be euthanized according to the normal procedures of the participating laboratory and necropsied. The method of humane killing should be recorded in the laboratory report.

The order in which the animals are necropsied should be designed such that 1 or 2 animals from each group (e.g., 1 per cage if there are 3 animals per cage) are necropsied to achieve a randomization of the groups. In this way, all the animals in the same treatment group are not necropsied at once, and any variation in the procedure over time should not unduly impact any particular group.

The 4 accessory sex tissues (VP, SV, LABC, COW) are mandatory measurements. Two additional tissues from the male reproductive tract, paired testes and paired epididymides, are optional measurements. All mandatory and optional tissues should be excised, carefully trimmed of excess adhering tissue and fat, and their fresh (unfixed) weights determined. Each tissue should be handled with particular care to avoid the loss of fluids and to avoid desiccation, which may introduce significant errors and variability by decreasing the recorded weights.

Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Previous work has indicated a range of coefficient of variations (CVs) that appears to differ based upon the proficiency of the laboratory. In a few cases, large differences in the absolute weights of the tissues, such as the VP and COW, have been observed within a particular laboratory. Therefore, it is important that technical staff responsible for the dissection of the accessory sex tissues be familiar with standard dissection procedures for these tissues. An SOP manual for dissection has been provided by the Lead Laboratory and was used in Phase-1 and Phase-2. This manual will remain the SOP reference for Phase-3. Careful training according to the SOP guide will minimize a potential source of variation in the study.

Each of the 4 mandatory accessory sex tissues, as well as the paired testes and paired epididymis, should be weighed without blotting to the nearest 0.1 mg, and the weights recorded for each identified animal.

Liver, paired kidney, and paired adrenal weights are other optional measurements. Again, tissues should be trimmed free of any adhering fascia and fat. The liver should be weighted to the nearest 0.1 g, and the paired kidneys and paired adrenals weighed to the nearest 0.1 mg. All weights should be recorded for each identified animal.

If the evaluation of each chemical requires necropsy of more animals than is reasonable for a single day, the starting date may be staggered on 2 consecutive days, so the necropsy can be staggered and the work burden of a single day reduced. 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 1 day for the second half of the animals. That is, each group should be split so that half of the animals are necropsied on each day in order to control variability among the groups.

Carcasses should be disposed of in an appropriate manner following necropsy.

REPORTING

Data

Data should be reported individually (i.e., BW, accessory sex tissue weights, optional measurements, and other responses and observations) and for each group of animals (means and SD). The data should be summarized in tabular form and show the number of animals at the start of the test, the number of animals found dead during the test, the number of animals showing signs of toxicity, and a description of the signs of toxicity observed, including time of onset, duration, and severity.

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

Test Report

The test report must include the following information:

- **Laboratory identification**
  - Name and location of laboratory
  - Principal investigator and other personnel, and their roles in the study
  - Dates study began and ended
- **Test substance**
  - Physical nature and, where relevant, physicochemical properties
  - Identification data and source
  - Purity
- **Vehicle identity and supplier**
- **Test animals and procedures**
  - Species/strain used
  - Source or supplier of animals, including full address
  - Number, age and sex of animals
  - Housing conditions (temperature, lighting, etc.), diet used, lot of diet, source of diet, bedding, and source of bedding
  - Caging conditions and number of animals per cage
  - Age at receipt, age at start of test substance administration and time of acclimatization
  - Individual weights of animals at the start of the study (to nearest 0.1 g)
  - Randomization process and a record of the assignment to vehicle, reference, and test substance groups
  - Mean and SD of the BWs for each group at the start of the study;
  - Necropsy procedures, including means of exsanguination and any anesthesia
- **Results**
  - Daily observations during administration, including:
    - Daily BWs (to the nearest 1 g)
    - Clinical signs (if any)
    - Test substance treatment (Yes or No) and the identify of that test substance

  - Dose level and volume administered each day
  - Time of dosing each day
  - Notes on food consumption or measurement of actual consumption each day
- On the day of necropsy, individual necropsy data on each animal, including absolute accessory sex tissue weights, liver and BWs including the following:
  - Date of necropsy
  - Animal ID
  - Home cage number or ID
  - Prosector
  - Time of day necropsy performed
  - Animal age
  - Order of animal killing and dissection at necropsy
- Weights of all four mandatory accessory sex tissues and glands
  - VP (fresh weight to the nearest 0.1 mg)
  - SVCG, including fluid (fresh weight - paired, to nearest 0.1 mg)
  - LABC (fresh weight to nearest 0.1 mg)
  - COW (fresh weight paired, to nearest 0.1 mg)
- Weights of additional male reproductive tissues
  - Testes (fresh weight, paired, to nearest 0.1 mg)
  - Epididymides (fresh weight, paired, to nearest 0.1 mg)
- Weights of optional tissues, if performed
  - Liver (optional - to nearest 0.1 g)
  - Kidney (optional - paired, to nearest 0.1 mg)
  - Adrenal (optional - paired, to nearest 0.1 mg)
- General remarks and comments

- **Discussion**

- **Conclusions**

Statistics and Interpretation of Results

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

LITERATURE IN THE PROTOCOL (also in Chapter 9.0 References)

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

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

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

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

5. Hershberger, L, Shipley, E, Meyer, R. (1953): Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method. Proc. Soc. Exp. Biol. Med. 83:175-180.

6. OECD (2002): Summary record of the sixth meeting of the Task Force on Endocrine Disrupters Testing and Assessment (EDTA 6). ENV/JMn\3/EDTA/M(2002)2

7. Stoker, TE, Parks, LG, Gray, LE, Cooper, RL. (2000): Endocrine-disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations. Crit. Rev. Toxicol. 30(2): 197-252.

8. Ojeda, SR, Urbanski, HF. (1994): Puberty in the rat. In *The physiology of reproduction. Second edition.* E. Knobil and J. D. Neill, eds. Raven Press, NY. pp.363-409.

9. Monosson, E, Kelce, WR, Lambright, CR, Ostby, J, Gray, LE, Jr. (1999): Peripubertal exposure to the antiandrogenic fungicide, VIN, delays puberty, inhibits the development of androgen-dependent tissues, and alters androgen receptor function in the male rat. Toxicol. Ind. Health 15:65-79.

10. Ashby, J, Lefevre, P. (2000): The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti-androgens, oestrogens and metabolic modulators. J. Appl. Toxicol. 20:35-47.

11. OECD (2002): Final OECD report of the work towards the validation of the rat Hershberger assay: Phase-1, Androgenic response to testosterone propionate and anti-androgenic effects of flutamide. ENV/JMATG/EDTA(2002)l/REV2. October, 2002. Paris, France.

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

CHAPTER 4.0 RESULTS OF STAGE 1A WITH TP AND STAGE 1B WITH TP PLUS FLU

4.1 Phase 1A: TP Dose-response Studies

Stage 1 of the OECD validation program, of the intact stimulated weanling version of the Hershberger bioassay, was divided into Phases 1A and 1B. In Phase 1A, the dose response of the weights of the following androgen-dependent tissues –VP, SVCG, LABC, COW, paired testes, and paired epididymides as mandatory weights; and liver, paired adrenal glands, and paired kidneys as optional weights – were assessed against a series of TP (CAS No. 1255-49-8; the reference androgen) doses as follows: 0 (vehicle), 0.4, 0.8, 1.0, 1.2, or 1.6 mg/kg/day, administered by SC injection once daily for 10 consecutive days in corn oil, stripped corn oil, or peanut oil vehicle. Three of the six testing laboratories (Chapter 2.0, Table 2.1) participated in Phase 1A: Laboratories A, B and C. Laboratories B and C represented government, and the Lab A represented industry. The lead laboratory (Dr. L. Earl Gray, Jr., Reproductive Toxicology Division, NHEERL, U.S. EPA, Research Triangle Park, NC) did not perform any Phase 1 studies.

The objectives of this Phase 1A study were to: (1) identify the dose response to TP (the reference androgen), and (2) select the appropriate TP dose to use as the stimulatory dose in the subsequent anti-androgen studies (Phase 1B).

It should be noted that, compared to the adult castrate version of the Hershberger bioassay, the dissection and weighing of the GP is problematic in immediately postwean prepubertal males due to no, or incomplete, PPS by the time of necropsy on PND 31-34. PPS occurs in rats at approximately PND 42-45 (the lack of PPS at necropsy was confirmed by the performing test laboratories). Therefore, the GP was not dissected or weighed for the weanling version. Since the weanling rats are intact, the paired testes and epididymides (which were surgically removed in the castrated adult males) are available and are optional organs for dissection and weighing in the intact stimulated weanling males, since they are both susceptible to direct androgenic control or indirect control through the feedback loop from the anterior pituitary by changes in circulating levels of LH. The HPG axis in the intact male is active at low levels in the weanling and slowly increases incrementally until puberty (in contrast with the female which exhibits a “spurt,” with the HPG hormones peaking at puberty).

Table 4.1 presents an overview of the Phase 1A protocol.

Table 4.1. Androgen Protocol Summary

	Factor	Protocol Specifications
Animals	Species	Rat
	Strain	No strain preference (one exclusion, not Fischer 344)
	Age at initiation of treatment	20-22 days
	Weight at time of treatment	Not specified; groups should be withi ± 20%
Animal Husbandry	Diet	Laboratory preference
	Bedding	Laboratory preference
	Caging [a]	Laboratory preference
Treatment Regimen	Animals per dose group	6
	Vehicle	Corn/peanut oil
	Volume of administration SC	1.0 ml/kg bw/day
	Dosing regimen (mg/kg bw/day)	10 consecutive daily administrations
	Sacrifice	24 hrs after last treatment
Measurements	Mandatory weights [b]	Total bw (to 0.1 g) VP (fresh) SVCG LABC COW
	Optional weights and measurements [b]	Testes Epididymides Liver weight (to 0.1 g) Adrenal gland (paired) weight Kidney (paired) weight

[a] For the immature animals, two or three animals per cage are recommended.

[b] Unless otherwise specified, the weight should be measured to 0.1 mg

Table 4.2 presents the laboratory parameters and conditions for the Phase 1A validation program.

Table 4.2. Laboratory Parameters and Conditions for the Phase 1A Validation

Lab	**Rat Strain & Supplier**		**Age on Study/ Necropsy**	**Vehicle**	**Diet**	**Animals per Cage**	**Bedding**
Lab A	Alpk:APfSD	Breeding facility on site	22-23 days/ 32-33/days	Tocopherol-stripped corn oil	R&M1, Special Diet Services, Witham, Essex, UK [a]	3	Sawdust, paper
Lab B	Sprague Dawley	Breeding facility on site	22 days/ 32 days	corn oil	PMI 5001, Agribrands Purina, Lab B	3	Willow shavings
Lab C	Crj:CD IGS (SD)	Charles River Canada, St Constant, Quebec, Canada	22 days/ 32 days	peanut oil	PMI 5001, Agribrand Purina,	2	Pro chip maple (dust-free hardwood chips), PWI Industries, St-Hyacinthe, Quebec, Canada

[a] Purified (semi synthetic) diet

Table 4.3 presents the mean ± SD for BW and all of the tissues from each of the three performing laboratories at the various TP doses. It is clear that the vehicle control data from all three performing laboratories were similar for all parameters (body and organ weights) evaluated. The intralaboratory assessment indicates that the SVCG and LABC were the most sensitive to TP stimulation, exhibiting statistically significant increases in all three testing laboratories at 0.4 through 1.6 mg/kg/day TP. VP and COWS weights were also statistically significantly increased at 0.4 through 1.6 mg/kg/day TP at Lab A and Lab B. Lab C did not detect significant increases at 0.4 mg/kg/day TP for either organ. The paired testes weights were statistically significantly decreased at 0.4 mg/kg/day only at Lab A. At 0.8 mg/kg/day, both Lab A and Lab C detected significant decreases (but not Lab B) in paired testes weight. All three laboratories detected significant decreases in paired testes weights at 1.0, 1.2, and 1.6 mg/kg/day TP. The least sensitive responding tissue were the paired epididymides, with the three laboratories detecting significant increases at 0.8-1.6 mg/kg/day TP but not at 0.4 mg/kg/day TP. The BW, liver, paired adrenal glands, and paired kidney weights were unaffected across all groups.

Table 4.3. Dose Response of Body and Tissue Weights to TP

Endpoint		TP (mg/kg-bw/day)					
(Unit)	**Laboratory**	**Vehicle**	**0.4 TP**	**0.8 TP**	**1.0 TP**	**1.2 TP**	**1.6 TP**
BW (g)	Lab A	123.8 ± 13.7	122.1 ± 17.9	126.4 ± 16.9	120.9 ± 22.9	129.0 ± 14.6	121.3 ± 10.0
	Lab B	124.6 ± 4.6	131.0 ± 5.4	133.4 ± 4.3	132.5 ± 5.9	135.2 ± 3.2	132.5 ± 4.7
	Lab C	130.0 ± 6.9	132.3 ± 8.6	136.7 ± 7.4	137.3 ± 4.9	136.3 ± 4.8	138.9 ± 3.8
VP (mg)	Lab A	63.4 ± 13.1	81.3 ± 15.8 *	99.1 ± 14.9 *	97.5 ± 20.0 *	121.0 ± 11.5 *	116.7 ± 10.0 *
	Lab B	75.6 ± 3.3	105.9 ± 2.2 *	111.2 ± 7.2 *	121.1 ± 15.9 *	124.4 ± 15.1 *	139.4 ± 16.8 *
	Lab C	84.2 ± 10.1	98.8 ± 3.1	131.9 ± 21.7 *	131.5 ± 14.9 *	140.0 ± 10.7 *	157.5 ± 17.0 *
SVCG (mg)	Lab A	71.2 ± 6.7	135.1 ± 17.5 *	194.18 ± 20.77 *	210.5 ± 53.1 *	236.7 ± 39.0 *	273.0 ± 45.6 *
	Lab B	53.9 ± 3.3	147.2 ± 15.2 *	194.13 ± 5.50 *	240.6 ± 18.5 *	263.2 ± 30.6 *	282.6 ± 17.4 *
	Lab C	38.5 ± 11.6	130.6 ± 25.3 *	222.08 ± 27.63 *	253.0 ± 32.9 *	273.4 ± 46.7 *	308.4 ± 26.6 *
LABC (mg)	Lab A	82.6 ± 8.6	119.42 ± 23.1 *	137.9 ± 22.8 *	142.7 ± 33.7 *	173.0 ± 25.4 *	168.5 ± 24.2 *
	Lab B	125.0 ± 4.4	234.32 ± 13.8 *	231.8 ± 8.3 *	250.3 ± 32.7 *	252.0 ± 31.9 *	297.9 ± 2.9 *
	Lab C	119.1 ± 27.6	215.88 ± 15.3 *	269.5 ± 43.3 *	277.7 ± 45.4 *	290.2 ± 30.6 *	296.7 ± 39.1 *
COW (mg)	Lab A	6.0 ± 0.7	14.9 ± 2.3 *	21.0 ± 2.5 *	21.3 ± 4.9 *	24.8 ± 4.6 *	25.6 ± 2.9 *
	Lab B	9.5 ± 0.3	22.6 ± 2.5 *	23.7 ± 2.0 *	26.1 ± 3.1 *	25.0 ± 3.8 *	27.7 ± 2.1 *
	Lab C	7.8 ± 1.0	18.8 ± 5.1	26.1 ± 3.7 *	28.4 ± 3.6 *	31.8 ± 4.9 *	33.5 ± 3.9 *
Testes (mg)	Lab A	1066.7 ± 129.3	758.9 ± 183.5 *	462.4 ± 67.6 *	416.0 ± 59.0 *	458.0 ± 111.5 *	523.3 ± 62.1 *
	Lab B	937.3 ± 4.7	928.3 ± 12.6	900.5 ± 0.4	621.9 ± 37.3 *	469.4 ± 45.6 *	454.1 ± 34.7 *
	Lab C	976.9 ± 123.8	760.8 ± 227.9	530.4 ± 77.0 *	509.3 ± 87.0 *	559.4 ± 88.0 *	629.3 ± 60.1 *
Epididymides (mg)	Lab A	103.5 ± 11.2	115.6 ± 21.5	124.5 ± 11.0 *	129.7 ± 18.7 *	139.2 ± 20.1 *	144.5 ± 14.6 *
	Lab B	114.3 ± 8.0	130.1 ± 7.9	135.5 ± 9.7 *	146.0 ± 12.3 *	148.0 ± 12.0 *	149.9 ± 8.4 *
	Lab C	126.2 ± 11.1	139.4 ± 5.3	167.8 ± 21.4 *	172.9 ± 15.8 *	168.5 ± 20.5 *	184.4 ± 18.2 *
Liver (g)	Lab A	6.1 ± 0.8	5.9 ± 0.8	6.3 ± 1.0	6.3 ± 1.1	6.4 ± 0.6	6.0 ± 0.7
	Lab B	5.8 ± 0.3	6.0 ± 0.1	6.2 ± 0.2	6.1 ± 0.5	6.3 ± 0.2	6.3 ± 0.3
	Lab C	6.2 ± 0.5	6.6 ± 0.6	6.6 ± 1.0	6.7 ± 0.4	6.6 ± 0.5	6.5 ± 0.4
Adrenals (mg)	Lab A	25.1 ± 2.4	22.7 ± 5.2	27.0 ± 3.2	23.5 ± 2.3	25.0 ± 5.5	23.1 ± 0.6
	Lab B	24.6 ± 1.7	26.1 ± 2.3	27.4 ± 1.1	23.9 ± 2.8	24.5 ± 2.8	24.3 ± 2.9
	Lab C	29.5 ± 3.7	31.9 ± 6.2	29.5 ± 4.4	26.0 ± 3.7	28.2 ± 2.8	27.6 ± 3.6
Kidneys (mg)	Lab A	1194.6 ± 118.2	1187.5 ± 166.0	1266.3 ± 176.5	1175.5 ± 197.1	1287.7 ± 140.0	1240.1 ± 177.1
	Lab B	1237.4 ± 38.4	1249.3 ± 30.7	1270.4 ± 102.7	1246.7 ± 32.6	1277.7 ± 52.8	1354.2 ± 75.6
	Lab C	1488.8 ± 127.1	1529.1 ± 105.0	1550.5 ± 149.8	1569.8 ± 63.8	1567.6 ± 117.7	1566.7 ± 90.2

* - Statistically significant by Dunnett's ($p < 0.05$)

The interlaboratory assessment indicates that Lab A detected the largest number of significant organ weight changes at 0.4 mg/kg/day TP (the lowest dose) and up for VP, SVCG, LABC, COW, and testes. Lab B detected significant organ weight changes at 0.4 mg/kg/day TP and up for VP, SVCG, LABC, and COW (but not the paired testes). Lab C detected the least number of significant organ weight changes at 0.4 mg/kg/day and up for SVCG and LABC. Lab C detected significant weight changes at 0.8 mg/kg/day for VP, SVCG, LABC, COW, paired testes (Lab B did not), and for paired epididymides.

There were no effects on BW at any TP dose evaluated. Significantly increased VP weights were detected by Lab A and Lab B but not by Lab C at 0.4 mg/kg/day TP and by all three laboratories at 0.8-1.6 mg/kg/day TP. Significantly increased weights of SVCG and LABC were detected by all three laboratories at 0.4-1.6 mg/kg/day TP. Significantly decreased paired testes weights were detected only by Lab A at 0.4 mg/kg/day TP and by all three laboratories at 0.8-1.6 mg/kg/day TP. Significantly increased paired epididymides weights were detected by all three laboratories at 0.8-1.6 mg/kg/day TP, with none of the laboratories reporting significant increases at 0.4 mg/kg/day TP. None of the laboratories reported significant changes in the weights of the liver, paired adrenal glands, or paired kidneys at any TP dose.

Table 4.4 presents the CVs for the body and organ weights in the Phase 1A TP dose-response study. For vehicle control group BWs and VP weights, Lab A exhibited the highest CVs; for SVCG and LABC, Lab C exhibited the highest CVs. The CVs were approximately equal for the control group weights of the COWS, paired testes, epididymides, liver, paired adrenal glands, and paired kidneys. For the TP-dosed groups, Lab A exhibited the highest CVs for BW and paired kidney weights across all doses and for most of the doses for the liver, epididymides, LABC, and VP.

Table 4.4. CVs for Body and Tissue Weights in TP Studies

Dose (mg/kg-bw/d)	Laboratory	BW	VP	SVCG	LABC	COWS	Testes	EPID	Liver	Adrenals	Kidneys
Vehicle	Lab A	11.03%	20.70%	9.44%	10.45%	12.44%	12.12%	10.79%	12.78%	9.74%	9.90%
	Lab B	3.71%	4.35%	6.06%	3.48%	2.78%	0.51%	7.00%	5.59%	6.71%	3.10%
	Lab C	5.33%	12.04%	30.06%	23.14%	12.76%	12.67%	8.79%	7.34%	12.52%	8.54%
0.4 TP	Lab A	14.62%	19.43%	12.92%	19.32%	15.37%	24.18%	18.62%	13.27%	22.84%	13.98%
	Lab B	4.10%	2.10%	10.32%	5.75%	11.28%	1.35%	6.06%	1.78%	9.01%	2.45%
	Lab C	6.51%	3.12%	19.38%	7.09%	27.08%	29.95%	3.78%	8.52%	19.39%	6.87%
0.8 TP	Lab A	13.40%	15.01%	10.70%	16.52%	11.74%	14.61%	8.86%	15.70%	11.76%	13.94%
	Lab B	3.22%	6.47%	2.83%	3.60%	8.28%	0.04%	7.14%	3.95%	4.12%	8.08%
	Lab C	5.38%	16.49%	12.44%	16.08%	14.26%	14.52%	12.74%	15.27%	15.06%	9.66%
1.0 TP	Lab A	18.92%	20.52%	25.23%	23.61%	23.24%	14.18%	14.41%	17.40%	9.92%	16.77%
	Lab B	4.42%	13.16%	7.70%	13.06%	11.82%	6.01%	8.44%	7.69%	11.63%	2.62%
	Lab C	3.58%	11.35%	12.99%	16.35%	12.63%	17.08%	9.17%	6.54%	14.39%	4.06%
1.2 TP	Lab A	11.31%	9.51%	16.49%	14.69%	18.59%	24.35%	14.45%	9.63%	22.14%	10.88%
	Lab B	2.38%	12.10%	11.62%	12.66%	15.02%	9.72%	8.10%	3.65%	11.59%	4.13%
	Lab C	3.55%	7.67%	17.09%	10.53%	15.43%	15.73%	12.20%	7.25%	9.92%	7.51%
1.6 TP	Lab A	13.98%	8.56%	16.71%	14.37%	11.33%	11.87%	10.09%	11.56%	2.66%	14.28%
	Lab B	3.52%	12.04%	6.14%	0.96%	7.58%	7.64%	5.59%	5.00%	11.91%	5.58%
	Lab C	2.75%	10.82%	8.61%	13.17%	11.58%	9.56%	9.87%	5.75%	13.02%	5.75%

Table 4.5 presents the statistical analyses of the target tissue weights presented in Table 4.3. All three testing laboratories detected statistically significant increases at all TP doses (beginning at 0.4 mg/kg/day) for SVCG, LABC, and COWS. For the VP, Lab A and Lab B detected statistically significant increased weights at all TP doses, but Lab C did not detect a statistically significant change (although it was elevated) at 0.4 mg/kg/day TP. For paired testes, Lab A and Lab C detected statistically significantly decreased weights at all TP doses. Lab B did not detect statistically significant decreases at 0.4 or 0.8 mg/kg/day TP. For paired epididymal weights, none of the three laboratories detected statistically significant increased weights at 0.4 mg/kg/day TP, but all three detected statistically significant increases at 0.8-1.6 mg/kg/day TP. There were no statistically or biologically significant differences among TP doses for the weights of the optional systemic organs, liver, paired adrenal glands, or paired kidneys in any of the three participating laboratories.

At the completion of Phase 1A, the dose of 1.0 mg/kg/day TP was selected as the reference stimulatory dose to be used in Phase 1B, to evaluate the appropriate inhibitory dose for FLU, a potent anti-androgen.

Table 4.5. TP Relative Dose Response of the Target Tissues[a]

	Lab A Mean (mg)	**SD**	**Relative Change[b]**	**Lab B Mean (mg)**	**SD**	**Relative Change**	**Lab C Mean (mg)**	**SD**	**Relative Change**
VP									
Vehicle	63.4	13.13	1.00	75.6	3.3	1.00	84.20	10.1	1.00
0.4 TP	81.3*	15.81*	1.28*	105.9*	2.2	1.40*	98.80	3.1	1.17
0.8 TP	99.1*	14.88*	1.56*	111.2*	7.2	1.47*	131.90*	21.7	1.57*
1.0 TP	97.5*	19.99*	1.54*	121.1*	15.9	1.60*	131.45*	14.9	1.56*
1.2 TP	121.0*	11.51*	1.91*	124.4*	15.1	1.64*	140.0*	10.7	1.66*
1.6 TP	116.7*	9.99*	1.84*	139.4*	16.8	1.84*	157.53*	17.0	1.87*
SVCG									
Vehicle	71.2	6.72	1.00	53.9	3.3	1.00	38.50	11.6	1.00
0.4 TP	135.1*	17.46*	1.90*	147.2*	15.2	2.73*	130.60*	25.3	3.39*
0.8 TP	194.2*	20.77*	2.73*	194.1*	5.5	3.60*	222.08*	27.63	5.77*
1.0 TP	210.5*	53.11*	2.96*	240.6*	18.5	4.47*	253.03*	32.9	6.57*
1.2 TP	236.7*	39.03*	3.32*	263.2*	30.6	4.89*	273.35*	46.7	7.10*
1.6 TP	273.0*	45.61*	3.83*	282.6*	17.4	5.25*	308.40*	26.6	8.01*
LABC									
Vehicle	82.6	8.63	1.00	125.0	4.4	1.00	119.08	27.6	1.00
0.4 TP	119.4*	23.07*	1.45*	234.3*	13.8	1.87*	215.88*	15.3	1.81*
0.8 TP	137.9*	22.79*	1.67*	231.8*	8.3	1.85*	269.48*	43.3	2.26*
1.0 TP	142.7*	33.69*	1.73*	250.3*	32.7	2.00*	277.72*	45.4	2.33*
1.2 TP	173.0*	25.42*	2.09*	252.0*	31.9	2.02*	290.20*	30.6	2.44*
1.6 TP	168.5*	24.22*	2.04*	297.9*	2.9	2.38*	296.65*	39.1	2.49*

Table 4.5 (continued)

	Lab A Mean (mg)	SD	Relative Change	Lab B Mean (mg)	SD	Relative Change	Lab C Mean (mg)	SD	Relative Change
COW									
Vehicle	6.0	0.74	1.00	9.5	0.3	1.00	7.85	1.0	1.00
0.4 TP	14.9*	2.28*	2.48*	22.6*	2.5	2.38*	18.77*	5.1	2.39*
0.8 TP	21.0*	2.46*	3.51*	23.7*	2.0	2.49*	26.12*	3.7	3.33*
1.0 TP	21.3*	4.95*	3.56*	26.1*	3.1	2.75*	28.35*	3.6	3.61*
1.2 TP	24.8*	4.61*	4.14*	25.0*	3.8	2.63*	31.82*	4.9	4.05*
1.6 TP	25.6*	2.90*	4.28*	27.7*	2.1	2.92*	33.52*	3.9	4.27*
Testes									
Vehicle	1066.7	129.26	1.00	937.3	4.7	1.00	976.93	123.8	1.00
0.4 TP	758.9*	183.48*	0.71*	928.3	12.6	0.99	760.78*	227.9	0.78*
0.8 TP	462.4*	67.57*	0.43*	900.5	0.4	0.96	530.38*	77.0	0.54*
1.0 TP	416.0*	59.00*	0.39*	621.9*	37.3	0.66*	509.28*	87.0	0.52*
1.2 TP	458.0*	111.53*	0.43*	469.4*	45.6	0.50*	559.43*	88.0	0.57*
1.6 TP	523.3*	62.13*	0.49*	454.1*	34.7	0.48*	629.33*	60.1	0.64*
Epididymides									
Vehicle	103.5	11.16	1.00	114.3	8.0	1.00	126.18	11.1	1.00
0.4 TP	115.6	21.52	1.12	130.1	7.9	1.14	139.37	5.3	1.10
0.8 TP	124.5*	11.03*	1.20*	135.5*	9.7	1.18*	167.83*	21.4	1.33*
1.0 TP	129.7*	18.69*	1.25*	146.0*	12.3	1.28*	172.92*	15.8	1.37*
1.2 TP	139.2*	20.12*	1.35*	148.0*	12.0	1.29*	168.47*	20.5	1.34*
1.6 TP	144.5*	14.57*	1.40*	149.9*	8.4	1.31*	184.42*	18.2	1.46*
Liver									
Vehicle	6.1	0.8	1.00	5.8	0.33	1.00	6.2	0.45	1.00
0.4 TP	5.9	0.8	0.97	6.0	0.11	1.03	6.6	0.56	1.06
0.8 TP	6.3	1.0	1.03	6.2	0.25	1.07	6.6	1.00	1.06
1.0 TP	6.3	1.1	1.03	6.1	0.47	1.05	6.7	0.44	1.08
1.2 TP	6.4	0.6	1.05	6.3	0.23	1.09	6.6	0.48	1.06
1.6 TP	6.0	0.7	0.98	6.3	0.31	1.09	6.5	0.38	1.05
Paired Adrenal Glands									
Vehicle	25.1	2.4	1.0	26.4	1.65	1.00	29.5	3.69	1.00
0.4 TP	22.7	5.2	0.9	26.1	2.35	0.99	31.9	6.18	1.08
0.8 TP	27.0	3.2	1.08	27.4	1.13	1.04	29.5	4.44	1.00
1.0 TP	23.5	2.3	0.94	23.9	2.78	0.90	26.0	3.73	0.88
1.2 TP	25.0	5.5	1.00	24.5	2.84	0.93	28.2	2.80	0.96
1.6 TP	23.1	0.6	0.92	24.3	2.90	0.92	27.6	3.59	0.94
Paired Kidneys									
Vehicle	1194.6	118.2	1.00	1237.4	38.35	1.0	1488.8	127.14	1.00
0.4 TP	1187.5	166.0	0.99	1249.3	30.66	1.01	1529.1	104.98	1.03
0.8 TP	1266.3	176.5	1.06	1270.4	102.68	1.03	1550.5	149.81	1.04
1.0 TP	1175.5	197.1	0.98	1246.7	32.64	1.01	1569.8	63.78	1.05
1.2 TP	1287.7	140.0	1.08	1277.7	52.82	1.03	1567.6	117.67	1.05
1.6 TP	1240.1	177.1	1.04	1354.2	75.56	1.09	1566.7	90.16	1.05

[a] Dunnett's multiple group comparison using an ANCOVA adjusted for BW; * = $p < 0.05$

[b] Relative change is done by normalizing the organ weight data to the vehicle control value, set 1.00, with increases above or decreases below the control value presented in the column, as values greater than or less than 1.00.

4.2 Phase 1B FLU Dose-Response Studies

In Phase 1B, the inhibition of the weight response of these tissues to a selected dose of TP (determined from Phase 1A) was assessed against a series of doses of FLU (CAS No. 13311-84-7; a potent reference androgen antagonist) as follows: 0, vehicle + TP (positive stimulatory control), and 0.1, 0.3, 1.0, 3.0, and 10.0 mg/kg/day FLU (approximately one-half log increments) by oral gavage, administered with concurrent SC injections of TP at 1.0 mg/kg/day once daily for ten consecutive days. The same three testing laboratories participated in both Phases 1A and 1B. Again, the lead laboratory did not perform any studies. The objectives for Phase 1B were to: (1) identify the dose response for FLU (the reference androgen antagonist) and (2) select the appropriate dose to use as the inhibitory dose for subsequent evaluations in Phases 3 and 4.

Table 4.6 presents an overview of the Phase 1B protocol. Note that the same endpoints are measured (mandatory and optional) as in Phase 1A, that the reference androgen, TP, is used as the stimulatory dose, administered by SC injection at 1.0 ml/kg/day in corn/peanut oil once daily for ten days, and that the reference androgen antagonist (anti-androgen) FLU is co-administered by oral gavage at 5.0 ml/kg/day in corn oil once daily for ten days. Also note that the GP is not dissected or weighed, and both the paired testes and paired epididymides (present in the intact weanling) are dissected and weighed.

Table 4.2 presents the laboratory parameters and conditions for the Phase 1B validation program as well.

Table 4.6. Anti-androgen Protocol Summary

	Factor	Protocol Specifications
Animals	Species	Rat
	Strain	No strain preference (one exclusion, not Fischer 344) [a]
	Age at initiation of treatment	20-22 days
	Weight at time of treatment	Not specified; groups should be within ± 20%
Animal Husbandry	Diet	Laboratory preference
	Bedding	Laboratory preference
	Caging [b]	Laboratory preference
Treatment Regimen	Animals per dose group	6
	Vehicle	Corn/peanut oil
	Volume of administration subcutaneous gavage	 1.0 ml/kg BW/day 5.0 ml/kg BW/day
	Dosing regimen (mg/kg-bw/day)	10 consecutive daily administrations
	TP stimulating dose	SC administration, 1 mg/kg-bw/d
	FLU	Oral gavage administration
	Sacrifice	24 hrs after last treatment
Measurements	Mandatory weights [c]	Total BW (to 0.1 g) VP (fresh) SVCG LABC COW Testes (paired) Epididymides (paired)
	Optional weights and measurements [c]	Liver weight (to 0.1 g) Adrenal gland (paired) weight Kidney (paired) weight

[a] The Fischer 344 (F344) rat is also specifically excluded in the EPA developmental neurotoxicity testing guidelines and excluded by implication in other testing guidelines because of its proclivity for and early onset of reproductive problems (functional and structural).

[b] For the immature animals, two or three animals per cage are recommended.

[c] Unless otherwise specified, the weight should be measured to 0.1 mg.

Table 4.7 presents the mean ± SD dose response from each of the three performing laboratories for BW and all of the tissues for TP alone at 1.0 mg/kg/day or for TP at 1.0 mg/kg/day plus FLU at 0.1, 0.3, 1.0, 3.0, or 10.0 mg/kg/day. It is clear that the TP control data from all three laboratories were similar for BW and all organs weighed, in that all laboratories exhibited increased weights of all the accessory sex organs, VP, SVCG, LABC, COW, and paired epididymides, and decreases in paired testes weights as compared to the vehicle control values. There were no statistical analyses performed for the weights in the TP control group versus the vehicle control group. The TP control SVCG weight was low for the Canadian laboratory relative to the TP control SVCG weight values for Lab A and Lab B.

The intralaboratory assessment indicates that the SVCG is the most sensitive organ, with detection of a significant difference at TP +0.1 mg/kg/day FLU for Lab C and at TP + 0.3 mg/kg/day FLU for Lab A. Lab B did not detect a significant decrease in SVCG weight until TP + 1.0 mg/kg/ FLU. The least sensitive organ weight was the testes weight. The testes weights were unaffected at 1.0 mg/kg/day TP plus 0.1 through 1.0 mg/kg/day FLU, and significantly increased (versus the 1.0 TP group) at 1.0 mg/kg/day TP plus 3.0 or 10.0 mg/kg/day FLU at Lab A and Lab B, but not at Lab C. All organs for all three laboratories, except for the testes, exhibited significant increases in weight at low FLU doses (+ TP) and significant decreases in weight at higher FLU doses (+ TP).

