INTEGRATED SUMMARY REPORT

for

VALIDATION OF 15-DAY INTACT ADULT MALE RAT ASSAY AS A POTENTIAL SCREEN IN THE ENDOCRINE DISRUPTOR SCREENING PROGRAM TIER-1 BATTERY

August 29, 2007

U.S. Environmental Protection Agency
Office of Science Coordination and Policy
Washington, D.C.

This page left blank intentionally.

TABLE OF CONTENTS

APPENDICES

TABLES

1 **ABBREVIATIONS**
2

Abbreviation	Definition
ADME	Absorption, distribution, metabolism, and excretion
ANOVA	Analysis of variance
AR	Androgen receptor
ASG	Accessory sex gland
AWA	Animal Welfare Act
C8	Ammonium perfluorooctanoate
CFR	Code of Federal Regulations
CO_2	Carbon dioxide
CRO	Contract research organization
CV	Coefficient of variation
DDE	Dichlorodiphenyldichloroethylene
DHT	Dihydrotestosterone
DRP	Detailed Review Paper
DSL	Diagnostic Systems Laboratory
EAC	Endocrine-active compound
EAT	Estrogen, androgen, and thyroid
ECVAM	European Centre for the Validation of Alternative Methods
EDMVAC	Endocrine Disruptor Methods Validation Advisory Committee
EDMVS	Endocrine Disruptor Methods Validation Subcommittee
EDSP	Endocrine Disruptor Screening Program
EDSTAC	Endocrine Disruptor Screening and Testing Advisory Committee
EPA	Environmental Protection Agency (U.S.)
ER	Estrogen receptor
FFDCA	Federal Food Drug and Cosmetic Act
FIFRA	Federal Insecticide, Fungicide, Rodenticide Act
FRN	Federal Register Notice
FSH	Follicle-stimulating hormone
g	Gram(s)
GLP	Good Laboratory Practices
GMP	Good Manufacturing Practices
GnRH	Gonadotropin-releasing hormone
H&E	Hematoxylin and eosin
HPG	Hypothalamus-pituitary-gonadal
HPT	Hypothalamus-pituitary-thyroidal
ICCVAM	Interagency Coordinating Committee for the Validation of Alternative Methods
IP	Intraperitoneal
ISR	Integrated Summary Report
kg	Kilogram(s)
LH	Luteinizing hormone
MAFF	Ministry of Agriculture, Forestry and Fisheries (Japan)
MOA	Mode(s) or mechanism(s) of action

Abbreviation	Definition
MSDS	Material Safety Data Sheet
MTD	Maximum tolerated dose
ND	Not determined
NIEHS	National Institute of Environmental Health Sciences
OECD	Organisation for Economic Co-Operation and Development
PDF	Portable Document Format (Adobe Systems®)
PRL	Prolactin
PTI	1-methyl-3-propylimidazole-2-thione
QA	Quality assurance
QAPP	Quality Assurance Project Plan
QC	Quality control
RIA	Radioimmunoassay
SAP	Scientific Advisory Panel (U.S. EPA)
SD	Sprague-Dawley (rat strain); Standard deviation (statistical analysis)
SE	Standard error
SOP	Standard Operating Procedure
SVCG	Seminal vesicles and coagulating glands (with fluid)
T_3	Triiodothyronine
T_4	Thyroxine
TD	Test day
TSCA	Toxic Substances Control Act
TSH	Thyroid stimulating hormone
U.S.C.	United States Code
UDP	Uridine diphosphate

This page left blank intentionally.

1.0 Introduction

1.1 Purpose of the EDSP

Section 408(p) of the Federal Food Drug and Cosmetic Act (FFDCA) requires the U.S. Environmental Protection Agency (EPA) to:

> *develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have an effect in humans that is similar to an effect produced by a naturally occurring estrogen, or other such endocrine effect as the Administrator may designate [21 U.S.C. 346a(p)].*

Subsequent to passage of the Act in 1996, the EPA formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), a committee of scientists and stakeholders that was charged by the EPA to provide recommendations on how to implement its Endocrine Disruptor Screening Program (EDSP). The EDSP is described in detail at the following website: http://www.epa.gov/scipoly/oscpendo/. Upon recommendations from the EDSTAC (EPA, 1998a), the EPA expanded the EDSP using the Administrator's discretionary authority to include the androgen and thyroid hormonal systems as well as wildlife.

1.2 Definition of an Environmental Endocrine Disruptor

An EPA Risk Assessment Forum was established to promote scientific consensus on risk assessment issues and to ensure that any consensus is incorporated into appropriate risk assessment guidance. The Forum released a report in 1997 (EPA, 1997) that addressed the hypothesis that certain chemicals may disrupt the endocrine system. In the report, an environmental endocrine disruptor was defined as:

> *an exogenous agent that interferes with the synthesis, secretion, transport, binding, action or elimination of natural hormones in the body that are responsible for the maintenance of homeostasis, reproduction, development, and/or behavior*.

1.3 EDSP Tiered Approach

The EPA accepted the EDSTAC's recommendations for a two-tier screening program as proposed in a Federal Register Notice (FRN) in 1998 (EPA, 1998b). The purpose of Tier 1 is to identify the potential of chemicals to interact with the estrogen, androgen, or thyroid (EAT) hormonal systems. A negative result in Tier 1 would be sufficient to put a chemical aside as having low to no potential to cause endocrine disruption, whereas a positive result would require further testing in Tier 2. The purpose of Tier 2 is to definitively identify and characterize the potential hazard on the endocrine system and to provide risk assessment based, in part, on dose-response relationships. Tier 2 is expected to comprise multigeneration tests in species representative of various taxa (i.e., mammals, birds, fish, amphibians, and invertebrates).

1.4 The Tier-1 Battery

The EDSTAC (EPA, 1998a) concluded that a Tier-1 battery should be comprised of a suite of complementary screening assays having the following characteristics:

- Maximum sensitivity to minimize false negatives while permitting an as yet undetermined, but acceptable, level of false positives.
- Range of organisms representing known or anticipated differences in metabolic activity and include assays from representative vertebrate classes to reduce the likelihood that important pathways for metabolic activation or detoxification of parent substances or mixtures are not overlooked.
- Capacity to detect all known modes of action (MOAs) for the endocrine endpoints of concern. All chemicals known to affect the action of EAT hormones should be detected.
- Range of taxonomic groups among the test organisms. There are known differences in endogenous ligands, receptors, and response elements among taxa that may affect the endocrine activity of chemical substances or mixtures.
- Diversity among the endpoints and within and among assays to reach conclusions based on "weight-of-evidence" considerations. Decisions based on the screening battery results will require weighing the data from several assays.
- Inexpensive, quick, and easy to perform.

To detect chemicals that may affect the EAT hormonal systems through any one of the known MOAs (e.g., interruption of hormone production or metabolism, binding of the hormone with its receptor, interference with hormone transport) the EDSTAC recommended the *in vitro* and *in vivo* assays shown in Table 1 for inclusion in the Tier-1 screening battery. In addition, the EDSTAC recognized there were other combinations of screening assays that may be suitable for a Tier-1 battery and, therefore, recommended that the EPA validate the alternative screening assays shown in Table 2. Note, a Tier-1 battery is expected to be proposed by the EPA after proposed screening assays have completed the peer review process.

1.5 Validation

As noted, Section 408(p) of the FFDCA requires the EPA to use validated test systems. Validation has been defined as *"the process by which the reliability and relevance of a test method is evaluated for a particular use"* (OECD, 1996; NIEHS, 1997).

> *Reliability* is defined as the reproducibility of results from an assay within and between laboratories.
>
> *Relevance* describes whether a test is meaningful and useful for a particular purpose (OECD, 1996). For Tier-1 EDSP assays, relevance can be defined as the ability of an assay to detect chemicals with the potential to interact with the EAT hormonal pathways.

Federal agencies are also instructed by the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) Authorization Act of 2000 to ensure that new and revised test methods are valid prior to their use.

Table 1. Tier-1 *in vitro* and *in vivo* screening assays recommended by the EDSTAC.

Assays	Reasons for consideration
Estrogen receptor (ER) binding or transcriptional activation	A sensitive *in vitro* assay to detect chemicals that may affect the endocrine system by binding to the ER.
Androgen receptor (AR) binding or transcriptional activation	A sensitive *in vitro* assay to detect chemicals that may affect the endocrine system by binding to the AR.
In vitro steroidogenesis	A sensitive *in vitro* assay to detect chemicals that interfere with the synthesis of the sex steroid hormones.
Uterotropic (rat)	An *in vivo* assay to detect estrogenic chemicals. It offers the advantage over the binding assay of incorporating absorption, distribution, metabolism, and excretion (ADME)
Hershberger (rat)	An *in vivo* assay to detect androgenic and anti-androgenic chemicals. It offers the advantage over the binding assay of incorporating ADME and differentiating between AR agonists and antagonists.
Pubertal female (rat)	An assay to detect chemicals that act on estrogen or through the hypothalamus-pituitary-gonadal (HPG) axis that controls the estrogen and androgen hormone systems. It is also enhanced to detect chemicals that interfere with the thyroid system.
Frog metamorphosis	A sensitive assay for detection of chemicals that interfere with the thyroid hormone system.
Fish screen	Fish are the furthest removed from mammalians among vertebrates both from the standpoint of evolution—their receptors and metabolism are different from mammals—and exposure/habitat, since they would be subject to exposure through the gills, whole body, and diet. Thus, the fish assay would augment information found in the mammalian assays and would be more relevant than the mammalian assays in triggering concerns for fish.

Table 2. Alternative *in vitro* and *in vivo* assays recommended for the Tier-1 Screening Battery by the EDSTAC.

Assays	Reasons for consideration
In vitro aromatase	The aromatase assay detects chemicals that inhibit aromatase and would be needed if either of the two following assays using males were substituted for the female pubertal assays. The male is not believed to be as sensitive to alterations in aromatase as the female and would not therefore be sufficient to detect interference with aromatase in the screening battery.
Pubertal male (rat)	The assay detects chemicals that act on androgen or through the HPG axis that controls the estrogen and androgen hormone systems. It is also enhanced to detect chemicals that interfere with the thyroid system. This assay could in part substitute for the female pubertal assay.
Adult male (rat)	The assay is also designed to detect chemicals that act on androgen or through the HPG axis that controls the estrogen and androgen hormone systems. It is also enhanced to detect chemicals that interfere with the thyroid system. This assay could in part substitute for the female pubertal assay.

In general, the EPA is following a five-part or stage validation process outlined by the ICCVAM (NIEHS, 1997). The EPA believes that it is essential to recognize that this process was specifically developed for *in vitro* assays intended to replace *in vivo* assays. The fundamental problem confronting the EPA is how to adapt and work with this process for rodent and ecological *in vivo* assays in Tiers 1 and 2 that have no suitable *in vitro* substitute at this time.

Nonetheless, the stages of the *Validation Process* outlined by the ICCVAM are as follows:

First Stage - *Test Development*, an applied research function which culminates in an initial protocol. As part of this phase, the EPA prepares a Detailed Review Paper (DRP) or an analogous document (e.g., Background Review Document) to explain the purpose of the assay, the context in which it will be used, and the scientific basis upon which the assay's protocol, endpoints, and relevance rest. The DRP reviews the scientific literature for candidate protocols and evaluates them with respect to a number of considerations, such as whether the candidate protocols meet the assay's intended purpose, the costs and other practical considerations. The DRP also identifies the developmental status and questions related to each protocol; the information needed to answer the questions; and, when possible, recommends an initial protocol for the initiation of the second stage of validation.

Second Stage - *Prevalidation* in which the protocol is refined, optimized, standardized and initially assessed for transferability and performance. Several different types of studies are conducted during this second phase depending upon the state of development of the method and the nature of the questions that the protocol raises. The initial assessment of transferability is generally a trial in a second laboratory to determine that another laboratory besides the lead laboratory can follow the protocol and execute the study.

Third Stage - *Inter-laboratory Validation* studies are conducted in independent laboratories with an optimized, standardized protocol. The results of these studies are used to primarily determine inter-laboratory variability and to set or cross-check performance criteria.

Fourth Stage - *Peer Review*, an independent scientific review by qualified experts.

Fifth Stage - R*egulatory Acceptance*, adoption for regulatory use by an agency. The EPA has developed extensive guidance on the conduct of peer reviews because the Agency believes that peer review is an important step in ensuring the quality of science that underlies its regulatory decisions (EPA, 2006).

Considering that the 15-day intact adult male rat assay as summarized in this report was initially developed by the chemical industry (DuPont) to identify MOAs of chemicals in support of product registration, Stage 1 of the validation process did not involve a DRP. However, test development is discussed throughout Section 2 and in more detail in a published review article (O'Connor *et al.*, 2002c). Stages 2 and 3 of the validation process are summarized in this report, and Stage 4 is the ongoing peer review process.

Criteria for the validation of alternative test methods *(in vitro* methods designed to replace animal tests in whole or in part) have generally been agreed upon in the United States by the

ICCVAM, in Europe by the European Centre for the Validation of Alternative Methods (ECVAM), and internationally by the Organization for Economic Co-Operation and Development (OECD). These criteria as stated by ICCVAM (NIEHS, 1997) are as follows:

1. The scientific and regulatory rationale for the test method, including a clear statement of its proposed use, should be available.
2. The relationship of the endpoints determined by the test method to the *in vivo* biologic effect and toxicity of interest must be addressed.
3. A formal detailed protocol must be provided and must be available in the public domain. It should be sufficiently detailed to enable the user to adhere to it and should include data analysis and decision criteria.
4. Within-test, intra-laboratory and inter-laboratory variability and how these parameters vary with time should have been evaluated.
5. The test method's performance must have been demonstrated using a series of reference chemicals preferably coded to exclude bias.
6. Sufficient data should be provided to permit a comparison of the performance of a proposed substitute test to that of the test it is designed to replace.
7. The limitations of the test method must be described (e.g., metabolic capability).
8. The data should be obtained in accordance with Good Laboratory Practices (GLPs).
9. All data supporting the assessment of the validity of the test methods including the full data set collected during the validation studies must be publicly available and, preferably, published in an independent, peer-reviewed publication.

The EPA has adopted these various validation criteria for the EDSP as described (EPA, 2007b). Although attempts have been made to thoroughly comply with all validation criteria, the various *in vitro* and *in vivo* screening assays are not replacement assays (Validation Criterion No. 6). Many of them are novel assays; consequently, large data bases do not exist as a reference to establish their predictive capacity (e.g., determination of false positive and false negative rates). It is expected that review of results from testing of the first group of 73 chemicals (EPA, 2007a) that was recommended by the Scientific Advisory Panel (SAP) (EPA, 1999) will allow a more complete assessment of the performance of the screening assays in a Tier-1 battery over time.

For technical guidance in developing and validating the various Tier-1 screens and Tier-2 tests, the EPA chartered two federal advisory committees: the Endocrine Disruptor Methods Validation Subcommittee, or EDMVS (from 2001 to 2003), and the Endocrine Disruptor Methods Validation Advisory Committee, or EDMVAC (from 2004 to 2006). These committees, composed of scientists from government, academia, industry, and various interest groups, were charged to provide expert advice to the EPA on development and validation of assay protocols. The EPA also cooperates with member countries of the OECD to develop and validate assays of mutual interest to screen and test for endocrine effects.

It should be remembered that even though assays are being developed and validated individually and peer reviewed on an individual basis (i.e., their strengths and limitations are being evaluated as stand-alone assays), Tier-1 assays will be used in a battery of complementary screens. An individual assay may serve to strengthen the weight of evidence in a determination (e.g., positive results in an ER binding assay in conjunction with positive results in the uterotropic and pubertal

female assays would provide a consistent signal for estrogenicity) or to provide coverage of MOAs not addressed by other assays in the battery. Information supporting the validation of an individual assay may be used at a later date by the Federal Insecticide, Fungicide, Rodenticide Act (FIFRA) SAP for peer review of the EPA's recommendations for a Tier-1 battery. It is expected that peer review of the Tier-1 battery will focus, in part, on the extent of coverage and overlap the suite of assays will have with one another in detecting endocrine-related effects associated with the EAT hormonal systems.

1.6 Purpose of this Integrated Summary Report

This integrated summary report (ISR) for the 15-day intact adult male rat assay was prepared by the EPA with contributions provided by scientists from the chemical industry. The ISR will be reviewed by a panel of independent scientists who will comment on the assay's biological strengths, weaknesses, and practicability in accordance with specific charges submitted by the EPA with the understanding that this bioassay may be one of a suite of *in vitro* and *in vivo* assays selected by the EPA to be considered in an EDSP Tier-1 screening battery.

As part of the peer review package, this ISR is the main focus of peer review of the intact adult male assay and is expected to facilitate the process by: 1) providing a historical overview of the bioassay (Section 2), 2) presenting key prevalidation efforts used to establish relevance of the bioassay (Section 3), 4) introducing the final, standardized assay protocol (Section 4) and 5) presenting and discussing the results of an EPA-sponsored inter-laboratory validation study used to determine reliability and feasibility of the bioassay (Section 5). The last section of the body of the report (Section 6) addresses the bioassay in relation to the various validation criteria prescribed by the EDSP.