Interlaboratory comparisons indicate that no laboratory detected changes in BW in any group. Significant VP weight changes were detected at 1.0 and 3.0 mg/kg/day FLU (+ TP) by Lab A and Lab C and by all three laboratories at 10.0 mg/kg/day FLU (+ TP). Significant SVCG weight changes were detected by Lab C at 0.1 mg/kg/day FLU (+ TP) by Lab C and Lab A at 0.3 mg/kg/day FLU (+ TP) and by all three laboratories at 1.0, 3.0, and 10.0 mg/kg/day FLU (+ TP). Significant LABC weight changes were detected by all three laboratories at 1.0, 3.0, and 10.0 mg/kg/day FLU (+ TP) but not at 0.1 or 0.3 mg/kg/day FLU (+ TP). Significant COWS weight changes were detected by Lab A and Lab B (but not Lab C) at 1.0 mg/kg/day FLU (+ TP) and by all three laboratories at 3.0 and 10.0 mg/kg/day FLU (+ TP). Significant testes weight changes were detected by Lab A and Lab B (but not Lab C) at 3.0 and 10.0 mg/kg/day FLU (+ TP). It appears from Table 4.9 that the statistical comparison was not made for the values in the vehicle control group versus the values in the TP control group for any parameter evaluated. The mean values were clearly increased in the TP control group for VP, SVCG, LABC, COW, and paired epididymides and clearly decreased in the TP control group for the paired testes weights.

Table 4.8 presents the weights of the optional systemic organs in the vehicle control, TP control, and FLU doses (+ TP). There were no statistically or biologically significant differences among any of the groups for the optional systemic organ weights (liver, paired adrenal glands, or paired kidneys).

The CVs for all of the parameter assessments in Phase 1B are presented in Table 4.9. For the vehicle control group, Lab C exhibited the highest CVs for BW, testes, epididymides, and adrenals; Lab B had the highest CVs for VP, SVCG, COW, liver, and kidney; and Lab A had the highest CV for only LABC. For the TP control group, Lab C exhibited the highest CVs for BW, VP, LABC, liver, and adrenal glands; Lab B had the highest CVs for SVCG, COW, testes, and epididymides; Lab A had the highest CV for kidneys. For the FLU (+ TP) groups, the CVs for almost all of the weights were relatively low and relatively close. However, Lab C exhibited a CV of 103.97% for COWS at 10.0 FLU (+ TP).

Table 4.10 presents the statistical analyses of the target tissue weights presented in Table 4.7. All three participating laboratories detected clear increases in weights of the VP, SVCG, LABC, and COW, and clear decreases in the weight of the paired testes in the TP only positive control group, relative to the vehicle control group. The organ weights in the TP only control group were not compared statistically to the vehicle control group values, so the above assessment is based only on the observation of the values.

For the TP + FLU groups (relative to the TP alone group), VP weight was significantly reduced at 1, 3, and 10 mg/kg/day FLU (+ P) for Lab A and Lab C, and significantly reduced only at 10 mg/kg/day FLU (+ TP) for Lab B. SVCG weight was significantly increased at 0.3, 1, and 3 mg/kg/day FLU (+ TP) and significantly reduced at 10 mg/kg/day FLU (+ TP) for Lab A. For Lab B, SVCG weight was significantly

reduced at 1, 3, and 10 mg/kg/day FLU (+ TP). For Lab C, SVCG weight was significantly increased at 0.1, 0.3, 1, and 3 mg/kg/day and significantly reduced at 10 mg/kg/day FLU (+ TP). LABC weight was significantly increased at 1 and 3 mg/kg/day FLU (+ TP), and significantly decreased at 10 mg/kg/day FLU (+ P) for Lab A. For both Lab B and Lab C, LABC weights were significantly increased at 1, 3, and 10 mg/kg/day FLU (+ TP). For COW, the weights were significantly increased at 1 and 3 mg/kg/day FLU (+ TP) and significantly reduced at 10 mg/kg/day FLU (+ TP) for Lab A. Lab B detected significant increases in COW weight at 1, 3, and 10 mg/kg/day FLU (+ TP), and Lab C detected significant increases in COW weight at 3 and 10 mg/kg/day FLU (+ TP). For paired testes weights, both Lab A and Lab B reported significant reductions at 3 and 10 mg/kg/day FLU (+ TP). Lab C reported no significant differences in paired testes weights at any FLU + TP dose. For the weights of the paired epididymides, Lab A, Lab B, and Lab C all reported significant reductions at 1, 3, and 10 mg/kg/day FLU (+ TP).

For the androgen-dependent ASO, TP (1 mg/kg/day) and TP (1 mg/kg/day) plus low doses of FLU resulted in increased organ weights, with a downward, dose-response curve and then decreased organ weights at TP (fixed at 1 mg/kg/day) plus higher doses of FLU. For the paired testes weights, the fixed TP plus and low doses of FLU resulted in severely decreased organ weights. At the fixed TP plus higher FLU doses, the organ weights (still below TP alone) were higher. For the paired epididymal weights, TP alone and the fixed TP plus low doses of FLU (0.1 and 0.3 mg/kg/day) resulted in higher organ weights than the vehicle for Lab A; for Lab B, the organ weights were higher than the vehicle for TP alone and TP + FLU at 0.1, 0.3, 1, and 3 mg/kg/day, and slightly (but statistically significantly) reduced at 10 mg/kg/day FLU (+ TP). For Lab C, the organ weights were significantly lower at 1, 3, and 10 mg/kg/day FLU (+ TP) and unaffected at TP alone (versus the vehicle control) and TP + FLU at 0.1 or 0.3 mg/kg/day.

For the optional systemic organs (liver, paired adrenal glands, and paired kidneys), there were no differences among groups (vehicle control, TP only control, and TP + FLU doses) for weights of these organs in any of the three participating laboratories.

Table 4.7. Dose Response of Body and Reproductive Tissue Weights to TP and TP + FLU

Endpoint	Laboratory	Vehicle	TP (1 mg/kg-bw/d) + Dose Series of FLU (mg/kg-bw/d)					
			1 TP	1 TP + 0.1 FLU	1 TP + 0.3 FLU	1 TP + 1 FLU	1 TP + 3 FLU	1 TP + 10 FLU
Body Wt (mg)	Lab A	116.7 ± 5.8[a]	121.0 ± 8.4	123.7 ± 9.4	112.0 ± 7.7	124.3 ± 4.5	115.2 ± 5.6	119.7 ± 6.4
	Lab B	131.4 ± 7.8	148.8 ± 4.9	144.3 ± 5.2	138.6 ± 7.7	141.6 ± 10.3	139.9 ± 11.0	138.6 ± 9.0
	Lab C	118.3 ± 9.2	124.2 ± 8.7	121.6 ± 10.9	125.4 ± 10.9	121.4 ± 6.8	119.2 ± 6.6	116.0 ± 8.1
VP (mg)	Lab A	62.0 ± 8.1	99.2 ± 15.4	88.2 ± 8.8	84.5 ± 10.9	73.8 ± 11.7 *	57.2 ± 7.0 *	48.5 ± 10.1 *
	Lab B	73.8 ± 13.8	122.1 ± 17.1	137.2 ± 8.4	114.6 ± 10.0	109.9 ± 4.9	102.1 ± 11.5	73.1 ± 14.6 *
	Lab C	66.6 ± 9.9	120.4 ± 19.3	102.2 ± 18.9	106.4 ± 23.1	82.5 ± 15.1 *	75.0 ± 12.6 *	50.9 ± 12.1 *
SVCG (mg)	Lab A	72.1 ± 5.8	245.9 ± 26.5	232.72 ± 23.0	190.9 ± 38.1 *	146.9 ± 9.2 *	101.2 ± 8.2 *	70.0 ± 9.6 *
	Lab B	50.9 ± 17.6	224.7 ± 34.6	234.57 ± 52.5	224.2 ± 39.3	154.9 ± 22.8 *	125.8 ± 17.5 *	77.0 ± 12.7 *
	Lab C	34.4 ± 7.3	212.6 ± 11.6	174.63 ± 29.7 *	163.6 ± 9.7 *	115.7 ± 19.4 *	75.0 ± 5.0 *	32.1 ± 5.6 *
LABC (mg)	Lab A	82.6 ± 17.0	159.50 ± 17.5	149.5 ± 16.7	145.7 ± 17.7	129.8 ± 8.7 *	106.2 ± 16.5 *	78.6 ± 11.4 *
	Lab B	125.7 ± 12.9	279.58 ± 13.8	277.7 ± 16.2	256.5 ± 7.6	228.7 ± 12.2 *	219.5 ± 11.0 *	185.9 ± 25.9 *
	Lab C	106.0 ± 10.9	250.85 ± 29.2	225.0 ± 38.9	218.6 ± 41.7	173.5 ± 14.7 *	154.5 ± 10.3 *	122.3 ± 18.3 *
COW (mg)	Lab A	6.4 ± 1.7	19.5 ± 1.8	18.7 ± 2.8	14.8 ± 3.0	12.9 ± 2.1 *	7.0 ± 1.2 *	4.5 ± 1.3 *
	Lab B	7.5 ± 4.2	21.7 ± 3.8	25.8 ± 2.8	20.6 ± 2.7	16.7 ± 2.4 *	14.6 ± 1.1 *	9.8 ± 2.7 *
	Lab C	8.4 ± 2.3	23.6 ± 3.3	21.1 ± 2.8	21.9 ± 2.8	17.5 ± 2.9	11.8 ± 1.5 *	15.0 ± 15.6 *
Testes (mg)	Lab A	1087.2 ± 109.0	409.2 ± 68.0	437.3 ± 146.5	397.0 ± 28.0	537.6 ± 202.8	775.1 ± 133.6 *	960.5 ± 89.4 *
	Lab B	1105.5 ± 71.7	593.8 ± 111.8	496.1 ± 54.4	554.9 ± 129.7	667.1 ± 105.6	956.4 ± 163.2 *	1015.05 ± 80.6 *

	Lab C	937.1 ± 139.3	509.7 ± 52.6	443.4 ± 29.8	459.0 ± 61.8	585.8 ± 181.4	565.9 ± 172.9	822.5 ± 132.3
EPID (mg)	Lab A	124.8 ± 17.8	159.5 ± 7.7	143.3 ± 13.0	132.7 ± 14.7	121.3 ± 20.9 *	104.2 ± 11.0 *	93.2 ± 16.4 *
	Lab B	112.8 ± 4.4	154.1 ± 12.9	151.1 ± 5.6	137.7 ± 13.9	121.2 ± 8.0 *	118.1 ± 8.3 *	102.4 ± 12.8 *
	Lab C	156.5 ± 23.1	167.1 ± 9.9	149.8 ± 24.7	152.8 ± 20.5	134.4 ± 18.5 *	129.1 ± 12.0 *	112.1 ± 20.6 *

[a] Data presented as mean ± SD

* Statistically significant by Dunnett's Test; * = $p < 0.05$

Table 4.8. Dose Response of Optional Tissue Weights to TP and TP + FLU

Endpoint	Laboratory	Vehicle	TP (1 mg/kg-bw/d) + Dose Series of FLU (mg/kg-bw/d)					
			1 TP	1 TP + 0.1 FLU	1 TP + 0.3 FLU	1 TP + 1 FLU	1 TP + 3 FLU	1 TP + 10 FLU
Liver (g)	Lab A	5.6 ± 0.4	6.0 ± 0.6	6.2 ± 0.5	5.6 ± 0.6	6.1 ± 0.4	5.9 ± 0.4	6.0 ± 0.6
	Lab B	5.7 ± 0.6	6.6 ± 0.5	6.4 ± 0.7	5.9 ± 0.6	6.1 ± 0.9	5.9 ± 0.5	5.8 ± 0.5
	Lab C	6.0 ± 0.5	6.4 ± 0.8	6.4 ± 0.9	6.4 ± 0.6	6.0 ± 0.3	6.0 ± 0.4	6.3 ± 0.7
Adrenals (mg)	Lab A	24.6 ± 2.4	30.9 ± 2.1	24.8 ± 5.7	21.9 ± 2.8	27.1 ± 2.3	24.2 ± 5.2	26.6 ± 4.1
	Lab B	19.5 ± 1.3	22.4 ± 1.9	22.3 ± 1.3	21.5 ± 1.9	21.7 ± 2.3	23.9 ± 0.6	23.6 ± 3.2
	Lab C	33.8 ± 6.5	27.2 ± 5.2	33.6 ± 6.8	28.4 ± 3.7	32.7 ± 6.9	30.9 ± 6.7	25.7 ± 2.0
Kidneys (mg)	Lab A	1162.8 ± 50.1	1232.4 ± 175.4	1166.0 ± 70.6	1140.7 ± 79.0	1217.1 ± 77.6	1177.4 ± 107.2	1148.8 ± 100.3
	Lab B	1270.1 ± 121.3	1424.4 ± 132.4	1346.7 ± 122.5	1335.1 ± 90.4	1340.0 ± 120.5	1402.0 ± 149.3	1328.5 ± 163.7
	Lab C	1286.9 ± 105.2	1295.8 ± 121.5	1333.5 ± 158.7	1360.2 ± 215.2	1318.0 ± 124.4	1278.2 ± 92.6	1230.9 ± 109.5

Table 4.9. CVs for Body and Tissue Weights in TP + FLU Studies

Doses (mg/kg-bw/d)		BW	VP	SVCG	LABC	COWS	Testes	EPID	Liver	Adrenals	Kidneys
Vehicle	Lab A	4.99%	13.14%	8.05%	20.52%	26.43%	10.02%	14.26%	6.88%	9.88%	4.30%
	Lab B	5.97%	18.70%	34.69%	10.25%	56.06%	6.48%	3.87%	11.22%	6.71%	9.55%
	Lab C	7.77%	14.80%	21.22%	10.31%	28.02%	14.87%	14.77%	9.08%	19.33%	8.18%
1 TP	Lab A	6.95%	15.56%	10.78%	10.96%	9.44%	16.63%	4.81%	10.11%	6.95%	14.23%
	Lab B	3.27%	13.97%	15.41%	4.92%	17.49%	18.83%	8.39%	7.64%	8.46%	9.29%
	Lab C	6.98%	16.05%	5.46%	11.66%	13.96%	10.31%	5.90%	12.73%	19.01%	9.37%
1 TP + 0.1 FLU	Lab A	7.60%	9.98%	9.89%	11.17%	14.69%	33.50%	9.05%	8.77%	22.85%	6.05%
	Lab B	3.59%	6.16%	22.37%	5.83%	10.98%	10.96%	3.68%	11.10%	5.76%	9.09%
	Lab C	9.00%	18.51%	16.98%	17.30%	13.03%	6.71%	16.47%	13.65%	20.09%	11.90%
1 TP + 0.3 FLU	Lab A	6.87%	12.96%	19.95%	12.16%	20.29%	7.05%	11.08%	11.04%	12.83%	6.93%
	Lab B	5.54%	8.76%	17.52%	2.97%	13.10%	23.37%	10.10%	10.76%	8.88%	6.77%
	Lab C	8.72%	21.67%	5.93%	19.08%	12.70%	13.48%	13.43%	9.71%	13.00%	15.82%
1 TP + 1 FLU	Lab A	3.62%	15.85%	6.26%	6.72%	16.20%	37.72%	17.21%	6.31%	8.49%	6.37%
	Lab B	7.25%	4.43%	14.70%	5.35%	14.41%	15.83%	6.63%	14.38%	10.68%	8.99%
	Lab C	5.56%	18.30%	16.76%	8.47%	16.45%	30.96%	13.78%	4.95%	21.15%	9.44%
1 TP + 3 FLU	Lab A	4.89%	12.23%	8.11%	15.53%	17.50%	17.24%	10.57%	7.55%	21.69%	9.10%
	Lab B	7.89%	11.24%	13.91%	5.01%	7.61%	17.06%	7.07%	8.26%	2.45%	10.65%
	Lab C	5.56%	16.82%	6.69%	6.64%	12.75%	30.55%	9.27%	5.92%	21.77%	7.24%
1 TP + 10 FLU	Lab A	5.30%	20.83%	13.70%	14.44%	28.86%	9.30%	17.60%	10.01%	15.55%	8.73%
	Lab B	6.47%	19.97%	16.50%	13.91%	27.00%	7.94%	12.48%	9.04%	13.58%	12.32%
	Lab C	6.99%	23.83%	17.46%	14.95%	103.97 %	16.09%	18.37%	10.61%	7.89%	8.90%

Table 4.10. Relative Dose Response of Body and Organ Weights to TP and TP + FLU[a]

Organ	Lab A Mean ± SD	Lab A Relative Change[b]	Lab B Mean ± SD	Lab B Relative Change[b]	Lab C Mean ± SD	Lab C Relative Change[b]
VP (mg)						
Vehicle	62.0 ± 8.1	1.00	73.8 ± 13.8	1.00	66.6 ± 9.9	1.00
1 TP[c]	99.2 ± 15.4	1.60	122.1 ± 17.1	1.65	120.4 ± 19.3	1.81
1 TP + 0.1 FLU[d]	88.2 ± 8.8	1.42	137.2 ± 8.4	1.86	102.2 ± 18.9	1.53
1 TP + 0.3 FLU	84.5 ± 10.9	1.36	114.6 ± 10.0	1.55	106.4 ± 23.1	1.60
1 TP + 1 FLU	73.8 ± 11.7*	1.19	109.9 ± 4.9	1.49	82.5 ± 15.1*	1.24
1 TP + 3 FLU	57.2 ± 7.0*	0.92	102.1 ± 11.5	1.38	75.0 ± 12.6*	1.13
1 TP + 10 FLU	48.5 ± 10.1*	0.78	73.1 ± 14.6*	0.99	50.9 ± 12.1*	0.76
SVCG (mg)						
Vehicle	72.1 ± 5.8	1.00	50.9 ± 17.6	1.00	34.4 ± 7.3	1.00
1 TP	245.9 ± 26.5	3.41	224.7 ± 34.6	4.41	212.6 ± 11.6	6.18
1 TP + 0.1 FLU	232.72 ± 23.0	3.23	234.57 ± 52.5	4.61	174.63 ± 29.7*	5.08
1 TP + 0.3 FLU	190.9 ± 38.1*	2.65	224.2 ± 39.3	4.40	163.6 ± 9.7*	4.76
1 TP + 1 FLU	146.9 ± 9.2*	2.04	154.9 ± 22.8*	3.04	115.7 ± 19.4*	3.36
1 TP + 3 FLU	101.2 ± 8.2*	1.40	125.8 ± 17.5*	2.47	75.0 ± 5.0*	2.18
1 TP + 10 FLU	70.0 ± 9.6*	0.97	77.0 ± 12.7*	1.51	32.1 ± 5.6*	0.93
LABC (mg)						
Vehicle	82.6 ± 17.0	1.00	125.7 ± 12.9	1.00	106.0 ± 10.9	1.00
1 TP	159.50 ± 17.5	1.93	279.58 ± 13.8	2.22	250.85 ± 29.2	2.37
1 TP + 0.1 FLU	149.5 ± 16.7	1.81	277.7 ± 16.1	2.21	225.0 ± 38.9	2.12
1 TP + 0.3 FLU	145.7 ± 17.7	1.76	256.5 ± 7.6	2.04	218.6 ± 41.7	2.06
1 TP + 1 FLU	129.8 ± 8.7*	1.57	228.7 ± 12.2*	1.82	173.5 ± 14.7*	1.64
1 TP + 3 FLU	106.2 ± 16.5*	1.29	219.5 ± 11.0*	1.75	154.5 ± 10.3*	1.46
1 TP + 10 FLU	78.6 ± 11.4*	0.95	185.9 ± 25.9*	1.48	122.3 ± 18.3*	1.15
COW (mg)						
Vehicle	6.4 ± 1.7	1.00	7.5 ± 4.2	1.00	8.4 ± 2.3	1.00
1 TP	19.5 ± 1.8	3.05	21.7 ± 3.8	2.89	23.6 ± 3.3	2.81
1 TP + 0.1 FLU	18.7 ± 2.8	2.92	25.8 ± 2.8	3.44	21.1 ± 2.8	2.51
1 TP + 0.3 FLU	14.8 ± 3.0	2.31	20.6 ± 2.7	2.75	21.9 ± 2.8	2.61
1 TP + 1 FLU	12.9 ± 2.1*	2.02	16.7 ± 2.4*	2.23	17.5 ± 2.9	2.08
1 TP + 3 FLU	7.0 ± 1.2*	1.09	14.6 ± 1.1*	1.95	11.8 ± 1.5*	1.40
1 TP + 10 FLU	4.5 ± 1.3*	0.70	9.8 ± 2.7*	1.31	15.0 ± 15.6*	1.79
Paired Testes (mg)						
Vehicle	1087.2 ± 109.0	1.0	1105.5 ± 71.7	1.0	937.1 ± 139.3	1.0

1 TP	409.2 ± 68.0	0.38	593.8 ± 111.8	0.54	509.7 ± 52.6	0.54
1 TP + 0.1 FLU	437.3 ± 146.5	0.40	496.1 ± 54.4	0.44	443.4 ± 29.8	0.47
1 TP + 0.3 FLU	397.0 ± 28.0	0.36	554.9 ± 129.7	0.50	459.0 ± 61.8	0.49
1 TP + 1 FLU	537.6 ± 202.8	0.49	667.1 ± 105.6	0.60	585.8 ± 181.4	0.62
1 TP + 3 FLU	775.1 ± 133.6*	0.71	956.4 ± 163.2*	0.86	565.9 ± 172.9	0.70
1 TP + 10 FLU	960.5 ± 89.4*	0.88	1015.05 ± 80.6*	0.92	822.5 ± 132.3	0.88

Table 4.10 (continued)

Organ	Lab A Mean ± SD	Relative Change[b]	Lab B Mean ± SD	Relative Change[b]	Lab C Mean ± SD	Relative Change[b]
Paired Epididymides (mg)						
Vehicle	124.8 ± 17.8	1.00	112.8 ± 4.4	1.00	156.5 ± 23.1	1.00
1 TP	159.5 ± 7.7	1.28	154.1 ± 12.9	1.37	167.1 ± 9.9	1.07
1 TP + 0.1 FLU	143.3 ± 13.0	1.15	151.1 ± 5.6	1.34	149.8 ± 24.7	0.96
1 TP + 0.3 FLU	132.7 ± 14.7	1.06	137.7 ± 13.9	1.22	152.8 ± 20.5	0.98
1 TP + 1 FLU	121.3 ± 20.9*	0.97	121.2 ± 8.0*	1.07	134.4 ± 18.5*	0.86
1 TP + 3 FLU	104.2 ± 11.0*	0.84	118.1 ± 8.3*	1.05	129.1 ± 12.0*	0.82
1 TP + 10 FLU	93.2 ± 16.4*	0.75	102.4 ± 12.8*	0.91	112.1 ± 20.6*	0.72
Liver (g)						
Vehicle	5.6 ± 0.4	1.0	5.7 ± 0.6	1.0	6.0 ± 0.5	1.0
1 TP	6.0 ± 0.6	1.07	6.6 ± 0.5	1.16	6.4 ± 0.8	1.07
1 TP + 0.1 FLU	6.2 ± 0.5	1.11	6.4 ± 0.7	1.12	6.4 ± 0.9	1.07
1 TP + 0.3 FLU	5.6 ± 0.6	1.00	5.9 ± 0.6	1.04	6.4 ± 0.6	1.07
1 TP + 1 FLU	6.1 ± 0.4	1.09	6.1 ± 0.9	1.07	6.0 ± 0.3	1.00
1 TP + 3 FLU	5.9 ± 0.4	1.05	5.9 ± 0.5	1.04	6.0 ± 0.4	1.00
1 TP + 10 FLU	6.0 ± 0.6	1.07	5.8 ± 0.5	1.02	6.3 ± 0.7	1.05
Paired Adrenal Glands (mg)						
Vehicle	24.6 ± 2.4	1.0	19.5 ± 1.3	1.0	33.8 ± 6.5	1.0
1 TP	30.9 ± 2.1	1.26	22.4 ± 1.9	1.15	27.2 ± 5.2	0.80
1 TP + 0.1 FLU	24.8 ± 5.7	1.01	22.3 ± 1.3	1.14	33.6 ± 6.8	0.99
1 TP + 0.3 FLU	21.9 ± 2.8	0.89	21.5 ± 1.9	1.10	28.4 ± 3.7	0.84
1 TP + 1 FLU	27.1 ± 2.3	1.10	21.7 ± 2.3	1.11	32.7 ± 6.9	0.97
1 TP + 3 FLU	24.2 ± 5.2	0.98	23.9 ± 0.6	1.23	30.9 ± 6.7	0.91
1 TP + 10 FLU	26.6 ± 4.1	1.0	23.6 ± 3.2	1.21	5.7 ± 2.0	0.76
Paired Kidneys (mg)						
Vehicle	1162.8 ± 50.1	1.0	1270.1 ± 121.3	1.0	1286.9 ± 105.2	1.0
1 TP	1232.4 ± 175.4	1.06	1424.4 ± 132.4	1.12	1295.8 ± 121.5	1.01
1 TP + 0.1 FLU	1166.0 ± 70.6	1.00	1346.7 ± 122.5	1.06	1333.5 ± 158.7	1.04
1 TP + 0.3 FLU	1140.7 ± 79.0	0.98	1335.1 ± 90.4	1.05	1360.2 ± 215.2	1.06
1 TP + 1 FLU	1217.1 ± 77.6	1.05	1340.0 ± 120.5	1.06	1218.0 ± 124.4	1.02
1 TP + 3 FLU	1177.4 ± 107.2	1.01	1402.0 ± 149.3	1.10	1278.2 ± 92.6	0.99
1 TP + 10 FLU	1148.8 ± 100.3	0.99	1278.2 ± 92.6	1.01	1230.9 ± 109.5	0.96

[a] Statistical analysis by Dunnett's multiple group comparison using an ANCOVA adjusted for BW; * = $p \leq 0.05$

[b] Relative change is done by normalizing the organ weight data to the vehicle control value, set to 1.0, with increases above or decreases below the control value presented in the column as values greater than or less than 1.00

[c] TP – 1.0 mg/kg BW/day

[d] FLU – 0.1, 0.3, 1.0, 3.0, or 10.0 mg/kg BW/day

4.3 Interlaboratory Comparisons

A model was developed by Dr. Breda Munoz (Senior Research Statistician at RTI International) to determine if there is a statistically significant laboratory effect (i.e., if the laboratories differ significantly in their ability to detect effects on specific organ weights). Since the participating laboratories apparently used both "log10(BW)" (logarithm base 10 of body weight) and "BW" (body weight), the model was constructed to use both terms. The full model, which evaluated the interaction of all fixed effects is as follows:

> Log10 (organ weight) = log10 (BW) + (DoseGroup) + log10 (BW)*(DoseGroup) + Lab + Lab*log10 (BW) + Lab*(DoseGroup) + Lab*log10 (BW)*(DoseGroup), where "*" indicates the interaction term

The model was run for each organ of interest, with "log10 (BW)" as shown and with "BW" as well. If the three-way interaction term "Lab*log10 (BW)*(DoseGroup)" is statistically significant ($p<0.05$), then changes in log10 (organ weight) across different dose groups and for given body weights are different across labs (suggesting lab's abilities to detect changes vary across body weight and dose group combinations). If the three-way interaction term "log10 (BW)*DoseGroup*labs" was not significant ($p>0.05$), then a reduced model (without this three-way interaction term) was run to fit the data. If the "log10 (BW)*lab" probability was not significant ($p>0.05$), then the model was rerun to fit the data without this term as well. If the two-way interaction term "DoseGroup*Lab" was significant ($p<0.05$), then the effect of dose groups on log10 (organ weight) was significantly different across labs (i.e., a lab effect is present).

If the two-way interaction term "DoseGroup*Lab" probability was nonsignificant ($p>0.05$), then a reduced model (without this two-way interaction term) was run to fit the data. If the two-way interaction term "Lab*log10 (BW)" was significant ($p<0.05$), then the effect of "log10(BW)" on "log10 (organ weight)" was significantly different across labs (i.e., a lab effect is present). If the two-way interaction term "Lab*log10 (BW)" was nonsignificant ($p>0.05$), then a reduced model (without this two-way interaction term) fit the data. If the term "lab" was significant ($p<0.05$), then lab conditions have a statistically significant different effect in detecting differences in "log10 (organ weight)". Finally, if the term "lab" was nonsignificant ($p>0.05$), then there is no statistical difference in lab's ability to detect differences in "log10 (organ weight)".

The "R-Squared" term generated from the model analysis is equal to the percentage of the variability around the "log10 (organ weight)" that is explainable by the terms in the model, including laboratory. Therefore, we can deduce whether at least one laboratory is significantly different from the rest, and that it impacts on the variability around the organ weight. However, this model cannot identify which lab(s) is different. Identification of the laboratory or laboratories requires another series of comparisons ("contrast tests") of laboratory 1 versus 2, 1 versus 3, etc.

The full model evaluates all of the fixed effects and is fitted to the data which include the lab effect as well as effects of body weight and dose group. Fixed effects are "DoseGroup" and "lab". "Log10 (BW)" is a continuous covariate.

A more detailed explanation and the results of the analyses are presented in Appendix D (Stage I Analyses). The results of the model for interlaboratory comparisons of Stage 1 TP and FLU data are presented below.

4.3.1 Testosterone Proprionate

TP: Paired Adrenal Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.842, not significant (NS)

B. Model minus triple interaction term

Double interaction "log10 (BW) * lab" term : p=0.032, is significant, so there is a lab effect, and 59.1% (R-Squared = 0.591) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.828 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.340 NS

C. Model minus triple and double interaction terms: p=0.094 NS

D. Model minus triple, double and "BW * lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 50.0% (R-Squared = 0.500) of the variability is explained by the terms in the final model.

TP: Paired Cowper's Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.201 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.136 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab": p=0.000, is significant, so there is a lab effect, and 94.4% (R-Squared = 0.944) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.173 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.005, is significant, so there is a lab effect, and 94.6% (R-Squared = 0.946) of the variability is explained by the terms in the final model.

TP: Paired Epididymides

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.762 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.609 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.594 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 80.8% (R-Squared = 0.808) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.766 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.596 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.609 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 80.3% (R-Squared = 0.803) of the variability is explained by the terms in the final model.

TP: Paired Kidneys

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.699 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.045, is significant, so there is a lab effect, and 91.2% (R-Squared = 0.912) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.645 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.3411 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.007, is significant, so there is a lab effect, and 89.4% (R-Squared = 0.894) of the variability is explained by the terms in the final model.

TP: LABC

I. Using "log10 (BW)"

A. Full model

Triple interaction term: p=0.276 NS

B. Model minus triple interaction "log10 (BW) * DoseGroup * labs" term

Double interaction "log10 (BW)*lab" term: p=0.033, is significant, so there is a lab effect, and 94.6% (R-Squared = 0.946) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.284 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.05 (NS – larger than 0.05)

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.018, is significant, so there is a lab effect, and 93.3% (R-Squared = 0.933) of the variability is explained by the terms in the final model.

TP: Liver

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.393 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.388 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.884 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.004, is significant, so there is a lab effect, and 61.9% (R-Squared = 0.619) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.436 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.815 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.703 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.007, is significant, so there is a lab effect, and 61.8% (R-Squared = 0.618) of the variability is explained by the terms in the final model.

TP: Paired Seminal Vesicles

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.183 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.012, is significant, so there is a lab effect, and 96.8% (R-Squared = 0.968) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.195 NS

B. Model minus triple interaction term

Double interaction term "DoseGroup * Lab": p=0.000, is significant, so there is a lab effect, and 96.8% (R-Squared = 0.968) of the variability is explained by the terms in the final model.

TP: Paired Testes

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.280 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.348 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.0.0000, is significant, so there is a lab effect, and 85.9% (R-Squared = 0.859) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.274 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.000, is significant, so there is a lab effect, and 86.0% (R-Squared = 0.860) of the variability is explained by the terms in the final model.

TP: Ventral Prostate

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.606 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.788 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.104 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 82.8% (R-Squared = 0.828) of the variability is explained by the terms in the final model.

II. Using ("BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.516 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.182 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.394 NS

D. Model minus triple, double and BW*Lab terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 82.7% (R-Squared = 0.827) of the variability is explained by the terms in the final model.

TP: Conclusions

For all of the organ weights in animals exposed to TP, there were significant laboratory effects, and the majority (predominantly the vast majority) of the variability around the organ weights is explained by the terms in the model, including interactions across dose groups, among body weights, and especially laboratories. In addition, there is no difference between the analyses using "log10(BW)" or "(BW)" in the model for any of the organ weights.

4.3.2 Flutamide

FLU: Paired Adrenal Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.338 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * lab": p=0.132 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" term: p=0.007, is significant, so there is a lab effect, and 60.8% (R-Squared = 0.608) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.375 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.003 is significant, so there is a lab effect, and 67.8% (R-Squared = 0.678) of the variability is explained by the terms in the final model.

FLU: Paired Cowper's Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.998 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.967 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.008, is significant, so there is a lab effect, and 83.8% (R-Squared = 0.838) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.999 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.009, is significant, so there is a lab effect, and 83.9% (R-Squared = 0.839) of the variability is explained by the terms in the final model.

FLU: Paired Epididymides

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.044, is significant, suggesting that log10(BW) and dose group have a different effect on log10 (epididymides) in the different laboratories. Since the high order interactions are significant, we cannot proceed to interpret the low order interactions or fixed effects. The final model explains 82.4% (R-Squared = 0.824) of the variability around the paired epididymal weights.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.036, is significant, suggesting that the combination of BW and dose group have a different effect on log10 (epididymides) across the different laboratories. Again, the final model explains 82.4% (R-Squared = 0.824) of the variability around the paired epididymal weight.

FLU: Paired Kidneys

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.587 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.051 NS

C. Model minus triple and double interaction terms

"DoseGroup * Lab" term: p=0.458 NS

D. Model minus triple, double, and "DoseGroup * Lab" terms: p=0.000, is significant, so there is a significant lab effect. The final model explains 71.1% (R-Squared = 0.711) of the variability around paired kidney weight.

II. Using BW

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.645 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.348 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.334 NS

D. Model minus triple, double, and "(BW) * Lab" terms: p=0.000, is significant so there is a significant lab effect, and 72.2% (R-Squared = 0.702) of the variability is explained by the terms in the final model.

FLU: LABC

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.646 NS

B. Model minus triple interaction term

Both double interaction terms, "log10 (BW)*Lab" and "DoseGroup*Lab": p=0.000 and p=0.049, respectively, are both significant, so there is a lab effect for both "(BW)" and "DoseGroup". The final model explains 71.1% (R-Squared = 0.711) of the variability around LABC weight.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.645 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.0348 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.334 NS

D. Model minus triple, double, and "(BW) * Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 72.2% (R-Squared = 0.722) of the variability around LABC weight is explained by the terms in the final model.