2.0 Historical Overview of the 15-Day Intact Adult Male Rat Assay

This section provides a brief overview of the intact adult male rat assay. A comprehensive description is not provided since most individual studies (O'Connor *et al.*, 1998a,b; 1999a,b; 2000a,b; 2002a,b) and a review of the study rationale (O'Connor *et al.*, 2002c) are readily available as published articles in peer-reviewed scientific journals. This section introduces the 15-day intact adult male rat assay from the perspective of a potential alternative *in vivo* assay in the EDSP Tier-1 screening battery as proposed by the EDSTAC.

2.1 Overall Purpose

According to numerous publications as reviewed by O'Connor *et al.*, (2002c), the 15-day intact adult male rat assay has been developed to detect ER agonists/antagonists, AR agonists/antagonists, progesterone agonists/antagonists, steroid biosynthesis inhibitors, gonadotropin and thyroid modulators either directly or indirectly by altering the HPG or -hypothalamus-pituitary-thyroidal (HPT) axes, and prolactin (PRL) modulators through neuroendocrine pathways.

Briefly, the design of the intact adult male rat assay consists of multiple endpoints, principally, terminal weights of primary and secondary sex organs and thyroid gland; histomorphology of the

testes, epididymides, and thyroid; and serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones. Results of the comparisons of these endpoints between control and treated groups at three dose levels (n=15 rats/group) administered by oral gavage are evaluated using a weight-of-evidence approach within the bioassay to determine whether a chemical has a positive effect on the EAT hormonal systems. Criteria for interpretation of endocrine-mediated effects within the bioassay are presented in Section 3.3.

The extent of the diversity of this assay to detect effects on the EAT hormonal system using a variety of endocrine-active compounds (EACs) has been hypothesized, tested, and reported in published peer-reviewed scientific journals (O'Connor *et al.*, 1998a,b; 1999a,b; 2000a,b; 2002a,b,c). Thus, the purpose of the intact adult male screening assay is to detect various MOAs, especially AR agonists/antagonists, steroid biosynthesis inhibitors, gonadotropin and thyroid modulators either directly or indirectly through intact HPG or HPT axes using a weight-of–evidence approach within the bioassay. In addition, since this assay is a candidate for an EDSP Tier-1 battery, results from within the bioassay are expected to contribute to the results of other assays in the battery and, using a weight-of-evidence approach within the battery, determine whether a chemical substance has a positive or negative effect on the EAT hormonal systems.

2.2 Basis for Initial Development by Industry

The 15-day intact adult male rat assay was initially developed by the chemical industry (DuPont) to identify MOAs of several chemicals in support of product registration. Ammonium perfluorooctanoate (C8), a peroxisome proliferator, was shown to induce aromatase activity and increase serum estradiol concentrations in 2-week studies using intact adult male rats similar to the adult male assay proposed for the EDSP Tier-1 battery (Cook *et al*., 1992). In addition, linuron (a herbicide) was identified as being a weak AR antagonist, a MOA which explains its ability to induce Leydig cell tumors in long-term rodent studies (Cook *et al*., 1993). In a mechanistic study, 1-methyl-3-propylimidazole-2-thione (PTI) was shown to alter thyroid function by directly inhibiting thyroid hormone synthesis and by enhancing thyroid hormone excretion via uridine diphosphate (UDP)-glucuronyltransferase induction (Biegel *et al*., 1995). Changes in serum concentrations of thyroid stimulating hormone (TSH), triiodothyronine (T_3), and thyroxine (T_4) were seen as early as 1 week after the beginning of treatment and were also seen at 3 and 13 weeks post-treatment; thus, demonstrating that a 2-week study design could potentially detect compounds that alter thyroid function (Biegel *et al*., 1995). The MOA of the proprietary compound that produced thyroid follicular cell tumors in the 2-year rat study via inhibition of 5'-monodeiodinase and induction of UDP-glucuronyltransferase was also identified using a study design similar to the 15-day intact adult male rat assay. In addition, the MOA for three other proprietary compounds (i.e., an aromatase inhibitor, a mixed testosterone/aromatase inhibitor, and an AR antagonist) which also produced Leydig cell tumors in 2-year rat bioassays were identified with a 2-week study design using intact adult male rats, (O'Connor *et al.*, Personal communication).

2.3 Strengths and Challenges of the Bioassay

2.3.1 Strengths

- Allows for a high-order neuroendocrine assessment of male reproductive and thyroid function due to the use of an intact endocrine system (i.e., HPG and HPT axes).

- Advances scientific understanding through its MOA and, perhaps, mechanistic approach (i.e., measurement of serum concentrations of reproductive steroids, gonadotropins and thyroid hormones).

- Provides MOA data (e.g., differentiates between receptor and nonreceptor-mediated effects) that can be used to tailor the design of more definitive Tier-2 tests to focus on selective endpoints to accurately identify potential hazards, define dose responses, and determine the level of risk of potential endocrine disruptors.

- Allows for the maximum tolerated dose (MTD) to be readily defined since mature animals are less susceptible to marked changes in growth and less susceptible to nonspecific alterations in endpoints secondary to bodyweight changes.

- Flexible for modifying or adding apical, histological and hormonal endpoints in the context of a single assay to detect other potential endocrine-related effects as future application may dictate.

- Complies with the basic principles of good laboratory animal practice (i.e., three R's - Reduce, Refine, and Replace), specifically through the effective use of a minimal number of animals.

- Complies with the expected simplicity and rapidity of a screen prescribed by the EDSTAC since the in-life portion of the assay is readily applied and minimal in duration.

2.3.2 Technical Challenges

Considering that blood hormone measurements are an integral aspect of the adult male assay, one of the most technically challenging features of the bioassay is inclusion of serum hormone analyses, especially since hormone measurements are not typically done in *in vivo* toxicological studies. Hormone assay kits are commercially available from a variety of manufacturers [e.g., Amersham Corporation (Arlington Heights, IL); Diagnostic Products Corporation (Los Angeles, CA); Diagnostic Systems Laboratories (or DSL, Webster, TX); Linco Research, Inc. (St. Charles, MO); Peninsula Laboratories (Belmont, CA); Polymedco, Inc. (Cortlandt Manor, NY)], some of which have been specifically designed for use with rat serum. For those that have been specifically designed for human serum (e.g., many steroid assay kits), rat reference and quality control (QC) standards can be readily prepared in relevant assay media and adapted into the various kit designs. Each hormone assay kit is supplied with detailed instructions and performance criteria as well as contact information for technical support from the manufacturers. Hence, laboratories with individuals knowledgeable of the concept and capable of running

immunoassays are expected to competently perform the hormone assays under government GLP/Good Manufacturing Practices (GMP) conditions and interpret whether the kit reference standards and QC samples are in accordance with the performance criteria prescribed for a particular hormone according to the manufacturers. Alternatively, since most contract toxicology laboratories are not necessarily experienced or equipped to run hormone assays, an endocrinology laboratory that is GLP/GMP compliant and that routinely runs hormone assays with human or animal samples may need to be contracted to analyze the experimental rat serum samples.

The conditions encompassing blood collection on the last day of the study are also important to minimize the extent of the variation associated with serum hormone analyses. *First*, to minimize the impact of variability that may be associated with cage-transport stress-induced hormone changes (e.g., PRL), animals are moved from the in-life room to the necropsy room a minimum of 1 hour before euthanasia. *Second*, the process of euthanasia is also orchestrated to minimize stress by exposing the animals to carbon dioxide (CO_2) for a defined time sufficient for anesthesia prior to cardiac puncture or decapitation. *Third*, to minimize the impact of variability that may be associated with diurnal variations in serum hormone concentrations, the timing of terminal euthanasia is defined so that blood collection is completed within a 2- to 3-hour window after the last treatment during the early morning hours (e.g., 0700 to 1000 hours). *Fourth,* the variability of hormone assay results is also minimized by stratifying euthanasia across dose groups and by having a relatively large number of animals per dose group.

In regard to the latter, given the extrinsic (e.g., hormone assay kits and operator-related variability) and intrinsic (e.g., hormone pulsatility) variability in serum hormone concentrations, the number of animals per dose group considered sufficient to detect significant endocrine-mediated changes in hormone concentrations in a single sample collected at necropsy between treated and control groups was determined by using power analysis of dose-group size as determined by O'Connor *et al.,* (Personal communication). Using the intact adult male rat assay with 15 rats per dose group, there is a >99% chance of detecting a 50% change in serum concentrations for each of the hormones in the treated groups as statistically significant from the control group. Considering a 25% change in serum concentrations for luteinizing hormone (LH), follicle-stimulating hormone (FSH), TSH, T_3, and T_4, there is a 92 to 100% chance of detecting a statistically significant difference between control and treated groups and, for testosterone, dihydrotestosterone (DHT), estradiol and PRL, there is a 61 to 83% chance of detecting a statistically significant difference.

2.4 Recognized and Recommended for Standardization and Validation in the EDSP

The EDSTAC report (EPA, 1998a) recommended to the EPA a primary Tier-1 screening battery and two alternative batteries (Table 3). All three EDSTAC recommended batteries are expected to incorporate a suite of complimentary *in vitro* and *in vivo* mammalian and non-mammalian screening assays to determine endocrine-mediated effects (EAT) of potential endocrine disruptors. In each of the proposed batteries, the *in vitro* ER and AR binding or transcriptional activation assays as well as the non-mammalian assays, are required. The EDSTAC identified one *in vitro* and three *in vivo* assays as possible substitutes, if properly developed, standardized, and validated for some of the component assays in the primary battery. In the first alternate

screening battery, the EDSTAC suggested that the Hershberger, female pubertal and *in vitro* steroidogenesis assays in the primary battery could potentially be replaced by the intact adult male rat assay. The suite of validated assays that will constitute a Tier-1 battery has not yet been determined, but is expected to be addressed by the EPA following peer review of individual assays proposed for Tier 1. Because the intact adult male assay is a rat model system having an intact HPG and HPT axes which has been shown to detect a variety of endocrine activities its potential role within the EDSP Tier-1 battery would be to complement and expand on the results of other *in vivo* and *in vitro* screening assay results by identifying endocrine-specific effects using multiple endpoints (i.e., weights of primary and secondary sex organs and thyroid gland, histomorphology of the testes, epididymides and thyroid, and serum concentrations of reproductive steroids, gonadotropins and thyroid hormones).

Table 3. EDSTAC proposed *in vitro* and *in vivo* screening assays for EDSP

Primary Tier-1 Screening Battery	**Alternate Tier-1 Screening Battery No. 1**	**Alternate Tier-1 Screening Battery No. 2**
***In vitro* assays**	***In vitro* assays**	***In vitro* assays**
ER binding/transactivation	ER binding/transactivation	ER binding/transactivation
AR binding/transactivation	AR binding/transactivation	AR binding/transactivation
Steroidogenesis assay	Placental/Recombinant Aromatase	Placental/Recombinant Aromatase
***In vivo* assays**	***In vivo* assays**	***In vivo* assays**
Uterotropic (rat)	Uterotropic (rat)	Uterotropic (rat)
Hershberger (rat)	Intact adult male (rat)	Pubertal male (rat)
Pubertal female (rat)		
Frog metamorphosis	Frog metamorphosis	Frog metamorphosis
Fish screen	Fish screen	Fish screen

3.0 Prevalidation

Prevalidation in the context of the EDSP (EPA, 2007b) refers to the initial stages of the validation process where a semi-standard protocol is further developed and used in a limited number of laboratories to test the logistical and technical aspects of the bioassay, optimize conditions, and establish biological relevance by designating target tissues and defining other specific and reliable endpoints using various EACs with different MOAs (Section 1.5). In addition, criteria were developed for immediate and future interpretation of results to differentiate overt toxicity from endocrine-mediated effects. Hence, prevalidation with endocrine-positive and -negative test chemicals has led to an optimized, standardized protocol (Section 4) for the second stage of the validation process involving inter-laboratory studies (Section 5).

3.1 Relevance of the Bioassay

Numerous EACs (one negative and 28 positive test chemicals) with different MOAs have been examined in the 15-day intact adult male assay (Table 4) using a protocol initially developed by industry in which dosing was done by oral gavage (13 chemicals), intraperitoneal (IP) injection (11 chemicals) or both routes of administration (5 chemicals). Positive test chemicals with known endocrine activities considered relatively strong or weak were used during prevalidation

Table 4. Test chemicals run in the 15-day intact adult male rat assay administered orally or intraperitoneally.

No.	Chemical	Endocrine Activity[1]	Oral Route[2]	IP Route[2]	Laboratory[3]
1	17β-estradiol	ER agonist (full or potent)	--	Pos[4]	DuPont
2	Coumestrol	ER agonist (weak or partial)	--	Neg[4]	DuPont
3	Methoxychlor	ER agonist (weak or partial)	Pos	--	RTI
4	Genistein	ER agonist (weak or partial)	Neg[4]	--	Syngenta
5	Nonylphenol	ER agonist (weak or partial)	Pos	--	BASF
6	ICI-182,780	ER antagonist	--	Pos	DuPont
7	Testosterone	AR agonist	--	Pos	DuPont
8	Methyltestosterone	AR agonist	Pos	--	WIL
9	Flutamide	AR antagonist (full or potent)	Pos	Pos	DuPont, Dow
10	*p,p'*-DDE	AR antagonist (weak or partial)	Pos	Pos	DuPont, Dow
11	Vinclozolin	AR antagonist (weak or partial)	Pos	--	DuPont
12	Cyproterone Acetate	AR antagonist (weak or partial)	Pos	--	DuPont
13	Linuron	AR antagonist (weak or partial)	Pos	--	DuPont, RTI
14	Di-n-butyl phthalate	Anti-androgen (non-receptor mechanism)	Pos	--	DuPont
15	Progesterone	PR agonist	--	Pos	DuPont
16	Mifepristone (RU486)	PR antagonist	--	Pos	DuPont
17	Apomorphine	D_2 receptor agonist	--	Neg	DuPont
18	Haloperidol	D_2 receptor antagonist	--	Pos	DuPont
19	Reserpine	Dopamine depletory (catecholamine depletion)	--	Pos	DuPont
20	Phenobarbital	Thyroid hormone excretion enhancer	Pos	Pos	DuPont, Dow, RTI
21	Oxazepam	Thyroid hormone excretion enhancer	Pos	--	DuPont
22	Propylthiouracil	Thyroid hormone synthesis inhibitor	Pos	Pos	DuPont
23	Propylimidazole-2-thione (PTI)	Thyroid hormone synthesis inhibitor	Pos	--	DuPont
24	Finasteride	5α-Reductase inhibitor	--	Pos	DuPont
25	Ketoconazole	Testosterone biosynthesis inhibitor	Pos	Pos	DuPont
26	Anastrozole	Aromatase inhibitor	--	Pos	DuPont
27	Fadrozole	Aromatase inhibitor	Pos	--	DuPont
28	Ammonium perfluorooctanoate	Aromatase inducer	Pos	--	DuPont
29	Allyl Alcohol	Nonendocrine (hepatotoxin)	Neg	--	DuPont

[1] Based on results from the scientific literature and intact adult male rat assay.
[2] Positive and negative effects were determined by comparing the pattern of effects observed in the intact male assay with the expected pattern of effects for the known MOA of the positive test materials (O'Connor *et al.*, 2002c).
[3] Linuron, Phenobarbital, flutamide and *p,p'*-DDE were run at different times in DuPont, Dow and RTI laboratories.
[4] Positive and negative responses for some chemicals are discussed in Sections 3.1.1 and 3.1.2, respectively.

of the intact adult male assay. A negative test chemical known only to induce non-endocrine effects on the liver was also used to assess the specificity of the bioassay. Thus, throughout prevalidation, the intact adult male assay has been run with 29 different test chemicals at various times in six different laboratories (four chemical industry laboratories and two different contract research organizations, or CRO laboratories). In some instances the same chemicals were tested in more than one laboratory at different times as shown in Table 4.

3.1.1 Positive Test Chemicals

Estrogens and estrogen-like chemicals: The relatively potent ER agonist, 17β-estradiol (Table 4), was readily detected in the adult male assay (O'Connor *et al.*, 1998b). Rats given 50 μg/kg/day of 17β-estradiol by IP injection had statistically significant decreases in terminal body weights (12% lower than controls), absolute testes and epididymides weights, and relative seminal vesicle, prostate and accessory sex gland (ASG) weights. Histopathological changes were noted in the testis and epididymis, including spermatocyte degeneration, spermatid retention, and interstitial cell and epididymal atrophy. At 7.5 μg/kg/day of 17β-estradiol, relative seminal vesicle and ASG weights were significantly decreased in the absence of significant body weight changes (3%). Histopathologically, slight interstitial cell atrophy was seen. Statistically significant decreases in serum testosterone, DHT, and LH and increased PRL concentrations were seen at all dose levels of 17β-estradiol (≥1 μg/kg/day).