FLU: Liver

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.521 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.010, is significant so there is a lab effect, and 80.8% (R-Squared = 0.808) of the variability around liver weight is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.457 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.120 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.512 NS

D. Model minus triple, double and BW*Lab terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 76.5% (R-Squared = 0.765) of the variability is explained by the terms in the final model.

FLU: Paired Seminal Vesicles

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.255 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.309 NS

C. Model minus triple and double interaction terms

"DoseGroup * Lab" term: p=0,000, is significant, so there is a lab effect, and 95.1% (R-Squared = 0.951) of the variability is explained by the terms in the final model.

II. Using BW

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.161 NS

B. Model minus triple interaction term

Double interaction term "DoseGroup * Lab": p=0.000, is significant, so there is a lab effect, and 95.2% (R-Squared = 0.952) of the variability is explained by the terms in the final model.

FU: Paired Testes

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.631 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.613 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" term: p=0.036, is significant, so there is a lab effect, and 82.7% (R-Squared = 0.827) of the variability is explained by the terms in the final model.

II. Using BW

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.615 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.038, is significant, so there is a lab effect, and 82.8% (R-Squared = 0.828) of the variability is explained by the terms in the final model.

FLU: Ventral Prostate

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.496 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.157 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" term: = p=0.250 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 83.7% (R-Squared = 0.837) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.540 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*Lab": term: p=0.251 NS

C. Model minus triple and double interaction terms

"(BW)*Lab" term: p=0.079 NS

D. Model minus triple, double and BW*Lab terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 81.0% (R-Squared = 0.810) of the variability is explained by the terms in the final model.

FLU: Conclusions

For all of the organ weights in animals exposed to FLU, there were significant laboratory effects, and the majority (predominantly the vast majority) of the variability around the organ weights is explained by the terms in the model, including interactions across dose groups, among body weights, and especially laboratories. In addition, there is no difference between the results of the analyses using "log10(BW)" or "(BW)" in the model for any of the organ weights.

4.4 Overall Conclusions

From Phase 1A, the stimulatory dose for TP (SC injection in corn/peanut oil) was selected as 1.0 mg/kg BW/day. From Phase 1B, the inhibitory dose of FLU (oral gavage in corn/peanut oil) was selected as 3.0 mg/kg BW/day. TP (the reference androgen) exhibited a monotonic dose-response curve in all three laboratories for increasing androgen-dependent tissue weights with increasing TP dose from 0.4 to 1.6 mg/kg BW/day in Phase 1A. TP exhibited a monotonic dose-response curve in all three laboratories for decreasing paired testes weights, since increasing exogenous androgen reduces endogenous T production and causes reduced testes weights.

FLU (the reference androgen antagonist/anti-androgen) exhibited a biphasic curve, with the initial phase involving increased weights of the androgen-dependent tissues (and no change in the weights of the paired testes) at the low FLU doses (with the constant 1.0 mg/kg/day TP dose), due to the androgenic effect of TP at low anti-androgenic FLU doses. The second phase involved decreased weights of the androgen-dependent tissues (and increased weights of the paired testes) at the high FLU doses (with the constant 1.0 mg/kg/day TP dose). The most likely explanation is that the higher FLU doses overcame

("overwhelmed") the fixed TP stimulatory dose and resulted in the observed anti-androgen effects, which were reduced weights of the androgen-dependent organs and increased weights of the paired testes.

For both TP and FLU, all of the weights of the organs of interest exhibited significant laboratory effects, with predominantly the vast majority of the variability explained by the terms of the final model. There is no difference between the results of the analyses using "log10 (BW)" or "(BW)" in the model for any of the organ weights. See Table 4.11 for a summary of the interlaboratory statistical analyses.

Table 4.11. Summary of Interlaboratory Statistical Evaluations

Organ:	**Adrenals**	**COW**	**EPID**	**Kidneys**	**Liver**	**SVCG**	**LABC**	**Testes**	**VP**
TP (Phase 1A)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FLU (Phase 1B)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Yes = statistically significant interlaboratory effects on the weight(s) of specific organ(s) with TP or TP + FLU.

CHAPTER 5.0 RESULTS OF STAGE 2 STUDIES WITH WEAK AGONISTS AND WEAK ANTAGONISTS

The three laboratories contributing to the Stage 2 studies (Lab A, Lab B, and Lab C; Chapter 3.0, Table 3.2) and the laboratory parameters and conditions are the same as for Stage 1A and 1B. The lead laboratory (Dr. L. Earl Gray, Jr., Reproductive Toxicology Division, NHEERL, U.S. EPA, Research Triangle Park, NC) did not perform any of the studies. See Table 5.1 below for laboratory parameters and conditions.

Table 5.1. Laboratory Parameters and Conditions for the Phase 1 Validation

Lab	**Rat Strain & Supplier**		**Age on Study/ Necropsy**	**Vehicle**	**Diet**	**Animals per Cage**	**Bedding**
Lab A	Alpk:APfSD	Breeding facility on site	22-23 days/ 32-33 days	Tocopherol-stripped corn oil	R&M1, Special Diet Services, Witham, Essex, UK [a]	3	Sawdust paper
Lab B	Sprague Dawley	Breeding facility on site	22 days/ 32 days	corn oil	PMI 5001, Agribrands Purina, Lab B	3	Willow shavings
Lab C	Crj:CD IGS (SD)	Charles River Canada, St Constant, Quebec, Canada	22 days/ 32 days	peanut oil	PMI 5001, Agribrand Purina,	2	Pro chip maple (dust-free hardwood chips), PWI Industries, St-Hyacinthe, Quebec, Canada

[a] Purified (semi synthetic) diet

The three laboratories used Sprague Dawley rats. The age at study start was very similar (22 or 23 days of age) as was the age at necropsy (32 or 33 days of age). Two laboratories used corn oil as the vehicle, tocopherol-stripped corn oil by Lab A and corn oil by Lab B, and Lab C used peanut oil. Lab B and Lab C used the same Purina feed (PMI 5001) from different sites, and Lab A used a purified, semi-synthetic

diet designated as R&M1 (Special Diet Services, UK). Bedding type and source varied in all three laboratories (paper saw dust, willow shavings, or maple dust-free hardwood chips). The three test chemicals were TREN, LIN, and DDE (see below).

5.1 TREN

TREN (CAS No. 10161-34-9; a synthetic androgen agonist) was evaluated by the three performing laboratories at 1.5, 8, and 40 mg/kg BW/day by oral gavage in corn/peanut oil. The negative control was vehicle only, and the positive control was TP at 1.0 mg/kg/day by SC injection. The results are presented in Table 5.2.

The response of the androgen-responsive tissues to TP at 1 mg/kg/day (increased weight) was detected by all three laboratories for VP, SVCG, LABC, COW, and for decreased weight of the paired testes, and only by Lab A for increased weight of the paired epididymides. None of the three performing laboratories detected any changes in terminal BW at any dose of TREN or at 1 mg/kg/day TP. At 1.5 mg/kg/day TREN, only Lab A detected a significant increase in VP weight and a significant decrease in testes weight. At 8 mg/kg/day TREN, only Lab A and Lab B detected significant increases in LABC, Lab A and Lab C detected significant decreases in paired testes weights, and only Lab B detected a significant reduction in weight of the paired epididymides. At 40 mg/kg/day TREN, all three laboratories detected significant increases in the weights of the SVCG and LABC, and significant decreases in the weights of the paired testes. Lab A and Lab B detected a significant increase in the weight of the VP. Only Lab A detected significant increased weight of COW, and only Lab B detected a significantly reduced weight of the paired epididymides (also detected only by Lab B at 8 mg/kg/day TREN).

The dose response of the optional systemic organs (liver, paired adrenal glands, and paired kidneys) for TREN in the three performing laboratories is presented in Table 5.3. There were no significant changes in the weights of the optional systemic organs in any group for any of the three testing laboratories.

The CVs of the values for the BWs and all of the organs evaluated with TREN are presented in Table 5.4. The CVs were low and approximately equivalent for BWs at all doses for all three laboratories (except for a slight increase for Lab A for TREN at 1.5 mg/kg/day; 11.22%) versus CV values of 3.96% to 7.92% for all the other parameters for all the testing laboratories. CVs above 30% were exhibited by Lab B (34.25%) for SVCG at 8 mg/kg/day TREN and for Lab C (33.27%) for COWS at 40 mg/kg/day TREN. CVs above 20% were exhibited by Lab A for COW at 0 mg/kg/day, SVCG at 40 mg/kg/day TREN, and adrenal glands at 8 and 40 mg/kg/day TREN; by Lab B for COW and SVCG at 1.5 mg/kg/day, for SVCG at 8 mg/kg/day, for COW at 40 mg/kg/day, and for testes at 1 mg/kg/day TP; and by Lab C for SVCG at 0 mg/kg/day and for SVCG and COW at 8 and 40 mg/kg/day TREN.

Table 5.2. Dose Response of Body and Target Tissue Weights to TREN

			TREN (mg/kg-bw/d)			TP
Endpoint	**Laboratory**	**Vehicle**	**1.5**	**8**	**40**	**1 mg/kg-bw/d**
BW (g)	**Lab A**	121.2 ± 4.8	117.5 ± 13.19	123.5 ± 6.50	115.3 ± 9.14	126.2 ± 7.47
	Lab B	133.5 ± 9.9	138.1 ± 6.21	133.2 ± 8.37	126.7 ± 9.98	142.3 ± 5.98
	Lab C	132.8 ± 7.1	131.2 ± 5.47	139.3 ± 7.73	126.9 ± 8.02	142.0 ± 9.46
VP (mg)	**Lab A**	62.3 ± 11.8	80.5 ± 18.6 *	76.9 ± 9.8	85.6 ± 7.3 *	97.9 ± 11.4 *
	Lab B	78.9 ± 7.2	75.4 ± 10.3	72.7 ± 7.1	100.8 ± 12.1 *	121.7 ± 6.8 *
	Lab C	89.3 ± 13.2	84.7 ± 13.9	100.4 ± 16.3	105.4 ± 20.4	152.3 ± 24.8 8 *
SVCG (mg)	**Lab A**	70.0 ± 10.5	80.6 ± 14.2	87.03 ± 9.56	115.8 ± 24.5 *	248.5 ± 33.1 *
	Lab B	54.5 ± 9.9	56.35 ± 12.06	62.7 ± 21.5	86.4 ± 16.4 *	299.0 ± 27.9 *
	Lab C	38.1 ± 8.0	40.1 ± 6.7	45.30 ± 10.10	69.7 ± 20.0 *	247.2 ± 32.2 *
LABC (mg)	**Lab A**	107.6 ± 18.0	100.52 ± 19.57	146.1 ± 14.5 *	173.9 ± 11.0 *	195.5 ± 17.8 *
	Lab B	133.7 ± 22.9	160.5 ± 23.1	196.6 ± 21.2 *	299.9 ± 19.9 *	319.42 ± 25.16 *
	Lab C	131.1 ± 18.3	142.58 ± 16.89	172.1 ± 28.7	239.8 ± 37.1 *	264.8 ± 27.6 *
COW (mg)	**Lab A**	7.8 ± 2.1	6.3 ± 1.3	7.5 ± 0.8	10.5 ± 0.8 *	24.5 ± 3.6 *
	Lab B	9.1 ± 1.4	10.0 ± 2.7	7.7 ± 1.2	13.3 ± 3.7	24.0 ± 3.8 *
	Lab C	9.3 ± 1.4	8.5 ± 1.4	8.9 ± 2.0	13.0 ± 4.3	30.1 ± 3.4 *
Testes (mg)	**Lab A**	1116.8 ± 85.5	926.6 ± 105.7 *	717.2 ± 98.4 *	695.4 ± 57.7 *	467.8 ± 70.1 *
	Lab B	1124.4 ± 97.9	1121.8 ± 99.1	931.6 ± 127.9	691.9 ± 26.1 *	458.2 ± 115.0 *
	Lab C	1055.8 ± 131.8	1031.8 ± 72.6	944.6 ± 164.2 *	709.4 ± 96.8 *	568.7 ± 100.3 *
EPID (mg)	**Lab A**	122.3 ± 13.0	113.2 ± 18.3	98.7 ± 14.0	116.9 ± 19.1	166.5 ± 19.9 *
	Lab B	124.8 ± 9.4	120.3 ± 13.5	108.3 ± 7.9 *	99.3 ± 12.3 *	151.3 ± 15.3
	Lab C	146.8 ± 27.1	134.5 ± 10.7	133.3 ± 18.5	125.4 ± 17.4	158.4 ± 18.8

* Statistically significant by Dunnett's ($p < 0.05$) using ANCOVA with terminal BW.

Table 5.3. Dose Response of Optional Tissue and Organ Weights to TREN

Endpoint	Laboratory	Vehicle	TREN (mg/kg-bw/d)			TP
			1.5	8	40	1 mg/kg-bw/d
Liver (g)	Lab A	6.6 ± 0.7	6.6 ± 0.8	6.9 ± 0.3	6.5 ± 0.5	6.8 ± 0.4
	Lab B	6.1 ± 0.9	6.6 ± 0.6	6.3 ± 0.6	6.1 ± 0.8	6.6 ± 0.6
	Lab C	7.0 ± 0.5	6.8 ± 0.6	7.2 ± 0.4	6.9 ± 0.5	7.2 ± 0.5
Adrenals (mg)	Lab A	37.4 ± 5.4	37.0 ± 2.9	35.1 ± 7.5	37.5 ± 8.2	39.2 ± 4.2
	Lab B	23.8 ± 2.7	22.0 ± 2.2	21.8 ± 2.8	21.4 ± 1.3	23.7 ± 2.0
	Lab C	29.6 ± 4.2	27.1 ± 2.3	25.8 ± 3.5	26.6 ± 2.9	28.0 ± 2.1
Kidneys (mg)	Lab A	1197.5 ± 61.8	1183.7 ± 117.2	1206.3 ± 98.7	1166.8 ± 92.1	1273.2 ± 76.2
	Lab B	1256.3 ± 70.4	1249.2 ± 72.8	1164.6 ± 18.7	1263.0 ± 43.4	1276.5 ± 44.5
	Lab C	1476.5 ± 111.2	1415.0 ± 133.6	1492.1 ± 177.1	1401.2 ± 97.7	1495.9 ± 174.1

Table 5.4. CVs for Body and Organ Weights With TREN

Doses (mg/kg-bw/d)		BW	VP	SVCG	LABC	COWS	Testes	EPID	Liver	Adrenals	Kidneys
	Lab A	3.96%	18.95%	15.02%	16.75%	26.75%	7.66%	10.66%	9.87%	14.36%	5.16%
Vehicle	**Lab B**	7.43%	9.09%	18.18%	17.14%	15.56%	8.71%	7.56%	14.32%	11.43%	5.61%
	Lab C	5.35%	14.75%	21.08%	13.93%	14.92%	12.48%	18.46%	6.76%	14.30%	7.53%
	Lab A	11.22%	23.06%	17.57%	19.47%	19.96%	11.40%	16.13%	11.46%	7.81%	9.91%
Tren 1.5	**Lab B**	4.50%	13.64%	21.40%	14.41%	26.82%	8.84%	11.19%	8.86%	10.21%	5.83%
	Lab C	4.17%	16.38%	16.69%	11.85%	16.79%	7.04%	7.97%	9.14%	8.44%	9.44%
	Lab A	5.27%	12.76%	10.99%	9.94%	10.76%	13.72%	14.15%	4.98%	21.29%	8.18%
Tren 8	**Lab B**	6.28%	9.78%	34.25%	10.76%	15.18%	13.73%	7.27%	9.00%	13.01%	1.61%
	Lab C	5.55%	16.27%	22.29%	16.66%	22.74%	17.38%	13.85%	5.99%	13.73%	11.87%
	Lab A	7.92%	8.55%	21.17%	6.30%	7.30%	8.30%	16.32%	8.20%	21.75%	7.89%
Tren 40	**Lab B**	7.88%	11.99%	18.98%	6.63%	27.89%	3.78%	12.34%	13.02%	6.30%	3.43%
	Lab C	6.32%	19.35%	28.73%	15.49%	33.27%	13.65%	13.90%	7.21%	10.91%	6.97%
	Lab A	5.92%	11.69%	13.31%	9.09%	14.85%	14.98%	11.94%	5.81%	10.69%	5.98%
TP 1	**Lab B**	4.20%	5.63%	9.32%	7.88%	15.83%	25.09%	10.11%	9.36%	8.50%	3.49%
	Lab C	6.66%	16.26%	13.04%	10.42%	11.26%	17.65%	11.87%	7.43%	7.34%	11.64%

TREN – at 1.5, 8, or 40 mg/kg/day; TP – at 1 mg/kg/day

5.2 LIN

LIN (CAS No. 330-55-2), a weak anti-androgen (androgen antagonist), was evaluated by all three laboratories, but Lab A did not evaluate the vehicle only group, and Lab B and Lab C did not evaluate the groups with 1 mg/kg/day TP plus 3 mg/kg/day LIN. Statistical comparisons were not provided for the TP positive control versus the vehicle control values, since evaluation of the vehicle control group was optional. Results are presented in Table 5.5.

There were no changes in terminal BWs in any of the six dose groups versus the vehicle control group by any laboratory. All laboratories detected increased weights of the VP, SVCG, LABC, COW, and paired epididymides, and decreased weights of the paired testes in the TP-positive control group versus the vehicle only group values. Lab A (the only laboratory that evaluated the group with 1 mg/kg/day TP and 3 mg/kg/day LIN) and all the other laboratories reported no significant changes in the weights of any of the organs evaluated at 3, 10, or 30 mg/kg/day LIN, all with TP at 1 mg/kg/day. At 100 mg/kg/day LIN (plus TP at 1 mg/kg/day), all three laboratories detected decreases in LABC weights versus the TP positive control group values; Lab B and Lab C detected decreased VP weight versus TP positive control group; and Lab A and Lab C detected decreased SVCG weight versus TP positive control. Only Lab A reported decreased weight of COW versus TP positive control, and only Lab B reported a significant decrease in epididymal weight versus the TP positive control group value. None of the laboratories detected any statistically significant changes in the weight of the paired testes in any of the LIN + TP dosed groups versus the TP positive control group values.

The CVs for the body and organ weights with LIN + TP only or TP + FLU are presented in Table 5.6. Again, note that Lab A did not evaluate the vehicle control group, and Lab B and Lab C did not evaluate the 1 mg/kg/day TP + 3 mg/kg/day LIN group. The CVs were relatively low and similar across laboratories for BW, paired epididymides, liver, and paired kidneys. CVs were above 30% for Lab B for SVCG (92.58%), LABC (36.40%), and COW (62.00%) in the vehicle control group; for paired testes (47.51%) at TP + 30 mg/kg/day LIN, and for COW (36.78%) and paired testes (32.56%) at TP + 3.0 mg/kg/day FLU. Lab A exhibited CVs above 30% for VP weight at TP + 3.0 mg/kg/day LIN (34.65%), TP + 10 mg/kg/day LIN (30.94%), and TP + 30 mg/kg/day LIN (43.05%). Lab C exhibited a CV above 30% for COW (35.59%) in the vehicle control group.

The responses of the optional organ weights (liver, paired adrenal glands, and paired kidneys) to TP only, TP + LIN at 3, 10, 30, or 100 mg/kg/day, or to TP + FLU are presented in Table 5.7. There were no effects reported on the weights of any optional organs at any dose group versus the TP only positive control group values.

Table 5.5. Dose Response of Body and Tissue Weight to TP and TP + LIN

Endpoint	Laboratory	Vehicle	1 TP	1 TP + 3 LIN	1 TP + 10 LIN	1 TP + 30 LIN	1 TP + 100 LIN	1 TP + 3 FLU
BW (g)	**Lab A**	ND	128.3 ± 9.6	129.7 ± 8.62	125.3 ± 4.76	128.5 ± 9.31	117.7 ± 4.76	131.5 ± 10.46
	Lab B	125.5 ± 12.1	131.4 ± 6.35	ND	135.7 ± 7.53	129.5 ± 8.84	128.0 ± 7.07	139.4 ± 16.29
	Lab C	132.7 ± 13.8	139.3 ± 13.06	ND	129.5 ± 10.80	134.8 ± 11.96	120.6 ± 9.49	133.7 ± 8.93
VP (mg)	**Lab A**	ND	82.1 ± 21.9	90.9 ± 31.5	92.1 ± 28.5	71.9 ± 30.9	83.8 ± 8.2	71.6 ± 18.6
	Lab B	94.5 ± 21.8	125.8 ± 12.1	ND	116.9 ± 11.9	109.7 ± 8.6	95.9 ± 9.4 *	78.8 ± 21.0 *
	Lab C	76.0 ± 13.3	145.8 ± 27.8	ND	147.5 ± 25.8	125.0 ± 16.5	106.5 ± 8.8 **	91.3 ± 13.0 *
SVCG (mg)	**Lab A**	ND	215.3 ± 21.0	199.9 ± 22.4	203.47 ± 12.07	191.6 ± 25.2	158.7 ± 12.9 *	125.9 ± 22.4 *
	Lab B	79.7 ± 73.8	256.8 ± 33.9	ND	220.17 ± 35.29	198.7 ± 56.4	171.3 ± 54.3	84.9 ± 19.1 *
	Lab C	35.7 ± 8.9	240.9 ± 46.0	ND	244.77 ± 32.19	228.0 ± 34.9	171.7 ± 50.3 **	79.7 ± 8.2 *
LABC (mg)	**Lab A**	ND	166.9 ± 21.9	168.48 ± 29.93	161.5 ± 11.3	148.9 ± 33.2	132.9 ± 10.3 **	120.8 ± 19.3 *
	Lab B	129.8 ± 47.3	241.38 ± 36.91	ND	224.7 ± 17.2	235.1 ± 26.4	186.0 ± 21.1 *	159.6 ± 25.9 *
	Lab C	124.7 ± 16.9	262.88 ± 32.68	ND	272.7 ± 23.9	259.6 ± 13.5	208.2 ± 27.4 **	179.9 ± 17.6 *
COW (mg)	**Lab A**	ND	19.4 ± 1.9	19.3 ± 2.6	19.3 ± 2.9	17.0 ± 4.1	13.9 ± 3.4 **	9.4 ± 2.6 *
	Lab B	10.9 ± 6.7	21.1 ± 2.9	ND	19.7 ± 3.6	20.0 ± 2.6	15.3 ± 3.6	10.4 ± 3.8 *
	Lab C	8.4 ± 3.0	28.0 ± 3.6	ND	28.0 ± 7.5	24.9 ± 5.1	21.5 ± 5.4	13.7 ± 3.0 *
Testes (mg)	**Lab A**	ND	485.9 ± 83.1	500.2 ± 87.4	451.3 ± 74.6	544.9 ± 138.5	529.1 ± 105.2	908.3 ± 162.9 *
	Lab B	1039.9 ± 300.1	588.7 ± 141.4	ND	552.4 ± 153.4	505.4 ± 240.1	628.0 ± 258.5	960.0 ± 312.6 **
	Lab C	1085.2 ± 102.8	539.2 ± 88.9	ND	469.0 ± 24.3	455.1 ± 78.1	443.8 ± 73.8	798.2 ± 202.2 *
EPID (mg)	**Lab A**	ND	131.7 ± 15.6	132.2 ± 9.9	132.4 ± 13.8	132.7 ± 9.0	118.4 ± 10.8	108.4 ± 16.0 *
	Lab B	134.9 ± 10.6	163.1 ± 11.7	ND	150.6 ± 11.9	141.9 ± 12.6	132.7 ± 11.6 *	107.0 ± 18.4 *
	Lab C	149.1 ± 19.0	178.5 ± 12.9	ND	166.5 ± 10.1	159.3 ± 13.2	153.1 ± 26.8	125.9 ± 12.7 *

* Statistically significant by Dunnett's ($p < 0.05$) using ANCOVA with terminal BWs

** Statistically significant by Dunnett's ($p < 0.05$) using ANCOVA with starting BWs

TP – at 1 mg/kg/day by sc injection; LIN – at 3, 10, 30, or 100 mg/kg/day; FLU – at 3.0 mg/kg/day by oral gavage; ND – not done

Table 5.6. CVs of Body and Organ Weights with LIN

		BW	VP	SVCG	LABC	COWS	TEST	EPID	Liver	Adrenals	Kidneys
Vehicle	**Lab A**	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Lab B	9.61%	23.07%	92.58%	36.40%	62.00%	28.86%	7.86%	16.01%	11.65%	9.80%
	Lab C	10.37%	17.51%	24.82%	13.54%	35.59%	9.48%	12.71%	19.15%	19.10%	13.52%
1 TP	**Lab A**	7.47%	26.68%	9.73%	13.13%	10.02%	17.10%	11.88%	6.72%	27.58%	7.26%
	Lab B	4.83%	9.60%	13.22%	15.29%	13.88%	24.01%	7.17%	13.58%	21.88%	6.63%
	Lab C	9.37%	19.10%	19.09%	12.43%	12.83%	16.48%	7.22%	12.01%	21.57%	9.07%
1 TP + 3 LIN	**Lab A**	6.65%	34.65%	11.21%	17.76%	13.27%	17.48%	7.53%	11.70%	10.29%	11.97%
	Lab B	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Lab C	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
1 TP + 10 LIN	**Lab A**	3.80%	30.94%	5.93%	6.97%	14.98%	16.54%	10.40%	6.05%	20.28%	3.67%
	Lab B	5.55%	10.21%	16.03%	7.65%	18.28%	27.78%	7.91%	8.09%	10.66%	7.77%
	Lab C	8.35%	17.48%	13.15%	8.77%	26.82%	5.18%	6.09%	9.13%	15.21%	9.76%
1 TP + 30 LIN	**Lab A**	7.25%	43.05%	13.16%	22.33%	24.24%	25.42%	6.78%	9.23%	21.47%	12.62%
	Lab B	6.83%	7.82%	28.40%	11.23%	12.80%	47.51%	8.87%	8.91%	8.55%	7.15%
	Lab C	8.87%	13.21%	15.30%	5.22%	20.47%	17.16%	8.27%	10.67%	10.65%	8.73%
1 TP + 100 LIN	**Lab A**	4.05%	9.81%	8.13%	7.77%	24.42%	19.89%	9.13%	8.77%	10.93%	3.42%
	Lab B	5.52%	9.83%	31.66%	11.36%	23.72%	41.16%	8.78%	10.60%	14.04%	10.23%
	Lab C	7.87%	8.27%	29.27%	13.16%	25.28%	16.63%	17.50%	6.41%	11.77%	7.43%
1 TP + 3 FLU	**Lab A**	7.96%	26.00%	17.78%	15.94%	27.88%	17.93%	14.78%	10.30%	24.65%	7.23%
	Lab B	11.69%	26.69%	22.50%	16.23%	36.78%	32.56%	17.18%	15.34%	16.71%	13.03%
	Lab C	6.68%	14.18%	10.24%	9.77%	22.00%	25.33%	10.12%	13.21%	10.81%	6.65%

83

Table 5.7. Dose Response of Optional Organ Weights to LIN Relative to TP Only Dose

Endpoint	Laboratory	Vehicle	1 TP	1 TP + 3 LIN	1 TP + 10 LIN	1 TP + 30 LIN	1 TP + 100 LIN	1 TP + 3 FLU
Liver (g)	Lab A	ND	6.5 ± 0.4	6.5 ± 0.8	6.2 ± 0.4	6.6 ± 0.6	6.4 ± 0.6	6.7 ± 0.7
	Lab B	5.5 ± 0.9	6.0 ± 0.8	ND	6.4 ± 0.5	6.0 ± 0.5	6.2 ± 0.7	6.8 ± 1.0
	Lab C	5.8 ± 1.1	7.0 ± 0.8	ND	6.6 ± 0.6	6.6 ± 0.7	5.8 ± 0.4	6.4 ± 0.8
Adrenals (mg)	Lab A	ND	21.3 ± 5.9	27.1 ± 2.8	25.2 ± 5.1	24.9 ± 5.3	26.3 ± 2.9	29.8 ± 7.4
	Lab B	20.7 ± 2.4	21.6 ± 4.7	ND	20.9 ± 2.2	21.7 ± 1.9	21.6 ± 3.0	22.9 ± 3.8
	Lab C	21.0 ± 4.0	27.6 ± 6.0	ND	23.9 ± 3.6	24.0 ± 2.6	23.6 ± 2.8	24.3 ± 2.6
Kidneys (mg)	Lab A	ND	1251.8 ± 90.8	1214.4 ± 145.4	1167.9 ± 42.9	1224.8 ± 154.5	1131.2 ± 38.7	1243.8 ± 89.9
	Lab B	1172.2 ± 114.9	1268.2 ± 84.1	ND	1289.7 ± 100.2	1214.2 ± 86.8	1250.5 ± 127.9	1351.0 ± 176.0
	Lab C	1309.2 ± 177.0	1422.9 ± 129.0	ND	1352.9 ± 132.1	1398.7 ± 122.1	1233.4 ± 91.6	1288.0 ± 85.7

TP – at 1 mg/kg/day; LIN – at 3, 10, 30, or 100 mg/kg/day; FLU – at 3 mg/kg/day; ND – Not done

5.3 DDE

DDE, an anti-androgen (androgen antagonist; Kelce et al., 1995, 1997; Leavens et al., 2002), was administered by oral gavage at 5, 16, 50, or 160 mg/kg/day, concurrent with TP at 1.0 mg/kg/day, with the negative control corn/peanut oil group, the positive stimulatory control group, TP alone at 1.0 mg/kg/day, and the positive inhibitory control group, TP (1.0 mg/kg/day) plus FLU (at 3 mg/kg/day). The results of the dose-response for the androgen-sensitive organs are presented in Table 5.8. Again, Lab A did not evaluate the vehicle control group, and Lab B and Lab C did not evaluate the TP (1.0 mg/kg/day) + DDE (5.0 mg/kg/day) group, the lowest DDE dose group. There were no statistical comparisons between the vehicle control and the TP positive control group values. All statistical comparisons were against the TP positive control group value.

BW was unaffected in any of the dose groups in any of the three laboratories. At 5.0 mg/kg/day DDE (+ TP), Lab A (the only laboratory participating at this dose) detected no significant changes in any of the androgen-sensitive organs evaluated. At 16 mg/kg/day DDE (+ TP), none of the three laboratories detected any significant differences for any of the organs evaluated. At 50 mg/kg/day DDE (+ TP), significantly decreased SVCG weights were detected by Lab B and Lab C. No other organ weights differed significantly from the positive control values at this dose in any of the laboratories. At 160 mg/kg/day DDE (+ TP), all three laboratories detected significant reductions in the weights of the VP, SVCG, LABC, and COW relative to the TP only positive control group values. Only Lab A detected a statistically significant decrease in the paired testes weights, and Lab B and Lab C (but not Lab A, although the value was lower) detected statistically significant reductions in the weights of the paired epididymides. All three laboratories detected significant reductions in all the organs evaluated in the TP + FLU group versus the organ weights from the TP only positive control group.

The results of the evaluation of DDE on the optional systemic organs are presented in Table 5.9. Neither DDE at 5, 16, 50, or 160 mg/kg/day nor TP + FLU (the positive inhibitory dose group) exhibited any significant changes in weights of any of the optional systemic organs evaluated (liver, paired adrenal glands, or paired kidneys).

The CVs for the body and organ weight values with DDE are presented in Table 5.10. CVs for BWs were low and approximately equivalent (below 10%) across all laboratories and dose groups except for Lab B with a CV of 17.57% for BW at 160 mg/kg/day DDE (+ TP). For the organs, only one laboratory (Lab C) exhibited a CV above 40% (42.25%) for paired testes at 50 mg/kg/day DDE (+ TP). Three laboratories exhibited CVs between 30 and 40%: Lab A (33.38%) and Lab B (32.97%) for paired testes, Lab B for COW (33.64%) and Lab C for SVCG (31.40%). All laboratories exhibited CVs between 20 and 30%; Lab B eight times and Lab C and Lab A both seven times. Most of the high CVs (above 20%) occurred for SVCG (eight times) and paired testes (nine times), with no CVs above 20% for BW, liver, or paired kidney weights. It is obviously not just organ size (although body, liver, paired testes, and kidneys are large). The high variances may be due to variable responsiveness of the SVCG and paired testes from laboratory to laboratory, even at the same dose.

Table 5.8. Dose Response of Body and Organ Weights to TP and TP + DDE

Endpoint	Laboratory	Vehicle	1 TP	1 TP + 5 DDE	1 TP + 16 DDE	1 TP + 50 DDE	1 TP + 160 DDE	1 TP + 3 FLU
BW (g)	Lab A	ND	129.0 ± 9.0	128.7 ± 6.98	130.7 ± 5.32	124.8 ± 8.82	118.7 ± 9.03	126.3 ± 5.82
	Lab B	130.0 ± 12.0	144.3 ± 13.17	ND	137.4 ± 6.05	141.0 ± 9.70	131.5 ± 23.10	138.0 ± 8.73
	Lab C	132.3 ± 10.1	143.3 ± 9.32	ND	133.2 ± 4.19	133.7 ± 4.77	130.1 ± 11.82	136.4 ± 7.85
VP (mg)	Lab A	ND	102.9 ± 11.0	102.8 ± 9.0	112.9 ± 14.9	96.0 ± 16.0	83.5 ± 15.9 **	70.5 ± 10.4 *
	Lab B	67.8 ± 10.6	120.5 ± 25.0	ND	113.6 ± 16.4	122.0 ± 38.1	76.0 ± 11.5 *	90.6 ± 10.3 **
	Lab C	101.9 ± 16.9	173.9 ± 27.0	ND	149.0 ± 16.9	136.9 ± 12.6	101.3 ± 17.1 *	112.0 ± 30.0 *
SVCG (mg)	Lab A	ND	235.3 ± 16.6	248.6 ± 23.6	238.02 ± 28.38	210.8 ± 27.9	167.1 ± 44.9 *	126.4 ± 7.1 *
	Lab B	47.1 ± 5.4	241.3 ± 35.3	ND	231.47 ± 15.10	178.2 ± 47.0 *	117.8 ± 33.7 *	97.3 ± 20.6 *
	Lab C	35.8 ± 6.6	279.9 ± 57.9	ND	223.98 ± 26.11	185.5 ± 58.3 **	90.5 ± 24.0 *	103.5 ± 28.4 *
LABC (mg)	Lab A	ND	178.3 ± 26.6	177.92 ± 24.22	162.9 ± 18.0	155.0 ± 20.1	144.6 ± 30.1 **	111.6 ± 7.2 *
	Lab B	140.5 ± 19.4	270.43 ± 33.20	ND	247.6 ± 12.1	260.6 ± 34.0	172.6 ± 40.8 *	200.7 ± 15.7 *
	Lab C	156.4 ± 16.8	290.47 ± 49.81	ND	248.6 ± 42.3	230.5 ± 25.7	157.8 ± 30.5 *	193.5 ± 22.1 *
COWS (mg)	Lab A	ND	21.8 ± 3.1	21.4 ± 1.7	23.3 ± 3.5	19.5 ± 2.5	11.5 ± 2.5 *	10.5 ± 1.3 *
	Lab B	7.3 ± 0.9	23.5 ± 4.2	ND	23.2 ± 2.8	21.2 ± 1.0	15.2 ± 5.1 *	13.7 ± 2.2 *
	Lab C	11.2 ± 2.2	29.4 ± 6.3	ND	24.5 ± 2.4	22.9 ± 3.5	14.3 ± 3.1 *	17.3 ± 2.9 *
Testes (mg)	Lab A	ND	473.6 ± 87.0	433.0 ± 106.5	406.6 ± 39.7	358.4 ± 68.6	499.6 ± 166.8 *	671.9 ± 141.0 *
	Lab B	1081.1 ± 87.5	693.1 ± 228.5	ND	500.9 ± 112.0	608.5 ± 156.9	631.4 ± 187.6	973.0 ± 154.1 *
	Lab C	1134.5 ± 61.4	602.4 ± 102.7	ND	582.6 ± 82.5	626.4 ± 264.7	781.5 ± 228.2	896.0 ± 152.6 *
EPID (mg)	Lab A	ND	136.4 ± 9.0	142.0 ± 9.3	133.8 ± 12.1	122.6 ± 6.0	97.7 ± 7.5	88.9 ± 8.4 *
	Lab B	124.0 ± 11.1	153.0 ± 17.5	ND	145.4 ± 9.3	134.4 ± 20.2	106.8 ± 26.7 *	119.6 ± 18.0 **
	Lab C	167.1 ± 22.6	171.2 ± 16.4	ND	154.2 ± 18.6	150.1 ± 29.7	127.2 ± 22.8 **	135.7 ± 10.5 **

* Statistically significant by Dunnett's ($p < 0.05$) using ANCOVA with terminal BWs

** Statistically significant by Dunnett's ($p < 0.05$) using ANCOVA with starting BWs

Table 5.9. Dose Response of Optional Tissue and Organ Weights to DDE

Endpoint	Laboratory	Vehicle	1 TP	1 TP + 5 DDE	1 TP + 16 DDE	1 TP + 50 DDE	1 TP + 160 DDE	1 TP + 3 FLU
Liver (g)	Lab A	ND	6.1 ± 0.5	6.6 ± 0.5	8.0 ± 0.6	8.4 ± 0.9	9.4 ± 1.0	6.3 ± 0.3
	Lab B	5.5 ± 0.8	6.4 ± 0.9	ND	7.4 ± 0.6	8.7 ± 0.9	9.7 ± 1.6	6.2 ± 0.4
	Lab C	6.4 ± 0.4	7.1 ± 0.7	ND	7.7 ± 0.6	9.3 ± 0.8	11.2 ± 1.5	6.8 ± 0.4
Adrenals (mg)	Lab A	ND	27.0 ± 1.8	25.4 ± 6.6	27.7 ± 5.7	24.4 ± 4.4	22.4 ± 5.7	23.8 ± 3.1
	Lab B	23.3 ± 6.5	24.2 ± 2.9	ND	22.7 ± 2.7	23.4 ± 2.4	24.7 ± 6.0	23.4 ± 3.6
	Lab C	30.1 ± 1.2	30.0 ± 3.4	ND	32.4 ± 4.1	29.7 ± 5.0	33.2 ± 2.6	31.3 ± 4.4
Kidneys (mg)	Lab A	ND	1237.7 ± 88.9	1242.5 ± 73.5	1285.3 ± 84.0	1259.7 ± 93.9	1240.0 ± 129.8	1150.3 ± 79.9
	Lab B	1279.9 ± 142.1	1307.3 ± 126.6	ND	1345.5 ± 80.4	1353.7 ± 127.9	1285.8 ± 230.1	1288.4 ± 71.1
	Lab C	1417.4 ± 107.2	1551.1 ± 101.8	ND	1429.6 ± 82.5	1533.6 ± 83.5	1526.7 ± 172.4	1448.1 ± 60.3

ND – not done

Table 5.10. CVs of Body and Organ Weights With DDE

Doses (mg/kg-bw/d)		BW	VP	SVCG	LABC	COWS	TEST	EPID	Liver	Adrenals	Kidneys
Vehicle	Lab A	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Lab B	9.20%	15.70%	11.38%	13.84%	11.95%	8.09%	8.98%	13.66%	27.74%	11.10%
	Lab C	7.67%	16.63%	18.45%	10.72%	19.94%	5.41%	13.51%	6.67%	4.05%	7.57%
1 TP	Lab A	6.99%	10.69%	7.03%	14.90%	14.11%	18.36%	6.58%	8.54%	6.83%	7.18%
	Lab B	9.13%	20.75%	14.63%	12.28%	18.03%	32.97%	11.47%	13.59%	11.98%	9.69%
	Lab C	6.51%	15.55%	20.69%	17.15%	21.55%	17.04%	9.59%	9.35%	11.23%	6.56%
1 TP + 5 DDE	Lab A	5.42%	8.78%	9.51%	13.61%	8.09%	24.60%	6.53%	6.86%	25.95%	5.91%
	Lab B	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Lab C	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
1 TP + 16 DDE	Lab A	4.07%	13.20%	11.92%	11.07%	15.03%	9.76%	9.07%	7.28%	20.50%	6.53%
	Lab B	4.41%	14.42%	6.52%	4.88%	12.22%	22.36%	6.39%	7.69%	11.71%	5.97%
	Lab C	3.14%	11.34%	11.66%	17.01%	9.63%	14.16%	12.08%	7.93%	12.64%	5.77%
1 TP + 50 DDE	Lab A	7.06%	16.71%	13.24%	12.99%	12.74%	19.13%	4.90%	10.81%	18.03%	7.45%
	Lab B	6.88%	31.18%	26.38%	13.03%	4.53%	25.78%	15.03%	9.86%	10.12%	9.45%
	Lab C	3.57%	9.18%	31.40%	11.14%	15.28%	42.25%	19.76%	8.92%	16.70%	5.44%
1 TP + 160 DDE	Lab A	7.61%	19.00%	26.86%	20.83%	22.18%	33.38%	7.70%	10.86%	25.63%	10.47%
	Lab B	17.57%	15.17%	28.64%	23.62%	33.64%	29.72%	24.99%	16.27%	24.43%	17.90%
	Lab C	9.08%	16.84%	26.49%	19.32%	21.69%	29.20%	17.91%	13.60%	7.74%	11.30%
1 TP + 3 FLU	Lab A	4.61%	14.76%	5.61%	6.47%	12.15%	20.98%	9.41%	5.21%	12.92%	6.95%
	Lab B	6.33%	11.34%	21.16%	7.82%	16.25%	15.84%	15.05%	6.56%	15.54%	5.51%
	Lab C	5.76%	26.75%	27.44%	11.40%	17.09%	17.03%	7.75%	6.40%	13.94%	4.16%

TP – at 1.0 mg/kg/day; DDE – at 5, 16, 50, and 160 mg/kg/day; FLU – at 3 mg/kg/day; ND – Not Done

5.4 Interlaboratory Comparisons

A model was developed by Dr. Breda Munoz (Senior Research Statistician at RTI International) to determine if there is a statistically significant laboratory effect (i.e., if the laboratories differ significantly in their ability to detect effects on specific organ weights). Since the participating laboratories apparently used both "log10(BW)" (logarithm base 10 of body weight) and "BW" (body weight), the model was constructed to use both terms. The full model, which evaluated the interaction of all fixed effects is as follows:

> Log10 (organ weight) = log10 (BW) + (DoseGroup) + log10 (BW)*(DoseGroup) + Lab + Lab*log10 (BW) + Lab*(DoseGroup) + Lab*log10 (BW)*(DoseGroup), where "*" indicates the interaction term

The model was run for each organ of interest, with "log10 (BW)" as shown and then with "BW" as well. If the three-way interaction term "Lab*log10 (BW)*(DoseGroup)" was statistically significant ($p<0.05$), then changes in log10 (organ weight) across different dose groups and for given body weights were different across labs (suggesting the laboratories' abilities to detect changes varied across body weight and dose group combinations). If the three-way interaction term "log10 (BW)*DoseGroup*labs" was not significant ($p>0.05$), then a reduced model (without this three-way interaction term) was run to fit the data. If the "log10 (BW)*lab" probability was not significant ($p>0.05$), then the model was rerun to fit the data without this term as well. If the two-way interaction term "DoseGroup*Lab" was significant ($p<0.05$), then the effect of dose groups on Log10 (organ weight) was significantly different across laboratories (i.e., a laboratory effect was present).

If the two-way interaction term "DoseGroup*Lab" probability was nonsignificant ($p>0.05$), then a reduced model (without this two-way interaction term) was run to fit the data. If the two-way interaction term "Lab*log10 (BW)" was significant ($p<0.05$), then the effect of "log10(BW)" on "log10 (organ weight)" was significantly different across labs (i.e., a lab effect was present). If the two-way interaction term "Lab*log10 (BW)" was nonsignificant ($p>0.05$), then a reduced model (without this two-way interaction term) fit the data. If the term "lab" was significant ($p<0.05$), then lab conditions had a statistically significant different effect in detecting differences in "log10 (organ weight)". Finally, if the term "lab" was nonsignificant ($p>0.05$), then there was no statistical difference in the laboratories' ability to detect differences in "log10 (organ weight)".

The "R-Squared" term generated from the model analysis is equal to the percentage of the variability around the "log10 (organ weight)" that is explainable by the terms in the model, including laboratory. Therefore, we can deduce whether at least one laboratory is significantly different from the rest, and that it impacts on the variability around the organ weight. However, this model cannot identify which lab(s) is different. Identification of the laboratory or laboratories requires another series of comparisons ("contrast tests") of laboratory 1 versus 2, 1 versus 3, etc.

The full model evaluates all of the fixed effects and is fitted to the data which include the lab effect as well as effects of body weight and dose group. Fixed effects are "DoseGroup" and "lab". "Log10 (BW)" is a continuous covariate.

A more detailed explanation and the results of the analyses are presented in Appendix E (Stage 2 Analyses). The results of the model for interlaboratory comparisons of Stage 2 TREN, DDE, and LIN data are presented below.

5.4.1 TREN

Paired Adrenal Glands

I. Using "log10 (BW)"
   A. Full model
      Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.297, not significant (NS)
   B. Model minus triple interaction term
      Double interaction "log10 (BW) * lab" term : p=0.899 NS
   C. Model minus triple and double interaction terms
      "DoseGroup*lab" term: p=0.865 NS
   D. Model minus triple, double and "DoseGroup*lab" terms
      "Lab" term: p=0.000; is significant, so there is a lab effect, and 79.0% (R-Squared = 0.790) of the variability is explained by the terms in the final model.

II. Using "(BW)"
   A. Full model
      Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.283 NS
   B. Model minus triple interaction term
      Double interaction "DoseGroup*lab" term: p=0.862 NS
   C. Model minus triple and double interaction terms:
      "(BW)"*lab" term: p=0.992 NS
   D. Model minus triple, double and "BW * lab" terms:
      "Lab" term: p=0.000, is significant, so there is a lab effect, and 79.0% (R-Squared = 0.790) of the variability is explained by the terms in the final model.

Paired Cowper's Glands

I. Using "log10 (BW)"
   A. Full model
      Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.266 NS
   B. Model minus triple interaction term
      Double interaction "log10 (BW)*lab" term: p=0.429 NS
   C. Model minus triple and double interaction terms
      "DoseGroup*Lab": p=0.654 NS
   D. Model minus triple, double and "DoseGroup*lab" terms
      "Lab" term: p=0.006; is significant, so there is a lab effect, and 85.1% (R-Squared = 0.851) of the variability is explained by the terms in the final model.

II. Using "(BW)"
   A. Full model
      Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.252 NS
   B. Model minus triple interaction term
      Double interaction "DoseGroup*lab" term: p=0.800 NS
   C. Model minus triple and double interaction terms
      "(BW)*lab" term: p=0.223 NS
   D. Model minus triple, double and "(BW)*lab" terms
      "Lab" term: p=0.008; is significant, so there is a lab effect, and 85.1% (R-Squared = 0.851) of the variability is explained by the terms in the final model.

Paired Epididymides

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.427 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.031; is significant so the effect of body weight on paired epididymides weight is different across laboratories. This model explains 69.8% (R-Squared = 0.698) of the variability around the paired epididymides weight.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.326 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.272 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.066 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 61.9% (R-Squared = 0.619) of the variability is explained by the terms in the final model.

Paired Kidneys

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.383 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.573 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.572 NS

D. Model minus triple, double and "DoseGroup*lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 73.7% (R-Squared = 0.737) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.500 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.650 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.401 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 73.3% (R-Squared = 0.733) of the variability is explained by the terms in the final model.

LABC

I. Using "log10 (BW)"

A. Full model

Triple interaction term: p=0.135 NS

B. Model minus triple interaction "log10 (BW) * DoseGroup * labs" term

Double interaction "log10 (BW)*lab" term: p=0.801 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.326 NS

D. Model minus triple, double and "DoseGroup*lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 86.7% (R-Squared = 0.867) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.130 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.326 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.990 NS

D. Model minus triple, double and "(BW)*lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 86.6% (R-Squared = 0.866) of the variability is explained by the terms in the final model.

Liver

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.286 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.172 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.226 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 61.7% (R-Squared = 0.617) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.251 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.251 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.269 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 61.8% (R-Squared = 0.618) of the variability is explained by the terms in the final model.

Paired Seminal Vesicles

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.391 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.252 NS

C. Model minus triple and double interactions

"DoseGroup*lab" term: p=0.000, is significant, so there is a lab effect, and 94.7% (R-Squared = 0.947) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.406 NS

B. Model minus triple interaction term

Double interaction term "DoseGroup * Lab": p=0.000, is significant, so the effect of dose group on seminal vesicle weight is different across laboratories (i.e., a lab effect). This model also explains 94.9% (R-Squared = 0.949) of the variability around seminal vesicle weight.

Paired Testes

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.360 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.971 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.789 NS

D. Model minus triple, double and "DoseGroup * lab": terms

"Lab" term: p=0.189 NS

E. Model minus triple, double and "DoseGroup * lab" and "lab": terms

"log10 (BW) * DoseGroup" term: p=0.435 NS

F. Model minus triple, double, "DoseGroup * lab", "lab", and "log10 (BW) * DoseGroup" terms

"DoseGroup" term: p=0.000 and "log10(BW)" term: p=0.014. Both are significant and both are significant causes of variability, around paired testes weight and the model explains 68.2% (R-Squared = 0.682) of the variability.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.411 NS

B. Model minus triple interaction term

Double interaction "DoseGroup * lab" term: p=0.808 NS

C. Model minus triple and double interaction terms

"(BW)"*lab" term: p=0.934 NS

D. Model minus triple, double, and "(BW)*lab" terms

"Lab" term: p=0.188 NS

E. Model minus triple, double, "(BW)*lab" and "lab" terms

"DoseGroup" term: p=0.000

"BodyWeight" term: p=0.016

Both terms are statistically significant and indicate that the variability in both effects explains the variability of log10 (testes) weight and that the labs are different. This model explains 68.1% (R-Squared = 0.681) of the variability around paired testes weight.

Ventral Prostate

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.124 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.303 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.246 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 52.6% (R-Squared = 0.526) of the variability is explained by the terms in the final model.

II. Using ("BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.134 NS

B. Model minus triple interaction term

Double interaction "DoseGroup * lab" term: p=0.156 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.710 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 52.5% (R-Squared = 0.525) of the variability is explained by the terms in the final model.

Conclusions

All of the organ weights in animals exposed to trenbolone (a synthetic androgen), except for paired testes, exhibited significant laboratory effects, and the majority (predominantly the vast majority) of the variability around the organ weights is explained by the terms in the model. For the paired testes weights, both dose group and body weight "log10(BW)" or "(BW)" were significant causes of the variability and explained 68.1 "(BW)" – 68.2 "log10(BW)" % of the variability around testes weights.

5.4.2 DDE

Paired Adrenal Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.389 (NS)

B. Model minus triple interaction term

Double interaction "log10 (BW) * lab" term : p=0.366 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.268 NS

D. Model minus triple, double and "DoseGroup*lab" terms

"Lab" term: p=0.000; is significant, so there is a lab effect, and 58.5% (R-Squared = 0.585) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.393 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.389 NS

C. Model minus triple and double interaction terms:

"(BW)"*lab" term: p=0.228 NS

D. Model minus triple, double and "DoseGroup*lab" and "BW * lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 58.4% (R-Squared = 0.584) of the variability is explained by the terms in the final model.

Paired Cowper's Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.516 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.976 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab": p=0.009 is significant, so there is a lab effect, and 88.8% (R-Squared = 0.888) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.522 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.017 is significant, so there is a lab effect (i.e., the effect of different dose groups on Cowper's glands is different across laboratories), and 88.5% (R-Squared = 0.885) of the variability is explained by the terms in the final model.

Paired Epididymides

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.154 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.032; is significant so the effect of body weight on paired epididymides weight is different across laboratories. This model explains 80.7% (R-Squared = 0.807) of the variability around the paired epididymides weight.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.171 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.078 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.035, is significant, so the effect of body weight on paired epididymides weight is different across laboratories (i.e., a lab effect), and 76.1% (R-Squared = 0.761) of the variability around the paired epididymides weight.

Paired Kidneys

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.920 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.348 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.104 NS

D. Model minus triple, double and "DoseGroup*lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 78.8% (R-Squared = 0.788) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.929 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.015 is significant, so there is a lab effect, and 82.4% (R-Squared = 0.824) of the variability is explained by the terms in the final model.

LABC

I. Using "log10 (BW)"

A. Full model

Triple interaction term: p=0.443 NS

B. Model minus triple interaction "log10 (BW) * DoseGroup * labs" term

Double interaction "log10 (BW)*lab" term: p=0.651 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.000 is significant, indicating a lab effect (i.e., the effect of dose group on LABC weight is different across laboratories), and 90.3% (R-Squared = 0.903) of the variability around LABC weight is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.387 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.000, is significant, so there is a lab effect, and 90.3% (R-Squared = 0.903) of the variability around LABC weight is explained by the terms in the final model.

Liver

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.684 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.764 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.608 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 91.8% (R-Squared = 0.918) of the variability around liver weight is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.584 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.577 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.939 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 91.6% (R-Squared = 0.916) of the variability is explained by the terms in the final model.

Paired Seminal Vesicles

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.316 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.936 NS

C. Model minus triple and double interactions

"DoseGroup*lab" term: p=0.000, is significant, so there is a lab effect, and 92.3% (R-Squared = 0.923) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.332 NS

B. Model minus triple interaction term

Double interaction term "DoseGroup * Lab": p=0.000, is significant, so the effect of dose group on seminal vesicle weight is different across laboratories (i.e., a lab effect). This model also explains 92.3% (R-Squared = 0.923) of the variability around seminal vesicle weight.

Paired Testes

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.023; is significant, so that the combined effect of both "DoseGroup" and "log10 (BW)" on paired testes weight is different across laboratories. This model explains 83.0% (R-Squared = 0.830) of the variability of paired testes weight.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.021; is significant, so that the combined effect of both "DoseGroup" and "BW" on paired testes weight is different across laboratories. This model explains 82.9% (R-Squared = 0.829) of the variability of paired testes weight.

Ventral Prostate

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.355 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.643 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.100 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 72.1% (R-Squared = 0.721) of the variability of VP weight is explained by the terms in the final model.

II. Using ("BW")

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.354 NS

B. Model minus triple interaction term

Double interaction "DoseGroup * lab" term: p=0.145 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.389 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 71.9% (R-Squared = 0.719) of the variability is explained by the terms in the final model.

DDE: Conclusions

For all of the organ weights in animals exposed to DDE, there is a laboratory effect; the effect of dose group and/or body weight on organ weight is different across laboratories. There is no difference between the analyses using "log10(BW)" versus "(BW)".

5.4.3 LIN

Paired Adrenal Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.297 (NS)

B. Model minus triple interaction term

Double interaction "log10 (BW) * lab" term : p=0.899 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.865 NS

D. Model minus triple, double and "DoseGroup*lab" terms

"Lab" term: p=0.000; is significant, so there is a lab effect, and 79.0% (R-Squared = 0.790) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.283 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.862 NS

C. Model minus triple and double interaction terms:

"(BW)"*lab" term: p=0.992 NS

D. Model minus triple, double and "DoseGroup*lab" and "BW * lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 79.0% (R-Squared = 0.790) of the variability is explained by the terms in the final model.

Paired Cowper's Glands

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.266 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.429 NS

C. Model minus triple and double interaction terms:

Double interaction "DoseGroup*Lab" term: p=0.654 NS

D. Model minus triple and double interaction terms

"Lab": p=0.006 is significant, so there is a lab effect, and 85.1% (R-Squared = 0.851) of the variability is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.252 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.800 NS

C. Model minus triple and double interaction terms:

Double interaction "(BW)*lab" term: p=0.223 NS

D. Model minus triple and double interaction terms

"Lab": p=0.008 is significant, so there is a lab effect, and 85.1% (R-Squared = 0.851) of the variability is explained by the terms in the final model.

Paired Epididymides

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.427 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.031; is significant so the effect of body weight on paired epididymides weight is different across laboratories. This model explains 69.8% (R-Squared = 0.698) of the variability around the paired epididymides weight.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.326 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.272 NS

C. Model minus triple and double interaction terms

Double interaction "(BW)*lab" term: p=0.066 NS

D. Model minus triple and double interaction terms

"Lab" term: p=0.000, is significant, so there is a significant lab effect, and 61.9% (R-Squared = 0.619) of the variability is explained by the terms in the final model.

Paired Kidneys

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.383 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.573 NS

C. Model minus triple and double interaction terms

"DoseGroup*lab" term: p=0.572 NS

D. Model minus triple, double and "DoseGroup*lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 73.7% (R-Squared = 0.737) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.500 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.650 NS

C. Model minus triple and double interaction terms

Double interaction "(BW)*lab" term: p=0.401 NS

D. Model minus triple and double interaction terms

"Lab" term: p=0.000, is significant, so there is a significant lab effect, and 73.3% (R-Squared = 0.733) of the variability is explained by the terms in the final model.

LABC

I. Using "log10 (BW)"

A. Full model

Triple interaction term: p=0.135 NS

B. Model minus triple interaction "log10 (BW) * DoseGroup * labs" term

Double interaction "log10 (BW)*lab" term: p=0.801 NS

C. Model minus triple and double interaction terms

Double interaction "DoseGroup*lab" term: p=0.326 NS

D. Model minus triple and both double interaction terms

"Lab" term: p=0.000, is significant, so there is a significant lab effect, and 86.7% (R-Squared = 0.867) of the variability is explained by the terms in the final model.

II. Using "BW"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.130 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.326 NS

C. Model minus triple and double interaction terms

Double interactions "(BW)*lab" term: p=0.990 NS

D. Model minus triple and both double interaction terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 86.6% (R-Squared = 0.866) of the variability is explained by the terms in the final model.

Liver

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.286 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.172 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.226 NS

D. Model minus triple, double and "DoseGroup*Lab" terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 61.7% (R-Squared = 0.617) of the variability of liver weight is explained by the terms in the final model.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.233 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.251 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.269 NS

D. Model minus triple and both double interaction terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 61.8% (R-Squared = 0.618) of the variability is explained by the terms in the final model.

Paired Seminal Vesicles

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.391 NS

B. Model minus triple interaction term

Double interaction term "log10 (BW) * Lab": p=0.252 NS

C. Model minus triple and double interactions

"DoseGroup*lab" term: p=0.000, is significant, so there is a lab effect (i.e., the effect of dose groups on seminal vesicle weight is different across labs). The model explains 94.7% (R-Squared = 0.947) of the variability of seminal vesicle weight.

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.406 NS

B. Model minus triple interaction term

Double interaction term "DoseGroup * Lab": p=0.000, is significant, so there is a significant lab effect. This model explains 94.9% (R-Squared = 0.949) of the variability around seminal vesicle weight.

Paired Testes

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.360 NS

B. Model minus triple interaction term

Double interaction "log10 (BW) * lab" term: p=0.971 NS

C. Model minus triple and double interaction terms

Double interaction "DoseGroup*lab" term: p=0.789 NS

D. Model minus triple and both double interaction terms

"Lab" term: p=0.189 NS

E. Model consisting of "log10(BW)" + "DoseGroup" + "log10(BW)*DoseGroup"

Interaction term: p=0.435 NS

F. Model consisting of "log10(BW)" + "DoseGroup"

"DoseGroup" term: p=0.000

"log10(BW)" term: p=0.014

Both terms are statistically significant, so both are significant factors in the variability of the paired testes weight. This final model explains 68.2% (R-Squared=0.682) of the variability. There is no significant laboratory effect (although the source of 31.8% of the variability remains unexplained).

II. Using "(BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.411 NS

B. Model minus triple interaction term

Double interaction "DoseGroup*lab" term: p=0.808 NS

C. Model minus triple and double interaction terms

Double interaction "(BW)*lab" term: p=0.934 NS

D. Model minus triple and both double interaction terms

"Lab" term: p=0.188 NS

E. Model consisting of "(BW)" + "DoseGroup" + "(BW)*DoseGroup"

Interaction term: p=0.425 NS

F. Model consisting of "(BW)" + "DoseGroup"

"DoseGroup" term: p=0.000

"(BW)" term: p=0.016

Both terms are significant, so both are significant factors in the variability of the paired testes weight. This final model explains 68.1% (R-Squared=0.681) of the variability. There is no significant laboratory effect (although the source of 31.9% of the variability remains unexplained).

Ventral Prostate

I. Using "log10 (BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.124 NS

B. Model minus triple interaction term

Double interaction "log10 (BW)*lab" term: p=0.303 NS

C. Model minus triple and double interaction terms

"DoseGroup*Lab" = p=0.246 NS

D. Model minus triple and double interaction terms:

"Lab" term: p=0.000, is significant, so there is a lab effect, and 52.6% (R-Squared = 0.526) of the variability of VP weight.

II. Using ("BW)"

A. Full model

Triple interaction "log10 (BW) * DoseGroup * labs" term: p=0.134 NS

B. Model minus triple interaction term

Double interaction "DoseGroup * lab" term: p=0.156 NS

C. Model minus triple and double interaction terms

"BW*Lab" term: p=0.710 NS

D. Model minus triple, double and "BW*Lab" terms

"Lab" term: p=0.000, is significant, so there is a lab effect, and 52.5% (R-Squared = 0.525) of the variability is explained by the terms in the final model.

LIN: Conclusions

For all of the organ weights (except paired testes) in animals exposed to LIN (a weak anti-androgen), there is a significant laboratory effect; with the majority (predominantly the vast majority) of the variability explained by the terms of the final model. For the paired testes weight, both dose group and body weight ("log10(BW)" or "(BW)") were significant causes of the variability, explaining 68.1% ("BW"), 68.2% ("log10(BW)") of the variability of paired testes weights. There is no difference in the analyses using "log10(BW)" or "(BW)".

5.5 Chapter Conclusions

5.5.1 TREN

Effects of TREN, a synthetic androgen agonist, were detected in the intact stimulated weanling version by all three laboratories as increases in androgen-dependent ASO weights and decreased paired testes weights. At the highest dose level tested, Lab A and Lab B detected statistically significant increases in all five ASO; Lab C, in three of the five. For all three laboratories, the response of all ASO to TREN at 40 mg/kg/day was less (lower increases) than the response of all organs (greater increases) at 1.0 mg/kg/day TP (the reference androgen stimulatory positive control) and less (lower decreases) in the paired testes weights in TREN-dosed groups than in the TP positive control group (greater decreases). Optional systemic organ weights and BWs were unaffected at all doses in all three laboratories.

5.5.2 LIN

Effects of LIN, a weak androgen antagonist (anti-androgen), were detected in the intact stimulated weanling version by all three laboratories in three of the six organs at 100 mg/kg/day LIN (+ TP), the highest LIN dose employed. At this dose, all three performing laboratories detected significant weight reductions (versus the TP only positive control group) for LABC, two laboratories detected significant weight reductions for VP and SVCG, and one laboratory detected significant weight reductions for COW and for paired epididymides. There were no effects on BW detected in any of the laboratories for any of the dose groups. For the positive inhibitory group (TP + FLU), all three laboratories detected significant reductions in weights of the SVCG, LABC, COW and paired epididymides, and significant increases in paired testes weights. Two laboratories also detected significantly reduced VP weights in this group. There were no effects on weights of the optional systemic organs at any LIN dose, at 1.0 mg/kg/day TP (positive stimulatory control), or at TP + FLU (positive inhibitory control).

5.5.3 DDE

Effects of DDE, a weak anti-androgen (androgen antagonist) were detected in the intact weanling male version at 50 mg/kg/day DDE (+ TP) only for SVCG weight and only in two of the three laboratories. At 160 mg/kg/day DDE (+ TP), the highest DDE dose, all three laboratories detected significant weight reductions in five of the six organs, including the VP, SVCG, LABC, and COW. Two laboratories detected significantly reduced weights of the paired epididymides, and one laboratory detected significantly reduced weights of the paired testes at this DDE dose (which may, in fact, be due to testicular toxicity, per se). For the positive inhibitory control group (TP + FLU), all three laboratories detected significant reductions in the weights of all the androgen-dependent organs evaluated. BW and the weights of the optional systemic organs were unaffected at any DDE dose or in the TP + FLU group.

5.6 Overall Conclusions From the Interlaboratory Analyses

The intact stimulatory weanling version of the Hershberger bioassay did not appear to be able to consistently detect effects on androgen-dependent organ weights from weak anti-androgens at the doses tested. There was considerable variability in detection of effects on the androgen-dependent organs across the participating laboratories.

The statistical model, developed by Dr. Breda Munoz of RTI International, is able to isolate laboratory effects from BW and/or dose group effects or to identify their interactions (i.e., “lab*body weight”, or “lab*DoseGroup”, or “bodyweight * DoseGroup * lab”). For all of the organs of interest (paired adrenal glands, COW, paired epididymides, paired kidneys, LABC, liver, seminal vesicles, paired testes, and VP) in animals exposed to DDE, there is a laboratory effect (i.e., the effect of dose group and/or BW on organ weight is different among the performing laboratories).

For all of the organs of interest, except for the paired testes, in animals exposed to TREN or LIN, there also is a laboratory effect (i.e., the effect of dose group and/or BW on organ weight is different among the performing laboratories). For the paired testes weights in animals exposed to TREN or to LIN, both dose group and body weight (“log10(BW)” or “BW”) were significant causes of the variability, and the model with these terms explained 68.1-68.2% of the variability of paired testes weights for both chemicals. There was no identified, significant effect of laboratory on paired testes weights for these two chemicals, but the source of the remaining variability (31.8-31.9%) remains unexplained. See Table 5.11 for an overall summary.

The explanation as to why only the paired testes weights from animals exposed to TREN or LIN were unaffected by laboratory is not known, but it is highly likely that it is due to:

- The paired testes are large organs relative to the other androgen-sensitive organs (the liver is also large, but it is a systemic organ).
- The paired testes are very sensitive to exogenous TP or FLU (strong androgen and anti-androgen) and also to the weaker androgen (TREN) and anti-androgens (LIN and DDE), so changes in weights of these relatively large organs is likely easier to detect than changes in very small organs across all laboratories.

The explanation as to why the paired testes weights from animals exposed to DDE were affected by laboratory is also not known, but it is highly likely that it is due to:

- For the statistical evaluation of paired testes weights in animals exposed to DDE, the triple interaction term "lab * log10 (BW)" or "(BW) * DoseGroup" was statistically significant.
- The interpretation is that the combined effects of both dose group and body weight ("log10(BW)" or "(BW)") on paired testes weight are different across performing laboratories.
- So, the model did detect the effects of dose group and BW on testes weights for DDE (as it did for TREN and LIN), but it also detected the interactive effect of laboratory (i.e., the triple interaction term was statistically significant for DDE but not statistically significant for TREN or LIN for paired testes weight).
- The nature of the paired testes, relatively large and sensitive to exogenous androgens and anti-androgens, likely enabled the laboratories to detect differences in organ weights and to enable the model to detect the interaction of all three terms (BW, dose group, and laboratory).

Table 5.11. Summary of Interlaboratory Statistical Evaluations

Organ:	Paired Adrenals	COW	Paired Epidid.	Paired Kidneys	Liver	Seminal Vesicles	LABC	Paired Testes	VP
TREN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
LIN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
DDE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Yes = statistically significant interlaboratory effects on the weight(s) of specific organ(s) with TREN, LIN, or DDE.

CHAPTER 6.0 STAGE 3 STUDIES WITH CODED CHEMICALS

6.1 Pilot Study

Prior to the start of Phase 3, the testing of coded samples (under the OECD VMG mammalian validation of the intact simulated weanling version of the Hershberger bioassay), Lab E conducted a pilot study with simulated weanling males using known test chemicals (Schladt, 2008; Freyberger et al., 2007). The weanling male Wistar rats were of the HsdCpb:WU strain (from Harlan Winkelmann GmbH, Borschen, Germany), with six animals/group and five groups:

(1) Vehicle (corn oil); oral gavage

(2) Corn oil + TP (1.0 mg/kg/day; SC injection), designated Combination 1

(3) FLU (30 mg/kg/day by oral gavage) + TP, designated Combination 2

(4) DDE (160 mg/kg/day by oral gavage] + TP, designated Combination 3

(5) LIN (100 mg/kg/day by oral gavage) + TP, designated Combination 4

TP was injected subcutaneously at a dosing volume of 0.5 ml/kg. All the other chemicals were administered by oral gavage at 5 ml/kg; corn oil was the vehicle alone or the vehicle used for test chemicals for all of the groups. The male weanlings were 22 days of age at first dose. They were housed three/cage from the same dose group, were dosed for ten consecutive days, and terminated approximately 24 hours after the tenth (last) dose. The in-life portion of the study was conducted from June 16-30, 2006. In-life and necropsy observations and measurements were as presented in Chapter 3.0 (the protocol) and as previously performed in the studies described in Chapters 4.0, 5.0, and subsequently in the rest of Chapter 6.0. The organs weighed were as described previously: VP, SVCG, LABC, COW, paired testes, paired epididymides, and liver; reported as absolute and relative to terminal BW (mg/100 g BW).

Results are presented in Tables 6.1 and 6.2 (derived from the final report tables). Table 6.1 presents the absolute and relative (organ weight relative to terminal BW as mg/100 g BW) weights of the organs for the vehicle and the vehicle + TP (designated Combination 1) groups. No statistically significant difference was found among groups for mean absolute BWs at the 5% significance level. Mean paired testes weight (absolute and relative) exhibited a statistically significant reduction across groups ($p<0.01$). All other androgen-sensitive organs exhibited statistically significant increases in mean weight (paired epididymides, VP, SVCG, LABC, and COW) across groups ($p<0.05$). No significant difference was detected for the mean liver weight across groups.

Table 6.2 presents the terminal body and organ weights for Combination 1 (vehicle [corn oil] + TP), Combination 2 (FLU + TP), Combination 3 (DDE + TP), and Combination 4 (LIN + TP) groups. The FLU+TP and DDE + TP combination groups exhibited statistically significant decreases in the mean weights of EPID, VP, SVCG, LABC, and COW compared to those from the vehicle + TP (corn oil + TP) combination group ($p<0.01$). The LIN + TP combination group exhibited statistically significant reductions of the mean absolute weights of TESTES ($p<0.01$), SVCG ($p<0.01$), LABC ($p<0.01$), and COW ($p<0.05$) compared to corn oil + TP combination group. The LIN+TP combination group also exhibited a statistically significant reduction in the EPID mean relative weight compared to the vehicle + TP mean value. The mean paired testes relative and absolute weights for the FLU + TP combination group, and the mean paired testes relative weight for LIN + TP and DDE+TP combinations, were not statistically significantly different from the corresponding means for vehicle + TP combination group.