Methoxychlor and nonylphenol, two xenobiotics, and genistein and coumestrol, two phytoestrogens that are reported to have relatively weak estrogen-like activity, also were evaluated in the adult male assay (Table 4). At 50 mg/kg/day of methoxychlor, rats had statistically significant decreases (12% relative to controls) in feed consumption and terminal body weight (EPA, 2005). At necropsy, decreased prostate and seminal vesicle sizes were observed grossly in two animals, and statistically significant decreases in relative seminal vesicle and ASG weights were recorded in the 50 mg/kg/day dose group. These organ weight effects can be attributed to endocrine-mediated effects rather than nonspecific body weight effects on the basis of: 1) a similar decrease in final body weight (12%) without an effect on organ weight endpoints in the methoxychlor group treated with 37.5 mg/kg/day; and 2) that relative (organ-to-final body weight ratio) seminal vesicle and ASG weights were not affected even with a 26% decrease in terminal body weight in feed-restricted rats compared to rats fed *ad libitum* (O'Connor *et al.*, 2000a) as discussed in Section 3.3.

Nonylphenol was administered to rats at 5, 20, 80, and 200 mg/kg/day for 15 day (Mellert, 2003). At 200 mg/kg/day, there was statistically significant decreases in feed consumption, final body weight (93% of control), relative prostate, seminal vesicle, and ASG weights, and serum testosterone and estradiol concentrations. Histopathologically, a portion of animals exhibited multiple degenerative foci in the seminiferous epithelium. Relative liver weights were increased and coupled with centrilobular hepatic hypertrophy. At 80 mg/kg/day, there were increased relative liver weights, decreased serum testosterone (34% less than controls) and a statistically significant decrease in estradiol concentrations.

The weak phytoestrogen, genistein, was administered by oral gavage at doses of 0, 50, 120, 400, or 1000 mg/kg/day (Milburn, 2004). While high-dose rats had statistically significant decreases

in body weights (96% of control) and feed consumption, treatment-related effects on clinical signs, target organ weights and histomorphology were not detected. Serum hormone concentrations were also not significantly altered. Although there was a significant decrease in final body weight relative to controls (4%), it was concluded that the limit dose of 1000 mg/kg/day was not enough to achieve a MTD.

Another relatively weak phytoestrogen, coumestrol, was administered by IP injection at doses of 0, 0.1, 0.5, 1.0, or 2.5 mg/kg/day (O'Connor *et al.*, 2000a). Coumestrol did not significantly alter final body weights, organ weights, or histomorphology of target organs; again, it was concluded that an MTD was not achieved. However, coumestrol produced hormonal changes consistent with an ER agonist MOA (i.e., statistically significant decreases in testosterone, estradiol, and DHT, and increased PRL) as reported by O'Connor *et al.*, (2002c). Although equivocal, these hormonal changes were consistent with estrogen-like compounds. Coumestrol has been positively detected in the ER binding assay (Yamasaki *et al*., 2003a) and uterotropic assay through interactions with uterine ERs and ARs (Schmidt and Katzenellenbogen, 1979; O'Connor *et al.,* 2002d; Yamasaki *et al*., 2003a), which are potential Tier-1 screening assays.

The combined results in the previous sections seem to indicate that the intact adult male assay has the ability to detect relatively strong- and, perhaps, weak-acting ER agonists, especially xenobiotics. While the apparent negative or equivocal results with genistein and coumestrol raise a question regarding the sensitivity of the assay to detect weak ER agonists. Although genistein was tested at the limit dose, neither genistein nor coumestrol seemed to achieve an MTD (generally defined as an approximate decrease in final body weight in the high-dose group compared to the control group of 10%). Based on the demonstrated ability of the adult male assay to positively detect relatively weak xenobiotics (methoxychlor and nonylphenol), the intact adult male assay could possibly detect the estrogenic effects of phytoestrogens (genistein and coumestrol) if tested at the appropriate MTD. Although MTD exposure to genistein and coumestrol has not been documented in the adult male assay, the estrogen-like activity of these and other estrogen-like compounds have been positively identified with the uterotropic assay (Kanno *et al.*, 2003a,b; Ashby *et al.*, 1999; Tinwell *et al.*, 2000; Kang *et al.*, 2005).

Androgens: Two relatively strong AR agonists (methyltestosterone and testosterone) have been evaluated in the intact male assay (Table 4). Methyltestosterone at doses of 0, 10, 30, 100, or 300 mg/kg/day was administered by oral gavage such that rats administered 300 mg/kg/day had statistically significant decreases in final body weights (88% of control), feed consumption, absolute testes weights, and serum testosterone, LH, and FSH concentrations (Stump, 2002). Relative prostate, seminal vesicles and coagulating glands with fluid (SVCG), and ASG weights were significantly increased by $\geq$35% relative to controls. Interstitial cell atrophy was observed during histomorphological examination of the testes. Some of these effects also were seen at 100 mg/kg/day, including statistically significant decreases in final body weight (96% of control), serum testosterone, LH, and FSH concentrations, and testicular histopathology. At 30 mg/kg/day, only serum testosterone and LH concentrations were significantly decreased. Serum hormone changes were associated with increased relative prostate, SVCG, and ASG weights and histolopathology results were similar to those reported by O'Connor *et al.* (2000a) when rats were treated with 0, 0.5, 2, 10, or 20 mg/kg/day of testosterone by IP injection. Although final body weights were not affected at any dose of testosterone (O'Connor *et al.* (2000a), testicular

interstitial cell atrophy was seen at all dose levels with spermatid retention observed at doses of 10 and 20 mg/kg/day. Statistically significant decreases in serum LH and FSH concentrations were seen at all doses and, at the high dose, PRL concentrations were significantly increased. Contrary to the results with methyltestosterone, increases in serum concentrations of testosterone, DHT, and estradiol (high dose only) were observed. In regard to the latter, exogenous testosterone likely contributed to the endogenous increases in DHT and estradiol as a result of conversions by 5α-reductase and aromatase, respectively. It is not known if environmental androgen-like agonists would be substrates for these enzymes since no androgen-like agonists have apparently been identified. If environmental xenobiotics were metabolized by these enzymes, these materials would be unlikely to yield a hormonal product (DHT or estradiol) measurable by the intact adult male assay.

Anti-androgen-like chemicals: The inter-laboratory reproducibility of the adult male assay at different times was initially examined by the chemical industry (Dow and DuPont, Table 4) using flutamide (anti-androgen) at doses of 0, 0.5, 1, 10, or 50 mg/kg/day. At Dow (Marty *et al.* 2002), 50 mg/kg/day of flutamide significantly decreased final body weights (91% of control) and increased relative liver weights. At 10 and 50 mg/kg/day, statistically significant decreases in absolute epididymides and relative prostate and seminal vesicle weights were observed, and interstitial cell hyperplasia and hypertrophy were seen in the testes. Serum LH, FSH, and testosterone concentrations were significantly increased. Thyroid parameters (organ weight, histomorphology and hormone concentrations) were not significantly altered with flutamide treatment. At DuPont, (O'Connor *et al.* 1998a; 2002a) the anti-androgenic effects of flutamide were detected as statistically significant decreases in absolute epididymal weight, and decreased relative prostate and seminal vesicle weights at doses $\geq$5 mg/kg/day. Relative liver weights were significantly increased at the highest dose in both studies. Interstitial cell hypertrophy and hyperplasia also were observed at doses $\geq$5 mg/kg/day. Serum hormone changes and overall conclusions were consistent in both studies; thus, demonstrating transferability of the protocol and reproducibility of the results across laboratories as a prevalidation exercise.

There have been some inconsistencies in detecting the effects of two relatively weak anti-androgens, *p,p'*-DDE (dichlorodiphenyldichloroethylene) and linuron due, in part, to the sensitivity of different strains of rats. The ability of the intact adult male assay to detect *p,p'*-DDE was readily observed in Long-Evans rats, but not in Sprague-Dawley (SD) rats which was attributable to suspected pharmacokinetic differences between these rat strains (O'Connor *et al.*, 1999a). Strain differences in responsiveness to *p,p'*-DDE have been reported previously (You *et al.*, 1998). *In utero* and lactational exposure to *p,p'*-DDE produced greater anti-androgenic responses in Long-Evans rats than in SD rats, which was attributed to higher serum and brain levels of *p,p'*-DDE in the Long Evans compared to SD rats. As discussed in more detail in Section 4.2, the SD rat is recommended over other rat strains in the intact adult male assay because it is often the rodent model of choice for general toxicological studies, commonly used to examine specific endocrine-mediated effects of natural and synthetic compounds on reproductive and thyroid function in single and multigeneration studies, and relatively large amounts of reference data are available in historical data bases.

The weak anti-androgen, linuron, also was tested in the intact adult male assay using SD rats during prevalidation (Table 4). In one study (EPA, 2005), linuron affected feed consumption at

all dose levels and significantly decreased final body weights at the 3 highest doses (50, 75, and 100 mg/kg/day) with a 12% decrease relative to controls in final body weight at the highest dose level. At the 100 mg/kg/day dose level, the most prevalent clinical sign was lethargy and, at necropsy, absolute epididymal weights were significantly decreased. Relative thyroid weights were significantly decreased at 25, 50, and 75 mg/kg/day, but not at 100 mg/kg/day. There was a dose-related increase in serum estradiol and decrease in serum T_4 levels, both of which were statistically significant at all doses of linuron. TSH and T_3 concentrations were not altered relative to controls. Serum PRL was decreased significantly only at 100 mg/kg/day. These results are consistent with previous studies using linuron in which the adult male assay detected the anti-androgen effects of linuron in both immature (32 to 45 days of age) and mature rats when dose levels produced a greater than 10% change in final body weight relative to controls (Cook *et al.*, 1993; O'Connor *et al.*, 2002a). Similar effects on absolute epididymal weight were observed at 150 mg/kg/day of linuron (O'Connor *et al.*, 2002a) as were significant decreases in relative prostate and ASG weights. In the latter study, 150 mg/kg/day of linuron produced a similar 11% decrease in terminal body weight compared to controls, whereas final body weight was decreased only 9% at 100 mg/kg/day. Spermatid retention was observed at 100 and 150 mg/kg/day in that study. Similar effects were observed with serum estradiol, T_4, and PRL concentrations, and serum T_3 concentrations also were significantly decreased at $\geq$50 mg/kg/day of linuron. Relative thyroid weights and thyroid histomorphology were not affected. These results confirmed findings reported by Cook *et al.* (1993) and, notably, a pair-fed control group was included in the study to differentiate endocrine-mediated effects from effects secondary to decreased body weight gain. Feed-restriction studies conducted more recently by O'Connor *et al.* (1999b; 2000b), supported the interpretation by Cook *et al.* (1993). Moreover, the latest feed restriction studies provided a basis for establishing and standardizing criteria for differentiating and interpreting endocrine-related effects in the intact adult male assay as described in Section 3.3. The results of linuron in the intact adult male assay are further discussed in Section 5 since it was used as a test chemical in the inter-laboratory validation exercise.

Thyroid modulating chemicals: Two thyroid-active agents were evaluated during development of the intact adult male assay at DuPont (O'Connor *et al.,* 1999b; O'Connor *et al.,* 2002b). Both phenobarbital and propylthiouracil were readily detected as thyroid-active agents (Table 4) based on significantly increased thyroid gland weight, histopathology and alterations in serum hormone concentrations (increased serum TSH, decreased serum T_4 and T_3). Phenobarbital perturbs thyroid homeostasis in rats through hepatic enzyme induction and enhanced clearance of thyroid hormones, whereas propylthiouracil inhibits thyroid hormone synthesis and 5'-deiodinase. Since phenobarbital was used in the inter-laboratory validation exercise presented and discussed in Section 5, further discussion of this and other test chemicals on the thyroid is reserved for that section.

3.1.2 Negative Test Chemical

Allyl alcohol, a well-characterized industrial chemical that is a known hepatotoxin, was tested in the 15-day intact adult male assay as a negative control for demonstrating specificity of the adult male assay by differentiating between endocrine-mediated effects from nonendocrine-mediated effects. According to the results of a dose-range finding study, allyl alcohol was administered by oral gavage at doses of 0, 10, 30, 40, or 50 mg/kg/day (O'Connor *et al.* 2007). Rats administered

50 mg/kg/day of allyl alcohol had statistically significant decreases in final body weights (90% of control), and significant increases in relative liver weights. At 50 mg/kg/day, serum testosterone and DHT concentrations were significantly decreased. At 40 mg/kg/day, relative liver weight was significantly increased and serum PRL concentrations were significantly decreased. There was no detectable histopathology of the testes, epididymides and thyroid gland and no other treatment-related effects were observed. Although there were statistically significant effects on some of the serum hormones, most likely as a result of increased hepatic clearance due to liver enzyme induction, these effects alone would not be sufficient to consider allyl alcohol as a potential endocrine disruptor since there were no chemical-related effects on target organ weights or histology (see Section 3.3 for interpretation of results). Thus, the results of allyl alcohol in the intact adult male assay were considered to be negative; no direct effect on the EAT hormonal systems.

3.2 Standardization of Endpoints

A prevalidation exercise was conducted to standardize the multiple endpoints used in the intact adult male assay (O'Connor *et al.*, 1998a,b; 1999a,b; 2000a,b; 2002a,b). Two primary goals of the exercise were to, first, determine which of the many endpoints evaluated in the exercise were not relevant and, therefore, could be excluded to optimize and standardize the assay protocol and, second, test the hypothesis that chemical-responsive "fingerprints" could be developed for many endocrine-related events and facilitate identifying a MOA for an EAC. To accomplish these goals, 15 EACs were identified that included ER binding agonists and antagonists, AR binding agonists and antagonists, progesterone receptor modulators, thyroid modulators, steroid biosynthesis inhibitors (aromatase, 5α-reductase, and testosterone biosynthesis), and PRL modulators. Additional chemicals have since been tested as presented in Table 4. As a result of this prevalidation exercise, which is summarized in a review by O'Connor *et al.* (2002c), key endocrine-specific endpoints were emphasized to optimize and standardize the intact adult male assay protocol as described in Appendix C of the ISR. Although no endpoints were necessarily removed, serum concentrations of DHT and PRL as well as liver microsome activity (UDP-glucuronyltransferase) may be considered on a need for basis dependent on the initial results of the bioassay.

Secondarily, the exercise demonstrated that chemical-responsive "fingerprints" could be developed to identify EACs and their MOAs and aid in the characterization of their underlying mechanisms of action. However, since the concept of developing chemical-responsive "fingerprints" using the intact adult male assay is a relatively novel approach in toxicological studies combining multiple apical endpoints (reproductive organ weights) and histology with systemic changes in serum concentrations of reproductive steroids, gonadotropins and thyroid hormones, many more studies with a variety of chemical substances will be required to establish this concept as a reliable and practical method to identify potential endocrine disruptors and predict their modes or mechanisms of action. Nevertheless, the example discussed below highlights and supports the possibilities that can result from further developing this methodology through the use of the intact adult male assay. Chemical-responsive "fingerprints" were obtained for flutamide and ketoconazole from studies conducted by O'Connor *et al*. (1998a and 2000c). Flutamide, an AR antagonist, competes with testosterone and DHT for binding to the AR. As flutamide blocks the recognition of testosterone and DHT, androgen-dependent organs such as

ASG weight decreased. Correspondingly, the ability of flutamide to block the negative feedback effect of testosterone and DHT at the hypothalamic and pituitary levels resulted in the secretion of gonadotropin-releasing hormone (GnRH) and LH, respectively, and the subsequent production of testosterone by the Leydig cells of the testes. Thus, the chemical-responsive "fingerprint" of an AR antagonist such as flutamide is a decrease in ASG weight and increased serum concentrations of testosterone and LH. Ketoconazole, an androgen steroid biosynthesis inhibitor, inhibits testosterone production by binding to the heme iron of the 3-cytochrome P_{450} isozymes of the androgen biosynthetic pathway. Similar to flutamide, ketoconazole resulted in a decrease in ASG weight but the mechanism of action is different. Ketoconazole acts directly at the testicular level to inhibit testosterone production. As a result, serum concentrations of testosterone decrease and, secondarily, serum concentrations of LH increase due to the lack of a negative feedback effect of testosterone at the hypothalamic/pituitary level. Hence, the chemical-responsive "fingerprint" of an androgen steroid biosynthesis inhibitor such as ketoconazole is a decrease in ASG weight and testosterone concentration and an increase in LH concentration and, for an AR antagonist such as flutamide, there is a decrease in ASG weight and an increase in testosterone and LH concentrations.