The author concluded that "the study showed that androgen-sensitive tissues of juvenile intact male rats responded to TP, and that the anti-androgenic potential of FLU, DDE, and LIN could be detected in T-stimulated intact juvenile male rats... The results of the study suggest that the juvenile intact male may

have the potential to replace the peripubertal castrated rat in the Hershberger bioassay, thereby contributing to a refinement of the assay in terms of animal welfare" (Schladt, 2008; p. 9).

6.2 Final Validation Tests

For this last Stage (or phase) 3 of the validation program for the intact stimulated weanling version of the Hershberger bioassay, all six testing laboratories participated; their parameters and specific conditions are presented in Table 6.3. These six laboratories included Labs A, B, and C (who participated in Stages 1A, 1B, and 2), and Labs D, E and F (all industrial entities). The lead laboratory did not perform any of these validation studies.

External sources of variability included differences across the six testing laboratories regarding the rat strain and source, the diet and bedding, as well as minor differences in the age of the males at study start (22 or 23 days of age) and at necropsy (32 or 33 days of age). Another source of variability was the identity and source (suppliers) of the vehicle. The vehicle was corn oil in all but one laboratory (Lab C used peanut oil), and the corn oil was tocophenol-stripped in two of the five laboratories using corn oil (Labs A and D). There were also differences in the staff and/or techniques (and in expertise and experience) across the six laboratories.

Although all six laboratories participated, specific laboratories did or did not evaluate specific groups as follows:

- Labs A, E, and F did not evaluate the vehicle control group (it was optional)
- The two sets of data from the Lab B for Stage 3 coded samples were from the same laboratory facility. The laboratory used a block design so that the necropsy could be spread over two days (information provided by Dr. J. William Owens, The Procter and Gamble Company, Secondment to OECD during the validation program). The first block evaluated 2,4-dinitrophenol (DNP) at 10 mg/kg/day, TP and FLU, but not NP. The second block (one day later), evaluated NP at 160 mg/kg/day and TP, but not DNP or FLU.

For the six chemicals employed in Phase 3 of the validation effort, their codes, identities, CAS numbers, routes of administration, and dose(s) employed are listed in Table 6.4 (information provided by Dr. J.W. Owens, Secondment to OECD). The single TP (1.0 mg/kg/d) and FLU (3.0 mg/kg/d) doses were determined in Stages 1A and 1B, respectively (Chapter 4.0). Two doses each of DDE and LIN were selected from the studies in Stage 2 (Chapter 5.0), and the single doses for DNP and NP were selected from the literature.

Table 6.1. Body (g) and Organ (mg) Weights (Mean ± SD) at Terminal Sacrifice for Vehicle and Vehicle Plus TP Groups

	BW	Liver	Testes	EPID	VP	SVCG	LABC	COWS
Vehicle Control								
Absolute	94 ± 15.3	4781 ± 856.3	761 ± 182.7	92 ± 6.7	29 ± 8.6	30 ± 7.6	94.9 ± 16.87	3.8 ± 1.02
Relative[a]	--	5089 ± 131.1	803 ± 87.0	100 ± 15.7	31 ± 6.5	32 ± 4.2	101 ± 5.08	4.0 ± 0.67
Vehicle + TP								
Absolute	104 ± 6.1	5444 ± 332.2	423 ± 51.6**	145 ± 13.7**	92 ± 6.0**	200 ± 24.4**	231 ± 9.87**	17.4 ± 2.56**
Relative[b]	--	5229 ± 237.8	408 ± 60.0**	140 ± 19.8**	88 ± 6.0**	193 ± 29.9*	222.4 ± 9.19**	16.7 ± 1.96**

[a] Relative weight = organ weight relative to terminal BW as mg/100 g BW

†, †† = $p < 0.05$; $p \leq 0.01$ statistically significant difference between organ weights (absolute and/or relative) and in the vehicle group versus those in the vehicle + TP group

* Change in mean with respect to vehicle control mean is significant at the $p < 0.05$ level

** Change in mean with respect to vehicle control mean is significant at the $p < 0.01$ level

Table 6.2. Body (g) and Organ (mg) Weights (Mean ± SD) at Terminal Sacrifice for the Combination Dose Groups

	BW	Liver	Testes	EPID	VP	SVCG	LABC	COWS
Combination 1: Corn Oil + TP								
Absolute	104 ± 6.1	5444 ± 332.2	423 ± 51.6	145 ± 13.7	92 ± 6.0	200 ± 24.4	231 ± 98.7	17.4 ± 2.56
Relative[a]	--	5229 ± 237.8	408 ± 60.0	140 ± 19.8	88 ± 6.0	193 ± 29.9	222.4 ± 9.19	16.7 ± 1.96
Combination 2: FLU + TP								
Absolute	99 ± 9.6	5176 ± 359.0	502 ± 118.7	82 ± 9.2**	46 ± 12.1**	60 ± 8.4**	144 ± 10.56**	8.1 ± 1.64**
Relative[a]	--	5230 ± 342.4	514 ± 142.7	83 ± 8.1**	46 ± 12.2**	60 ± 5.9**	145 ± 11.82**	8.1 ± 1.45**
Combination 3: DDE + TP								
Absolute	94 ± 6.6	7371 ± 799.6**	286 ± 109.0*	76 ± 8.1**	42 ± 3.9**	56 ± 19.9**	126.7 ± 23.00**	6.2 ± 1.65**
Relative[a]	--	7826 ± 367.2**	305 ± 117.3	81 ± 6.0**	44 ± 3.5**	59 ± 18.0**	134.9 ± 23.11**	6.5 ± 1.39**
Combination 4: LIN + TP								
Absolute	99 ± 6.8	5393 ± 396.9	344 ± 62.6**	113 ± 5.9	73 ± 22.9	140 ± 15.1**	193.2 ± 23.72**	13.9 ± 2.10*
Relative[a]	--	5450 ± 241.9	347 ± 60.4	115 ± 10.6*	80 ± 20.2	142 ± 19.2	195.8 ± 26.70	14.0 ± 1.49

Note: Sample size is 6 subjects per group

[a] Relative weight = organ weight relative to terminal BW as mg/100 g BW

Mean organ weight values are significantly different from the mean organ weight values in the positive control group (corn oil + TP) at * = $p<0.05$ or ** = $p<0.01$

Table 6.3. Laboratory Parameters and Conditions for the Phase 3 Validation[a]

Lab	Rat Strain & Supplier		Age on Study/ Necropsy	Vehicle	Diet	Animals per Cage	Bedding
A	Alpk:APfSD	Breeding facility on site	22-23 days/ 32-33 days	Tocopherol-stripped corn oil	R&M1 Special Diet Services Witham, Essex, UK	3	Sawdust. paper
B	Sprague Dawley	Breeding facility on site	22 days/ 32 days	corn oil	PMI 5001, Agribrands Purina, Korea	3	Willow shavings
C	Crj:CD IGS (SD)	Charles River Canada, St Constant, Quebec, Canada	22 days/ 32 days	Peanut oil	PMI 5001, Agribrand Purina	2	Pro chip Maple (dust-free hardwood chips), PWI Industries, St-Hyacinthe, Quebec, Canada
D	Sprague-Dawley Crl:CD (SD)	Charles River, St Germain-sur-l'Abresle, France	23 days/ 33 days	Tocopherol stripped corn oil	Scientific Animal Food and Engineering, Augy, France	6	Wood chips (fir tree origin)
E	Wistar rats; Crl:WI(Han)	Charles River, Sulzfeld, Germany	23 days/ 33 days	Corn oil, Sigma C-8267	Provimi Kliba SA, Kaiseraugst, Switzerland	3	SNIFF type 3/4 fibres, dust-free bedding
F	HsdCpb:WU	Harlan Winkelmann Borchen, Germany	22-23 days/ 32-33/days	Corn oil Caesar & Loretz	Provimi Kliba S.A., Kaiseraugst, Switzerland	3	SNIFF low-dust wood granulate Soest, Germany

[a] These conditions apply to both the agonist and antagonist studies.

Table 6.4. Phase 3 Coded Samples for Intact Stimulated Weanling Evaluation

Chemical:	**DDE**		**LIN**		**DNP**	**NP**	**FLU**	**TP**
Dose:[a]	**16**	**160**	**10**	**100**	**10**	**160**	**3**	**1.0**
Laboratory	**Code**	**Code**	**Code**	**Code**	**Code**	**Code**	**Code**	**Code**
A	2048	6833	3156	9199	3886	5915	6066	Not Coded
E	2799	6974	3698	9328	4854	5241	5971	Not Coded
D	1892	7351	3577	9771	4415	5423	6347	Not Coded
C	1491	8213	3576	9846	5035	5747	6280	Not Coded
B	2638	8878	3628	9929	4366	5856	6583	Not Coded
F	1654	9007	3713	9946	4467	5277	6704	Not Coded

[a] Doses are in mg/kg BW/day; all chemicals except TP are administered by oral gavage in stripped corn oil at 5 ml/kg; TP is administered by subcutaneous injection in stripped corn oil at 1 ml/kg

6.2.1 Evaluation of the Coded Negative Chemicals (DNP and NP) and the Coded Positive Control Chemicals (TP and FLU)

Table 6.5 presents the mean weights (± SD) of the androgen-dependent organs and the results of Dunnett's multiple comparisons test between mean organ weights for the different chemicals and the corresponding means of the TP-only positive control group.

As anticipated, there were no statistically or biologically significant effects observed within each participating laboratory between any organ mean weight at 10 mg/kg/day DNP (+ 1.0 mg/kg/day TP) and the corresponding organ mean weight values at 1.0 mg/kg/day TP. For NP at 160 mg/kg/day (+ TP), there were no statistically significant effects observed for VP, SVCG, COW, or paired epididymides in any laboratory. For LABC, only Lab E detected a reduction in weight in NP (+ TP) at 160 mg/kg/day at $p \leq 0.08$ and $p \geq 0.05$ (termed "a marginally insignificant change") when the Dunnett's Test with ANCOVA (analysis of covariance) used the starting (but not terminal) BW as the covariate. Lab F detected a statistically significant reduction in paired mean testes weight in NP (+ TP) at 160 mg/kg/day ($p \leq 0.05$, ANCOVA Dunnett's using the terminal weight as the covariate). For the inhibitory positive control group (TP + FLU), all participating laboratories detected statistically significant reduction in mean weights of organs VP, SVCG, LABC, COW, and paired epididymides. For the paired testes, three of the six laboratories detected significant reductions in weight. No laboratories detected any statistically significant changes in BWs between the DNP or NP groups and the TP control group.

Table 6.6 presents the weights of the optional systemic organs exposed to DNP + TP, NP + TP, TP alone, or TP + FLU. Within each laboratory, no statistically significant changes in mean weight of the liver, paired adrenal glands, or paired kidneys were detected between the chemical groups and the vehicle group values.

Table 6.7 presents the CVs for all of the organ weights in all of the test groups. Three parameters exhibited CVs above 40%: COW vehicle only by Lab B and COW FLU + TP by Lab E, and testes at FLU + TP by Lab D. Five parameters exhibited CVs at 30-40%: SVCG by Lab F at TP positive control, COW by Lab B at FLU + TP, by Lab F at TP positive control and at FLU + TP, and paired testes by Lab F at DNP, 10 mg/kg/day + TP. Thirty-two parameters exhibited CVs at 20-30%: seven laboratories for VP, five laboratories for SVCG, three laboratories for LABC, six laboratories for COW, eight laboratories for paired testes, and no laboratories for paired epididymides. Therefore, there was a total of 40 CVs greater than 20%.

Table 6.5. Body and Target Organ Weights in Phase 3 With Negative and Reference Positive Test Chemicals

Parameter	Laboratory	Vehicle only	TP[a] 1 mg/kg/d	DNP 10 mg/kg/d + TP	NP 160 mg/kg/d + TP	FLU 3 mg/kg/d + TP
BW (g)	Lab A	Not done	85.9 ± 6.8	86.8 ± 5.65	82.2 ± 11.24	84.0 ± 7.56
	Lab B	129.7 ± 9.6	140.8 ± 6.53	137.4 ± 11.23 [b]	Not done	139.0 ± 9.46
		131.0 ± 11.0	150.2 ± 6.79	Not done	134.0 ± 5.18	Not done
	Lab C	Not done	134.9 ± 9.7	132.4 ± 4.25	129.3 ± 6.45	130.9 ± 10.10
	Lab D	105.2 ± 7.8	103.9 ± 8.43	109.7 ± 10.59	97.9 ± 3.94	111.0 ± 11.33 [c]
	Lab E	Not done	132.2 ± 17.9	128.5 ± 6.98	113.7 ± 12.34	130.5 ± 9.14
	Lab F	Not done	120.0 ± 9.6	118.4 ± 7.07	111.4 ± 7.90	116.4 ± 4.39
VP (mg)	Lab A	Not done	74.9 ± 12.5	82.4 ± 13.9	69.7 ± 13.9	47.7 ± 10.9*
	Lab B	79.2 ± 9.5	141.3 ± 16.5	121.8 ± 12.0	Not done	88.2 ± 16.4*
		71.5 ± 12.1	135.2 ± 11.9	Not done	127.7 ± 22.0	Not done
	Lab C	Not done	132.2 ± 26.9	118.4 ± 12.5	125.0 ± 14.2	94.3 ± 10.1*
	Lab D	56.3 ± 14.1	117.8 ± 24.5	123.5 ± 28.6	107.2 ± 14.44	74.0 ± 22.0*
	Lab E	Not done	114.8 ± 15.2	103.6 ± 14.9	101.6 ± 7.3	59.8 ± 11.9*
	Lab F	Not done	103.2 ± 11.9	98.2 ± 9.74	86.8 ± 14.5	55.5 ± 14.45*
SVCG (mg)	Lab A	Not done	250.4 ± 57.0	210.4 ± 22.9	234.0 ± 33.4	107.6 ± 24.7*
	Lab B	47.5 ± 5.1	354.2 ± 46.5	377.4 ± 33.5	Not done	170.2 ± 22.1*
		68.8 ± 13.7	406.2 ± 82.0	Not done	349.4 ± 42.9	Not done
	Lab C	Not done	219.7 ± 24.1	219.1 ± 28.9	210.52 ± 20.37	84.5 ± 15.4*
	Lab D	81.5 ± 12.1	290.7 ± 57.8	311.0 ± 60.9	284.5 ± 53.71	126.2 ± 16.4*
	Lab E	Not done	308.9 ± 58.0	297.4 ± 17.1	250.68 ± 27.39	109.8 ± 18.8*
	Lab F	Not done	141.8 ± 45.5	129.0 ± 27.46	120.30 ± 18.09	35.4 ± 7.71*
LABC (mg)	Lab A	Not done	144.9 ± 33.7	144.3 ± 17.3	156.5 ± 12.9	96.0 ± 15.6*
	Lab B	119.4 ± 14.4	303.57 ± 34.84	285.3 ± 41.2	Not done	201.3 ± 11.8*
		123.4 ± 22.0	296.62 ± 30.44	Not done	262.3 ± 33.5	Not done
	Lab C	Not done	260.3 ± 25.6	245.0 ± 9.3	249.5 ± 14.1	161.9 ± 11.6*
	Lab D	76.2 ± 16.9	165.83 ± 18.18	186.7 ± 25.3	165.0 ± 31.27	116.6 ± 21.2*
	Lab E	Not done	284.9 ± 41.0	254.92 ± 24.73	240.5 ± 21.9††	182.7 ± 26.5*
	Lab F	Not done	165.2 ± 15.0	162.4 ± 29.40	171.0 ± 42.0	102.8 ± 13.26*

Table 6.5 (continued)

Parameter	Laboratory	Vehicle only	TP[a] 1 mg/kg/d	DNP 10 mg/kg/d + TP	NP 160 mg/kg/d + TP	FLU 3 mg/kg/d + TP
COW (mg)	Lab A	Not done	16.2 ± 3.2	15.7 ± 1.5	19.3 ± 3.1	8.4 ± 1.8*
	Lab B	7.3 ± 3.3	26.6 ± 2.7	27.3 ± 4.0	Not done	15.6 ± 6.0*
		6.6 ± 1.3	26.5 ± 4.0	Not done	24.8 ± 3.1	Not done
	Lab C	Not done	25.8 ± 5.1	25.3 ± 4.0	26.0 ± 4.6	14.7 ± 3.6*
	Lab D	7.0 ± 1.4	20.3 ± 3.4	22.0 ± 1.1	19.0 ± 2.61	10.0 ± 2.6*
	Lab E	Not done	17.0 ± 4.9	15.4 ± 2.7	14.1 ± 1.1	7.9 ± 3.6*
	Lab F	Not done	15.3 ± 5.2	13.9 ± 3.35	11.7 ± 1.2	5.2 ± 2.07*
Testes (mg)	Lab A	Not done	228.2 ± 59.4	227.1 ± 60.9	223.4 ± 41.0	213.2 ± 75.3
	Lab B	1114.0 ± 78.9	534.3 ± 57.4	515.2 ± 48.3	Not done	808.1 ± 125.7*
		1106.5 ± 162.8	489.5 ± 89.0	Not done	455.7 ± 65.2	Not done
	Lab C	Not done	512.0 ± 22.8	511.4 ± 49.0	521.3 ± 39.5	839.4 ± 83.6*
	Lab D	708.5 ± 196.7	394.8 ± 90.4	375.3 ± 78.6	395.0 ± 88.70	552.6 ± 225.7
	Lab E	Not done	655.6 ± 139.2	642.5 ± 134.7	583.2 ± 77.5	928.1 ± 159.1*
	Lab F	Not done	882.8 ± 235.7	703.4 ± 237.30	450.0 ± 89.8*	1050.3 ± 262.16
Epididymides (mg)	Lab A	Not done	128.6 ± 11.3	132.6 ± 14.1	124.4 ± 19.3	76.4 ± 10.0*
	Lab B	137.1 ± 17.3	165.0 ± 11.6	178.8 ± 16.0	Not done	131.4 ± 18.5*
		135.3 ± 14.6	168.6 ± 9.8	Not done	166.3 ± 18.0	Not done
	Lab C	Not done	168.0 ± 16.2	164.1 ± 19.1	175.5 ± 22.0	138.8 ± 13.5*
	Lab D	96.2 ± 16.5	145.0 ± 27.7	139.3 ± 16.6	147.5 ± 16.93	91.2 ± 15.4*
	Lab E	Not done	174.5 ± 19.5	165.3 ± 13.7	164.2 ± 7.3	118.9 ± 20.5*
	Lab F	Not done	145.3 ± 16.2	132.7 ± 16.75	123.6 ± 12.5	96.9 ± 15.71*

[a] TP (1 mg/kg/d) is the stimulatory reference positive control. It is administered alone in the positive stimulatory control group, along with DNP and NP, and with FLU as the positive inhibitory control. FLU (3 mg/kg/d) is the inhibitory reference positive control.

[b] One animal in the 10 mg/kg DNP group had no left testes or left epididymis; n=5 for those tissues.

[c] One animal in the specified treatment group died; n=5.

* Statistically significant by Dunnett's at $p \leq 0.05$ using ANCOVA with terminal BW.

†† "Marginally insignificant change" by Dunnett's using ANCOVA with starting BW ($p < 0.08$ and ≥ 0.05).

Table 6.6. Optional Organ Weights in Phase 3 With Negative and Reference Positive Test Chemicals

Parameter	Laboratory	Vehicle only	TP[a] 1 mg/kg/d	DNP 10 mg/kg/d + TP	NP 160 mg/kg/d + TP	FLU 3 mg/kg/d + TP
Liver (g)	Lab A	Not done	4.0 ± 0.3	3.8 ± 0.4	4.3 ± 0.6	3.8 ± 0.4
	Lab B	5.9 ± 0.6	6.6 ± 0.4	6.2 ± 0.8	Not done	6.9 ± 0.5
		6.0 ± 0.7	7.1 ± 0.3	Not done	7.6 ± 0.9	Not done
	Lab C	Not done	7.4 ± 0.6	7.0 ± 0.2	7.9 ± 0.5	6.9 ± 0.4
	Lab D	4.8 ± 0.7	5.3 ± 0.2	5.0 ± 0.7	5.9 ± 0.53	5.0 ± 0.4
	Lab E	Not done	6.6 ± 1.2	6.4 ± 0.4	6.6 ± 1.4	6.8 ± 0.5
	Lab F	Not done	3.9 ± 0.3	3.8 ± 0.18	4.1 ± 0.6	4.0 ± 0.58
Adrenals (mg)	Lab A	Not done	19.1 ± 5.2	18.4 ± 6.0	16.6 ± 3.3	16.5 ± 6.3
	Lab B	21.5 ± 2.9	23.9 ± 3.0	21.7 ± 2.7	Not done	24.0 ± 2.2
		24.1 ± 1.0	25.6 ± 2.1	Not done	23.4 ± 3.0	Not done
	Lab C	Not done	27.1 ± 2.1	24.0 ± 3.1	27.4 ± 3.8	25.6 ± 2.2
	Lab D	21.7 ± 2.5	21.7 ± 4.9	20.5 ± 4.5	22.7 ± 4.27	23.6 ± 3.0
	Lab E	Not done	Not done	Not done	Not done	Not done
	Lab F	Not done	30.4 ± 2.8	29.2 ± 3.70	29.4 ± 4.0	30.7 ± 3.61
Kidneys (mg)	Lab A	Not done	797.9 ± 82.4	746.3 ± 26.8	775.2 ± 75.2	724.4 ± 30.8
	Lab B	1276.9 ± 160.5	1434.1 ± 105.3	1348.1 ± 192.8	Not done	1361.5 ± 84.4
		1315.1 ± 141.4	1468.2 ± 147.7	Not done	1387.0 ± 129.7	Not done
	Lab C	Not done	1578.3 ± 99.1	1525.4 ± 85.3	1613.8 ± 127.1	1531.2 ± 71.0
	Lab D	1048.8 ± 72.3	1029.8 ± 82.2	1104.2 ± 139.4	1206.7 ± 173.75	1089.4 ± 81.5
	Lab E	Not done	Not done	Not done	Not done	Not done
	Lab F	Not done	1133.0 ± 73.4	1118.6 ± 70.09	1254.8 ± 104.9	1118.6 ± 17.31

[a] TP (1 mg/kg/d) is the stimulatory reference positive control. It is administered alone in the positive stimulatory control group, along with DNP and NP, and with FLU as the positive inhibitory control. FLU (3 mg/kg/d) is the inhibitory reference positive control.

Table 6.7. CVs for Body and Target Organ Weights in Phase 3 With Negative and Reference Positive Test Chemicals

Parameter	Laboratory	Vehicle only	TP[a] 1 mg/kg/d	DNP 10 mg/kg/d + TP	NP 160 mg/kg/d + TP	FLU 3 mg/kg/d + TP
BW (g)	Lab A	Not done	7.90%	6.51%	13.68%	9.00%
	Lab B	7.39%	4.64%	8.17%	Not done	6.80%
		8.37%	4.52%	Not done	3.87%	Not done
	Lab C	Not done	7.18%	3.21%	4.99%	7.72%
	Lab D	7.41%	8.12%	9.65%	4.02%	10.20%
	Lab E	Not done	13.52%	5.43%	10.86%	7.00%
	Lab F	Not done	8.03%	5.98%	7.10%	3.77%
VP (mg)	Lab A	Not done	16.74%	16.89%	19.87%	22.93%
	Lab B	11.94%	11.69%	9.86%	Not done	18.57%
		16.90%	8.84%	Not done	17.25%	Not done
	Lab C	Not done	20.36%	10.59%	11.39%	10.68%
	Lab D	24.94%	20.77%	23.15%	13.48%	29.71%
	Lab E	Not done	13.27%	14.37%	7.16%	19.92%
	Lab F	Not done	11.57%	9.92%	16.74%	26.05%
SVCG (mg)	Lab A	Not done	22.77%	10.88%	14.28%	22.94%
	Lab B	10.77%	13.12%	8.87%	Not done	12.99%
		19.93%	20.18%	Not done	12.29%	Not done
	Lab C	Not done	10.95%	13.20%	9.67%	18.25%
	Lab D	14.90%	19.88%	19.57%	18.88%	13.01%
	Lab E	Not done	18.78%	5.74%	10.93%	17.12%
	Lab F	Not done	32.10%	21.29%	15.04%	21.78%

(continued)

Table 6.7 (continued)

Parameter	Laboratory	Vehicle only	TP[a] 1 mg/kg/d	DNP 10 mg/kg/d + NP 160 mg/kg/d + TP	FLU 3 mg/kg/d + TP
LABC (mg)	Lab A	Not done	23.30%	12.00%	8.22%	16.27%
	Lab B	12.03%	11.48%	14.44%	Not done	5.85%
		17.84%	10.26%	Not done	12.76%	Not done
	Lab C	Not done	9.85%	3.79%	5.67%	7.16%
	Lab D	22.15%	10.96%	13.53%	18.95%	18.22%
	Lab E	Not done	14.40%	9.70%	9.09%	14.51%
	Lab F	Not done	9.05%	18.10%	24.57%	12.89%
COW (mg)	Lab A	Not done	20.00%	9.38%	15.99%	21.31%
	Lab B	45.00%	10.28%	14.61%	Not done	38.09%
		20.09%	15.08%	Not done	12.71%	Not done
	Lab C	Not done	19.94%	15.82%	17.55%	24.75%
	Lab D	20.20%	16.65%	4.98%	13.72%	26.46%
	Lab E	Not done	28.64%	17.85%	7.47%	45.26%
	Lab F	Not done	33.73%	24.10%	10.42%	39.77%
Testes (mg)	Lab A	Not done	26.03%	26.80%	18.36%	35.29%
	Lab B	7.08%	10.74%	9.37%	Not done	15.55%
		14.71%	18.19%	Not done	14.30%	Not done
	Lab C	Not done	4.45%	9.57%	7.59%	9.96%
	Lab D	27.76%	22.89%	20.93%	22.45%	40.85%
	Lab E	Not done	21.24%	20.97%	13.28%	17.14%
	Lab F	Not done	26.70%	33.74%	19.95%	24.96%
Epididymides (mg)	Lab A	Not done	8.82%	10.66%	15.51%	13.08%
	Lab B	12.64%	7.03%	8.95%	Not done	14.11%
		10.77%	5.81%	Not done	10.83%	Not done
	Lab C	Not done	9.63%	11.61%	12.52%	9.69%
	Lab D	17.13%	19.10%	11.95%	11.48%	16.89%
	Lab E	Not done	11.20%	8.26%	4.45%	17.27%
	Lab F	Not done	11.13%	12.62%	10.11%	16.21%

[a] TP (1 mg/kg/d) is the stimulatory reference positive control. It is administered alone in the positive stimulatory control group, along with DNP and NP, and with FLU as the positive inhibitory control. FLU (3 mg/kg/d) is the inhibitory reference positive control.

6.2.2 Evaluation of the Positive Coded (DDE and LIN) Chemical Groups and the Positive Stimulatory (TP) and Positive Inhibitory (FLU) Control Groups

Table 6.8 presents the results on the androgen-sensitive organ weights after exposure to TP only, DDE at 16 or 160 mg/kg/day + TP, and LIN at 10 or 100 mg/kg/day + TP. The vehicle control group was optional and not evaluated by Labs A, C, E, or F. The organ weight values are compared against the TP only (positive control) group organ weight values.

None of the laboratories detected statistically or biologically significant changes in BWs in any group. At 16 mg/kg/day DDE + TP, no laboratories detected significant organ weight changes for VP, LABC, or COW. Two of six laboratories detected statistically significantly reduced weights of the SVCG using ANCOVA with starting (not terminal) BWs. One laboratory detected a "marginally insignificant" change ($P < 0.08$ and ≥ 0.05) using ANCOVA with terminal (but not starting) BW at 16 mg/kg/day DDE. Only Lab F detected a significant reduction in paired testes weight, and Labs E and F detected "marginally insignificant" reduced weights of the epididymides at 16 mg/kg/day DDE. At 160 mg/kg/day DDE + TP, all laboratories detected significantly reduced weights of the VP, SVCG, LABC, and COW. Labs C and F detected significantly increased testes weights, and Lab A detected significantly reduced testes weights. The remaining three laboratories detected no change in testes weights. Five of the six laboratories detected significantly reduced weights of the epididymides; the sixth laboratory, Lab C, detected a "marginally insignificant" change in the paired epididymides.

For LIN at 10 mg/kg/day + TP, none of the six laboratories detected significant changes in the weights of any of the organs evaluated. At 100 mg/kg/day LIN + TP, one of the six laboratories detected reduced VP weight; four of the six detected significantly reduced SVCG (and one more detected a "marginally insignificant" reduction); three of six laboratories detected significantly reduced LABC, and the other three detected "marginally insignificant" reduced LABC; two of the six laboratories detected significantly reduced paired testes weights; and three of the six detected significantly reduced paired epididymides weights, with two more detecting "marginally insignificant" reductions in paired epididymides weights.

The weights of the optional systemic organs, liver, paired adrenal glands, and paired kidneys after exposure to DDE (at 16 or 160 mg/kg/day) + TP or LIN (at 10 or 100 mg/kg/day) + TP versus the TP only (1 mg/kg/day) positive stimulatory control group) values, for the participating laboratories, are presented in Table 6.9. Please note that Lab E did not evaluate the paired adrenal glands or the paired kidneys, and neither Labs C nor F evaluated the vehicle control group (it was optional). There were no statistically or biologically significant differences in any of the participating laboratories for any of the optional organ weights after exposure to DDE + TP or LIN + TP versus the TP only group values.

The CVs of the BW and all of the androgen-dependent organ weights are presented in Table 6.10. Only one laboratory exhibited a CV greater than 50% (Lab C for COW at 16 mg/kg/day DDE). Six parameters exhibited CVs at 40-50%: Lab F for VP at 16 mg/kg/day DDE, Lab D for SCVG at 160 mg/kg/day DDE, Lab B (first block) for COW at vehicle only, Lab B (second block) at 160 mg/kg/day DDE, Lab D also for COW at 10 mg/kg/day DDE, and Lab F for paired testes at 10 mg/kg/day LIN. Nine parameters exhibited CVs at 30-40%: Lab F for VP at 16 mg/kg/day DDE, Lab E for VP at 160 mg/kg/day DDE, Lab F for SVCG at TP positive control and D for SVCG at 100 mg/kg/day LIN, Lab F for COW at TP positive control, Lab D for COW at 160 mg/kg/day DDE, Lab C for COW at 100 mg/kg/day LIN, Lab F for paired testes by Lab F at 16 mg/kg/day DDE and by Lab E at 160 mg/kg/day DDE. Thirty-nine (39) parameters exhibited CVs at 20-30%: with organ rank order: paired testes (13), COWS (11), VP (8), LABC (6), VP (5), SCVG (4), paired epididymides (1), and BW (1). For all CVs greater than 20%, Lab E exhibited 15, Lab F 14, Lab D 10, Lab C 8, Lab B (second block) 6, Lab A 3, and Lab B (first block) 2.