Regardless of the concept of chemical-responsive "fingerprints", by including measurements for multiple hormones in the intact adult male assay as a means to support the results of key apical and histomorphological endpoints, not only is the MOA more readily identified, but one can potentially distinguish between different mechanisms of action associated with the anti-androgenic effects of potential endocrine disruptors on the male reproductive system. Moreover, they may add to the weight of evidence within the bioassay and, perhaps, within a Tier-1 battery.

3.3 Historical Reference Data Bases

The results from 29 test chemicals (one negative and 28 positives, Table 4, above) with known MOAs run in the intact male assay from studies sponsored by the chemical industry and the EPA have been used to establish relevance of the bioassay as discussed in Section 3.1. The information was also used to create reference data bases (O'Connor *et al.*, 1999b; 2000a,b; 2002a,b,c) for the inter-laboratory validation studies presented in Section 5. From 28 of these studies (O'Connor *et al.*, 2002a), the weights of primary and secondary sex organs and thyroid gland, histomorphology of the testes, epididymides and thyroid, and serum concentrations of reproductive steroids, gonadotropins and thyroid hormones served as historical control reference data to compare the expected results with observed results in the vehicle-control group presented in Section 5.2 (Table 9). In addition, the results from prevalidation studies, done with linuron and phenobarbital, served as expected historical results to compare the observed results of the inter-laboratory studies with linuron and phenobarbital presented in Sections 5.3 (Table 11) and 5.4 (Table 14), respectively.

3.4 Interpretation of Endocrine-Mediated Effects Within the Bioassay

3.4.1 Effect of final body weight on target organ weight and hormone concentrations

As with all of the *in vivo* screening assays proposed for the EDSP Tier-1 battery, interpretation of the data to differentiate between compound-related effects and effects that may be due to

acute toxicity or overexposure secondary to an extreme decrease in final body weight during treatment will be a challenge. In advance of discussing the standardized protocol for the intact adult male assay in Section 4 and analysis of the inter-laboratory validation results in Section 5, a rationale for the criteria for interpretation of organ weight, histological and hormonal effects in the intact adult male rat assay is presented.

Changes in target organ weights and histomorphology as well as serum hormone concentrations are expected to be interpreted in the context of final body weight decrements according to results obtained in dietary restriction experiments conducted during prevalidation of the intact adult male rat (O'Connor *et al.*, 1999b; 2000b).

The first consideration in this series of studies was to determine the dependency of target organ weight on final body weight. As shown in Table 5, relative (organ-to-body weight ratio) testis and epididymal weights significantly increased in association with a ≥10% decrease in final body weight in the feed-restricted animals compared to the *ad libitum*-fed controls, whereas absolute testis and epididymal weights were not significantly different between the feed-restricted animals and the *ad libitum*-fed control animals until a body weight decrement of 26% was reached. In contrast, the thyroid, ASG (total prostate plus SVCG), SVCG and prostate were considered body-weight dependent since relative organ weights did not change significantly between feed-restricted animals and the *ad libitum*-fed control animals throughout a 26% decrement in final body weight. While both absolute and relative liver weights were affected by dietary restriction, relative liver weight corrected for most of the body weight decrement. This was in keeping with the generally accepted theory that liver weight is body weight dependent and that expression on a relative to body weight basis will correct for body weight decrements (Feron *et al.*, 1973). Thus, when evaluating target organ weight data following chemical exposure using the 15-day intact adult male rat assay, weights of the testes and epididymides should be evaluated on an absolute organ weight basis, and weights of the liver, thyroid, ASG, SVCG, prostate glands should be evaluated on a relative to final body weight basis in order to optimize interpretation of endocrine-related effects.

A second consideration in this series of studies was to determine the degree of body weight loss that can occur before target organ weights and serum hormone concentrations are secondarily affected by an extreme decrease in final body weight that may be indicative of acute toxicity or overexposure to chemical treatment (O'Connor *et al.*, 1999b; 2000b). As shown in Table 5, absolute weight of the testes and epididymides and relative weights of the liver, thyroid, ASG, SVCG, and prostate were not significantly different between feed-restricted and *ad libitum*-fed control animals until a decrement in final body weight of ≥26% was reached. As shown in Table 6, serum hormone concentrations were not significantly different between feed-restricted and *ad libitum*-fed control animals until a final body weight decrement of 15% was reached for T_3 and T_4, 21% for estradiol and DHT, and ≥26% for PRL, FSH, LH and TSH. Although targeting a final body weight decrement in the high-dose group in the intact adult male rat assay of around 10% of control at the time of euthanasia minimizes the potential for confounding secondary effects due to acute toxicity or overexposure of treatment, final body weight decrements from 15 to 20% relative to controls may be acceptable for interpretation of endocrine-mediated effects on some target organs, histomorphology and serum hormones.

Table 5. Mean (± SE) effect of dietary restriction on final body and target organ weights in the intact adult male rat assay (O'Connor *et al.*, 1999b; 2000b).

Feed/day (grams)	Final body (grams)	Final body weight (% control)	Liver	Thyroid	Testes	Epididymides	Accessory sex gland	Seminal vesicles	Prostate
Absolute organ weights (g)									
ad libitum[a]	414 ± 6[b]	100	16.0 ± 0.4	0.025 ± 0.001	3.3 ± 0.1	1.14 ± 0.02	2.3 ± 0.1	1.6 ± 0.1	0.617 ± 0.021
22	373 ± 4*	90	13.4 ± 0.1*	0.021 ± 0.001*	3.2 ± 0.0	1.11 ± 0.03	2.0 ± 0.1	1.5 ± 0.1	0.555 ± 0.031
19	351 ± 3*	85	12.0 ± 0.2*	0.019 ± 0.001*	3.3 ± 0.1	1.08 ± 0.02	1.8 ± 0.1*	1.2 ± 0.1*	0.529 ± 0.034
16	328 ± 3*	79	10.5 ± 0.2*	0.019 ± 0.001*	3.2 ± 0.1	1.11 ± 0.01	1.8 ± 0.1*	1.3 ± 0.1*	0.524 ± 0.039
13	307 ± 2*	74	9.8 ± 0.1*	0.019 ± 0.001*	3.2 ± 0.1	1.06 ± 0.02*	1.6 ± 0.1*	1.1 ± 0.1*	0.454 ± 0.029*
Relative organ weights (% body weight)									
ad libitum[a]	414 ± 6[b]	100	3.9 ± 0.1	0.006 ± 0.0003	0.79 ± 0.02	0.276 ± 0.006	0.552 ± 0.018	0.396 ± 0.017	0.149 ± 0.005
22	373 ± 4*	90	3.6 ± 0.1*	0.006 ± 0.0003	0.86 ± 0.01*	0.296 ± 0.007*	0.548 ± 0.020	0.394 ± 0.016	0.149 ± 0.008
19	351 ± 3*	85	3.4 ± 0.1*	0.006 ± 0.0003	0.94 ± 0.02*	0.308 ± 0.005*	0.504 ± 0.020	0.350 ± 0.020	0.150 ± 0.009
16	328 ± 3*	79	3.2 ± 0.0*	0.005 ± 0.0003	0.97 ± 0.02*	0.338 ± 0.004*	0.561 ± 0.036	0.411 ± 0.026	0.160 ± 0.012
13	307 ± 2*	74	3.2 ± 0.0*	0.006 ± 0.0003	1.04 ± 0.02*	0.344 ± 0.006*	0.516 ± 0.022	0.364 ± 0.018	0.148 ± 0.010

[a] *Ad libitum* control rats consumed 25.8 g/day.
[b] Mean ± standard error.
* Significantly different ($p < 0.05$) from control by Dunnett's Test. (n=15 animals/feed group)

Table 6. Mean (±SE) effect of dietary restriction on serum hormone concentrations in the intact adult male rat assay (O'Connor *et al.*, 1999b; 2000b).

Feed/ day (grams)	Final body (% of control)	Estradiol (pg/ml)	Testos-terone (ng/ml)	Dihydro-testosterone (pg/ml)	Prolactin (ng/ml)	Follicle stimulating hormone (ng/ml)	Luteiniz-ing hormone (ng/ml)	Thyroid stimulating hormone (ng/ml)a	T_3 (ng/dl)[a]	T_4 (µg/dl)[a]
ad libitum[b]	100	3.5 ± 0.5[c]	11.1 ± 1.4	162.3 ± 25.4	17.9 ± 2.9	13.1 ± 0.7	4.4 ± 0.3	17.3 ± 1.3	80.7 ± 4.0	4.3 ± 0.2
22	90	3.9 ± 0.6	11.9 ± 1.2	175.3 ± 19.6	11.8 ± 1.5	14.9 ± 0.9	5.2 ± 0.4	17.0 ± 1.8	79.9 ± 3.6	4.0 ± 0.2
19	85	3.7 ± 0.8	12.9 ± 1.2	176.3 ± 32.6	16.5 ± 2.3	13.4 ± 0.6	4.8 ± 0.3	16.7 ± 1.5	68.1 ± 3.7#	3.6 ± 0.2#
16	79	1.6 ± 0.4#	12.8 ± 1.6	81.3 ± 14.0#	9.9 ± 1.4	13.9 ± 0.6	5.1 ± 0.3	14.1 ± 1.1	70.5 ± 3.7#	3.2 ± 0.2#
13	74	0.9 ± 0.3#	ND	60.6 ± 12.9#	10.1 ± 2.1#	12.8 ± 0.7	5.1 ± 0.3	10.8 ± 1.5#	60.8 ± 2.8#	3.1 ± 0.2#

[a] Data from O'Connor *et al.* (1999b).
[b] *Ad libitum* control rats consumed 25.8 g/day.
[c] Mean ± standard error.
ND – not determined due to a lack of serum for analysis.
Significantly different ($p < 0.05$) from control by Jonckheere's test for trend. (n=15 animals/feed group).

Thus, interpretation of whether the results of chemical exposure are endocrine-related involves consideration of whether weight changes of target organs are affected on an absolute or relative basis and whether the final body weight decrement is within the limits of interpretation of an endocrine-related effect rather than an acute toxic effect secondary to an extreme decrease in final body weight during treatment.

3.4.2 Priority of endpoints for interpretation of results

Weight changes and histopathology of target organs are expected to carry a heavier weight of evidence within the intact adult male assay than changes in serum hormone concentrations alone to indicate whether a substance affects the EAT hormonal system. That is, hormonal changes alone are of insufficient weight within the bioassay to make a conclusion. An increased incidence of histopathologic alterations of the testes, epididymides or thyroid gland in treated animals compared to controls would be an indication of a compound-induced effect independent of effects on target organ weights or serum hormone concentrations. However, statistically significant changes in respective organ weights and related hormones between treated and control groups would add weight-of-evidence within the assay to the histopathological results, and also allow differentiation of MOA based on the pattern of the effects. Statistically significant target organ weight changes alone would also be considered compound-related with a relatively high degree of confidence if the results correspond to a significant linear trend indicating that the results are dose-dependent. If the linear trend analysis is not significant, it is possible that a significant difference between treated and control groups at any dose level is spurious and not compound-related; however, a weight-of-evidence approach among the multiple endpoints within the assay combined with biological plausibility can help distinguish compound-related from spurious alterations of an endpoint result.

Statistically significant changes in serum hormone concentrations are expected to support target organ weight and histopathological changes as well as provide additional information to differentiate between various MOAs for unknown chemicals. Instances when only serum hormone concentrations are significantly altered will not be considered sufficient evidence alone within the assay to identify a positive endocrine test result but, perhaps, may be considered relevant in a weight-of-evidence approach between or among assays when interpreting the entire EDSP Tier-1 screening battery. In addition, if the results among the endpoints for organ weights and histomorphology are equivocal with respect to an effect on the endocrine system within the bioassay, they too, perhaps, may be considered relevant in a weight-of-evidence approach between or among assays in the Tier-1 screening battery. An approach to interpretation of the results of a Tier-1 battery has not been thoroughly defined since the EPA has not yet proposed what suite of assays will constitute a battery.

An example with ketoconazole demonstrates that data interpretation based solely on changes in serum thyroid hormone concentrations are not sufficient for identifying potential endocrine disruptors that may affect the thyroid. In the 15-day intact adult male assay, the steroid biosynthesis inhibitor ketoconazole significantly decreased serum T_3 and T_4 concentrations, with no corresponding effects on TSH concentration, thyroid organ weight, or histomorphology (O'Connor *et al.*, 2002b). It has been shown that long-term studies with ketoconazole in rodents does not induce thyroid tumors (Physician's Desk Reference, 2004). This example in the intact

adult male assay is consistent with previous reports (Döhler *et al.,* 1979) that numerous factors can acutely affect thyroid hormone concentrations either directly or indirectly by a wide variety of chemicals and that a large proportion of these chemicals will not be thyroid toxicants in long-term studies. In this regard, of the 29 compounds that have been evaluated in the adult male assay (Table 4), 27 of the compounds caused a statistically significant change in at least one of the thyroid hormones (i.e., TSH, T_3, or T_4).

Thus, the relevance of the intact adult male assay according to its multiple endpoints discussed throughout Sections 2 and 3 allows for a comprehensive weight-of-evidence approach within the bioassay to detect effects on the EAT hormonal systems by first assessing final body weight and thereafter evaluating the weights of primary and secondary sex organs and the thyroid gland, histomorphology of the testes, epididymides, and thyroid, and serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones. If included in a Tier-1 screening battery, the results of the intact adult male assay are expected to be taken into account with the results of other complimentary *in vitro* and *in vivo* assays using a weight-of-evidence approach among the entire suite of assays within the battery to make a determination of whether a test chemical is a potential endocrine disruptor. However, before the intact adult male assay can be considered eligible for consideration in a Tier-1 screening battery, the reliability and feasibility of the assay protocol needs to be critically evaluated. Reliability and feasibility of a standardized protocol are the main topics presented and discussed in the next sections beginning with an overview of the standardized assay protocol in Section 4 followed by an inter-laboratory validation exercise using the standardized protocol with linuron and phenobarbital as test chemicals in Section 5.

4.0 Standardized Protocol for Inter-Laboratory Validation

Considering the 15-day intact adult male rat assay protocol as used in numerous prevalidation studies (review, O'Connor *et al.,* 2002c), a more standard protocol was developed for inter-laboratory validation that was sponsored by the EPA. An abbreviated version of the protocol is presented in this section with emphasis on key technical aspects that were optimized based on prevalidation results. It will be emphasized that some aspects of the protocol were controlled to enhance and compare the results of the intact adult male assay run concurrently across three different CRO laboratories (RTI, WIL, and Charles River) with the same two positive test chemicals (linuron and phenobarbital) as presented in Section 5. A detailed version of the protocol used in the inter-laboratory studies can be found in the individual laboratory reports included in the peer review package and the final standardized protocol can be found in Appendix C of this ISR.

4.1 Objective

The objectives of the inter-laboratory validation exercise were primarily to evaluate the reliability and transferability of the standardized protocol for the 15-day intact adult male rat assay and, secondarily, to continue to assess its relevance. Two test chemicals known to affect the endocrine system (linuron and phenobarbital, affecting the androgen and thyroid hormonal pathways, respectively) were run concurrently in three different CRO laboratories to determine:

Reliability, defined herein as the ability of the assay to detect endocrine-mediated effects on prescribed target organs within laboratories with the expectation of consistency among laboratories and that the observed results in the present studies would be comparable to expected results reported in previous studies.

Relevance, defined herein as the ability of the assay, within and among laboratories, to detect known effects of linuron and phenobarbital on the endocrine system primarily by measuring changes in the weight of primary and secondary sex organs and thyroid gland, histomorphology of the testes, epididymides and thyroid, and serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones.

Transferability, defined herein as the feasibility of the assay protocol to be conducted in various CRO laboratories in a logistical and practical manner to be compliant with standard operating procedures (SOPs) within laboratories, the study protocol, and government GLP conditions in such a manner not to jeopardize study results.

4.2 Animals

The test animals were adult outbred-derived albino rats (Sprague-Dawley Crl:CD® or SD rats) obtained from the same supplier by all laboratories to minimize source variation among laboratories. The animals were obtained so that they were approximately 10 weeks of age at the start of dose administration.

The basis for selecting the SD rat over other rat strains (e.g., Long Evans and Wistar) is that it has often been the animal model of choice for determining general toxicological and, to a lesser extent, endocrinological effects. More recently, SD rats have been used to examine specific endocrine-mediated effects of natural and synthetic compounds on reproduction and thyroid function in intact rodent models. Many laboratories use SD rats for multigeneration studies, including the two-generation reproduction toxicity test currently proposed for the EDSP Tier-2 battery and, therefore, this model will allow for an examination of reproducibility of endpoints common to Tiers 1 and 2 in the same strain of rats. Furthermore, relatively large historical data bases are available for reference.

Intact adult male rats were chosen initially during prevalidation because they were shown to be a more stable and sensitive model than immature rats for detecting compound-induced hormonal changes (Cook *et al.,* 1993). However, in a series of recent experiments using several chemicals with well-characterized endocrine activity (vinclozolin, flutamide, di-n-butyl phthalate, linuron, phenobarbital, and propylthiouracil), O'Connor *et al*. (2005) has shown that the sensitivity of adult versus immature animals for detecting these EACs is comparable when looking at both compound-induced organ weight changes and compound-induced hormonal changes. Thus, adult animals were preferred not because they are more or less sensitive than immature animals but, in part, because the change in body weight due to growth is relatively small in adult animals; thus, the potential for confounding effects of relatively large changes in body weight due to accelerated growth from weaning to puberty are minimized. Also, adult animals provide a larger volume of blood for analyzing the numerous serum hormones involved in this rat bioassay and, again, there is a larger historical data base for adult animals than immature animals for reference.

4.3 Housing, Environment, Feed, and Water

Animals were housed individually in solid-bottom, polycarbonate cages fitted with stainless steel wire lids with Sani-Chip® cage bedding or wire-mesh cages. Water was available *ad libitum* through plastic bottles with stainless steel sipper tubes or an automatic watering system.

Animal rooms were maintained on a 12:12 hours light:dark cycle. Target conditions for temperature and relative humidity in the animal rooms were between 64 and 79°F and between 30 and 70%, respectively.

There was some concern that phytoestrogens in commercial rat feed may be present at a concentration high enough to interfere in the interpretation of the results. A brief review of the literature indicated considerable doubt as to the likelihood that such levels could affect the interpretation of the results in adult male rats over a 2-week period. Nevertheless, powdered feed with low phytoestrogen content was available *ad libitum* and obtained from the same supplier by all laboratories to minimize source variation among laboratories. The diet was analyzed by the supplier for characterization of concentrations of phytoestrogens (e.g., genistein, daidzein, and glycitein) with the expectation not to exceed 300 μg/g.

4.4 Study Design, Test Chemicals and Dose Selection, and Duration

As presented in Table 7, the animals were dosed by oral gavage with vehicle (i.e., methylcellulose) or chemical formulations at a dose volume of 5 ml/kg body weight daily from Test Day 1 through Test Day 15 (TD 1 to 15). Animals were dosed beginning early in the morning so that at termination blood collection and necropsy could be completed within a 2- to 3-hour window after the last dose on TD 15 before the afternoon hours. Typical necropsy times used in previous experiments were from 0700 to 1000 hours and were recommended herein to compare with published historical control data. The studies in each of the laboratories were initiated in a staggered manner across dose groups to accommodate the number of animals scheduled for necropsy within a defined time (2 to 3 hours) after administration of last dose on TD 15.

Table 7. Study design of the 15-day intact adult male rat assay.

Group	Chemical[a]	Dose level (mg/kg/day)[b]	Animals
1	Vehicle[c]	0	15
2	Linuron	50	15
3	Linuron	100	15
4	Linuron	150	15
5	Phenobarbital	25	15
6	Phenobarbital	50	15
7	Phenobarbital	100	15

[a] Vehicle and test chemical formulations were coded to minimize bias during data collection.
[b] Test compounds administered once daily by gavage on TDs 1 to 15.
[c] Vehicle only (0.25% methylcellulose) at a dose volume of 5 ml/kg.

Linuron, a relatively weak anti-androgen that competitively binds to the AR, and phenobarbital, a relatively weak thyroid toxicant that alters thyroid function indirectly through enhanced liver

metabolism and excretion of thyroid hormones, were selected, in part, to challenge the assay through these different MOAs. Each of the test chemicals has previously been run in the adult male assay in different laboratories (Table 4) during prevalidation with many of the results documented in published in peer-reviewed scientific journals and cited in a review article by O'Connor *et al.* (2002c) or in a final study report by the EPA (EPA, 2005). In addition, these compounds have been run previously in the pubertal male assay, Hershberger assay, and other *in vitro* and non-mammalian assays under consideration in the EDSP Tier-1 screening battery.

Dose selection was based on the results of previous studies (O'Connor *et al*., 1999a, 2002a,b; EPA, 2005) using linuron and phenobarbital as test chemicals. In this regard, a dose range finding study to establish an MTD was not justified as a reasonable use of animals, time, and other resources, especially since dose levels were available from previous studies. Moreover, a dose range finding study to establish an MTD was not considered a validation exercise since it is common practice among toxicology laboratories and would likely be established well before any Tier-1 screening assay is run.

Duration of treatment was adopted after an internal methods development study identified 2 weeks or 15 days as sufficient to detect endocrine-mediated effects of proprietary chemicals (Section 2.2). These in-house results were supported in a documented study (O'Connor *et al.,* 1999b) indicating that there was little improvement in assay sensitivity when dosing duration was extended to 4 weeks. Consequently, the added expense to extend treatment with little value was not justified.

4.5 Dose Analyses

Analytical chemistry for assessing stability of the formulated suspensions was established before the start of the study, and formulation homogeneity and dose concentration verification were done at the start and end of the study. In compliance with government GLP, these tests were done to assure that the animals had been treated and exposed accordingly to the test chemicals.

4.6 Clinical Observations, Body Weights, and Feed Consumption

Clinical observations of animals were documented at least twice daily, at dosing and at a designated time post-dosing.

Body weights were measured every day prior to each day's dosing, in part, to adjust daily dosing volume to most recent body weight. TD 14 body weights were used for dosing animals on TD 15. TD 15 body weights were collected at the time of necropsy (live weights before euthanasia).

Feed consumption was measured weekly.

4.7 Euthanasia

Animals were not fasted prior to euthanasia.

CO_2 for approximately 60 seconds was used to anesthetize the animals prior to decapitation, which was done in a manner to avoid disturbing the integrity and collection of the thyroid gland.

Considering the concern that undue stress associated with pre- and post-administration of anesthesia could interfere with some hormones (e.g., PRL), decapitation was incorporated immediately after timed CO_2 exposure with the expectation to optimize a more accurate measurement of serum hormone concentrations.

4.8 Blood Collection

Trunk blood from each animal was collected after decapitation and immediately cooled and held cold (e.g., on ice) until centrifugation and processing of serum.

Serum from each animal was put into aliquots based on the number of different hormone assays that were expected to be run in a day to minimize potential freeze/thaw effects on serum hormone concentrations. Aliquots were labeled and stored frozen until analysis.

4.9 Macroscopic Examination and Organ Weight Interpretation

Gross examination of the animals was conducted at necropsy. Body weights were taken and target organs were collected and weighed as follows:

- Body (live weight before euthanasia)
- Liver
- Testes (left and right weighed separately and combined for paired weight; inter-laboratory study only)
- Epididymides (paired weight)
- Prostate (combined dorsolateral and ventral)
- SVCG
- Thyroid gland (weighed after fixation and final dissection by one individual)
- ASG.

Note: ASG weight represented the combined weights of the entire prostate plus SVCG. The testes and epididymides were preserved in Bouin's and the thyroid gland was preserved in formalin for subsequent histological examination.

4.10 Microscopic Examination and Interpretation

The testes, epididymides, and thyroid gland were evaluated histologically. The embedded tissues were sectioned (2 to 5 microns) transversely for the testes, longitudinally for the epididymides, and according to laboratory SOP for the thyroid and, subsequently, stained with hematoxylin and eosin (H&E). A minimum of two sections for the thyroid and a sufficient number of sections for each testis and epididymis were prescribed for examination.

Interpretation of the histomorphological alterations was done by a board-certified veterinary pathologist knowledgeable of the control and high-dose groups for each of the test chemicals but not the nature of the chemicals. The pathologist examined the sections for histopathology and

potential treatment-related effects using conventional methods. Sections for respective organs from the low- and middle-dose groups were not examined.

4.11 Hormone Analyses

Blood serum hormone concentrations were analyzed in a standardized sequence for all laboratories according to the nature of the test chemicals.

- Testosterone
- LH
- TSH
- T_4
- T_3
- FSH
- Estradiol
- PRL
- DHT

Note, serum DHT and PRL measurements were considered necessary in the inter-laboratory validation exercise but may be considered optional in the final standardized protocol.

Serum hormones were measured using commercially available radioimmunoassay (RIA) kits. All steroid and the T_3 and T_4 assays were prescribed for use with human serum according to the manufactures and were not modified for use with rat serum, whereas the other assays (FSH, LH, PRL and TSH) were prescribed for use with rat serum. The same kit suppliers were used by the laboratories (Biotrak™, Amersham Biosciences, and DSL) except that RTI used DSL Kit No. 39100 for estradiol analyses, whereas WIL inadvertently used DSL Kit No. 4400; the kits differed in the level of assay sensitivity.

Each hormonal assay included all experimental samples from the control group and each treated group for linuron and phenobarbital. If all serum samples for a particular test chemical could not be included in one assay, samples within a group were randomized and balanced across the different hormone assays. Each experimental sample was run in duplicate. QC standard samples were also run in duplicate and, in some instances, in replicates of duplicates within the assay. The basis for the QC standard samples was primarily an attempt to determine the performance (i.e., within- and between-assay coefficients of variation, CV) of each hormonal assay within and between laboratories. Details for preparing the QC standard samples and tabulating and analyzing the CV results are documented in a separate report in Appendix A of the ISR.

In the event that the amount of serum from each animal was insufficient to conduct all hormone assays, hormone analyses were prioritized according to initial study results or any previous knowledge of the nature of the test chemical. To facilitate interpretation of the QC standard results, the precision (within-assay) and repeatability (between-assay) of each hormone assay was considered acceptable if the CVs were ≤10%, reasonable if the CVs were 11 to 15%, questionable if the CVs were 16 to 20%, and unacceptable if the CVs were >20%. By convention, these criteria reflect a level of confidence in the performance of a hormonal assay

from relatively high to low, respectively, that can be used accordingly to interpret the hormonal results of the experimental samples.

4.12 Statistical Analyses

A common statistical plan was prepared and used within each of the three CRO laboratories. A detailed account of the statistical approach and results from within each laboratory is available in the individual laboratory final reports included in the peer review package. In general, the analyses included:

- Tests for outliers to identify extreme values for possible exclusion from analyses
- Tests for homogeneity of variance to determine whether or not to normalize data prior to analyses
- One-way analysis of variance (ANOVA) to determine an effect of treatment
- T-tests (two-sided; $P<0.05$) and Dunnett's adjusted ($P<0.006$) to determine an effect of each dose level relative to the control
- Linear trend analysis across dose levels for determining dose response

4.13 Retention of Specimens and Records

All records and specimens were handled according to specifications in the protocol, respective laboratory SOPs, and government GLPs.

4.14 Quality Control and Assurance

QC and quality assurance (QA) procedures followed those outlined in the Quality Assurance Project Plan (QAPP) prepared for this inter-laboratory study by the primary contractor. In addition, the laboratories complied with specifications in the protocol, respective laboratory SOPs, and government GLPs.

5.0 Inter-Laboratory Validation Results and Discussion

A detailed account of data collection, tabulation, statistical analyses, and presentation of the inter-laboratory results in tables and figures discussed in this section are presented in Appendix A of the ISR: *"Inter-laboratory Validation of the 15-Day Intact Adult Male Rat Assay: Hormonal Assay Quality Control Standards Data"* and Appendix B of the ISR: *"Inter-laboratory Validation of the 15-Day Intact Adult Male Rat Assay: Statistical Analysis of Among-Laboratory Results"*.

In general, the inter-laboratory statistical analyses in Sections 5.3 and 5.4 for linuron and phenobarbital, respectively, were carried out to compare the results across the three CRO laboratories. Complete intra-laboratory statistical analyses are reported in the individual laboratory reports for RTI, WIL and Charles River which are included in the peer review package and summarized in the inter-laboratory report provided in Appendix B of the ISR.

The objective of the inter-laboratory analyses presented herein is to summarize the results concerning the consistency or variation across laboratories as a measure of the repeatability or reliability of the intact adult male rat assay. This was accomplished, in part, by statistically evaluating the ratios of treatment group means for linuron and phenobarbital to vehicle-control group means for each endpoint initially using a two-way ANOVA to determine the laboratory-by-dose interaction and main effect of laboratory for body weight and food consumption (Appendix B, Tables 7 and 8, respectively), target organ weights (Appendix B, Tables 9 and 10, respectively) and serum hormone concentrations (Appendix B, Tables 11 and 12, respectively). A separate one-way ANOVA was done thereafter to determine the effect of laboratory within each dose level for body weight and food consumption (Appendix B, Tables 1a-c and 2a-c, respectively), target organ weights (Tables 3a-c and 4a-c, respectively), and serum hormone concentrations (Appendix B, 5a-c and Tables 6a-c, respectively). For each endpoint within the tables, probability values were determined according to the likelihood ratio where statistical significance was defined as $P \leq 0.05$.

The variation across laboratories within each dose was also assessed according to the among-laboratory CVs in respective tables listed above and presented in Appendix B. Comparison of the ratios of the results between treated and control groups were used to adjust, in part, for operational differences among the CRO laboratories. The overall hypothesis was that the variation in mean responses among laboratories would not differ from zero or that the mean results are equal across laboratories. This was hypothesized to be true for final body and target organ weights and serum hormone concentrations. Histological results were not subjected to statistical analyses among laboratories but were assessed qualitatively.

The statistical approach also involved an initial examination of extreme values or outliers for all the endpoints. Outliers were detected in the results from RTI (see Tables 9-16 in the individual laboratory report) and WIL (see Appendix G/Appendix B in the individual laboratory report) but not Charles River. The results of linuron and phenobarbital were analyzed excluding the outliers (Appendix B, Tables 15 and 16 and Tables 13a-c and 14a-c, respectively) and found not to differ markedly from the analyses including the outliers. Hence, the results described in this integrated summary report involve analyses with the few outliers included.

As reported independently by each laboratory, there were no SOP, protocol or GLP deviations that were considered to affect the integrity of the studies, which adds support to the feasibility or transferability of the assay protocol.

5.1 Hormone Assay Performance with QC Standard Samples

The 15-day intact adult male rat assay is unique in that it involves the measurement of nine different serum hormones (reproductive steroids, gonadotropins, and thyroid hormones) at the end of the treatment period. These measurements are intended to provide support for the apical and histological effects as well as to provide information on MOA. While hormone analyses are not routinely done in *in vivo* toxicological studies, measurements are likely to become more commonplace, especially since thyroid hormone measurements (TSH, T_3, and T_4) are proposed for other *in vivo* assays in the EDSP Tier-1 battery and Tier-2 tests Given the novelty of hormonal measurements in these types of toxicological bioassays, validation of the intact adult

male assay also included an initial assessment of the performance of the commercially available hormone assay kits with QC-standard samples used within and between laboratories. This section provides an overview of the original results presented in Appendix A of the ISR. It describes the performances (i.e., within- and between-assay CVs) of the hormone assay kits run by RTI and WIL laboratories. Note, Charles River laboratory did not have the in-house capabilities to run any of the hormone assays. Their contribution was to collect, process and ship the experimental samples to RTI for analysis.

The RTI laboratory prepared and analyzed QC-standard samples as well as experimental serum samples in-house as did WIL laboratory, which is described in the final report in Appendix A of the ISR. Although there was an attempt within the intact adult male assay protocol to standardize the preparation of QC standards for both laboratories, RTI and WIL prepared and designated the standards differently. For RTI, those QC standards that were included with the kit and prepared in rat serum were designated "rat serum calibrator QCs" and those that were not included in the kit but acquired from an outside source were prepared in the kit-supplied zero calibrator (human or rat serum) and designated "zero calibrator QCs". For WIL, those QCs that were included with the kit and acquired from an outside source and subsequently prepared in either rat or human serum were designated "kit QCs" or "non-kit QCs," respectively.

For RTI, descriptive statistics (e.g., sample size, mean, SD, minimum and maximum) for the QC-standard data, as determined on each TD 15 (i.e., for each assay) for each QC standard, as well as the unweighted means of the test date means are presented in Table 1 in Appendix A of the ISR. The table also includes the pooled within-assay SD, as determined by the one-way ANOVA, and the estimates of the within-assay CVs and the among-assay CVs. Values for individual QC-standard samples are presented in Table 2 in Appendix A. For WIL, corresponding data and the within-assay CVs based on variation between duplicate samples are presented in Table 3 in Appendix A.

The within-assay CVs for WIL reflect only the measurement component of variation (e.g., pipette and/or operator variation) and so underestimate the extent of the within-assay variation as discussed below. In addition, the relatively small number of QC-standard duplicate samples at the beginning and lack of any replicates in the middle and end of each assay run by both RTI and WIL also underestimates the extent of the within-assay variation. Nonetheless, the results provide some assessment of hormone assay performance regarding precision (within-assay) and repeatability (between-assay).

The ranges of CVs across the various QC-standard levels reported by RTI and WIL for each hormone assay are summarized for qualitative purposes in Table 8. Details of the extent of the variation are displayed in Tables 1 and 3 in Appendix A of the ISR. Expectedly, the WIL CVs were less than the RTI CVs since the WIL CVs were based on variation between each sample of a duplicate, whereas the RTI CVs were based on variation between each sample mean of a duplicate for two or more replicates per QC-standard sample.