Table 6.8. Body and Target Organ Weights in Phase 3 With Coded Positive Test Chemicals

Parameter	Laboratory	Vehicle only	TP 1 mg/kg/d	DDE + TP 16 mg/kg/d	160 mg/kg/d	LIN + TP 10 mg/kg/d	100 mg/kg/d
BW (g)	Lab A	Not done	85.9 ± 6.8	81.2 ± 8.65	79.9 ± 7.53 [a]	86.7 ± 10.87	77.7 ± 2.91
	Lab B	129.7 ± 9.6	140.8 ± 6.53	Not done	Not done	143.2 ± 4.68	Not done
		131.0 ± 11.0	150.2 ± 6.79	143.9 ± 9.23	132.9 ± 11.88	Not done	130.0 ± 8.43
	Lab C	Not done	134.9 ± 9.7	133.3 ± 11.19	127.7 ± 7.01	132.0 ± 7.21	119.8 ± 3.14
	Lab D	105.2 ± 7.8	103.9 ± 8.43	107.8 ± 10.65	99.6 ± 7.05	109.9 ± 13.43	105.7 ± 4.36
	Lab E	Not done	132.2 ± 17.9	129.5 ± 9.20	104.7 ± 22.47	129.2 ± 5.88	121.3 ± 6.59
	Lab F	Not done	120.0 ± 9.6	114.2 ± 6.48	112.3 ± 9.45	115.7 ± 11.34	109.1 ± 7.64
VP (mg)	Lab A	Not done	74.9 ± 12.5	67.0 ± 7.5	45.0 ± 18.79[*]	77.7 ± 11.5	64.7 ± 26.80
	Lab B	79.2 ± 9.5	141.3 ± 16.5	Not done	Not done	139.5 ± 12.6	Not done
		71.5 ± 12.1	135.2 ± 11.9	134.5 ± 9.3	98.1 ± 21.2[*]	Not done	104.1 ± 8.9[**]
	Lab C	Not done	132.2 ± 26.9	110.0 ± 48.66	92.3 ± 16.72	129.1 ± 22.0	107.7 ± 25.9
	Lab D	56.3 ± 14.1	117.8 ± 24.5	118.3 ± 20.9	62.2 ± 10.50[*]	108.3 ± 14.60	97.7 ± 17.3
	Lab E	Not done	114.8 ± 15.2	95.4 ± 8.2	52.2 ± 18.3[*]	120.6 ± 14.95	95.7 ± 8.92
	Lab F	Not done	103.2 ± 11.9	94.0 ± 32.4	73.7 ± 15.8[†,**]	86.3 ± 13.1	76.6 ± 19.3
SVCG (mg)	Lab A	Not done	250.4 ± 57.0	189.0 ± 41.7[**]	124.6 ± 35.24[*]	222.85 ± 34.73	157.8 ± 14.36[*]
	Lab B	47.5 ± 5.1	354.2 ± 46.5	Not done	Not done	347.43 ± 30.52	Not done
		68.8 ± 13.7	406.2 ± 82.0	357.85 ± 4.61	193.1 ± 46.8[*]	Not done	212.4 ± 37.0[*]
	Lab C	Not done	219.7 ± 24.1	201.6 ± 31.59	88.1 ± 9.08[*]	203.4 ± 22.5	137.4 ± 21.2[*]
	Lab D	81.5 ± 12.1	290.7 ± 57.8	301.0 ± 53.7[**]	122.0 ± 26.60[*]	294.2 ± 43.61	233.00 ± 77.09[†]
	Lab E	Not done	308.9 ± 58.0	232.6 ± 37.3[†]	74.6 ± 31.7[*]	285.5 ± 22.25	210.7 ± 46.06[†**]
	Lab F	Not done	141.8 ± 45.5	105.9 ± 12.9	52.1 ± 11.4[*]	122.1 ± 14.2	108.0 ± 20.5
LABC (mg)	Lab A	Not done	144.9 ± 33.7	129.70 ± 15.50	91.0 ± 17.29[*]	148.9 ± 15.6	116.0 ± 17.95[~,**]
	Lab B	119.4 ± 14.4	303.57 ± 34.84	Not done	Not done	278.0 ± 33.7	Not done
		123.4 ± 22.0	296.62 ± 30.44	274.0 ± 20.2	194.3 ± 52.1[†,**]	Not done	215.7 ± 36.3[**]
	Lab C	Not done	260.3 ± 25.6	239.6 ± 24.71	146.0 ± 22.82[*]	239.6 ± 30.8	206.48 ± 17.48[*]
	Lab D	76.2 ± 16.9	165.83 ± 18.18	178.0 ± 45.0	98.6 ± 8.99[*]	176.5 ± 31.39	142.0 ± 27.0[†]
	Lab E	Not done	284.9 ± 41.0	257.3 ± 46.6	115.0 ± 26.4[*]	271.6 ± 17.15	235.3 ± 22.21[†]
	Lab F	Not done	165.2 ± 15.0	148.1 ± 7.5	97.75 ± 20.11[*]	167.4 ± 32.0	129.7 ± 23.6[††]

(continued)

Table 6.8 (continued)

Parameter	Laboratory	Vehicle only	TP 1 mg/kg/d	DDE + TP 16 mg/kg/d	160 mg/kg/d	LIN + TP 10 mg/kg/d	100 mg/kg/d
COW (mg)	Lab A	Not done	16.2 ± 3.2	15.0 ± 3.2	8.8 ± 1.46*	16.7 ± 2.5	12.5 ± 1.48†
	Lab B	7.3 ± 3.3	26.6 ± 2.7	Not done	Not done	27.9 ± 4.0	Not done
		6.6 ± 1.3	26.5 ± 4.0	25.2 ± 2.4	14.5 ± 6.0*	Not done	17.5 ± 3.3†,**
	Lab C	Not done	25.8 ± 5.1	32.0 ± 17.54	13.0 ± 2.61*	23.7 ± 1.1	17.6 ± 5.4†
	Lab D	7.0 ± 1.4	20.3 ± 3.4	23.3 ± 2.9	8.2 ± 2.59*	21.2 ± 9.02	15.8 ± 2.5†
	Lab E	Not done	17.0 ± 4.9	14.5 ± 3.1	4.9 ± 1.4*	16.1 ± 4.42	13.1 ± 2.96
	Lab F	Not done	15.3 ± 5.2	11.7 ± 3.4	7.0 ± 3.3*	12.9 ± 2.0	10.9 ± 2.7
Testes (mg)	Lab A	Not done	228.2 ± 59.4	196.3 ± 20.6	141.2 ± 38.5*	201.9 ± 44.7	173.0 ± 29.78*
	Lab B	1114.0 ± 78.9	534.3 ± 57.4	Not done	Not done	528.2 ± 103.2	Not done
		1106.5 ± 162.8	489.5 ± 89.0	437.9 ± 55.8	506.7 ± 128.7	Not done	363.2 ± 91.2
	Lab C	Not done	512.0 ± 22.8	489.3 ± 110.86	693.5 ± 196.05*	483.3 ± 37.1	453.0 ± 45.0
	Lab D	708.5 ± 196.7	394.8 ± 90.4	343.0 ± 34.6	324.6 ± 60.32	329.7 ± 59.40	270.2 ± 53.9*
	Lab E	Not done	655.6 ± 139.2	600.9 ± 179.9	570.6 ± 213.9	573.0 ± 119.01	556.5 ± 123.78
	Lab F	Not done	882.8 ± 235.7	504.6 ± 154.8*	1093.0 ± 153.8*	599.9 ± 245.9	798.9 ± 206.7
Epididymides (mg)	Lab A	Not done	128.6 ± 11.3	108.9 ± 8.7	76.0 ± 9.9*	123.8 ± 11.3	98.0 ± 13.59*
	Lab B	137.1 ± 17.3	165.0 ± 11.6	Not done	Not done	196.7 ± 48.8	Not done
		135.3 ± 14.6	168.6 ± 9.8	157.8 ± 12.2	118.2 ± 13.2*	Not done	132.6 ± 15.2†,**
	Lab C	Not done	168.0 ± 16.2	157.0 ± 16.59	146.9 ± 19.56†	169.0 ± 15.9	147.7 ± 14.6
	Lab D	96.2 ± 16.5	145.0 ± 27.7	146.0 ± 13.5	80.0 ± 12.29*	131.7 ± 19.99	106.3 ± 11.9†,**
	Lab E	Not done	174.5 ± 19.5	151.8 ± 19.4†	100.4 ± 15.0*	188.0 ± 15.79	156.0 ± 16.28~~
	Lab F	Not done	145.3 ± 16.2	126.2 ± 22.1†	107.1 ± 17.7*	138.8 ± 16.1	129.4 ± 9.5††

[a] One animal in the treatment group died, n=5

* Statistically significant by Dunnett's ($p < 0.05$) using ANCOVA with terminal BW.

** Statistically significant by Dunnett's ($p < 0.01$) using ANCOVA with starting BW.

†: "Marginally insignificant change" by Dunnett's using ANCOVA with terminal BW ($p < 0.08$ and > 0.05).

††: "Marginally insignificant change" by Dunnett's using ANCOVA with starting BW ($p < 0.08$ and > 0.05).

Table 6.9. Optional Organ Weights in Phase 3 With Coded Positive Test Chemicals

Parameter	Laboratory	Vehicle only	TP 1 mg/kg/d	DDE + TP 16 mg/kg/d	160 mg/kg/d	LIN + TP 10 mg/kg/d	100 mg/kg/d
Liver (g)	**Lab A**	Not done	4.0 ± 0.3	4.4 ± 0.5	6.0 ± 0.77	3.8 ± 0.6	3.6 ± 0.40
	Lab B	5.9 ± 0.6	6.6 ± 0.4	Not done	Not done	7.0 ± 0.5	Not done
		6.0 ± 0.7	7.1 ± 0.3	7.9 ± 0.8	11.1 ± 0.4	Not done	6.0 ± 1.0
	Lab C	Not done	7.4 ± 0.6	8.9 ± 1.02	11.6 ± 1.51	7.2 ± 0.4	6.6 ± 0.5
	Lab D	4.8 ± 0.7	5.3 ± 0.2	5.7 ± 0.7	8.5 ± 0.91	5.1 ± 0.91	5.4 ± 0.3
	Lab E	Not done	6.6 ± 1.2	7.8 ± 0.8	8.6 ± 1.5	6.6 ± 0.44	6.1 ± 0.44
	Lab F	Not done	3.9 ± 0.3	4.4 ± 0.2	6.4 ± 0.5	4.0 ± 0.6	3.8 ± 0.3
Adrenals (mg)	**Lab A**	Not done	19.1 ± 5.2	18.0 ± 4.1	20.3 ± 2.06	18.0 ± 3.9	16.0 ± 3.93
	Lab B	21.5 ± 2.9	23.9 ± 3.0	Not done	Not done	24.6 ± 4.1	Not done
		24.1 ± 1.0	25.6 ± 2.1	26.0 ± 1.6	25.9 ± 3.2	Not done	25.1 ± 2.2
	Lab C	Not done	27.1 ± 2.1	27.6 ± 3.39	26.9 ± 2.98	25.0 ± 6.3	23.4 ± 1.4
	Lab D	21.7 ± 2.5	21.7 ± 4.9	22.2 ± 3.8	23.6 ± 3.58	21.2 ± 3.87	24.0 ± 7.5
	Lab E	Not done	Not done	Not done	Not done	Not done	Not done
	Lab F	Not done	30.4 ± 2.8	28.5 ± 2.9	33.9 ± 3.6	28.8 ± 3.6	35.1 ± 6.0
Kidneys (mg)	**Lab A**	Not done	797.9 ± 82.4	752.6 ± 52.1	734.1 ± 35.13	740.9 ± 63.4	734.1 ± 67.40
	Lab B	1276.9 ± 160.5	1434.1 ± 105.3	Not done	Not done	1466.5 ± 129.3	Not done
		1315.1 ± 141.4	1468.2 ± 147.7	1490.9 ± 169.4	1405.4 ± 136.4	Not done	1326.3 ± 68.5
	Lab C	Not done	1578.3 ± 99.1	1603.9 ± 104.79	1603.9 ± 130.86	1535.6 ± 87.7	1410.3 ± 99.1
	Lab D	1048.8 ± 72.3	1029.8 ± 82.2	1157.7 ± 86.8	1078.4 ± 73.31	1206.7 ± 107.94	1117.0 ± 68.5
	Lab E	Not done	Not done	Not done	Not done	Not done	Not done
	Lab F	Not done	1133.0 ± 73.4	1141.2 ± 93.5	1258.1 ± 170.1	1081.1 ± 54.4	1080.7 ± 46.8

Table 6.10. CVs for Body and Target Organ Weights in Phase 3 With Coded Positive Test Chemicals

Parameter	Laboratory	Vehicle only	TP[a] 1 mg/kg/d	DDE + TP 16 mg/kg/d	160 mg/kg/d	LIN + TP 10 mg/kg/d	100 mg/kg/d
BW (g)	Lab A	Not done	7.90%	10.65%	9.43%	12.54%	3.75%
	Lab B	7.39%	4.64%	Not done	Not done	3.27%	Not done
		8.37%	4.52%	6.42%	8.94%	Not done	6.48%
	Lab C	Not done	7.18%	8.40%	5.49%	5.46%	2.63%
	Lab D	7.41%	8.12%	9.89%	7.09%	12.22%	4.12%
	Lab E	Not done	13.52%	7.11%	21.47%	4.55%	5.43%
	Lab F	Not done	8.03%	5.67%	8.42%	9.80%	7.00%
VP (mg)	Lab A	Not done	16.74%	11.15%	41.77%	14.74%	41.45%
	Lab B	11.94%	11.69%	Not done	Not done	9.01%	Not done
		16.90%	8.84%	6.89%	21.65%	Not done	8.53%
	Lab C	Not done	20.36%	44.23%	18.12%	17.01%	24.00%
	Lab D	24.94%	20.77%	17.68%	16.88%	13.47%	17.71%
	Lab E	Not done	13.27%	8.58%	35.04%	12.39%	9.32%
	Lab F	Not done	11.57%	34.51%	21.42%	15.19%	25.13%
SVCG (mg)	Lab A	Not done	22.77%	22.08%	28.28%	15.58%	9.10%
	Lab B	10.77%	13.12%	Not done	Not done	8.78%	Not done
		19.93%	20.18%	1.29%	24.24%	Not done	17.42%
	Lab C	Not done	10.95%	15.67%	10.31%	11.05%	15.41%
	Lab D	14.90%	19.88%	17.84%	21.80%	14.83%	33.09%
	Lab E	Not done	18.78%	16.03%	42.47%	7.79%	21.87%
	Lab F	Not done	32.10%	12.21%	21.91%	11.64%	19.00%

(continued)

Table 6.10 (continued)

Parameter	Laboratory	Vehicle only	TP 1 mg/kg/d	DDE + TP 16 mg/kg/d	DDE + TP 160 mg/kg/d	LIN + TP 10 mg/kg/d	LIN + TP 100 mg/kg/d
LABC (mg)	Lab A	Not done	23.30%	11.95%	19.00%	10.49%	15.47%
	Lab B	12.03%	11.48%	Not done	Not done	12.13%	Not done
		17.84%	10.26%	7.38%	26.84%	Not done	16.82%
	Lab C	Not done	9.85%	10.31%	15.62%	12.87%	8.46%
	Lab D	22.15%	10.96%	25.29%	9.12%	17.79%	19.04%
	Lab E	Not done	14.40%	18.13%	22.96%	6.31%	9.44%
	Lab F	Not done	9.05%	5.06%	20.58%	19.11%	18.20%
COW (mg)	Lab A	Not done	20.00%	21.54%	16.69%	14.88%	11.86%
	Lab B	45.00%	10.28%	Not done	Not done	14.48%	Not done
		20.09%	15.08%	9.71%	41.56%	Not done	18.78%
	Lab C	Not done	19.94%	54.80%	20.16%	4.54%	30.74%
	Lab D	20.20%	16.65%	12.62%	31.57%	42.62%	15.68%
	Lab E	Not done	28.64%	21.34%	27.93%	27.43%	22.58%
	Lab F	Not done	33.73%	29.46%	46.72%	15.46%	24.73%
Testes (mg)	Lab A	Not done	26.03%	10.50%	27.32%	22.15%	17.22%
	Lab B	7.08%	10.74%	Not done	Not done	19.53%	Not done
		14.71%	18.19%	12.74%	25.39%	Not done	25.12%
	Lab C	Not done	4.45%	22.66%	28.27%	7.68%	9.94%
	Lab D	27.76%	22.89%	10.08%	18.58%	18.02%	19.94%
	Lab E	Not done	21.24%	29.95%	37.49%	20.77%	22.24%
	Lab F	Not done	26.70%	30.68%	14.07%	40.99%	25.88%
Epididymis (mg)	Lab A	Not done	8.82%	7.98%	13.10%	9.11%	13.87%
	Lab B	12.64%	7.03%	Not done	Not done	24.80%	Not done
		10.77%	5.81%	7.75%	11.12%	Not done	11.46%
	Lab C	Not done	9.63%	10.57%	13.32%	9.42%	9.90%
	Lab D	17.13%	19.10%	9.25%	15.36%	15.18%	11.20%
	Lab E	Not done	11.20%	12.79%	14.91%	8.39%	10.43%
	Lab F	Not done	11.13%	17.52%	16.55%	11.59%	7.37%

6.3 Interlaboratory Statistical Analyses

A linear model, developed by Dr. Breda Munoz (Senior Research Statistician at RTI International), was used to determine the statistical significance of the laboratory effect (i.e., if there is a difference in the laboratories' ability to detect differences between the test chemical group and the positive control group mean organ weights). Since the participating laboratories used both "log10(BW)" (logarithm base 10 of body weight) and "BW" (body weight) as a covariate in their analysis, models using each one of these two terms were fitted to the data. Since the DDE and LIN groups have two dose combinations, the following full model, including all fixed effects and their interactions, was fitted for the DDE and LIN groups:

Full Model:
Log10 (organ weight) = log10 (BW) + DoseGroup + log10 (BW)*DoseGroup + Lab + Lab*log10 (BW) + Lab*DoseGroup + Lab*log10 (BW)*DoseGroup

where the symbol "*" indicates the interaction term.

Similarly, when using BW as the covariate, the expression for the full model is:

Full Model:
Log10 (organ weight) = BW + DoseGroup + BW*DoseGroup + Lab + Lab* BW + Lab*DoseGroup + Lab* BW*DoseGroup

where the symbol "*" indicates the interaction term.

The double interaction BW*DoseGroup explores if the effect of the different groups (chemical and dose groups) on the organ weight depends on the values of BW, similarly for log10 (BW)*DoseGroup.

The double interaction Lab*DoseGroup explores if the effect of the different dose groups varies across laboratories.

The double interaction Lab*BW explores if the effect of the BW on the organ weight varies across participating laboratories; similarly for Lab*log10(BW).

The three-way interaction, log10 (BW)*DoseGroup*labs, explores if the effect of the different dose groups on the organ weight across laboratories depends on the values of log10 (BW), similarly for BW*DoseGroup*lab.

The two models (using log10 (BW) and BW) were run for each organ of interest If the three-way interaction term "Lab*log10 (BW)*(DoseGroup)" (or Lab*BW*DoseGroup) was statistically significant ($p<0.05$), then changes in log10 (organ weight) across different dose groups and for given body weights were different across labs (suggesting laboratory abilities to detect changes varied across body weight and dose group combinations).

If the three-way interaction term "log10 (BW)*DoseGroup*labs" (or Lab*BW*DoseGroup) was not significant ($p>0.05$), then this three-way interaction term was dropped from the model and the following reduced model was fitted to the data (Reduced Model 1).

Reduced Model 1 (using log10(BW)):
Log10 (organ weight) = log10 (BW) + (DoseGroup) + log10 (BW)*(DoseGroup) + Lab + Lab*log10 (BW) + Lab*(DoseGroup)

Reduced Model 1 (using BW):
Log10 (organ weight) = BW + DoseGroup + BW*DoseGroup)+ Lab + Lab* BW + Lab*DoseGroup

The next step was the analysis of the significance of the double interaction terms. If the second order interactions Lab*BW (or Lab*log10(BW)) was significant ($p <0.05$), then the effect of BW (or log10(BW)) on "log10 (organ weight)" was different across participating laboratories. Similarly, if the second order interaction Lab*DoseGroup was significant ($p <0.05$), then the effect of DoseGroup on "log10 (organ weight)" was different across participating laboratories.

If the second order interaction term between log10(BW) (or BW) and Laboratory, "log10 (BW)*lab" (or BW*lab), was not significant ($p>0.05$), then a new reduced model (Reduced Model 2) without this interaction term was fitted to the data.

Reduced Model 2 (using log10(BW)):
Log10 (organ weight) = log10 (BW) + DoseGroup + log10 (BW)*DoseGroup + Lab + Lab*log10 (BW) + Lab*DoseGroup

Reduced Model 2 (using BW):
Log10 (organ weight) = BW + DoseGroup + BW*DoseGroup)+ Lab + Lab* BW + Lab*DoseGroup

If the two-way interaction term "DoseGroup*Lab" was nonsignificant ($p>0.05$), then a reduced model (Reduced Model 3) without this two-way interaction term was run to fit the data.

Reduced Model 3 (using log10(BW)):
Log10 (organ weight) = log10 (BW) + DoseGroup + log10 (BW)*DoseGroup + Lab

Reduced Model 3 (using BW):
Log10 (organ weight) = BW + DoseGroup + BW*DoseGroup+ Lab

If the two-way interaction term "Lab*log10 (BW)" was nonsignificant ($p>0.05$), then a reduced model (Reduced Model 4) that does not include this two-way interaction term was run to fit the data.

If the term "lab" was significant ($p<0.05$), then laboratory conditions had a statistically significant different effect in detecting differences in "log10 (organ weight)". Therefore, the significance of the lab term suggested that the effect of at least one laboratory was significantly different from the rest, and that this difference affected variability around the organ weight. However, this model cannot identify which lab(s) is different. Identification of the laboratory or laboratories requires between laboratory comparisons ("contrast tests"). Finally, if the term "lab" was nonsignificant ($p>0.05$), then there was no statistical difference in the laboratories' ability to detect differences in "log10 (organ weight)".

Detailed tables showing the consecutive models can be found in Appendix F (Stage 3 Analyses).

The groups DNP, FLU, TNP and TP have only one dose combination, then the Full Model for these groups is:

Full Model:
Log10 (organ weight) = log10 (BW) + Lab + Lab*log10 (BW)

where the symbol "*" indicates the interaction term.

Similarly, when using BW as the covariate, the expression for the full model is:

Full Model:
Log10 (organ weight) = BW + Lab + Lab* BW

where the symbol "*" indicates the interaction term.

The reduced models in these cases are of the form:

Reduced Model 1:
Log10 (organ weight) = log10 (BW) + Lab

Reduced Model 1:
Log10 (organ weight) = BW + Lab

And

Reduced Model 2:
Log10 (organ weight) = log10 (BW)

Reduced Model 2: Log10 (organ weight) = BW

The results of the model for interlaboratory comparisons of Stage 3 data on DDE, DNP, FLU, TP, NP, and vehicle are presented below.

Tables showing the statistical significance of each term in the model are presented for each organ (see Appendix F). The models, as specified above, include the fixed effects DoseGroup and Lab, and the continuous covariate BW or log10 (BW).

The models are fitted in a sequential hierarchical form, starting with the full model (also called saturated model), which has all model terms, and ending with a model that has a statistical significant effect for an interaction term. For instance, in the case of Drug=DDE and outcome = log10(adrenals) (denoted as log10ad), the sequence of model fitted to the data goes from the Full Model to Reduced Model 3. The Model Term column lists the fixed effects (Dose Group, log10BW, Lab), the double interactions (DoseGroup*log10BW, DoseGroup*Lab, log10BW*lab) and the triple interaction (DoseGroup*log10BW*Lab).

Statistics shown in the tables are the "F-statistic" (that evaluates the significance of each model term) and the corresponding "p-value". As explained below, terms are dropped sequentially and one by one from the models if their p-value is greater than 0.05 (see example below).

Drug=DDE

Outcome = log10ad								
Model	Full Model		Reduced Model 1		Reduced Model 2		Reduced Model 3	
Model Term	F-Statistic	p-value	F-Statistic	p-value	F-Statistic	p-value	F-Statistic	p-value
Dose_Group	0.213	0.647	0.207	0.652	0.016	0.900	1.909	0.173
log10bw	0.002	0.969	0.000	0.998	0.025	0.874	0.061	0.807
Dose Group*log10bw	0.243	0.625	0.237	0.629	0.024	0.876	1.685	0.200
Lab	1.352	0.269	2.055	0.104	10.084	0.000	10.571	0.000
Dose Group*Lab	0.091	0.985	1.169	0.338	0.693	0.986		
log10bw*Lab	1.283	0.294	1.979	0.986				
Dose Group*log10bw*Lab	0.087	0.986						
RSquare	0.719		0.717		0.663		0.643	

In the example above, for chemical DDE and adrenal gland weight, the triple interaction term (p=0.986, not significant) was dropped from the Full Model, resulting in the Reduced Model 1. The last term in Reduced Model 1, log10BW*Lab, was also not statistically significant (p=0.986). Therefore, it was also dropped and a reduced model, Reduced Model 2, was fitted to the data. The last term in Reduced Model 2, DoseGroup*Lab, was also not statistically significant (p=0.986) and was therefore also dropped from the model. The last model in the table (Reduced Model 3) had a Lab term which was statistically significant (p=0.000), suggesting an effect of Laboratory on the observed effects of DDE on the log10 adrenal gland weights.

Summary of Results:

6.3.1 DDE

Paired Adrenal Glands

Using "log10 (BW)", the laboratories differences have a statistically significant effect on the weight of the paired adrenal glands ($p<0.001$, Reduced Model 3).

Using “BW”, the laboratories differences have a statistically significant effect on the weight of the paired adrenal glands (p<0.001, Reduced Model 3).

Liver

Using “log(10) BW”, the interaction DoseGroup*lab has a statistically significant effect on the weight of the liver (p=0.008, Reduced Model 2).

Using “(BW)”, the interaction DoseGroup*lab has a statistically significant effect on the weight of the liver (p=0.004, Reduced Model 2).

Paired Kidneys

Using “log10 (BW)”, the differences between laboratories have a statistically significant effect on the weight of the paired kidneys (p<0.001, Reduced Model 3).

Using “BW”, the differences between laboratories have a statistically significant effect on the weight of the paired kidneys (p<0.001, Reduced Model 3).

Paired Cowper’s Glands

Using “log10 BW”, the differences between laboratories have a statistically significant effect on the weight of the paired kidneys (p<0.001, Reduced Model 3).

Using “(BW)”, the differences between laboratories have a statistically significant effect on the weight of the paired kidneys (p<0.001, Reduced Model 3).

Paired Epididymides

Using “log10 BW”, the interaction DoseGroup*LAB has a statistically significant effect on the weight of the paired epididymides (p=0.001, Reduced Model 2).

Using “(BW)”, the interaction DoseGroup*LAB has a statistically significant effect on the weight of the paired epididymides (p=0.001, Reduced Model 2). .

LABC

Using “log10 (BW)”, the differences between laboratories have a statistically significant effect on the weight of the LABC (p<0.001, Reduced Model 3).

Using “BW”, the differences between laboratories have a statistically significant effect on the weight of the LABC (p<0.001, Reduced Model 3).

Seminal Vesicles

Using “log10 BW”, the interaction DoseGroup*LAB has a statistically significant effect on the weight of the seminal vesicles (p=0.012 Reduced Model 2).

Using “BW”, the interaction DoseGroup*LAB has a statistically significant effect on the weight of the seminal vesicles (p=0.025, Reduced Model 2).

Ventral Prostate

None of the terms involving laboratory effects were statistically significant in either model (using log10BW or BW). These results suggest that changes in ventral prostate weight can be explained by the effects of body weight (or log10 BW), of dose group, and of the interaction between DoseGroup and body weight (or DoseGroup and log10 (BW)).

Paired Testes

Using “log10 BW”, the interaction DoseGroup*LAB has a statistically significant effect on the weight of the paired testes (p<0.001, Reduced Model 2).

Using “BW”, the interaction DoseGroup*LAB has a statistically significant effect on the weight of the paired testes (p<0.001, Reduced Model 2).

6.3.2 DNP

Paired Adrenal Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of the paired adrenal glands (p=0.020, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the paired adrenal glands (p=0.024, Reduced Model 1).

Paired Cowper's Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of the paired Cowper's glands (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the paired Cowper's glands (p<0.001, Reduced Model 1) weights.

Paired Epididymides

None of the terms involving laboratory effects were statistically significant in either model (using log10BW or BW). These results suggest that changes in ventral prostate weight are associated with body weight (or log10 BW).

Liver

Using "log10 (BW)", the differences between laboratories have a statistically significant effect on the weight of the liver (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the liver (p<0.001, Reduced Model 1).

LABC

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of LABC (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of LABC (p<0.001, Reduced Model 1).

Seminal Vesicles

Using "log(10) BW", the differences between laboratories have a statistically significant effect on the weight of the seminal vesicles (p<0.001, Reduced Model 1).

Using "(BW)", the differences between laboratories has a statistically significant effect on the weight of the seminal vesicles (p<0.001, Reduced Model 1).

Ventral Prostate

None of the terms involving laboratory effects were statistically significant in either model (using log10BW or BW). These results suggest that changes in ventral prostate weight can be explained by the effects of body weight (or log10 BW).

Paired Kidneys

Using "log10 BW", the differences of body weight across laboratories (lab*bodyweight) have a statistically significant effect on the weight of paired kidneys (p=0.011, Full Model).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the paired kidneys (p=0.002, Full Model).

Paired Testes

Using "log(10) BW", the differences between laboratories have a statistically significant effect on the weight of paired testes (p<0.001, Reduced Model 1).

Using "(BW)", the differences between laboratories have a statistically significant effect on the weight of the paired testes (p=0.001, Reduced Model 2).

6.3.3 FLU

Paired Adrenal Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands (p=0.009, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands (p=0.017, Reduced Model 1).

Paired Cowper's Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands (p<0.001, Reduced Model 3).

Paired Epididymides

Using "log10 BW", differences between laboratories have a statistically significant effect on the weight of the paired epididymides (p=0.045, Reduced Model 2).

Using "(BW)", differences between laboratories has a statistically significant effect on the weight of the paired epididymides (p=0.019, Reduced Model 2).

Paired Kidneys

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys ($p<0.001$ Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys ($p<0.001$, Reduced Model 1).

LABC

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of LABC ($p<0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of LABC ($p<0.001$, Reduced Model 1).

Liver

Using "log(10) BW", differences between laboratories have a statistically significant effect on the weight of the liver ($p<0.001$, Reduced Model 1).

Using "(BW)", differences between laboratories have a statistically significant effect on the weight of the liver ($p<0.001$, Reduced Model 1).

Seminal Vesicles

Using "log(10) BW", differences between laboratories have a statistically significant effect on the weight of seminal vesicles ($p<0.001$, Reduced Model 1).

Using "(BW)", differences between laboratories has a statistically significant effect on the weight of seminal vesicles ($p<0.001$, Reduced Model 1).

Paired Testes

Using "log(10) BW", the interaction BodyWeight*LAB (effect of body weight in organ weight varies across labs) has a statistically significant effect on the weight of the paired testes ($p<0.001$, Full Model).

Using "(BW)", the interaction BodyWeight*LAB (effect of body weight in organ weight varies across labs) has a statistically significant effect on the weight of the paired testes ($p<0.001$, Full Model).

Ventral Prostate

Using "log(10) BW", differences between laboratories have a statistically significant effect on the weight of ventral prostate ($p=0.005$, Reduced Model 1).

Using "(BW)", differences between laboratories have a statistically significant effect on the weight of ventral prostate ($p=0.004$, Reduced Model 1).

6.3.4 LIN

Paired Adrenal Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands ($p<0.001$, Reduced Model 3).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands ($p<0.001$, Reduced Model 3).

Paired Cowper's Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands ($p<0.001$, Reduced Model 3).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands ($p<0.001$, Reduced Model 3).

Paired Epididymides

Using "log(10) BW the differences between laboratories have a statistically significant effect on the weight of the paired epididymides ($p<0.001$, Reduced Model 3).

Using "(BW)", the differences between laboratories have a statistically significant effect on the weight of the paired epididymides ($p<0.001$, Reduced Model 3).

Paired Kidneys

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys ($p<0.001$, Reduced Model 3).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys ($p<0.001$, Reduced Model 3).

LABC

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of LABC ($p<0.001$, Reduced Model 3).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of LABC ($p<0.001$, Reduced Model 3).

Liver

Using "log(10) BW", the differences between laboratories have a statistically significant effect on the weight of the liver ($p<0.001$, Reduced Model 3).

Using "(BW)", the differences between laboratories have a statistically significant effect on the weight of the liver ($p<0.001$, Reduced Model 3).

Seminal Vesicles

Using "log(10) BW the differences between laboratories have a statistically significant effect on the weight of the seminal vesicles ($p<0.001$, Reduced Model 3).

Using "(BW)", the differences between laboratories have a statistically significant effect on the weight of the seminal vesicles ($p<0.001$, Reduced Model 3).

Paired Testes

Using "log(10) BW", the interaction DoseGroup*LAB has a statistically significant effect on the weight of the paired testes ($p=0.026$, Reduced Model 2).

Using "(BW)", the interaction DoseGroup*LAB has a statistically significant effect on the weight of the paired testes ($p=0.034$, Reduced Model 2).

Ventral Prostate

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate (p=0.006, Reduced Model 3).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate (p=0.005, Reduced Model 3).

6.3.5 NP

Paired Adrenal Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands (p=0.002, Reduced Model 1).

Using "BW", the differences between laboratories has a statistically significant effect on the weight of paired adrenal glands (p=0.004, Reduced Model 1).

Paired Cowper's Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands (p<0.001, Reduced Model 1).

Paired Epididymides

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired epididymides (p=0.002, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired epididymides (p=0.002, Reduced Model 1).

Paired Kidneys

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys (p<0.001, Reduced Model 1).

LABC

Using "log10 BW", the effect of the different body weights on LABC is statistically significant different across laboratories (p<0.001, Full Model).

Using "BW", the effect of the different body weights on LABC is statistically significant different across laboratories (p<0.001, Full Model).

Liver

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys (p<0.001, Reduced Model 1).

Seminal Vesicles

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of seminal vesicles (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of seminal vesicles (p<0.001, Reduced Model 1).

Paired Testes

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired testes ($p<0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired testes ($p<0.001$, Reduced Model 1).

Ventral Prostate

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate ($p=0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate ($p=0.002$, Reduced Model 1).

6.3.6 Group TP

Paired Adrenal Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands ($p=0.011$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired adrenal glands ($p=0.016$, Reduced Model 1).

Paired Cowper's Glands

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands ($p<0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired Cowper's glands ($p<0.001$, Reduced Model 1).

Paired Epididymides

Using "log10BW" and "Body Weight (BW)", no term involving Lab was statistically significant, suggesting that changes in paired epididymides can be explained by body weight.

Paired Kidneys

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys ($p<0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired kidneys ($p<0.001$, Reduced Model 1).

LABC

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of LABC ($p<0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of LABC ($p<0.001$, Reduced Model 1).

Liver

Using "log10 BW", the effect of the different body weights on the liver is statistically significant different across laboratories ($p<0.001$, Full Model).

Using "BW", the effect of the different body weights on the liver is statistically significant different across laboratories ($p<0.001$, Full Model).

Seminal Vesicles

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of seminal vesicles ($p<0.001$, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of seminal vesicles (p<0.001, Reduced Model 1).

Paired Testes

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired testes (p<0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired testes (p<0.001, Reduced Model 1).

Ventral Prostate

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate (p=0.007, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate (p<0.016, Reduced Model 1).

6.3.7 Vehicle

To expand our evaluation of the interlaboratory variability around the organ weights of interest, we also examined the data from the performing laboratories for the vehicle (corn oil; negative control) control group. The results of these interlaboratory statistical evaluations follow.

Paired Adrenal Glands

None of the effects in the model were significant, suggesting that any laboratory or body weight differences were not associated with changes in paired adrenal gland weights.

Paired Epididymides

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired epididymides (p=0.007, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired epididymides (p=0.011, Reduced Model 1).

Paired Kidneys

Using "log10 BW", the effect of log10BodyWeight has a statistically significant effect on the weight of paired kidneys (p=0.007, Reduced Model 2).

Using "BW", the effect of bodyweight has a statistically significant effect on the weight of paired kidneys (p=0.011, Reduced Model 2).

LABC

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of LABC (p=0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of LABC (p=0.002, Reduced Model 1).

Liver

Using "log10 BW", the effect of log10BodyWeight has a statistically significant effect on the weight of the liver (p<0.001, Reduced Model 2).

Using "BW", the effect of body weight has a statistically significant effect on the weight of the liver (p<0.001, Reduced Model 2).

Paired Cowper's Glands

None of the effects in the model were significant, suggesting that any laboratory and body weight differences were not associated with changes in paired Cowper's gland weights.

Seminal Vesicles

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of seminal vesicles (p=0.001, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of seminal vesicles (p=0.001, Reduced Model 1).

Paired Testes

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of paired testes (p=0.015, Reduced Model 1)

Using "BW", the differences between laboratories have a statistically significant effect on the weight of paired testes (p=0.021, Reduced Model 1).

Ventral Prostate

Using "log10 BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate (p=0.025, Reduced Model 1).

Using "BW", the differences between laboratories have a statistically significant effect on the weight of the ventral prostate (p=0.023, Reduced Model 1).

6.4 Conclusions

All participating laboratories identified DNP and NP (as coded samples) as negatives in the assay. All participating laboratories detected statistically significant differences in weights of the VP, SVCG, LABC, COW and paired epididymides; paired testes weights were significantly increased at FLU + TP for three of the six participating laboratories relative to the positive inhibitory control group (FLU + TP). Therefore, the laboratories can detect significant organ weight changes under appropriate conditions. These findings were appropriate (DNP and NP are negative, and FLU + TP is positive) and indicated that the assay is specific and selective.

Very few laboratories detected significant differences from the TP positive control group for androgen-sensitive organ weights at 16 mg/kg/day DDE + TP, whereas essentially all of the laboratories detected significant organ weight changes for all of the androgen-sensitive organs at 160 mg/kg/day DDE (Leavens et al., 2002; O'Connor et al., 1999). No laboratory detected significant changes in androgen-sensitive organ weights at 10 mg/kg/day LIN, but almost all of the laboratories detected significant weight changes for those organs at 100 mg/kg/day. Changes in BWs were not detected by any laboratory for either chemical at either dose.

These findings are appropriate since both DDE and LIN are weak androgen antagonists, and BWs are not affected with these chemicals at these doses, and indicate that the assay is reasonably sensitive.

The interlaboratory statistical evaluation resulted in consistent effects across chemicals, in that all organ weights for all chemicals (except for paired testes with DDE or LIN) exhibited a significant "laboratory effect" (see Appendix F). A summary of the interlaboratory statistical analyses is presented in Table 6.11.