Table 8. Summary of coefficients of variation (CV) for QC standard samples and assay sensitivity for each hormonal analysis using commercial assay kits in the RTI and WIL laboratories.

	RTI[1]			WIL[2]	
Hormone QC Samples	**Within-Assay CV%[3]**	**Between-Assay CV%[3]**	**Assay Sensitivity[4]**	**Within-Assay CV%[5]**	**Assay Sensitivity[4]**
Testosterone			0.04 ng/ml		0.2 ng/ml
High	5 - 4	5 - 4		4 - 3	
Medium	8	4		2	
Low	5 - 11	17- 12		1 - 10	
DHT			4.0 pg/ml		30 pg/ml
High	6 - 10	20 - 7		ND - 4	
Medium	11	18		6	
Low	30 - 12	35 - 6		2 - 1	
Estradiol			0.6 pg/ml		20 pg/ml
High	10 - 15	17 - 2		2 - 27	
Medium	9	14		ND	
Low	15 - 19	12 - 5		1 - 7	
LH			0.9 ng/ml		0.8 ng/ml
High	5 - 6	1 - 2		3 - 2	
Medium	7	7		ND	
Low	10 - 6	12 - 5		3 - 2	
FSH			0.9 ng/ml		3.1 ng/ml
High	4 - 8	0.3 - 1		1 - 3	
Medium	4	2		ND	
Low	10 - 3	14 - 4		0.1 - 1	
PRL			0.7 ng/ml		0.8 ng/ml
High	3 - 17	36 - 8		2 - 4	
Medium	4	46		ND	
Low	16 - 13	52 - 5		1 - 3	
TSH			0.5 ng/ml		2.0 ng/ml
High	3 - 7	4 - 4		7 - 2	
Medium	6	4		ND	
Low	7 - 3	0.4 - 1		1 - 2	
T3			7.0 ng/dl		20 ng/dl
High	4 - 5	0.1 - 7		0.1 - 0.2	
Medium	9	7		2	
Low	12 - 4	13 - 8		2 - 6	
T4			0.25 µg/dl		0.1 µg/dl
High	8 - 6	4 - 1		4 - 4	
Medium	11	4		1	
Low	12 - 8	24 - 5		2 - 5	

ND=not determined.

[1]Measurements between several replicates of duplicate samples within and between assays.

[2]Measurements between duplicate samples of one replicate within each assay.

[3]For the high and the low QCs, the first and second CV values represent those QCs that were included with the kit and prepared in rat serum or QCs that were acquired elsewhere and prepared in the kit-supplied zero calibrator (human or rat serum), respectively. The medium QC was the kit QC prepared in rat serum.

[4]Defined according to the manufacturer of each hormone assay kit for RTI or lowest standard according to WIL SOP.

[5]For the high and the low QCs, the first and second CV values represent those QCs that were included with the kit and acquired elsewhere, respectively, and subsequently prepared in either rat or human serum dependent upon the assay kit. The medium QC was the kit QC prepared in rat or human.

The basis for the differences in hormone assay sensitivities between the RTI and WIL laboratories is due, in part, to the way assay sensitivity was defined by each laboratory. RTI defined sensitivity according to the manufacturer of each hormone assay kit, whereas WIL defined sensitivity according to the lowest reference standard in compliance with WIL's SOPs since hormone assay sensitivity was not defined in the protocol.

Nonetheless, according to the ranges in CVs when averaged within and between assays and across the high, medium and low QC-standard samples, hormone assay performances in the RTI laboratory were considered acceptable for testosterone, LH, FSH, TSH, T_3, and T_4 (mean CVs 4 to 8%), reasonable for estradiol (mean CV 12%) and questionable for DHT and PRL (mean CVs 16 and 20%, respectively).

5.2 Vehicle Control—Observed and Expected Results

As discussed in Section 3.3, organ weight changes of the testes and epididymides are not necessarily linked as closely to final body weight changes as are the liver, prostate, SVCG, ASG, and thyroid in the intact adult male assay (O'Connor *et al.*, 1999b; 2000b); therefore, absolute weight changes of the testes and epididymides and relative weight changes (organ-to-body weight ratio) of the liver, prostate, SVCG, ASG, and thyroid were considered most appropriate for interpreting the observed results and comparing organ weight changes in the vehicle control group with expected results in historical controls. Qualitatively, the observed vehicle-control values for the relative weights of the liver, prostate, SVCG, ASG, and thyroid and the absolute weight of the testes and epididymides were comparable among the three laboratories and within the expected ranges according to historical control data as shown in Table 9. The variation within a laboratory for the various organs in the control group ranged from 7 to 9% for the liver, testes, and epididymides and from 10 to 23% for the prostate, SVCG, ASG, and thyroid. The extent of the variation was considered relatively consistent for the various organs across the three laboratories except that Charles River reported slightly higher values (23%) for the prostate, SVCG, and thyroid. Compared to historical control values which originated from 28 studies in one laboratory and where the CVs were <10% for all organs except the prostate (13%), the variation in the experimental control group ranged from 4 to 14% higher for the various organs across the CRO laboratories. Thus, the observed vehicle-control results for relative and absolute organ weights and corresponding variation within laboratories were considered consistent, not extremely different across the three laboratories, and comparable to the expected historical control results.

The observed mean absolute hormone concentration for testosterone, LH, FSH, TSH, T_3, and T_4 seemed reasonable within laboratories and relatively consistent among laboratories and, when averaged across laboratories, within the expected historical range of concentrations (Table 9). However, with respect to variation in hormone concentrations, the CVs were relatively high and inconsistent among laboratories and not necessarily in agreement with historical results, except for estradiol and LH (RTI and Charles River only) and FSH, T_3, and T_4 (all CROs). For testosterone, DHT, PRL, and TSH in the three CROs and estradiol in the WIL laboratory, the variation in concentrations ranged from 35 to 110% higher in the vehicle-control samples compared to respective historical control results. The basis for the variability in DHT and PRL concentrations in the vehicle-control samples within each laboratory may be more operational

Table 9. Vehicle control—summary of mean ± standard deviation (SD)[1] and coefficient of variation (CV) for organ weights, histology and hormones within and among laboratories and in relation to historical control data.

Key Endpoints[2]	Current Control[3]			Historical Control[4] (n=28 studies)
	RTI	WIL	Charles River	
Liver (rel)	3.8 ± 0.3 (7%)	3.7 ± 0.3 (8%)	3.5 ± 0.3 (8%)	3.9 ± 0.1 (3%)
Testes (abs)	3.3 ± 0.3 (9%)	3.3 ± 0.2 (6%)	3.3 ± 0.2 (7%)	3.2 ± 0.1 (3%)
Epididymides (abs)	1.1 ± 0.1 (9%)	1.0 ± 0.1 (8%)	1.2 ± 0.1 (9%)	1.2 ± 0.1 (5%)
Prostate (rel)	0.24 ± 0.04 (14%)	0.20 ± 0.04 (19%)	0.28 ± 0.06 (23%)	0.19 ± 0.02 (13%)
Seminal Vesicle & Coagulating Gland (rel)	0.30 ± 0.04 (13%)	0.43 ± 0.07 (17%)	0.33 ± 0.08 (23%)	0.39 ± 0.04 (9%)
Accessory Sex Gland (rel)	0.54 ± 0.06 (10%)	0.62 ± 0.09 (15%)	0.61 ± 0.11 (19%)	0.59 ± 0.03 (5%)
Thyroid (rel)	0.004 ± 0.001 (14%)	0.005 ± 0.001 (17%)	0.007 ± 0.001 (23%)	0.005 ± 0.000 (ND)
Testosterone (ng/ml)	3.4 ± 2.9 (87%)	6.1 ± 3.7 (60%)	9.9 ± 7.2 (73%)	3.1 ± 0.6 (20%)
DHT (pg/ml)	219.4 ± 126.9 (58%)	225.6 ± 153.6 (68%)	487.7 ± 245.1 (50%)	137.6 ± 44.8 (33%)
Estradiol (pg/ml)	23.0 ± 4.6 (20%)	38.4 ± 50.2 (131%)	25.4 ± 5.9 (23%)	9.9 ± 3.6 (36%)
LH (ng/ml)	1.3 ± 0.3 (20%)	0.7 ± 0.4 (55%)	2.2 ± 0.5 (23%)	3.7 ± 0.7 (18%)
FSH (ng/ml)	15.4 ± 2.0 (14%)	13.1 ± 2.3 (17%)	14.8 ± 2.2 (15%)	14.4 ± 3.1 (22%)
PRL (ng/ml)	10.8 ± 16.0 (148%)	2.9 ± 1.5 (51%)	36.5 ± 27.1 (74%)	11.5 ± 4.4 (38%)
TSH (ng/ml)	18.5 ± 12.3 (67%)	15.4 ± 6.6 (43%)	13.1 ± 6.5 (50%)	15.4 ± 3.1 (20%)
T3 (ng/dl)	87.5 ± 11.3 (15%)	80.0 ± 11.1 (14%)	81.6 ± 10.0 (12%)	75.0 ± 10.7 (14%)
T4 (ug/dl)	5.6 ± 0.6 (16%)	5.0 ± 0.8 (17%)	4.7 ± 0.6 (12%)	3.6 ± 0.7 (18%)

ND=not determined

[1]SD results converted from the SE results from the RTI, WIL and Charles River laboratory reports which are included in the peer review package.

[2]Absolute (abs) organ weight is expressed in grams (g) for testes and epididymides and relative (rel) organ weight is expressed as a percentage of final body weight for prostate, SVCG, ASG, and thyroid (n=15 animals/group).

[3]Vehicle control results are based on those reported in individual CRO laboratory reports included in the peer review package.

[4]Based on a summary of 28 studies with a comparable study design (15 animals/group) using SD rats 10 to 12 weeks of age (O'Connor *et al.,* 2002a**).**

(extrinsic) than biological (intrinsic) since the performances of these assays were considered questionable according to the extent of the variation associated with the QC-standard samples (Section 5.1). In contrast, the basis for the variability in testosterone and TSH concentrations may be more intrinsic than extrinsic since the performances of these assays were considered acceptable according to the QC-standard samples (Section 5.1). The discrepancy with estradiol concentrations, especially among laboratories, may be due, in part, to a different assay kit used by WIL (i.e., same supplier but different model) compared to the RTI and Charles River.

Thus, the repeatability of the hormonal assay results in the vehicle-control samples among laboratories and in relation to historical control results is dependent on the assay and whether the hormone concentration means or CVs are evaluated. Observed mean concentrations of testosterone, LH, FSH, TSH, T_3, and T_4 were relatively consistent across laboratories and comparable to expected results, whereas only the CVs for FSH, T_3, and T_4 were relatively low, consistent across laboratories, and comparable to expected results. It seems that hormone assays that performed in an acceptable manner according to QC-standard samples (testosterone, LH, FSH, TSH, T_3, and T_4; Table 8), are relatively inconsistent according to the results with vehicle-control samples (testosterone, TSH). The basis for this discrepancy is not known but may be attributable, in part, to the degrees of intrinsic and extrinsic variation associated with the operational and biological aspects of the hormonal assays and rat bioassay. Nonetheless, the interpretation and use of the hormonal results in the intact adult male assay are considered supplemental. They are intended to support interpretation of the principal apical and histological endpoints in the bioassay and should be weighted according to the degree of confidence one has in the performance of an assay, which is based, in part, on the results with QC standards and the biological plausibility of the hormonal results in relation to weight changes of the reproductive organs and histopathology in control and treated animals.

5.3 Linuron—Observed and Expected Results Among Laboratories

This section provides an overview of the results presented in Appendix B of the ISR following exposure to linuron and a qualitative interpretation of the results based on a weight-of-evidence approach among the multiple endpoints within the intact adult male assay using the criteria described in Section 3.3.

5.3.1 Body Weights and Food Consumption

There was no significant interaction of laboratory-by-dose or main effect of laboratory on body weight change, final body weight, and food consumption (Appendix B, Table 7). However, in a separate analysis, examination of the effect of laboratory within each dose group indicated significant effects for body weight change and food consumption for some intervals in the low- (50 mg/kg/day) and mid- (100 mg/kg/day) dose groups but not in the high (150 mg/kg/day) dose group (Appendix B, Tables 1a-c). The among-laboratory CVs in absolute values for final body weight and food consumption ranged from 2 to 8% across dose groups (Appendix B, Tables 1a-c). Thus, apart from some intervals of body weight change and food consumption, the statistical analyses indicated that other intervals of body weight change, final body weight, and food consumption were consistent across laboratories regardless of dose levels with minimal variability.

From a qualitative perspective, all treated animals showed dose-related decreases in body weight change (Appendix B, Figures 1 to 3), final body weight (Appendix B, Figure 4) and food consumption over TD 1-15 (Appendix B, Figures 5 to 7). In general toxicological studies, the degree of final body weight decrease relative to controls (Table 10, Studies 4, 5 and 6) combined with euthanasia of two moribund animals (RTI and WIL laboratories) would suggest that 150 mg/kg/day of linuron exceeded the MTD.

Table 10. Mean percent of final body weight relative to concurrent control group in the intact adult male assay using various concentrations of linuron in different laboratories.

			Linuron (mg/kg/day)				
Study	**Type**	**Laboratory ID**	**25**	**50**	**75**	**100**	**150**
1	Prevalidation	DuPont	97%	96%	ND	91%	89%
2	Prevalidation	Dow	ND	ND	ND	86%	82%
3	Prevalidation	RTI 2005	95%	90%	86%	88%	ND
4	Inter-laboratory	RTI 2006	ND	94%	ND	87%	81%
5	Inter-laboratory	WIL 2006	ND	90%	ND	83%	80%
6	Inter-laboratory	CRL 2006	ND	88%	ND	85%	80%

ND=not determined.
(n=15 animals/dose level)

In a prevalidation study (O'Connor *et al.,* 2002a), an 11% decrease in final body weight occurred at 150 mg/kg/day linuron compared with the 19 to 20% decrease observed herein across the three CRO laboratories (Table 10, Studies 4, 5 and 6). Moreover, there were no mortalities reported in the O'Connor study. Even in Studies 3 through 6 (Table 10) with analytical confirmation of dose level concentrations, there was some variability in the magnitude of the effects on final body weight. At 50 mg/kg/day, final body weight varied from 88 to 94%, whereas DuPont's value was 96% of concurrent controls. At 100 mg/kg/day, the final body weight ranged from 83 to 88% of concurrent control body weights in the CRO laboratories, whereas the DuPont value was 91%. There appeared to be less variability at 150 mg/kg/day, where final body weights ranged from 80 to 81% of control values; the DuPont value was 89% of the control final body weight. Nonetheless, the reason for the differential magnitude of these mean body weight effects in response to linuron in different laboratories at different times as well as the basis for two moribund animals is not known. Neither DuPont nor Dow conducted analytical characterization to confirm dose levels during prevalidation, and the same lot of linuron was only confirmed for the inter-laboratory validation studies. In a prevalidation study done by RTI (EPA, 2005), it was noted that the high-dose level of linuron in methylcellulose was difficult to keep well mixed in suspension during dosing and sampling and, in the inter-laboratory validation study herein, RTI, WIL, and Charles River used low phytoestrogen-containing feed which was not used by DuPont and Dow and, speculatively, may have impacted the sensitivity of body weight to linuron treatment.

5.3.2 Organ Weights and Histopathology Supported by Hormonal Changes

There was no significant interaction of laboratory-by-dose or main effect of laboratory on target organ weights, except for a significant ($P<0.04$) main effect of laboratory on relative liver weight (Appendix B, Table 9). In a separate analysis, examination of the effect of laboratory within each dose group on liver weight indicated a tendency that approached significance ($P<0.09$) for

the low and mid-dose groups but not for the high-dose group (Appendix B, Tables 3a-c). The among-laboratory variations across dose groups for all target organ weights including absolute and relative changes were ≤10% (Appendix B, Tables 3a-c). Thus, apart from liver weight in the lower dose groups, the statistical analyses indicated that target organ weights were relatively consistent across laboratories regardless of dose levels with minimal variability.

As indicated earlier (Sections 3.3), organ weight changes of the testes and epididymides are not necessarily linked as closely to final body weight changes as are the liver, prostate, SVCG, ASG, and thyroid (O'Connor *et al.*, 1999b; 2000b). Thus, absolute weight changes of the testes and epididymides (Appendix B, Figures 9 to 12) and relative weight changes (organ-to-body weight ratio) of the liver, prostate, SVCG, ASG, and thyroid (Appendix B, Figures 17 and 22 to 25) were considered key endpoints most appropriate for determining the endocrine effects of test materials on target organ weights in the intact adult male rat assay. From a qualitative perspective and considering that linuron is anti-androgenic, the observed absolute weight of the testes and testes histopathology as well as relative weights of the prostate, SVCG, and ASG were inconsistent across laboratories; conversely, a decrease in absolute weight of the epididymides was highly consistent across laboratories and in accordance with the expected historical results (Table 11). Although the anti-androgenic effects of linuron were not necessarily consistent among laboratories for many of the same endpoints, the observed responses within a laboratory were supportive of the expected anti-androgenic effects of linuron as follows:

1. In the RTI laboratory, there was a significant linear trend (i.e., dose response) and decrease in absolute weight of the epididymides and increased histopathology of the testes (21% of the animals).