Overall, these evaluations with coded chemicals, both negative and positive anti-androgens, indicate that the intact stimulated weanling male version of the Hershberger bioassay is reasonably sensitive, specific, and selective.

Table 6.11. Summary of Interlaboratory Statistical Evaluations

Organ	**ADRENALS**	**COWS**	**EPID**	**KIDNEYS**	**LIVER**	**SVCG**	**TESTES**	**VP**
Lab Effect: Yes or No								
TP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FLU	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TREN	Yes	Yes	Yes	Yes	Yes	Yes	NO	Yes
DDE	Yes	Yes	Yes	Yes	Yes	Yes	NO	Yes
LIN	Yes	Yes	Yes	Yes	Yes	Yes	NO	Yes

Yes = statistically significant effect of laboratory on the organ and chemical as indicated.

No = no statistically significant effect of laboratory on the organ and chemical as indicated.

CHAPTER 7.0 COMPARISON OF RESPONSES BY THE INTACT WEANLING MALE WITH THOSE OF THE CASTRATE ADULT MALE VERSION OF THE HERSHBERGER BIOASSAY

Direct comparisons can be made between the responses of the intact, stimulated weanling male version and those of the castrated adult male version of the Hershberger bioassay, based on the data generated and reported in the validation programs for the two versions. These comparisons can only be made for those organs which were removed and weighed in both versions of the assay and, therefore, include the VP, SCVG, LABC, and COW. The comparison does not include GP since the intact male weanling has not yet achieved PPS, so dissection and weighing of the GP is problematic in the weanling. The comparison also does not include the weights of the paired testes and paired epididymides since they are surgically removed to render the adult males castrated. The doses of TP (the reference stimulatory androgen) differ between the two versions; for the initial evaluation to set the reference stimulatory TP dose and for the subsequent tests, with 1 mg/kg/day used in the intact weanlings and 0.4 (or 0.8) mg/kg/day used in the castrate adults. The doses of FLU (the reference inhibitory anti-antigen) differ between the two versions for the initial evaluation to set the FLU dose, but are the same for the subsequent tests at 3 mg/kg/day, and are used for both the intact weanling and castrated adults. Given these "caveats," the following sections provide comparisons of the organ weight responses in the two versions of the Hershberger bioassay for Stages 1 (1A and 1B), 2 and 3.

7.1 Phase 1A: TP

A comparison of the intact weanling responses to the castrate adult responses to TP, to determine the appropriate TP dose for use as the stimulatory reference dose in both versions, is presented in Table 7.1.

Table 7.1. Comparison of Weanling and Castrate Responses to TP

TP (mg/kg-w/day)	**0**	**0.1**	**0.2**	**0.4**	**0.8**	**1**	**1.2**	**1.6**
VP-Weanling	1.00[a]			1.29	1.53	1.57	1.74	1.85
VP-Castrate	1.00	3.53	5.64	8.84	11.86			13.19
SVCG-Weanling	1.00			2.67	4.03	4.67	5.10	5.70
SVCG-Castrate	1.00	3.05	6.14	10.60	15.91			21.48
LABC-Weanling	1.00			1.71	1.93	2.02	2.18	2.30
LABC-Castrate	1.00	1.78	2.36	3.04	3.47			3.86
COWS-Weanling	1.00			2.42	3.11	3.31	3.61	3.82
COWS-Castrate	1.00	2.83	4.55	6.32	8.51			9.85

[a] Control value set to 1.00, with the values in the other groups presented as relative to the control value.

The TP doses of 0.1 and 0.2 mg/kg/day were evaluated in the castrate adult but not in the intact weanling version. The TP doses of 1.0 and 1.2 mg/kg/day were evaluated in the intact weanling but not in the castrate adult version. The four tissues common to both versions were compared: the VP, SVCG, LABC, and COW. The organ weight data are normalized to the vehicle control value, set to 1.00, with increases above or decreases below the control value presented in the table as the data. It is clear that the castrate adult male version is much more sensitive (in terms of exhibiting greater increases in organ weights) than the intact weanling male version for all four organs evaluated at all of the TP doses in common. However, both versions detected increases in all tissues evaluated at 0.4 mg/kg/day (the lowest dose common to both versions; the weanling version was not evaluated with 0.1 or 0.2 mg/kg/day TP), and both versions exhibited clear dose-response patterns in all tissues evaluated across the TP doses employed.

7.2 Phase 1B: FLU

A comparison of the intact weanling responses to the castrate adult responses to FLU, to determine the appropriate FLU dose for use as the inhibitory reference dose in both versions, is presented in Table 7.2. For the adult castrates, there were two subgroups; one provided the stimulatory TP at 0.2 mg/kg/day and the other provided the stimulatory TP at 0.4 mg/kg/day. The intact weanlings received the stimulatory TP dose of 1.0 mg/kg/day. The values for the four tissues common to both the adult castrates and intact weanlings were again compared: VP, SVCG, LABC, and COW. The organ weight data are normalized to the TP "control" with no FLU and set to 1.0 for the intact weanling and castrate adult tissues at each TP dose employed. The normalized data exhibiting decreases below 1.0 (the TP control) are presented in the table as the data.

Table 7.2. Comparison of Weanling and Castrate Responses to FLU

FLU Doses (mg/kg-bw/day)	**0**	**0.1**	**0.3**	**1**	**3**	**10**
VP-Weanling [a]	1.00	0.95	0.89	0.78	0.68	0.50
VP-Castrate 0.2 TP [b]	1.00	1.00	0.81	0.54	0.28	0.19
VP-Castrate 0.4 TP	1.00	0.84	0.78	0.55	0.27	0.14
SVCG-Weanling	1.00	0.94	0.85	0.61	0.44	0.26
SVCG-Castrate 0.2 TP	1.00	0.83	0.70	0.39	0.21	0.16
SVCG-Castrate 0.4 TP	1.00	0.84	0.77	0.47	0.20	0.10
LABC-Weanling	1.00	0.94	0.90	0.77	0.69	0.55
LABC-Castrate 0.2 TP	1.00	0.97	0.88	0.71	0.53	0.47
LABC-Castrate 0.4 TP	1.00	0.94	0.86	0.71	0.51	0.38
COWS-Weanling	1.00	1.01	0.88	0.72	0.51	0.44
COWS-Castrate 0.2 TP	1.00	0.97	0.77	0.67	0.39	0.30
COWS-Castrate 0.4 TP	1.00	0.91	0.80	0.64	0.41	0.23

[a] Weanlings received 1 mg TP/kg-bw/day by SC injection alone or with FLU by oral gavage

[b] Castrates had two subgroups, with 0.2 and 0.4 mg TP/kg-bw/day, by SC injection

It is clear that the adult castrate version at 0.2 and 0.4 mg/kg/day TP exhibited greater decreases in androgen-dependent tissues in the presence of all doses of FLU from 0.1 through 10.0 mg/kg/day FLU for SVCG and COW. For VP and LABC, the adult castrate version exhibited greater decreases at 0.3 through 10.0 mg/kg/day FLU. At 0.1 mg/kg/day FLU, the intact weanling version with 1.0 mg/kg/day TP was more sensitive (exhibited greater decreases) in VP and LABC than the castrate adult version at 0.2 mg/kg/day TP but not at 0.4 mg/kg/day TP. Both the intact weanling at 1.0

mg/kg/day TP and the castrate adult version at 0.2 and 0.4 mg/kg/day TP exhibited clear and consistent dose-response curves across the range of FLU doses employed.

7.3 Phase 2: TREN, LIN, Added

7.3.1 TREN

Comparison of the responses of the intact weanling and the castrate adult versions to TREN, a synthetic androgen agonist (at 1.5, 8, or 40 mg/kg/day), or to TP (at 1 mg/kg/day for the intact weanlings and at 0.4 mg/kg/day for the castrate adults), is presented in Table 7.3. The values at 0 mg/kg/day (vehicle control) are set to 1.00 mg/kg/day for both the intact weanling and castrate adult for the organs evaluated (VP, SVCG, LABC, and COW). The paired testes, paired epididymides, and GP are not included in this table since they were not evaluated in both versions of the assay. The intact weanling and adult castrate versions detected slight but demonstrable increases in VP, SVCG, and LABC at 1.5 mg/kg/day TREN. For COW, the castrate exhibited a slight increase at 1.5 mg/kg/day VP; the weanling exhibited increases only at 40 mg/kg/day TREN. At 40 mg/kg/day TREN and at 1 mg/kg/day TP, the castrates exhibited greater increases in the weights of all the organs evaluated than did the weanlings. Both the weanlings and the castrates exhibited clear monotonic upward dose-response curves for TREN.

Table 7.3. Comparison of Intact Weanling and Adult Castrate Responses to TREN as Relative Increase in Tissue Weight Versus Vehicle

TREN (mg/kg-bw/day)	0	1.5	8	40	TP[a]
VP-Weanling	1.00	1.07	1.09	1.28	1.61
VP-Castrate	1.00	1.17	1.35	2.18	7.85
SVCG-Weanling	1.00	1.08	1.19	1.69	5.18
SVCG-Castrate	1.00	1.04	1.09	2.25	8.55
LABC-Weanling	1.00	1.07	1.38	1.90	2.08
LABC-Castrate	1.00	1.08	1.22	1.95	2.57
COWS-Weanling	1.00	0.94	0.92	1.40	3.01
COWS-Castrate	1.00	1.04	1.04	1.99	4.70

[a] For weanlings, TP dose was 1 mg/kg-bw/day, and, for adult castrate, TP dose was 0.4 mg/kg-bw/day

7.3.2 LIN (+ TP) and FLU (+ TP)

Comparison of the responses of the intact weanling to the castrate adult rat to TP + LIN, a weak anti-androgen (at 3, 10, 30, or 100 mg/kg/day), or to TP + FLU (at 3 mg/kg/day) is presented in Table 7.4. The response to TP only (the positive control group at 1 mg/kg/day for weanlings and at 0.4 mg/kg/day for castrate adults) is set to 1.00 for the intact weanling and castrate adult for the organs compared: VP, SVCG, LABC, and COW. The paired testes, paired epididymides, and GP are not included since they were not evaluated in both versions of the assay. The data are the increases/decreases of the organs in the LIN + TP or FLU + TP groups, relative to the TP-only values set to 1.00.

The intact weanling was more sensitive than the castrate adult in terms of the lowest LIN dose at which a decrease was detected for all of the organs evaluated. The castrate adult was more sensitive than the intact weanling in terms of the degree of reduction in the organ weights (VP, SVCG, LABC) at the highest LIN dose (100 mg/kg/day) + TP and for all of the organs evaluated in the positive control group (TP + FLU). The most sensitive organ appears to be the SVCG in the weanling (reduced weight at lowest LIN dose). The adult castrate detected decreased weights in the SVCG, LABC, and COW at TP + 30 mg/kg/day LIN + TP and also in the VP at 100 mg/kg/day LIN + TP.

7.3.3 DDE (+ TP)

The comparison of the intact weanling and the castrate adult to TP and DDE (a weak anti-androgen) are presented in Table 7.5. The response to TP only (at 1.0 mg/kg/day for weanlings and at 0.4 mg/kg/day for adult castrates) is set to 1.00 for both versions of the assay for the organs compared (the VP, SVCG, LABC, and COW; the paired testes, paired epididymides and GP are not included since they were not evaluated in both the castrate adult and intact weanling versions). At 3 mg/kg/day DDE (+ TP), only the adult castrate version detected decreases in all of the organs compared (0.93-0.99); the intact weanling version did not (1.00-1.06). Both versions detected reductions in all the organs at 10 mg/kg/day DDE (+ TP) and up, but the reductions were always greater for each organ at each DDE dose in the adult castrate version.

Table 7.4. Comparison of Weanling and Castrate Responses to TP and LIN

	TP[a]	**TP + 3 LIN[b]**	**TP + 10 LIN**	**TP + 30 LIN**	**TP + 100 LIN**	**TP + 3 FLU**
VP-Weanling	1.00	1.11	1.02	0.87	0.84	0.71
VP-Castrate	1.00	1.13	1.18	1.00	0.60	0.33
SVCG-Weanling	1.00	0.93	0.94	0.87	0.71	0.42
SVCG-Castrate	1.00	1.13	1.04	0.86	0.52	0.24
LABC-Weanling	1.00	1.01	0.98	0.95	0.79	0.69
LABC-Castrate	1.00	1.03	1.00	0.91	0.65	0.52
COWS-Weanling	1.00	0.99	0.98	0.90	0.74	0.49
COWS-Castrate	1.00	1.06	1.04	0.94	0.74	0.44

[a] For weanling, TP dose was 1 mg/kg-bw/day, and for surgical castrate, TP dose was 0.4 mg/kg-bw/day

[b] Only 1 of 3 laboratories administered this dose to intact weanlings; all other points are the average of 3 laboratories for weanling and 4 laboratories for the adult castrate.

Table 7.5. Comparison of Intact Weanling and Adult Castrate Responses to TP and DDE Relative Decrease in Tissue Weight Versus TP Only Dose

Parameter						
TP (mg/kg-w/day)	**TP[a]**	**TP + 5 DDE[b]**	**TP + 16 DDE**	**TP + 50 DDE**	**TP + 160 DDE**	**TP + 3 FLU**
VP-Weanling	1.00	1.00	0.94	0.92	0.71	0.79
VP-Castrate	1.00	0.93	0.96	0.75	0.41	0.28
SVCG-Weanling	1.00	1.00	0.95	0.78	0.54	0.49
SVCG-Castrate	1.00	0.99	0.98	0.77	0.40	0.27
LABC-Weanling	1.00	1.06	0.93	0.88	0.63	0.65
LABC-Castrate	1.00	0.95	0.92	0.79	0.51	0.50
COWS-Weanling	1.00	1.00	0.91	0.85	0.65	0.60
COWS-Castrate	1.00	0.97	0.88	0.73	0.51	0.35

[a] For weanlings, TP dose was 1 mg/kg-bw/day, and, for adult castrate, TP dose was 0.4 mg/kg-bw/day

[b] Only 1 of 3 laboratories administered this dose to intact weanlings; all other points are the average of 3 laboratories for weanlings and 4 laboratories for adult castrates.

Both the intact weanling and the adult castrate exhibited monotonic downward dose-response curves across doses for all of the organs. In the inhibitory control group (TP + FLU), both versions responded maximally, exhibiting the greatest reduction in organ weights: 0.27-0.50 for the castrate adult and 0.49-0.79 for the intact weanling. However, the response of the intact weanling to the inhibitory control group (0.49-0.79) was approximately equivalent (0.54-0.71) to that at 100 mg/kg/day DDE (+ TP) for all four organs, while the response of the castrate to the inhibitory control group (0.27-0.50) was greater (more reduction) than that (0.40-0.51; less reduction) at 100 mg/kg/day DDE (+ TP).

The castrate adult version was more sensitive than the intact weanling model in terms of the lowest dose with a response (3 mg/kg/day DDE for the castrate adult versus 10 mg/kg/day DDE for the intact weanling) and in terms of the degree of the response in the castrate adult (lower organ weights) versus the responses in the intact weanling version (weights closer to the TP only positive stimulatory control group values).

7.4 Phase 3: Coded Samples

For Stage 3 with coded chemicals, the castrated adult male version (exposed to LIN at 100 mg/kg/day; coded) exhibited a linear dose response, with the SVCG most sensitive, then VP, LABC, COW, GP, and BW, with the last two endpoints the least sensitive, and the R^2=0.9534. Exposure to DDE at 160 mg/kg/day (coded) also resulted in a linear dose response, with the same order of tissue sensitivity: SVCG most sensitive, then VP, LABC and COW approximately equivalent, GP next, and BW least sensitive, and the R^2=0.9934 (data provided by J.W. Owens, presentation at ISRTP Workshop: Conducting and Assessing the Results of Endocrine Screening, February 19-20, 2008, Washington, DC. Title: Validation of Endocrine Screens: Status and Lessons Learned). Note that in the castrate male version, there are no paired testes or epididymides; the GP is present, already separated, dissected free, and weighed.

For Stage 3, the intact stimulated weanling male version exposed to LIN at 100 mg/kg/day (coded) and to DDE at 160 mg/kg/day (coded) exhibited a very similar order of androgen-sensitive organs from most to least sensitive, as did the castrated adult. Note that in the intact weanling, there are paired testes and epididymides; the GP has not yet separated, so it is not dissected free or weighed. The order for DDE at 160 mg/kg/day was SVCG most sensitive, then LABC, VP, COW, paired epididymides and least sensitive the paired testes. For LIN at 100 mg/kg/day, the order was the same. The only difference in order was the VP weight being the second most sensitive (and more sensitive than the LABC) for the castrate adult male, versus the intact weanling male where the second most sensitive organ weight was for LABC, while the VP was less sensitive (third in rank order). The testes and epididymides, not evaluated since they had been surgically removed in the castrate adult male, were present and least sensitive for both DDE and LIN (both relatively weak anti-androgens), with TP co-administered in the intact stimulated weanling.

7.5 Conclusions

The validation effort of the intact, stimulated weanling male version of the Hershberger bioassay (and of the completed validation of the intact weanling female version of the uterotrophic assay), in addition to the castrated adult male version or to the ovariectomized adult female version of these assays, is in response to animal welfare concerns for the surgical procedures and an attempt to include evaluation of all aspects of the intact HPG-target tissue axis in the prepubertal animals, rather than the much more limited assessment of receptor binding in target tissues in the castrate or ovariectomized adult. It may also obviate the need to perform the male and/or female pubertal assay in the EPA EDSTAC Tier 1 Screening Battery.

The intact weaning is considered less sensitive in that it requires more TP to get sufficient stimulatory response (1 mg/kg/day) versus the castrate adult (0.2 or 0.4 mg/kg/day). At equivalent doses of the anti-androgen FLU (3 mg/kg/day), the intact weanling also appears less sensitive to weak anti-androgens such as LIN or DDE (in the presence of TP) than the castrate adult. The reduced

sensitivity of the intact weanling versus the castrate adult is due to the differences in age and hormonal status of the two animal models. The intact weanling male has an intact HPG axis (testes and epididymides are present) at a stage when the HP are just beginning to produce the appropriate releasing, stimulatory, and inhibitory endocrine hormones. The target end organs (gonads and ASO) are just beginning to detect the endocrine signals and respond to them. There are also low levels of endogenous T (slowly increasing from weaning to puberty) from the intact gonads. This process of fine tuning proceeds slowly and incrementally to puberty and adulthood in males. In contrast, the castrate adult has previously achieved maximum/optimum HPG signaling with, and gonad/ASO responsivity to high endogenous levels of T. Castration (surgical removal of testes and epididymides) as an adult suddenly removes the source of T production and the site of receipt of the HP signaling, but presumably retains the target organ maximum/optimal recognition and sensitivity (not yet developmentally achieved in the intact weanling).

The greatest strength of the intact, stimulated weanling version of the Hershberger bioassay is its ability to detect effects on androgen-producing and androgen-sensitive organs from insults to and mechanism(s) mode(s) of action involving the hypothalamus, and/or pituitary, and/or testes, and/or target-end organs. The specific mechanism(s)/mode(s) of action cannot necessarily be identified, but this version of the assay provides greater confidence that all possible effects are evaluated. Please note that with an intact HPG, even indirect endocrine-relevant effects (e.g., HP-thyroid axis and T3/T4 directly or indirectly affecting the testes and steroidogenesis) can be detected.

CHAPTER 8.0 DISCUSSION AND CONCLUSIONS

Based on the data and analyses presented in Chapters 4.0 - 6.0, the castrated adult male version of the Hershberger bioassay is more sensitive than the intact stimulated weanling version, in terms of both a greater response at the same dose, i.e., greater increases or greater decreases in the ASO weights from TP (0.8 mg/kg/day) or FLU (3.0 mg/kg/day), and greater responsivity at a lower TP dose (0.2 or 0.4 mg/kg/day TP in castrated adults versus 1.0 mg/kg/day TP in intact stimulated weanlings). However, the castrated adult male version is responsive to and detects only those ASO responses that are mediated through the AR in the absence of the gonads (testes).

The intact, stimulated weanling version of the Hershberger bioassay showed lower sensitivity, requiring more TP in the intact weanling version to achieve a sufficient stimulatory response; 1.0 mg/kg/day TP versus 0.2 or 0.4 mg/kg/day TP in the adult castrate. The reduced sensitivity of the intact weanling versus the castrate adult is due to the differences in age and hormonal status of the two animal models. The presence of the intact HPG axis is an asset in that it reflects the normal, intact animal and can potentially detect any effects on ASO weights originating anywhere along the HPG axis; however, there are concerns with the intact weanling version:

- Because it is less sensitive than the adult castrate version of the assay, it requires higher doses of weaker endocrine-disrupting chemicals (which are not identified *a priori*) than does the adult castrate version. This raises the potential for obtaining false negatives with this assay relative to the adult castrate.

- The smaller sizes and lower weights of the ASO in the intact weanling male (even in the presence of stimulation by TP or another androgenic chemical), especially for the COW, require appropriately trained and experienced staff and careful, consistent dissection and weighing (in the participating laboratories) to produce reproducible values with low variances, to allow statistically significant differences to be detected, where appropriate, within and across the laboratories.

- Responses of hormone receptors and consequences of ligand binding during development (including prepuberty) may be different from the responses and consequences in adulthood. For example, the mode of action of the thyroid hormones on the developing testis is not yet fully understood and may be different than in the adult.
- Distinctions between direct endocrine effects and indirect endocrine-relevant effects may be difficult to discern. For example, chemicals which inhibit steroidogenesis in the testis will therefore affect spermatogenesis and ASO weights. However, these chemicals will be detected in the assay as anti-androgenic (for example, specific phthalates).
- It has been noted that the strength of the intact stimulated weanling version of the Hershberger bioassay is the intact HPG. For the detection of agonists, this is accurate. The HPG is intact and capable of responding. For the detection of antagonists, the bioassay requires use of TP (or another potent androgen such as MT or TREN) in all test groups with the positive control group receiving TP plus a potent anti-androgen (such as FLU). The "negative" control group receives only TP. However, the testes in TP-exposed males are clearly smaller (and statistically significantly lower), so that there is almost a "chemical castration" effect, and the HPG axis is not truly intact (due to the

effect on the gonads), and the testes may no longer be responsive to the signals from the hypothalamus and/or the pituitary.

Either version of this assay can provide screening level information to identify potential endocrine disruptors. The choice of the version employed should be made with a careful consideration of the advantages and disadvantages of each.

CHAPTER 9.0 REFERENCES

Ashby J, and Lefevre PA. 2000a. Preliminary evaluation of the major protocol variables for the Hershberger castrated male rat assay. Regul. Toxicol. Pharm. **31:**92-105.

Ashby J, and Lefevre PA. 2000b. The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti-androgens, oestrogens and metabolic modulators. J. Appl. Toxicol. **20:**35-47.

Ashby J, Owens W, Deghenghi R, and Odum J. 2000a. Concept evaluation: an assay for receptor-mediated and biochemical antiestrogens using pubertal rats. Reg. Tox. Pharm. **35:**393-397.

Ashby J, Owens W, and Lefevre PA. 2002b. Concept evaluation: androgen-stimulated immature intact male rats as an assay for antiandrogens. Reg. Tox. Pharm., **35:**280-285.

Bunyan PJ, Townsend MG, Taylor A. 1972. Pesticide-induced changes in hepatic microsomal enzyme systems. Some effects of 1,1-D1(p-chlorophenyl)-2,2,2 – trichloroethane (DDT) and 1,1-DI(p-chlorophenyl)-2,2-dichloroethylene (DDE) in the rat and Japanese quail. Chem. Biol. Interact. **5(1):** 3-26.

Campbell MA, Gyorkos J, Leece B, Homonko K, Safe S. 1983. The effects of twenty-two organochlorine pesticides as inducers of the hepatic drug-metabolizing enzymes. Gen. Pharmacol. **14(4):** 445-454.

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

Dorfman RI. 1962. Methods in Hormone Research Vol II, Part IV: Standard Methods Adopted by Official Organizations, (Dorfman RI, ed.) Academic Press, NY, 707-731.

EDSTAC (Endocrine Disruptor Screening and Testing Advisory Committee). Final Report: Volume I and Volume II, August, 1998. http://www.epa.gov/endo/pubs/edspoverview/finalrpt.htm

Eisenberg E, and Gordan GS. 1950. The levator ani muscle of the rat as an index of myotrophic activity of steroidal hormones. J Pharmacol Exp Therap **99:**38-44.

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

Freyberger, A. Schladt, L. Brockes, A. and Krötlinger, F. 2007. Performance of the Hershberger Assay Using Intact Juvenile Rats – A Pilot Study. Annual Spring Meeting of the German Society of Pharmacology and Toxicology. Mainz, 2007; Nauyn-Schmiedeberg's Arch. Pharmacol. **375 Suppl. 1**, 75.

Freyberger A, Hartmann E, Krotlinger F. 2005. Evaluation of the rodent Hershberger bioassay using three reference (anti)androgens. Arh. Hig. Rada. Toksikol. **56(2):**131-139.

George JD, Tyl RW, Hamby BT, Myers CB and Marr MC. 2003. Assessment of Pubertal Development and Thyroid Function in Juvenile Male CD (Sprague-Dawley) Rats After Exposure to Selected Chemicals Administered by Gavage on Postnatal Days 23 to 52/53. Final Report, Endocrine Disruptor Screening Program, U.S. Environmental Protection Agency. http://www.epa.gov/scipoly/oscpendo/docs/edmvs/malepubertalshortreportnov1103.pdf

Gray LE Jr, Furr J and Ostby JS. 2005. Hershberger assay to investigate the effects of endocrine disrupting compounds with androgenic and antiandrogenic activity in castrate-immature male rats. In: Current Protocols in Toxicology **16.9.1-16.9.15.** J Wiley and Sons Inc.

Hershberger L, Shipley E, and Meyer R. 1953. Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method. Proc Soc Exp Biol Med **83:**175-180.

Hilgar AG, and Hummel DJ. 1964. Endocrine bioassay data. Part III: The androgenic and myogenic evaluation of steroids and other compounds – Assay I. Endocrine Evaluation Branch, CCNSC, NCI, NIH, Bethesda, MD, pp. 15-35.

ICCVAM (Intraagency Coordinating Committee on the Validation of Alternative Methods). 1997. Validation and regulatory accepteance of toxicological test methods. A report of the Ad Hoc Interagency Coordinating Committee on the validation of Alternative Methods, NIH Report No. 97-3981. Research Triangle Park, NC: National Institute of Environmental Health Science.

Joubert Y, Tobin C and Lebart MC. 1994. Testosterone-induced Masculinization of the Rat Levator Ani Muscle during Puberty. Developmental Biology. **162:**104-110.

Kang IH, Kim HS, Shin JH, Kim TS, Moon HJ, Kim, IY, Choi KS, Kil KS, Park YI, Dong MS, Han SY., 2004. Comparison of anti-androgenic activity of flutamide, vinclozolin, procymidone, linuron, and p,p'-DDE in rodent 10-day Hershberger assay. Toxicology **199:**145-159.

Kanno J, Onyon L, Haseman J, Fenner-Crisp P, Ashby J, and Owens W. 2001. The OECD Program to Validate the Rat Uterotrophic Bioassay to Screen Compounds for *In Vivo* Estrogenic Responses: Phase 1. Environ. Health Perspect. **109:**785-794.

Kanno J, Onyon L, Haseman J, Peddada S, Ashby J, Jacob E, and Owens W. 2003a. The OECD Program to Evaluate the Rat Uterotrophic Bioassay: Phase 2: Coded Single-Dose studies. Environ. Health Perspect. **111(12):**1550-1558.

Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, and Owens W. 2003b. The OECD Program to Validate the Rat Uterotrophic Bioassay: Phase 2 – Dose-Response Studies. Environ. Health. Perspect. **111(12):**1550-1558.

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:**581-5.

Kelce, W.R., Lambright, C.R., Gray, L.E. and Roberts, K.P. 1997. Vinclozolin and p,p-DDE alter androgen-dependent gene expression: in vivo conformation of an androgen receptor-mediated mechanism. Toxicol. Appl. Pharmacol. **142:**192-200.

Kennel PF, Pallen CT, Bars RG. 2004. Evaluation of the rodent Hershberger assay using three reference endocrine disrupters (androgen and antiandrogens). Reprod. Toxicol. **18:**63-73.

Kim, H.S., Han, S.Y., Kim, T.S., Kwack, S.J., Lee, R.D., Kim, I.Y., Seok, J.H., Lee, B.M., Yoo, S.D. and Park, K.L. 2002. No androgenic/anti-androgenic effects of bisephenol-A in Hershberger assay using immature castrated rats. Toxicol. Lett. **135:**111-123.

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

Korenchevsky V, Dennison M and Kohn-Speyer A. 1933. Changes produced by testicular hormone in normal and in castrated rats, Biochem. J. 27, 557-579.

Korenchevsky V. 1932. The assay of testicular hormone preparations. Biochem J **26:**413-422.

Korenchevsky V, Dennison M, and Schalit R. 1932. The response of castrated male rats to the injection of the testicular hormone. Biochem J **26:**1306-1314.

Leavens, T.L., Sparrow, B.R. and Devito, M.J. 2002. Lack of antiandrogenic effects in adult male rats following acute exposure to 2,2-bis(4-chlorophenyl)-1,1-dicloroethylene (p,p'-DDE). Toxicology **174:**69-78.

Monosson E, Kelce WR, Lambright C, Ostby J, and Gray LE Jr. 1999. Peripubertal exposure to the antiandrogenic fungicide, vinclozolin, delays puberty, inhibits the development of androgen-dependent tissues, and alters androgen receptor function in the male rat. Toxicol. Ind. Health **15:**65-79.

O'Connor JC, Frame SR, Davis LG, and Cook JC. 1999. Detection of the environmental antiandrogen *p,p'*-DDE in CD and Long-Evans rats using a Tier 1 screening battery and a Hershberger assay. Toxicol. Sci. **51:**44-53.

O'Connor JC, Frame SR, and Ladics GS. 2002a. Evaluation of a 15-day screening assay using intact male rats for identifying steroid biosynthesis inhibitors and thyroid modulators. Toxicol. Sci. **69:**79-91.

O'Connor JC, Frame SR, and Ladics GS. 2002b. Evaluation of a 15-day screening assay using intact male rats for identifying antiandrogens. Toxicol. Sci. **69:**92-108.

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

OECD. 2007. Draft Report of the OECD Validation of the Rat Hershberger Bioassay: Phase 3: Coded Testing of Androgen Agonists, Androgen Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol. ENV/JM/TG/EDTA(2007)2; 16-FEB-2007. pp. 1-125.

OECD. 2006. Final OECD Report of the Initial Work Towards the Validation of the Rat Hershberger Assay: Phase-1, Androgenic Response to Testosterone Propionate, and Anti-androgenic Effects of Flutamide. ENV/JM/MONO(2006)30.

OECD. 2006. Environment Directorate: Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology. Task Force on Endocrine Disrupters Testing and Assessment (EDTA) of the Test Guidelines Programme. ENV/JM/TG/EDTA (2003)5/ANN. Annex 3 to the Draft Report of the OECD Validation of the Rodent Hershberger Bioassay: Phase 2; Hershberger Assay Interlaboratory Study: Statistical analysis of phase 2 data from 16 laboratories in the multi-chemical dose response study.

OECD. 2005. Draft Report of the OECD Validation of the Rat Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories. ENV/JM/TG/EDTA(2003)4/ ADD1.

OECD. 2003a. Final report of the 4th meeting of the validation management group for mammalian testing (VMG-Mamm) of the task force on endocrine disrupters testing and assessment (EDTA). ENV/JM/TG/EDTA/M(2003)3. OECD, Paris.

OECD. 2003b. OECD model protocols and guidance for the conduct of the rodent Hersheberger bioassay with the surgical castrated peripubertal male. For Phase-3 of the OECD program to validate the rodent Hershberger bioassay. Draft as of 17 Oct 2003. Paris, OECD

OECD. 2002a. Final report of the sixth meeting of the task force on endocrine disrupters testing and assessment (EDTA 6). ENV/JM/TG/EDTA/M(2003)4. OECD, Tokyo.

OCED. 2002b. Final report of the work towards the validation of the rat Hershberger assay: phase-1, androgenic response to testosterone propionate, and antiandrogenic effects of flutamide. ENV/JM/TG/EDTA(2002)1/REV2. OECD, Paris.

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

OECD. 2002d. Final OECD report of the work towards the validation of the rat Hershberger assay. Phase 1. Androgenic response to testosterone propionate and anti-androgenic effects of flutamide. ENV/JM/TG/EDTA(2002)1/REV2/ADD1. October 2002. Paris, France.

OECD. 2001a. Acute oral toxicity – up-and-down procedure. OECD Testing Guideline for the testing of chemicals No. 425.

OECD. 2001b. Guidance document on acute oral toxicity. Environmental Health and Safety Monograph Series on Testing and Assessment No 24. ENV/JM/MONO(2001)4.

OECD. 2001c. Third Meeting of the Validation Management Group for the Screening and Testing of Endocrine Disruptors (mammalian effects). OECD, Paris.

OECD, 2000d. Second Meeting of the Validation Management Group for the Screening and Testing of Endocrine Disruptors (mammalian effects). OECD, Paris.

OECD. 2000. Guidance document on the recognition, assessment and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Environmental Health and Safety 732 Monograph Series on Testing and Assessment No 19. ENV/JM/MONO(2000)7.

OECD. 1998a. Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10th-11th March 1998, ENV/MC/CHEM/RA(98)5.

OECD. 1998b. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, Organisation of Economic Cooperation and Development, Paris, 1998. OECD Principles of Good Laboratory Practice (as revised in 1997).

OECD. 1998. The Validation of Test Methods Considered for Adoption as OECD Test Guidelines. ENV/MC/CHEM(98)6, Paris:OECD.

OECD. 1996. Final report of the OECD workshop on the Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Seventh Meeting of the National Co-ordinators of the Test Guidelines Programme, 18-19 September 1996. ENV/MC/CHEM/TG(96)9. Paris: OECD.

Ojeda SR, and Urbanski HF. 1994. Puberty in the rat. In *The Physiology of Reproduction. 2nd*

Edition (E. Knobil and J.D. Neill, Eds.) Raven Press, NY. pp. 363-409.

Owens W, Zeiger E, Walker M, Ashby J, Onyon L, and Gray Jr, LE. 2006. The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: Use of a potent agonist and a potent antagonist to test the standardized protocol. Env. Health Persp. **114:**1259-1265.

Supplemental materials for Owens et al. 2006. The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: Use of a potent agonist and a potent antagonist to test the standardized protocol. Env. Health Persp. **114:**1265-1269. See, section II, The dissection guidance provided to the laboratories: http://www.ehponline.org/docs/2006/8751/suppl.pdf.

Owens W, Ashby J, Odum J, and Onyon L. 2003. The OECD program to validate the rat uterotrophic bioassay: phase two-dietary phytoestrogen analyses. Environ. Health Perspect **111(12):**1559-1567.

Owens JW, and Ashby J. 2002. Critical review and evaluation of the uterotrophic assay for the identification of possible estrogen agonists and antagonists. Crit. Rev. Toxicol. **32:**445-520.

Schladt, L. (Author). Rodent Hershberger Bioassay with the Stimulated Weanling Male, Bayer Health Care AG, Wuppertal, Germany. (Study No. T6077020; Report No. AT04352) Study Completion Date: January 18, 2008; 62 pages.

Shin JH, Moon HJ, Kang IH, Kim TS, Lee SJ Ahn JY, Bae H, Jeung EB and Han SY. 2007. OECD validation of the rodent Hershberger assay using three reference chemicals; 17α–methyltestosterone, procymidone, and p,p'-DDE, Arch. Toxicol. **81:**309-318.

Sloan SC, Tyl RW, George JD, Pearce SW, Vick KD, Hamby BT, Myers CB, and Marr MC. 2005. 15-Day Tier 1 Screen of Endocrine Active Compounds Administered by Gavage to Adult Male Sprague-Dawley CD® Rats. RTI International. Research Triangle Park, NC, Final Report 65U-8055.001.021.

Sloan C, Vick K, Fail PA, and Becker RA. 2002. Hershberger assay: dose-response and comparison of tissue responses to androgens given orally and subcutaneously. The Toxicologist **66 (1-S):**276 (Abstract 1351).

Stoker TE, Parks LG, Gray LE and Cooper RL. 2000. Endocrine-disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations. Crit. Rev. Toxicol. **30(2):**197-252.

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

Yamasaki K., Ohta R., and Okuda H. 2006. OECD valiadation of the Hershberger assay in Japan: Phase 3. Blind study using coded chemicals. Toxicology Letters **163:**121-129.

Yamasaki K, Sawaki M, Ohta R, Okuda H, Katayama S, Yamada T, Ohta T, Kosaka T, and Owens W. 2003a. OECD Validation of the Hershberger Assay in Japan: Phase 2 Dose Response of Methyltestosterone, Vinclozolin and *p,p'-DDE*. Environ. Health Perspect. **111(16):**1912-1919.

Yamasaki K, Takeyoshi M, Sawaki M, Imatanaka N, Shinoda K, Takatsuki M. 2003b. Immature rat uterotrophic assay of 18 chemicals and Hershberger assay of 30 chemicals. Toxicology **183:** 93-115.

You, L. Chan, S.K., Bruce, J.M., Archibeque-Engel, S., Casanova, M. Corton, J.C., and d'A Heck, H.

1999. Modulation of testosterone-metabolizing hepatic cytochrome p-450 enzymes in developing Sprague-Dawley rats following in utero exposure to p,p'-DDE. Toxicol. Appl. Pharmacol. **158:**197-205.

CHAPTER 10.0 ACKNOWLEDGEMENTS

The staff at RTI International, Research Triangle Park, NC, thank the following individuals:

- Mr. Gary E. Timm, EPA Task Order Project Officer, for the EPA Endocrine Disruptors Screening Program, Task Order 16, for allowing the staff at RTI the opportunity to produce the final report for the OECD Validation of Intact Stimulated Weanling Male Version of the Hershberger Bioassay, and for his helpful support and guidance during the performance of this task order
- Dr. J. William Owens, The Procter and Gamble Company and Secondment to OECD, during the validation program for his generation of the data files from the participating laboratories, his intralaboratory statistical analyses, and for his unfailing assistance to the authors, as needed
- Dr. L. Earl Gray, Jr., Senior Reproductive Toxicologist, Head of the Reproductive Toxicology Division (RTD), Lead Laboratory, U.S. EPA, NHEERL, who helped design the validation protocol and was available for consultation with the performing laboratories during the validation program
- CEFIC, the European Chemical Industry Trade Association, for funding the chemical purchases and the repository for the mammalian OECD program.
- A special thanks also to Ms. Nathalie Delure and Ms. Laurence Musset of OECD.
- Mr. Elard Jacob of BASF, the overall manager of the repository
- TNO, the contract repository

The following RTI staff is also thanked for their participation:

- Dr. Rochelle Tyl, author
- Ms. Carol Sloan, reviewer
- Dr. Julia George, task order leader
- Ms. Bonnie Hamby, task order manager and reviewer
- Ms. Gayle Bieler, statistician
- Dr. Breda Munoz, statistician
- Ms. Shu-yi Lin, statistical data support
- Ms. Michelle Oh, quality assurance auditor
- Ms. Cathee Winkie, administrative coordinator
- Ms. Ashley Jones, administrative support for references

APPENDIX A

DEFINITIONS

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

Antiandrogenic is the capability of a substance to suppress the action of TP on the growth of androgen-dependent tissues in a mammalian organism.

Date of birth is postnatal day (pnd) 0.

Dosage is a general term comprising of dose, its frequency, timing and the duration of dosing.

Dose is the amount of test substance administered. For the Hershberger bioassay, the dose is expressed as weight of test substance per unit body weight of test animal per day (e.g., mg/kg body weight/day).

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

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

Positive and negative results for the castrate peripubertal male version of the Hershberger bioassay are defined as follows: A positive result is defined as the percent weights of at least four of the five androgen-dependent organs exhibiting statistically specific differences from the negative control mean values (vehicle control for androgenic evaluation; TP control for anti-androgenic evaluation).

Positive and negative results for the intact stimulated weanling of the Hershberger bioassay are defined as follows: A positive result for a test substance is defined as the mean weights of at least five of the six androgen-dependent organs exhibiting statistically specific differences from the negative control mean values (vehicle control for androgenic evaluation: TP control for anti-androgenic evaluation).

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

Sensitivity is the capability of a test method to correctly identify chemicals having the property that is being tested for.

Validation is a scientific process designed to characterize the operational requirements and limitations of a test method and to demonstrate its sensitivity, specificity, reproducibility and relevance for a particular purpose.

Specificity is the capability of a test method to correctly identify chemicals not having the property that is being tested for.

APPENDIX B

OECD MODEL PROTOCOLS AND GUIDANCE FOR THE CONDUCT OF THE RODENT HERSHBERGER BIOASSAY WITH THE STIMUALTED WEANLING MALE

(17 October 2003)

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

LEAD LABORATORY

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

Lead Investigator: Dr. Earl GRAY

Study Manager and Director: Dr. J. William Owens

INTRODUCTION

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

2. The precursor of the rodent Hershberger bioassay was first developed in the 1930s and included various tissues of the male reproductive tract (1)(2)(3), including the ventral prostate, the seminal vesicles with coagulating glands, the Cowper's glands, the glans penis, and preputial glands. The measurement of the levator ani and bulbocavernosus muscles were subsequently investigated in the 1940s (3)(4). After publication of work with an extensive number of compounds by Hershberger *et. al.* in 1953 (5), the procedure has been commonly referred to as the Hershberger bioassay. The primary model for the Hershberger bioassay has been the surgically castrated pubertal male, and this model has been the basis for work in Phase-1 and Phase-2 of the Hershberger validation program. However, there are animal welfare concerns with the castration procedure, and an alternative is sought to avoid the castration step. This model protocol outlines the use of the uncastrated stimulated weanling as that possible alternative.

3. The Hershberger bioassay is an *in vivo* short-term assay whose conditions are designed to achieve low endogenous hormone levels and employ target tissues that are highly responsive to administration of exogenous hormones. The focus of the Hershberger bioassay is on male sex hormones (androgens and antiandrogens). The objective in Phase-3 of the validation with this model protocol is to assess the potential of the stimulated weanling as reliable and relevant alternative to the surgical castration.

4. The Hershberger is being validated by OECD as potential short-term screen (6). The information generated by the bioassay can be used to build on that information already available, e.g. from relevant *in vitro* screens, to narrow the field of chemicals that may need longer-term animal testing. This current stimulated weanling protocol is based largely on the standardisation and optimisation work performed in the Syngenta CTL laboratory.

BASIS FOR THE VALIDATION WORK WITH THE STIMULATED WEANLING

- 5. The OECD program on the Hershberger assay has previously demonstrated the response and the reliability of surgical castrate by measuring male sex accessory glands and tissues, i.e., ventral prostrate (VP), seminal vesicles with coagulating glands (SVCG), the levator ani and bulbocavernosus muscles (LABC), the Cowper's (or bulbourethral) glands (COWS), and glans penis (GP). In the case of the stimulated weanling, the immature animals have not undergone preputial separation due to their sexual immaturity. This lack of preputial separation compromises the dissection of the GP, so that the reliability of this particular tissue with the stimulated weanling is in question and will be addressed in the validation program. Therefore, the mandatory target tissues in the stimulated weanling will be reduced to four: the VP, SVCG, LABC, and COWS. However, two tissues removed during castration are available for investigation in this model, the testes (T) and epididymes (EP). Both of these tissues will be measured and their utility assessed as part of the validation program.

- 6. Biological plausibility would suggest that the stimulated weanling and the surgically castrated versions of the Hershberger bioassay should respond in a qualitatively similar manner. In both versions, growth of the target mandatory tissues is controlled through the androgen receptor, and testosterone is converted to dihydrotestosterone by the 5α-reductase enzyme. The same male reproductive tract tissues can be employed, with the exception of the GP, T, and EP as noted above.

- 7. There are, however, differences to be noted between the stimulated weanling and the surgically castrated male. As a result, comparative work is needed to define the quantitative differences that may exist between the two versions. The key differences in the stimulated weanling are low prepubertal circulating levels of testosterone produced by the testes and the potential ability of the intact hypothalamic-pituitary-gonadal axis to respond and compensate to some degree to chemical insult. These may diminish the responsiveness of the stimulated weanling version of the assay. For further background, there are reviews and recent manuscripts that trace the development of the male rat for both hormonal levels and the growth of reproductive tract tissues prior to puberty (7)(8)(9)(10). It should be noted that the post natal day (pnd) 20-35 time period in the male is not fully analogous to the female. There are no dramatic surges in the male in hormonal levels or tissue growth in this time period. The change and rearrangement of the hypothalamic-pituitary-gonadal axis is more gradual in circulating FSH, LH, testosterone, and 5α-reduced steroids. At the same time, just as this is a period of rapid overall growth in the animal, the weights of the testes and the relevant tissues of the reproductive tract are also increasing in absolute terms. The increase *relative to body weight* for these tissues is rather modest and constant, indicating no disproportionate surge in growth.

OUTLINE OF THE WORK FOR THE VALIDATION OF THE STIMULATED WEANLING

- 8. The program to assess the stimulated weanling as an alternative will proceed in the same fashion as that taken by the preceding surgical castrate model. Due to the low circulating testosterone levels and the intact HPG axis in the weanling, testosterone propionate (TP) (CASRN 57-85-2) and Flutamide (FLU) (CASRN 1311-84-7) standardization curves with the stimulated weanling male are necessary to ensure the proper TP dose is selected for the antagonist studies. Therefore, the first two stages, analogous to Phase-1A and Phase-1B (11)are:

  - In the first stage, a dose response curve with TP will be generated as the first step to select optimum candidates for the TP coadministration dose with antiandrogens. The basis for selection is intended to be the same as with the surgical castrate: to identify ED_{50}-ED_{70} doses of TP that are applicable to all the responsive tissues. The doses to be employed are 0.4, 0.8, 1.0, 1.2, and 1.6 mg/kg/d in order to provide a complete comparison with the surgical castrate data set. This dose series should adequately characterize the dose response for each of the four tissues (VP, SV, LABC, and COWS), its relative maximum response, and provide insight on the CVs that labs can achieve with the smaller weanling tissues. With the vehicle control, this comprises a total of six dose groups per laboratory (36 animals based on n = 6). This will allow the selection of two TP doses to proceed into the second stage, the FLU dose series experiments.

  - In the second stage, two selected doses of TP will be used as the reference dose for coadministration to assess the response of stimulated weanling tissues to a dose series of FLU. The doses of FLU to be employed are 0.3, 1.0, and 3.0 mg/kg/d FLU. Higher doses are judged to be wasteful of animals, based on the available data. With the two TP reference dose groups, this comprises a total of eight dose groups per laboratory (48 animals based on n = 6). Two additional groups are voluntary: a vehicle control and a fourth dose of FLU at 10 mg/kg/d. This will allow

the selection of the optimum TP reference dose in order to proceed into the third stage, the dose response experiments to weak agonists and antagonists.

- 9. The next two stages comprise studies analogous to the Phase-2 dose response studies and the Phase-3 coded studies for the surgical castrate:

- In the third stage, the stimulated weanling will be used to assess its capability to detect a weak androgen agonist, trenbolone, and to detect two weak androgen antagonists, linuron and *p,p'*-DDE. All three compounds challenged the surgical castrate in one or more laboratories where one or more of the five castrate mandatory tissues did not achieve statistical significance. In order to make a comparison, the same dose series will be used, but the lowest dose will be omitted in the interests of animal welfare and conserving resources:

    - Trenbolone (TREN) (CASRN 10161-33-8):
      1.5, 8, and 40 mg/kg/d (rationale: no lab reached statistical significance with the surgical castrate at 1.5 mg TREN/kg/d; therefore, the 0.3 mg/kg/d dose is not an apparent sound use of animals or resources)
    - Linuron (LIN) (CASRN 330-55-2):
      10, 30, 100 mg/kg/d (rationale: no lab reached statistical significance with the surgical castrate at 10 mg LIN/kg/d; therefore, the 3 mg/kg/d dose is not an apparent sound use of animals or resources)
    - *p,p'*-DDE (DDE) (CASRN 72-55-9):
      16, 50, 160 mg DDE /kg/d (rationale: no lab reached statistical significance with the surgical castrate at 10 or 16 mg/kg/d – doses inside and outside of Japan, respectively; therefore, the 5 mg/kg/d or similar dose is not an apparent sound use of animals or resources)

- The fourth stage is analogous to final Phase-3 of the surgical castrate validation. The weanling version's reliability will be assessed by its response to and ability to correctly identify coded positive and negative substances in both agonist and antagonist studies. In these studies, the doses of certain positive substances will be identical to that in stage 3, so that the reproducibility of the bioassay over time may be assessed. In addition, the relative sensitivity and effectiveness of the different sex accessory tissues and glands in the assay will continue to be assessed.

INITIAL CONSIDERATIONS AND PRINCIPLE OF THE ASSAY

10. The rodent Hershberger assay has been based historically on changes in the weights of androgen-responsive male sex accessory tissues largely in peripubertal, castrated male rats. Accessory sex tissues and glands depend upon androgen stimulation to gain and maintain weight during and after puberty. When endogenous sources of androgen are low (the weanling before puberty), the biological activity of exogenous substances can be assayed by the increase (agonist response) in the weights of these sex accessory tissues or by blocking (antagonist response) the activity of administered androgens and by preventing an increase in the weights of these sex accessory tissues. The rodent Hershberger assay modified to use the stimulated weanling then evaluates the ability of a chemical to show biological activities consistent with the agonism or antagonism of androgens (e.g., testosterone and dihydrotestosterone).

11. The available data indicates that the androgen-dependent sex accessory tissues of the weanling are sensitive to androgens. This is plausible as these animals have both androgen receptors and appropriate steroidogenic enzymes necessary for agonist, antagonist, and 5α-reductase inhibitor responses. In addition to sensitivity, the weanling rodent sex accessory tissues have a small relative weight.

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

- In stage 1, the dose response of the stimulated weanling version of the Hershberger bioassay to a series of Testosterone Propionate (TP) doses for comparisons with the surgical castrate and as the basis for selecting two doses for coadministration studies with Flutamide (FLU)
- In stage 2, the dose response of the stimulated weanling version of the Hershberger bioassay to a series of FLU doses using two selected TP coadministration doses for comparisons with the surgical castrate and as the basis for selecting a single TP dose for coadministration with antagonists in dose response studies
- In stage 3, the dose response of the stimulated weanling version of the Hershberger bioassay to a series of doses with a weak agonist, trenbolone, and two weak antagonists, linuron and DDE
- In stage 4, to assess the capability of the stimulated weanling version of the Hershberger bioassay to identify coded positive and negative substances.
- The relative effectiveness of the different sex accessory tissues and glands in the assay.
- The reproducibility of the stimulated weanling version of the bioassay over time by comparing appropriate data from stage 4 to that generated in stages, 1, 2 and 3.
- Continue the investigation of the value of the different accessory tissues and glands

13. The test substances will be coded and administered to groups of six animals (n = 6) for 10 consecutive days in all studies in all stages. The animals will then be necropsied approximately 24 hours later on the 11^{th} day (24 hours after the last test substance administration). After dissection, the weights of the mandatory sex accessory tissues will be measured as well as the T and EP weights.

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

Androgen agonists

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

Androgen antagonists

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

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. Where the screening assay may be preliminary to a repeated dose oral study, a reproductive and developmental study, or a long-term study, preferably animals from the same strain and source should be used in all studies. If a laboratory is planning to use an unusual rat strain, or one unique to their own facility, they should determine whether the sexual development criteria noted under the section, *INITIAL CONSIDERATIONS AND PRINCIPLE OF THE ASSAY*, are met.

Age and Acclimatisation

- 18. Young weanling animals must be employed in a relatively small time window between weaning and before puberty, i.e. postnatal days 20 to 34. Animals will be observed daily, and any animals with evidence of disease or physical abnormalities will be removed. The treatment with initiation of dosing (on study) may commence as early as pnd 21 days of age, but preferably not later than pnd 24. This allows a laboratory some flexibility 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 be 50 to 60%, but should not exceed limits of 30 to 70% except during room cleaning. Lighting should be artificial, the photoperiod being 12 hours light, 12 hours dark.

20. Laboratories participating in the validation should use the laboratory diet normally used in their chemical testing work. In previous phases, no effects or variability were observed that were attributable to the diet. The diet used will be recorded and a sample of the laboratory diet will be retained for possible future analysis. Both diet and drinking water will be supplied *ad libitum*.

21. Weanling animals should be caged in groups of no more than 6 similarly treated rats per cage, giving a minimum of 1 cage of 6 rats/cage per treatment group and a maximum of 2 cages of 3 rats/cage per treatment group. When cages are properly sized (~2000 square centimeters), six animals or less per cage avoids crowding. Cages should be thoroughly cleaned to remove possible contaminants and arranged in such a way that possible effects due to cage placement are minimised.

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

Body Weight and the selection of animals for the study

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

24. Experimental control of body weight is accomplished in two steps. The first step involves selection of animals with relatively small variation in body weight for the study cohort from the larger population of animals that have been supplied. Unusually small or large animals should be avoided and should not placed in the study cohort. A reasonable level of body weight variation within the study cohort should be tolerated. Here, ± 20% of the mean body weight for the cohort population is judged to be reasonable (e.g. 50g ± 10 g). The second step involves the assignment of animals to different treatment groups (n = 6) by a

randomised complete block approach. Under this approach animals are randomly assigned to treatment groups so that each group has the same mean and standard deviation in weight at the beginning of the study. The procedure used for block randomization should be recorded.

Non-routine health and safety requirements

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

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

PROCEDURE - VALIDATION OF THE STIMULATED WEANLING

Administration of doses

27. TP will be administered by subcutaneous (sc) injection in all stages of the validation program.

28. All other test substances will be administered by oral gavage.

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

30. Oral gavage will be the delivery of the test substance in vehicle by means such as intubation with an oral gavage syringe. The maximum limit on the volume administered per animal will be 5 ml/kg/day. As the weanling animals are small, the technical staff conducting the gavage should be experienced in order to avoid gavage errors that might lead to morbidity or mortality.

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

Good Laboratory Practice

32. Work should be conducted according to the principles of Good Laboratory Practice (OECD Good Laboratory Practice and Compliance Monitoring (12). In particular, data should have a full audit trail and be retained on file. Data will be collected in a manner that will allow independent peer review and written records maintained.

Vehicle

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

34. The following procedures are to be conducted in four consecutive stages.

STAGE 1 (Analogous to the Completed Phase 1A with the surgical castrate)

Test substance and doses in stage 1 of the stimulated weanling validation

35. The test substance for stage 1 of the stimulated weanling validation will be TP.

36. The reference androgen agonist, TP, will be administered in a series of doses comprising 0.4, 0.6, 1.0, 1.2, and 1.6 mg/kg/d.

Test groups in stage 1 of the stimulated weanling validation

37. 6 animals of the same age and cohort will be used for the vehicle, the TP doses, and any other control group or treatment used by the laboratory.

38. The response of the sex accessory tissues and glands to a reference agonist, Testosterone Propionate (TP), will be studied. This work will involve five test groups for increasing doses of TP and one vehicle control group. The test groups are illustrated in Table 1.

Table 1. Test Groups for Stage 1 of the Stimulated Weanling Validation: Testosterone Propionate Dose Series

•	• **Agonist response**
• Group A • Vehicle Control • **Voluntary**	• Vehicle only
• Group B	• 0.4 mg/kg/d Testosterone Propionate
• Group C	• 0.8 mg/kg/d Testosterone Propionate
• Group D	• 1.0 mg/kg/d Testosterone Propionate
• Group E	• 1.2 mg/kg/d Testosterone Propionate
• Group F	• 1.6 mg/kg/d Testosterone Propionate

STAGE 2 (Analogous to the completed Phase 1B with the surgical castrate)

Test substance and doses in stage 2 of the stimulated weanling validation

39. The test substances for stage 2 of the stimulated weanling validation will be TP and FLU.

40. Two doses of the reference androgen agonist, TP, will be selected based upon stage 1 data, and these selected doses will be coadministered with a series of FLU doses. The FLU doses are 0.3, 1.0, and 3.0 mg/kg/d. Laboratories may voluntarily extend the FLU dose series by adding test groups with the selected TP dose plus 10/mg/kg/d.

Test groups in stage 2 of the stimulated weanling validation

41. 6 animals of the same age and cohort will be used for the vehicle, the two selected TP doses, and two series of three groups each where a selected TP dose is coadministered with an increasing series of FLU doses.

42. The response of the sex accessory tissues and glands to selected doses of a reference agonist, Testosterone Propionate (TP), coadministered to a series of FLU doses will be studied. This work will involve TP control groups for each selected dose and three test groups for each selected dose where that TP dose is coadministered with FLU having a prescribed does of the test substance and one vehicle control group. The required and voluntary test groups are illustrated in Table 2.

Table 2. Test Groups for Stage 2 of the Stimulated Weanling Validation:

Two Selected Testosterone Propionate Doses Coadministered with Flutamide Dose Series

•	• **Antagonist response** • **Selected TP dose #1**	• **Antagonist response** • **Selected TP dose #2**
• Group A • Vehicle Control • **Voluntary**	• Vehicle • **Voluntary**	• Vehicle • **Voluntary**
• Group B • Negative Control	• Selected Dose of TP #1	• Selected Dose of TP #2
• Group C	• Selected Dose of TP #1 + • 0.3 mg/kg/d FLU	• Selected Dose of TP #2 + • 0.3 mg/kg/d FLU
• Group D	• Selected Dose of TP #1 + • 1.0 mg/kg/d FLU	• Selected Dose of TP #2 + • 1.0 mg/kg/d FLU
• Group E	• Selected Dose of TP #1 + • 3.0 mg/kg/d FLU	• Selected Dose of TP #2 + • 3.0 mg/kg/d FLU
• Group F • **Voluntary**	• Selected Dose of TP #1 + • 10 mg/kg/d FLU • **Voluntary**	• Selected Dose of TP #2 + • 10 mg/kg/d FLU • **Voluntary**

STAGE 3 (Analogous to the Phase 2 with the surgical castrate)

Test substances in stage 3 of the stimulated weanling validation

43. The reference androgen agonist will be TP. The reference androgen antagonist will be FLU. The appropriate doses of these substances will be selected based upon stage 2 data.

44. The test substances for Phase 3 stage 3 dose response studies for the stimulated weanling will be:

Trenbolone (17β-Hydroxyestra-4,9,11-trien-3-one)	CAS No 10161-33-8
Linuron	CAS No 330-55-2
p,p'-DDE (4,4'Dichlorodiphenyldichloroethylene)	CAS No 72-55-9

45. The agonist test substance doses for Phase 3 stage 3 dose response studies for the stimulated weanling will be:

Trenbolone	1.5, 8, and 40 mg/kg/d

The two weak antagonists will be administered along with a dose of TP selected from the stage 2 studies with each of the following doses:

Linuron	10, 30, 100 mg/kg/d
p,p'-DDE	16, 50, 160 mg/kg/d

Test groups in stage 3 of the stimulated weanling validation

46. 6 animals of the same age and cohort will be used for the vehicle, TP, and any other control group and for each treatment or test substance group.

47. The response of the sex accessory tissues and glands to a weak agonist, trenbolone, will be studied. This work will involve three test groups for each agonist having a prescribed does of the test substance and one vehicle control group. The required test groups are illustrated in Table 3.

48. The response of the sex accessory tissues and glands to two weak antagonists, linuron and *p,p'*-DDE, and will be studied. This work will involve three test groups for each antagonist and the coadministration of a dose of the reference agonist TP to each group, where that dose will be selected based on stages 1 and 2. Each test substance will have a prescribed dose for each of its groups. The negative control group will be the reference dose of TP from stage 2. The required test groups are illustrated in Table 3.

Table 3. Test Groups for Stage 3 of the Stimulated Weanling Validation: Agonist and Antagonist Dose Responses

•	• **Trenbolone response**	• **Linuron response**	• **DDE response**
• Group A • Vehicle Control	• Vehicle • Mandatory for agonist	• Vehicle • Voluntary for antagonist	• Vehicle • Voluntary for antagonist
• Group B • Negative Control	• Provided by vehicle control (no additional group needed for agonist)	• Selected TP dose from stage 2	• Selected TP dose from stage 2
• Group C	• 1.5 mg/kg/d Trenbolone	• Selected TP dose from stage 2 • + 10 mg/kg/d linuron	• Selected TP dose from stage 2 • + 16 mg/kg/d DDE
• Group D	• 8 mg/kg/d Trenbolone	• Selected TP dose from stage 2 • + 30 mg/kg/d linuron	• Selected TP dose from stage 2 • + 50 mg/kg/d DDE
• Group E	• 40 mg/kg/d Trenbolone	• Selected TP dose from stage 2 • + 100 mg/kg/d linuron	• Selected TP dose from stage 2 • + 160 mg/kg/d DDE

• Group F • Positive Treatment Group • **Voluntary**	• TP dose selected from stage 2 • **Voluntary**	• TP dose selected from stage 2 + FLU dose selected from stage 2 • **Voluntary**	• TP dose selected from stage 2 + FLU dose selected from stage 2 • **Voluntary**

STAGE 4 (Analogous to the current Phase 3 with the surgical castrate)

Test groups in stage 4 of the stimulated weanling validation

49. 6 animals of the same age and cohort will be used for the vehicle, TP, and any other control group and for each treatment or test substance group.

• 50. All participating laboratories will use the coded test substances supplied and follow the instructions for the preparation of the proper doses in order to achieve specified dosages that can be compared data generated in stage 3. There are two coded series. One series is for androgen agonists and requires a vehicle control. A second series is for androgen antagonists and requires a reference TP dose group using the TP dose selected from stage 2. In the latter series, a vehicle control group and a positive control of FLU + the selected TP dose are voluntary. The required test groups are illustrated in Table 4.

Table 4. Test Groups for Stage 4 of the Stimulated Weanling Validation:
Coded Dose Studies

•

•	• **Agonist Coded Samples**	• **Antagonist Coded Samples**
• Group A • Vehicle Control	• Vehicle	• Vehicle – **Voluntary** in antagonists series
• Group B • Negative Treatment Control	• Not applicable - provided by vehicle control (no additional group needed for agonist series)	• Selected TP from stage 2
• Group C	• Coded Agonist Test substance #1 [a]	• Selected TP from stage 2 + • Coded Antagonist Test substance #1 [a]
• Group D	• Coded Agonist Test substance #2 [a]	• Selected TP from stage 2 + • Coded Antagonist Test substance #2 [a]
• Group E	• Coded Agonist Test substance #3 [a]	• Selected TP from stage 2 + • Coded Antagonist Test substance #3 [a]
• Group F	• Coded Agonist	• Selected TP from stage 2 +

	Test substance #4 [a]	• Coded Antagonist Test substance #4 [a]
• Group G	• Not applicable	• Selected TP from stage 2 + • Coded Antagonist Test substance #5 [a]
• Group H	• Not applicable	• Selected TP from stage 2 + • Coded Antagonist Test substance #6 [a]
• Group I • Positive Treatment Group	• TP dose selected from stage 2 • **Voluntary**	• TP dose selected from stage 2 + FLU dose selected from stage 2 • **Voluntary**

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

OBSERVATIONS FOR ALL STAGES IN THE STIMULATED WEANLING VALIDATION

Clinical observations

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

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

Body weight and food consumption

53. Individual body weights will be recorded prior to start of treatment (to the nearest 1 g or to the nearest 0.1 g if that is normal practice of the laboratory with smaller animals), on each day of administration period and prior to necropsy. Group means and standard deviations will be calculated.

54. Food consumption should be generally observed and any significant changes recorded. It may be voluntarily recorded per cage, and an average value per animal calculated based on the number of animals per cage.

Necropsy

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

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

57. The four sex accessory tissues (VP, SV, LABC, COW) are mandatory measurements. Two additional tissue from the male reproductive tract, T and E, are optional measurements. All mandatory and optional tissues will be excised, carefully trimmed of excess adhering tissue and fat, and their fresh (unfixed) weights

determined. Each tissue should be handled with particular care to avoid the loss of fluids and to avoid desiccation, which may introduce significant errors and variability by decreasing the recorded weights.

58. Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Previous work has indicated a range of coefficient of variations that appears to differ based upon the proficiency of the laboratory. In a few cases, large differences in the absolute weights of the tissues such as the VP and COWS have been observed within a particular laboratory. Therefore, it is important that persons carrying out the dissection of the sex accessory tissues are familiar with standard dissection procedures for these tissues. A standard operating procedure (SOP) manual for dissection has been provided by the Lead Laboratory and was used in Phase-1 and Phase-2. This manual will remain the SOP reference for Phase-3. Careful training according to the SOP guide will minimise a potential source of variation in the study.

59. Each of the four mandatory sex accessory tissues as well as the optional paired testes and paired epididymes will be weighed without blotting to the nearest 0.1 mg, and the weights recorded for each identified animal.

60. Liver, paired kidney, and paired adrenal weights are other optional measurements. Again, tissues should be trimmed free of any adhering fascia and fat. The liver will be weighted to the nearest 0.1 g, and the paired kidneys and paired adrenals will be weighted to the nearest 0.1 mg. All weights will be recorded for each identified animal.

61. If the evaluation of each chemical requires necropsy of more animals than is reasonable for a single day, the starting date may be staggered on two consecutive days so that the necropsy can be staggered and the work burden of a single day reduced. In this case the work could be divided so that necropsy of 3 animals per treatment per day (1 cage) takes place on the first day with the dosing and necropsy being delayed by one day in the second half of the animals. That is, each group should be split so that half of the animals are necropsied on each day in order to control variability among the groups.

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

REPORTING

Data

63. Data will be reported individually (i.e. body weight, accessory sex tissue weights, optional measurements and other responses and observations) and for each group of animals (means and standard deviations). The data will be summarised in tabular form. The data will show the number of animals at the start of the test, the number of animals found dead during the test or found the test number of animals found showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration and severity.

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

Test report

65. The test report must include the following information:

Laboratory identification

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

Test substance:

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

Vehicle identity and supplier:

Test animals and procedures:

– Species/strain used;
– Source or supplier of animals, including full address;
– Number, age and sex of animals;
– Housing conditions (temperature, lighting, and so on), diet used, lot of diet, source of diet, bedding and source of bedding;
– Caging conditions and number of animals per cage;
– Age at receipt, age at start of test substance administration and time of acclimitisation;
– Individual weights of animals at the start of the study (to nearest 0.1 g);
– Randomization process and a record of the assignment to vehicle, reference, and test substance groups;
– Mean and standard deviation of the body weights for each group at the start of the study;
– Necropsy procedures, including means of exsanguination and any anesthesia; and

Results:

– Daily observations during administration, including:
  - Daily body weights (to the nearest 1 g),
  - Clinical signs (if any),
  - Test substance treatment (Yes or No) and the identify of that test substance,
  - Dose level and volume administered each day,
  - Time of dosing each day, and
  - Notes on food consumption or measurement of actual food consumption each day.
– On the day of necropsy, individual necropsy data on each animal including absolute sex accessory tissue weights, liver and body weights including the following:
  - Date of necropsy,
  - Animal ID,
  - Home Cage Number or ID,
  - Prosector,

  - Time of day necropsy performed,
  - Animal age, and
  - Order of animal killing and dissection at necropsy,
- Weights of all four mandatory sex accessory tissues and glands.
  - Ventral prostate (fresh weight – to the nearest 0.1 mg),
  - Seminal vesicles plus coagulating glands, including fluid (fresh weight – paired, to nearest 0.1 mg),
  - Levator ani plus bulbocavernosus muscle (fresh weight - to nearest 0.1 mg),
  - Cowper's glands (fresh weight – paired, to nearest 0.1 mg).
- Weights of additional male reproductive tissues.
  - Testes (fresh weight, paired, to nearest 0.1 mg),
  - Epididymes (fresh weight, paired, to nearest 0.1 mg),
- Weights of optional tissues, if performed.
  - Liver (optional – to nearest 0.1 g),
  - Kidney (optional – paired, to nearest 0.1 mg), and
  - Adrenal (optional – paired, to nearest 0.1 mg).

- General remarks and comments

Discussion

Conclusions

INTERPRETATION OF RESULTS

66. Statistical comparisons will be made for the different mandatory sex accessory tissues, the optional male reproductive tissues, and other optional tissues. Differences in weight will be considered statistically significant if $p < 0.05$. For androgen agonism, the test substance groups will be compared to the vehicle control. A statistically significant increase in tissue weight of the mandatory sex accessory tissues with the same tissue in the vehicle control will be considered consistent with the finding of 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 of the mandatory sex accessory tissues versus the same tissue in the positive control TP group will be considered consistent with a positive antagonist result. Statistically significant changes, positive or negative, in the tissues other than the mandatory sex accessory tissues will be noted and considered to be characteristic for the test substance, but not evidence for androgen agonism or antagonism.

LITERATURE

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

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

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

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

5. Hershberger, L., Shipley, E. and Meyer, R. (1953): Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method. Proc Soc Exp Biol Med. 83:175-180.

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

7. Stoker, T. E., Parks, L. G., Gray, L. E., Cooper, R. L. (2000): Endocrine-disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations. Crit. Rev. Toxicol. 30(2):197-252.

8. Ojeda, S. R., Urbanski, H. F. (1994): Puberty in the rat. In *The physiology of reproduction. Second edition.* E. Knobil and J. D. Neill, eds. Raven Press, NY. pp.363-409.

9. Monosson, E., Kelce, W. R., Lambright, C. R., Ostby, J., and Gray, L. E., Jr. (1999): Peripubertal exposure to the antiandrogenic fungicide, vinclozolin, delays puberty, inhibits the development of androgen-dependent tissues, and alters androgen receptor function in the male rat. Toxicol. Ind. Health 15:65-79.

10. Ashby, J., Lefevre, P. (2000): The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti-androgens, oestrogens and metabolic modulators. J. Appl. Toxicol. 20:35-47.

11. OECD (2002): Final OECD report of the work towards the validation of the rat Hershberger assay: Phase-1, Androgengic response to testosterone propionate and anti-androgenic effects of Flutamide. ENV/JM/TG/EDTA(2002)1/REV2. October, 2002. Paris, France

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