2. In the WIL laboratory, there was a significant decrease in absolute weight of the testes and epididymides and relative weight of the prostate, SVCG, and ASG.

3. In the Charles River laboratory, there was a significant dose response and a decrease in absolute weight of the epididymides associated with significant decreases in concentrations of testosterone and DHT in a dose-responsive manner.

Apart from the apparent anti-androgenic effects on target organs within laboratory, linuron appeared to have an indirect effect on the thyroid, probably because of its enhanced effect on liver metabolic activity. Relative liver weight increased in two of the three laboratories and was associated with a consistent decrease in T_3 and T_4 concentrations across all laboratories in agreement with historical results (Table 11). Correspondingly, there was no consistent increase in TSH concentrations or any observed histopathological effects associated with the thyroid; however, there was a consistent increase in relative thyroid weight across all laboratories which concurs with expected historical results.

5.3.3 Hormone Concentrations in Experimental Serum Samples

There was no significant interaction of laboratory-by-dose on hormone concentrations (Appendix B, Table 11), but there were significant main effects of laboratory for testosterone ($P<0.009$), LH, ($P<0.03$), T_3 ($P<0.04$), T_4 ($P<0.002$), and FSH ($P<0.04$). In a separate analysis, examination

of the effect of laboratory within each dose group on hormone concentration indicated a significant effect ($P<0.003$) for T_4 in the mid- and high-dose groups (Appendix B, Tables 5a-c) and a tendency ($P<0.1$) that approached significance for T_3, LH, and FSH in the low-dose group.

Table 11. Linuron—summary of observed results for organ weights, histology and hormones within and among laboratories and in relation to expected historical results.

Key Endpoints[1]	Observed[2]			Expected Historical[3]
	RTI	WIL	Charles River	
Liver (rel)	No Change	↑ L, M, H (LT)	↑ H (LT)	↑
Testes (abs)	No Change	↓ M	No Change	No Effect
Testis Histopath	Slight ↑ in Seminiferous Tubule Degeneration	No Change	No Change	Minimal Spermatid Retention
Epididymides (abs)	↓ M, H (LT)	↓ M (LT)	↓ M, H (LT)	↓
Epididymis Histopath	No Change	No Change	No Change	No Effect
Prostate (rel)	No Change	↓ M	No Change	↓
Seminal Vesicle & Coagulating Gland (rel)	No Change	↓ M	No Change	↓ (Not Significant)
Accessory Sex Gland (rel)	No Change	↓ M	No Change	↓
Testosterone	No Change	No Change	↓ L, M, H (LT)	↓
DHT	No Change	No Change	↓ H (LT)	↓
Estradiol	↑ L, M, H (LT)	No Change[4]	↑ L, M, H (LT)	↑
LH	No Change	No Change	↓ L, M	↓
FSH	↑ M (LT)	↑ L, M, H (LT)	No Change	No Effect
PRL	No Change	No Change[5]	↓ L, M, H (LT)	↓
Thyroid (rel)	↑ L, M, H (LT)	↑ M, H (LT)	↑ L	↑ L, M, H [6]
Thyroid Histopath	No Change	No Change	No Change	No Effect
TSH	↓ M	No Change	No Change	No Effect
T3	↓ M, H (LT)	↓ M, H (LT)	↓ L, M, H (LT)	↓
T4	↓ L, M, H (LT)	↓ L, M, H (LT)	↓ L, M, H (LT)	↓

[1]Organ weight change compared to the control is absolute (abs) for testes and epididymides or relative to final body weight (rel) for prostate, SVCG, ASG, and thyroid.

[2]The directional change in the observed results is based on a significant difference ($P\leq0.05$) from control at the low- (L, 50 mg/kg/d), mid- (M, 100 mg/kg/d) or high- (H, 150 mg/kg/d) dose levels and (LT) indicates a significant ($P\leq0.05$) linear trend as reported in the individual laboratories (n=15 animals/dose level). Histopathology is based on whether the findings are related to treatment relative to the high-dose group only.

[3]Based on a summary of numerous studies with the same compounds using a comparable study design as reviewed (O'Connor *et al.,* 1999b; 2002a,b,c).

[4]When outliers were removed, a significant increase was seen at all three dose levels.

[5]When outliers were removed, a significant decrease was seen at all three dose levels.

[6]These increases were not significant at the $P\leq0.05$ level.

The distribution of the among-laboratory CVs for various serum hormones in response to linuron are shown in Table 12. The extent of the means and variation within and among laboratories for each hormone are depicted in Figures 27 to 34 in Appendix B.

Table 12. Distribution of among-laboratory coefficients of variation in serum hormones in response to linuron.

Dose Levels (mg/kg/d)	Among-laboratory CVs		
	≤10%	11-17%	>17%
50	T4, T3, FSH, Estradiol	LH, TSH	T, DHT, PRL
100	T3, FSH, Estradiol	LH, TSH	T, DHT, T4, PRL
150	T3, FSH	LH, TSH, Estradiol	T, DHT, T4, PRL

(n=15 animals/dose level)

From a qualitative perspective, the most extreme among-laboratory CVs across dose groups ranged from 39 to 41% for testosterone, 42 to 44% for PRL, 22 to 25% for DHT and 9 to 42% for T_4 (Appendix B, Tables 5a-c). Within-laboratory, PRL concentrations were markedly lower in the Charles River laboratory than in the RTI and WIL laboratories. Although not as extreme, concentrations of testosterone and DHT were lower in the Charles River laboratory than in the RTI and WIL laboratories. Correspondingly, for testosterone, DHT, and PRL, the extent of the among-laboratory CVs was associated with relatively high within-laboratory variation for testosterone (range across dose groups and laboratories, 21 to 30%), DHT (16 to 25%) and PRL (24 to 52%) in the WIL and RTI laboratories compared to the Charles River laboratory for testosterone (range across dose groups, 9 to 13%), DHT (12 to 17%) and PRL (1 to 4%). The basis for the extent of the among-laboratory variability for T_4 in the mid- and high-dose groups is unknown since concentrations were comparable across laboratories and the within-laboratory variability was relatively low and consistent across dose groups and laboratories (1 to 5%). In addition, the directional changes in T_3 and T_4 were consistent across laboratories and in agreement with historical results as shown in Table 11.

Although not necessarily reflected in the among-laboratory CVs and concentrations for estradiol, the within-laboratory variation in response to treatment ranged across dose groups from 53 to 60% for the WIL laboratory, whereas the CVs for the other two laboratories ranged from 10 to 14%. Perhaps the different model estradiol kit used by WIL compared to RTI was more variable. Thus, from a statistical perspective, the hormonal results were inconsistent across laboratories for testosterone, LH, T_3, T_4 and FSH (main effect of laboratory, $P<0.04$) and, from a qualitative perspective, highly variable within (>20%) and among (>17%) laboratories primarily for testosterone, DHT, and PRL.

5.3.4 Interpretation of the Anti-Androgenic Results of Linuron Using a Weight-of-Evidence Approach within the Bioassay

The approach to interpretation of the results among the multiple endpoints in the 15-day intact adult male rat assay as to whether test chemical exposure induced an effect on the EAT hormonal systems has been described in Section 3.3. In the experiments conducted independently by RTI, WIL, and Charles River laboratories using linuron (anti-androgen) and summarized herein, terminal body weights were decreased approximately 20% in the high-dose (150 mg/kg/day) group compared to the control group. While the magnitude of body weight decrease exceeded

10% relative to controls, data from feed restriction studies by O'Connor *et al.* (1999b; 2000b) discussed in Section 3.3 indicated that body weight changes of this magnitude do not necessarily affect the targeted endocrine endpoints except, perhaps, for serum T_3 and T_4 concentrations. Hence, the linuron data from these CROs are considered interpretable at all dose levels. From Table 11, RTI, WIL and Charles River laboratories reported a decrease in absolute epididymal weight; additionally, RTI reported testicular histopathology (21% of the animals) and WIL reported a decrease in absolute testes and relative prostate, SVCG, and ASG weights. Hormonally, RTI reported an increase in estradiol and FSH, WIL reported an increase in FSH, and Charles River reported a decrease in testosterone, DHT, LH, and PRL and an increase in estradiol. Although statistically there was a relatively high degree of consistency across the CRO laboratories, especially for the apical endpoints, qualitatively, there were inconsistencies across laboratories for apical, histomorphological, and hormonal endpoints. However, considering the consistent change in observed epididymal weights among laboratories and concordance with expected results as well as the results within laboratory (RTI with the change in epididymal weights complemented with testicular histopathology; WIL with the change in testicular, epididymal, prostate, SVCG, and ASG weights; and Charles River with the change in epididymal weight complemented with relevant hormonal changes) that were in agreement with historical results on an individual laboratory basis, linuron would likely be flagged as having a positive endocrine effect on the androgen hormonal pathway, which was the conclusion reached independently by each CRO as indicated in the individual laboratory reports.

With respect to the thyroid endpoints, all laboratories had increased relative thyroid weight and decreased T_3 and T_4 concentrations but there was no complementary increase in TSH concentrations and no accompanying histomorphological changes in the thyroid. Even though the T_3 and T_4 results may have been confounded by decreased final body weight and, perhaps, a consequence of induced liver enzymes, there was no accompanying increase in TSH to account for the increase in thyroid weight and no thyroid histopathology. On the basis of biological plausibility, therefore, linuron would not likely be considered as having a direct positive effect on the thyroid hormonal system according to the design of the intact adult male assay and results in individual laboratories.

5.4 Phenobarbital—Observed and Expected Results Among Laboratories

This section provides an overview of the original results presented in Appendix B of the ISR following exposure to phenobarbital and a qualitative interpretation of the results based on a weight-of-evidence approach among the multiple endpoints within the intact adult male assay using the criteria described in Section 3.3.

5.4.1 Body Weights and Food Consumption

There was no significant interaction of laboratory-by-dose or main effect of laboratory on body weight change, final body weight, and food consumption (Appendix B, Table 8). In a separate analysis, examination of the effect of laboratory within each dose group indicated significant effects for body weight change and food consumption for some intervals in the high- (100 mg/kg/day) dose group but not in the low- (25 mg/kg/day) and mid- (50 mg/kg/day) dose groups (Appendix B, Tables 2a-c). The among-laboratory CVs in absolute values for final body weight and food consumption ranged from 2 to 7% across dose groups (Appendix B, Tables 2a-c).

Thus, apart from some intervals of body weight change and food consumption, the statistical analyses indicated that other intervals of body weight change, food consumption, and, especially, final body weight were consistent across laboratories regardless of dose levels with minimal variability.

From a qualitative perspective, there were no clear dose-related decreases in body weight change, final body weight, and food consumption over TDs 1 to 15 (Appendix B, Figures 3, 4, and 7). At 100 mg/kg/day, phenobarbital produced a consistent decrease across laboratories for body weight gain, final body weight, and food consumption. As shown in Table 13, a final body weight decrease was expected in the animals in the high-dose group (5 to 10%), due to the clinical manifestations seen previously with administration of phenobarbital as well as its action as a hepatic enzyme inducer that enhances the clearance of thyroid hormones (reviewed in Capen, 2001).

Table 13. Mean percent of final body weight relative to concurrent control group in the intact adult male assay using various concentrations of phenobarbital in different laboratories.

			Phenobarbital (mg/kg/day)			
Study	**Type**	**Laboratory ID**	**5**	**25**	**50**	**100**
1	Prevalidation	DuPont	100%	105%	101%	98%
2	Prevalidation	DuPont	100%	100%	101%	97%
3	Prevalidation	Dow	100%	105%	ND	ND
4	Prevalidation	Dow	100%	99%	ND	ND
5	Inter-laboratory	RTI 2006	ND	100%	100%	90%
6	Inter-laboratory	WIL 2006	ND	100%	98%	95%
7	Inter-laboratory	CRL 2006	ND	102%	101%	95%

ND=not determined
(n=15 animals/dose level)

In prevalidation studies by O'Connor *et al.* (1999b; 2002b), terminal body weights were decreased 2 to 3% with 100 mg/kg/day of phenobarbital (Table 13, Studies 1 and 2), which is comparable with the current results. In the present studies (Table 13, Studies 5, 6 and 7), however, two males were found dead (WIL and Charles River laboratories) and two were euthanized moribund (RTI laboratory) in the high-dose group, suggesting that 100 mg/kg/day exceeded the MTD, which is in contrast with no mortalities in the O'Connor *et al.* (2002b) study. In the studies by O'Connor *et al.* (1999b, 2002b), a range-finding study was conducted and dose levels were selected to produce the maximum pharmacological effect not to exceed an MTD. Dose levels in the inter-laboratory validation exercise were based on O'Connor's previous work. The basis for the differential effects concerning animal death in the high-dose group is not known. In the O'Connor *et al.* studies, analytical characterization to confirm dose levels was not done during prevalidation and the same lot of phenobarbital was only confirmed for the inter-laboratory exercise. In addition, the three CROs used low phytoestrogen-containing feed which was not used in the O'Connor *et al.* studies and, speculatively, may have impacted the sensitivity of body weight to phenobarbital treatment.

5.4.2 Organ Weights and Histology Supported by Hormonal Changes

There was no significant interaction of laboratory-by-dose or main effect of laboratory on target organ weights, except for a significant main effect of laboratory on absolute and relative liver weight (Appendix B, Table 10). In a separate analysis, examination of the effect of laboratory within each dose group indicated a significant ($P<0.002$) effect on absolute and relative liver weight and a tendency that approached significance ($P<0.1$) for relative testis, paired testes, and ASG weights in the high-dose group (Appendix B, Tables 4a-c). The among-laboratory variations across dose groups for all target organ weights including absolute and relative changes were ≤9% (Appendix B, Tables 4a-c). Thus, apart from liver weight in the high-dose group, the statistical analyses indicated that target organ weights were relatively consistent across laboratories regardless of dose levels with minimal variability.

Recall that absolute weight changes for the testes and epididymides (Appendix B, Figures 9 to 12) and relative (organ to final body weight) weight changes for the liver, prostate, SVCG, ASG, and thyroid (Appendix B, Figures 17 and 22 to 25) were considered most appropriate for determining the endocrine effect of test materials on target organ weights in the intact adult male rat assay because of the differential dependence that reproductive organ weights have with final body weight changes (O'Connor *et al.*, 1999b; 2000b). From a qualitative perspective and considering that phenobarbital is a thyroid toxicant that alters thyroid function indirectly through enhanced liver metabolism and excretion of thyroid hormones (McClain *et al.* 1989), relative liver weights significantly increased in a dose-responsive manner in association with significant decreases in T_3 and T_4 concentrations (Appendix B, Figures 17, 29, and 30) that were consistent across laboratories and with expected historical results (Table 14). Correspondingly, relative thyroid weights significantly increased in a dose-responsive manner in association with significant dose-related responses and significant increases in TSH concentrations (Appendix B, Figures 25 and 28) that were also consistent across laboratories and with historical results. Except for the RTI laboratory with no observed histopathology, histopathological results of the thyroid were observed in 100% of the animals at WIL and 87% of the animals at Charles River, which agrees with the expected historical results. It should be noted that histomorphological assessment of the thyroid was done only on the control and high-dose groups and that the WIL and Charles River laboratories evaluated two or more histological sections of the thyroid, whereas the RTI laboratory prepared and evaluated only one thyroid section.

Apart from the apical, histological and hormonal effects associated with the thyroid, phenobarbital appeared to have effects on reproductive organs and hormones. Although there were no absolute changes in weight of the testes and epididymides across the laboratories, which concur with expected historical results (Table 14), histopathologically there was spermatid retention associated with the testes in 36% of the animals from WIL. Although this result was supported by expected historical results, increased mononuclear cell infiltration associated with the epididymides in 54% of the animals from RTI and 100% of the animals from WIL was not supported by historical results (Table 14). In regard to the latter, the testicular and epididymal effects observed in the WIL laboratory appear to be associated with dose-dependent decreases in testosterone, DHT and LH concentrations that agree with the expected historical results.

Table 14. Phenobarbital—summary of observed results for organ weights, histology and hormones within and among laboratories and in relation to expected historical results.

Key Endpoints[1]	Observed[2]			Expected Historical[3]
	RTI	WIL	Charles River	
Liver (rel)	↑ L, M, H (LT)	↑ L, M, H (LT)	↑ L, M, H (LT)	↑
Testes (abs)	No Change	No Change	No Change	No effect – 2 studies
Testis Histopath	No Change	spermatid retention	No Change	No effect/slight spermatid retention+
Epididymides (abs)	No Change	No Change	No Change	No effect – 2 studies
Epididymis Histopath	↑ severity of mononuclear cell infiltration	↑ multifocal mononuclear cell infiltrates	No Change	No effect – 2 studies
Prostate (rel)	↑ H (LT)	↑ M	No Change	No effect – 2 studies
Seminal Vesicle & Coagulating Gland (rel)	↑ H (LT)	No Change (LT)	No Change	No effect – 2 studies
Accessory Sex Gland (rel)	↑ H (LT)	No Change (LT)	No Change	No effect – 2 studies
Testosterone	No Change	↓ M, H (LT)	↓ M, H (LT)	↓ - 2 studies (not significant)
DHT	No Change	↓ M, H (LT)	↓ M, H (LT)	No effect/↓
Estradiol	↑ M, H (LT)	↑ H (LT)	↑ L, M, H (LT)	No effect/↑
LH	No Change	No Change	↓ L, M, H (LT)	↓ - 2 studies
FSH	↓ L, M, H (LT)	↓ L, M	↓ M, H (LT)	↓/ No effect
PRL	No Change	No Change	↓ L, M, H (LT)	↓ - 2 studies
Thyroid (rel)	↑ L, M, H (LT)	↑ L, M, H (LT)	↑ L, M, H (LT)	↑ - 2 studies
Thyroid Histopath	No Change[4]	↑ follicular cell height ↓ colloid area ↑ mitotic figures	Hypertrophy and hyperplasia of follicular epithelium	Pale staining and/or depletion of colloid; follicular cell hypertrophy
TSH	↑ H	↑ L, M, H (LT)	↑ L, M, H (LT)	↑ - 2 studies
T3	↓ L, M, H (LT)	↓ M, H (LT)	↓ L, M, H (LT)	↓ - 2 studies
T4	↓ L, M, H (LT)	↓ L, M, H (LT)	↓ L, M, H (LT)	↓ - 2 studies

[1]Organ weight change compared to the control is absolute (abs) for testes and epididymides or relative to final body weight (rel) for prostate, SVCG, ASG, and thyroid.

[2]The directional change in the observed results is based on a significant difference ($P \leq 0.05$) from control at the low- (L, 25 mg/kg/d), mid- (M, 50 mg/kg/d) or high- (H, 100 mg/kg/d) dose levels and (LT) indicates a significant ($P \leq 0.05$) linear trend as reported in the individual laboratories (n=15 animals/dose level). Histopathology is based on whether the findings are related to treatment relative to the high-dose group only.

[3]Based on a summary of numerous studies with the same compounds using a comparable study design as reviewed (O'Connor *et al.,* 1999b; 2002a,b,c).

[4]Based on evaluation of one section as opposed to two or three sections in the other laboratories.

In association with the consistent decrease in FSH concentrations across laboratories, there was a consistent increase in estradiol concentrations across laboratories (Table 14), which also seems to agree with the expected historical results.

There was no change in relative weights of the prostate, SVCG, and ASG across two of the three laboratories, which concurs with the historical results (Table 14).

It is generally accepted that phenobarbital has a depressive effect on the central nervous system and enhanced effect on liver enzymatic activity. Speculatively, therefore, decreased serum concentrations of gonadotropins at the hypothalamic/pituitary level may have led to decreased steroidogenesis at the testicular level resulting in organ weight and histomorphological changes that were detected in some of the laboratories.

5.4.3 Hormone Concentrations in Experimental Serum Samples

There was no significant interaction of laboratory-by-dose on hormone concentrations (Appendix B, Table 12) but there were significant main effects of laboratory for testosterone ($P<0.05$), LH, ($P<0.01$), T_4 ($P<0.04$), and PRL ($P<0.004$). In addition, there were tendencies that approached significance for TSH ($P<0.08$) and DHT ($P<0.1$). In a separate analysis, examination of the effect of laboratory within each dose group on hormone concentration indicated a significant effect for LH ($P<0.018$) and T_4 ($P<0.002$) in the mid-dose group and T_4 ($P<0.003$) in the high-dose group (Appendix B, Tables 6a-c).

The distribution of the among-laboratory CVs for various serum hormones in response to phenobarbital are shown in Table 15. The extent of the means and variation within and among laboratories for each hormone are depicted in Figures 27 to 34 in Appendix B of the ISR.

Table 15. Distribution of among-laboratory coefficients of variation in serum hormones in response to phenobarbital.

Dose Levels (mg/kg/d)	Among-laboratory CVs		
	≤10%	11-19%	>19%
25	T4, T3, FSH	LH, TSH, Estradiol	T, DHT, PRL
50	T3, FSH, Estradiol	LH, TSH, T4	T, DHT, PRL
100	T3, FSH, Estradiol	LH, TSH, T4	T, DHT, PRL

(n=15 animals/dose level)

From a qualitative perspective, the extent of the among-laboratory variations in testosterone, DHT and PRL in response to phenobarbital (Appendix B, Tables 6a-c) is generally similar to the responses to linuron as detailed in Section 5.3.3. In addition, concentrations of testosterone, DHT, and PRL and within-laboratory variation were generally lower in the Charles River laboratory compared to respective values in the RTI and WIL laboratories. Directional changes in estradiol, FSH, TSH, T_3, and T_4 concentrations appeared consistent across laboratories and, generally, were in agreement with historical results (Table 14). Thus, from a statistical perspective, the hormonal results were inconsistent across laboratories for testosterone, LH, T_4, and PRL and, from a qualitative perspective, highly variable within laboratories for testosterone, DHT and PRL in the RTI and WIL laboratories. The source of the apparent discrepancy between the statistical and qualitative results regarding T_4 is not known.

5.4.4 Interpretation of the Thyroidogenic Results of Phenobarbital Using a Weight-of-Evidence Approach within the Bioassay

The approach to interpretation of the results among the multiple endpoints in the 15-day intact adult male rat assay as to whether test chemical exposure induced an effect on the EAT hormonal systems has been described in Section 3.3. Final body weight decreases were within 10% of control values for each CRO at the highest dose level (100 mg/kg/day) indicating that all target endocrine endpoints were interpretable at all dose levels. As shown in Table 14, all of the laboratories had increased relative thyroid weights coupled with increased TSH and decreased T_3 and T_4 concentrations. Thyroid histopathological changes were also noted at the WIL and Charles River laboratories. Each laboratory also reported increased relative liver weights. These results are consistent with historical data and consistent with thyroid perturbation by phenobarbital. Changes in serum concentrations for other hormones were noted, a finding that may be related to metabolic induction by the liver. Thus, based on the statistical and qualitative results within and among laboratories relative to expected historical results, phenobarbital exposure to intact adult male rats would likely be flagged as having a positive endocrine effect on the thyroid hormonal pathway and, perhaps, on the androgen pathway which was the conclusion reached independently by each CRO as indicated in the individual laboratory reports.

5.5 Conclusion of Inter-laboratory Validation

There were no statistically significant effects due to the testing laboratory when looking at final body weight, absolute weight of the testes and epididymides, relative weight of the prostate, SVCG, ASG, and thyroid gland with either linuron or phenobarbital. Relative liver weight was the only non-endocrine target organ that was significantly different among laboratories for both linuron and phenobarbital. For both test compounds, the among-laboratory CVs for final body weight and target organ weights including the liver were ≤10%. For linuron, there was no significant effect due to the testing laboratory when looking at serum concentrations of DHT, estradiol, PRL, and TSH and, for phenobarbital, there was no significant effect due to the testing laboratory when looking at DHT, estradiol, FSH, TSH, and T_3 concentrations. However, there was a significant effect of laboratory for the other hormones according to respective test chemicals.

Despite the effect or lack of effect of the testing laboratory on serum hormone concentrations, the among-laboratory CVs sometimes seemed contradictory (Table 16) with the statistical analyses. That is, mean DHT concentrations were not significantly different among laboratories, yet the among-laboratory CVs were >17% with linuron and >19% with phenobarbital. Conversely, mean T_3 concentrations were significantly different among laboratories with linuron, yet the among-laboratory CVs were ≤10%. Although the basis for this contradiction is not clear, it seems that for some hormonal changes the results among laboratories can be extremely variable as indicated by the extent of the CVs but yet statistically consistent across laboratories or slightly variable and inconsistent across laboratories. Nonetheless, within the intact adult male assay, significant changes in serum hormone concentrations alone are not considered a positive indicator of endocrine disruption, but are considered in context with organ weight and histological changes. The apparent contradictions betweens CVs and statistical results for

Table 16. Among-laboratory coefficients of variation (CVs) for hormone concentrations taken at termination of the intact adult male assay following linuron and phenobarbital treatments.

Linuron			
Dose Levels (mg/kg/d)	**Among-laboratory CVs**		
	≤10%	**11-17%**	**>17%**
50	T4, T3, FSH, Estradiol	LH, TSH	T, DHT, PRL
100	T3, FSH, Estradiol	LH, TSH	T, DHT, T4, PRL
150	T3, FSH	LH, TSH, Estradiol	T, DHT, T4, PRL
Phenobarbital			
Dose Levels (mg/kg/d)	**Among-laboratory CVs**		
	≤10%	**11-19%**	**>19%**
25	T4, T3, FSH	LH, TSH, Estradiol	T, DHT, PRL
50	T3, FSH, Estradiol	LH, TSH, T4	T, DHT, PRL
100	T3, FSH, Estradiol	LH, TSH, T4	T, DHT, PRL

(n=15 animals/dose level)

hormonal changes were not observed with the apical (organ weight) endpoints, except for liver weight.

In the linuron group, a decrease in final body weights averaged across laboratories ranged from 10 to 20% from the low- to the high-dose groups compared to the control group. Despite the degree of final body weight decrease in the treated groups relative to the control group, there were specific endocrine-related changes in the treated groups consistent with the anti-androgenic activity of linuron. Although not necessarily consistent across laboratories, within each laboratory these effects included significant dose-related responses and a significant decrease in absolute weight of the epididymides and histopathology of the testes (RTI); a significant decrease in absolute weight of testes and epididymides and relative weight of the prostate, SVCG, and ASG (WIL); and a significant decrease in absolute weight of the epididymides associated with significant decreases in serum concentrations of testosterone and DHT (Charles River). Although these effects were seen when final body weights in the treated groups were decreased >10% relative to the control group, the effects were considered to be endocrine-mediated based on the results of feed restriction studies that indicated non-specific alterations in these endocrine endpoints did not occur until final body weights decreased 26% relative to control-animal body weight (Tables 5 and 6). Qualitatively, the observed results within laboratory (androgen-dependent organ weights, histology, and all but the thyroid hormones) corresponded to the expected results 46% (5/13), 46% (5/13) and 69% (9/13) of the time in the RTI, WIL, and Charles River laboratories, respectively.

Although linuron may not have had a direct effect on the thyroid hormonal system, all laboratories observed significant decreases in T_3 and T_4 concentrations and a significant increase in relative thyroid weight without any significant corresponding changes in TSH concentration or detectable thyroid histopathology, which were consistent with expected historical results.

In the phenobarbital group, final body weights averaged across laboratories ranged from 0 to 10% from the low- to the high-dose groups relative to the control group. Despite the minimal change in final body weights in the treated groups compared to the control group, there were significant dose-related responses and a significant increase in relative liver weights, which was

likely due to enhanced hepatic metabolizing enzymes. As a result, there was a significant decrease in serum T_3 and T_4 concentrations, a significant increase in serum TSH concentrations and a significant increase in relative thyroid weight within each laboratory. In addition, WIL and Charles River but not RTI detected histopathological changes in the thyroid gland. Qualitatively, the observed thyroidogenic results within and among laboratories (organ weight, histopathology, and hormone concentrations) corresponded to the expected results 80% (4/5), 100%, (5/5) and 100% (5/5) of the time in the RTI, WIL, and Charles River laboratories, respectively.

In conclusion, the statistical comparison of the apical endpoints among the three CRO laboratories indicated that target organ weight responses to linuron and phenobarbital in the intact adult male assay were highly consistent, with relatively low variability across laboratories. Statistical comparison of hormonal endpoints and qualitative comparison of apical, histopmorphological, and hormonal endpoints among laboratories, however, indicated lack of consistency across laboratories and concordance with historical results for some of the multiple endpoints following linuron and phenobarbital exposure. The basis for the inconsistencies among laboratories is not fully known but is likely attributable, in part, to the degrees of operational and biological variation associated with various aspects of the rat bioassay and hormonal assays conducted within laboratories. Nonetheless, the results were interpreted using a weight-of-evidence approach combined with biological plausibility among the multiple endpoints within the intact adult male assay; with this approach, linuron and phenobarbital were identified independently by each CRO laboratory as having an effect on the androgen and thyroid hormonal systems, respectively. Hence, the end results were consistent across all laboratories and concurred with the known MOA of each test chemical.

6.0 EPA Validation Criteria

6.1 Validation Criteria Prescribed by the EDSP

This section is meant to provide a synopsis and reference to relevant sections within the text of this document in support of the level at which the 15-day intact adult male rat assay meets each of the validation criteria prescribed by the EDSP (EPA, 2007b) as initially presented in Section 1.

6.1.1 The scientific and regulatory rationale for the test method, including a clear statement of its proposed use, should be available.

> The scientific basis for the intact adult male assay is discussed in Sections 2 through 4. Briefly, the 15-day intact adult male rat assay was developed by the chemical industry (DuPont) to identify EACs and their MOAs in support of product registration. The bioassay consists of multiple endpoints, principally, terminal weights of primary and secondary sex organs and thyroid gland, histomorphology of the testes, epididymides and thyroid, and serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones. Results of the comparisons of these endpoints between control and treated groups at three dose levels are evaluated on a weight-of-evidence basis within the bioassay (Section 3.3) to determine whether a chemical has a positive effect on the EAT hormonal systems.

The intact adult male rat assay has been used to detect ER binding agonists/antagonists, AR binding agonists/antagonists, progesterone binding agonists/antagonists, steroid biosynthesis inhibitors, gonadotropin and thyroid modulators either directly or indirectly by altering the HPG or HPT axes, and PRL modulators through neuroendocrine pathways.

The extent of the diversity of this assay to detect effects, especially on the EAT hormonal systems using a variety of EACs has been hypothesized, tested, and reported in numerous studies published in peer-reviewed scientific journals as cited throughout Sections 2 to 4.

The potential regulatory basis for the intact adult male assay within the EPA is primarily discussed in Sections 1 and 2. The purpose of the intact adult male screening assay is to detect various MOAs, especially AR agonists/antagonists, steroid biosynthesis inhibitors, gonadotropin and thyroid modulators either directly or indirectly through intact HPG or HPT axes using a weight-of–evidence approach within the bioassay. In addition, since this assay is a candidate for an EDSP Tier-1 battery, results from within the bioassay are expected to contribute to the results of other assays in the battery and, using a weight-of-evidence approach within the battery, determine whether a chemical substance has a positive or negative effect on the EAT hormonal systems.

6.1.2 The relationship of the endpoints determined by the test method to the in vivo biologic effect and toxicity of interest must be addressed.

The comprehensive and unique approach of using terminal weights of primary and secondary sex organs and thyroid gland, histomorphology of the testes, epididymides and thyroid, and serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones in the intact adult male assay to detect effects on the EAT hormonal pathways is, in part, based on prevalidation exercises discussed throughout Section 3.

Two primary goals of the simulation were to, first, determine which of the many endpoints evaluated in the exercise were not relevant and, therefore, could be excluded to standardize the assay protocol and, second, test the hypothesis that chemical-responsive "fingerprints" could be developed for endocrine-related events. The results of this simulation are discussed in Section 3.2. Briefly, endpoints were removed that failed to improve the assay and those that were kept are used in the current standardized protocol (Appendix C). Most of the apical endpoints (target organ weights) are primarily androgen dependent, and are sensitive to androgen agonists and antagonists but also to estrogen agonists and antagonists, as discussed in Section 3.1.1. Thyroid organ weight is also included to detect direct and indirect effects on the thyroid gland. Histomorphology of the testes, epididymides, and thyroid gland are included as more sensitive endpoints to detect effects on the EAT hormonal system. Serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones are included to add to the weight of evidence involving organ weight and histomorphological changes and to get initial information on potential MOAs. Finally, terminal or final body weight and liver weight are included in the overall list of endpoints to differentiate endocrine- from

nonendocrine-mediated effects and facilitate interpretation of the results as direct or indirect effects on the endocrine system, as discussed in Section 3.3.

Secondarily, the exercise demonstrated that chemical-responsive "fingerprints" could be developed to identify EACs and their MOAs and aid in the characterization of their underlying mechanisms of action. However, since the concept of developing chemical-responsive "fingerprints" using the intact adult male assay is a relatively novel approach in toxicological studies combining multiple apical endpoints (reproductive organ weights) and histology with systemic changes in serum concentrations of reproductive steroids, gonadotropins and thyroid hormones, many more studies with a variety of chemicals will be required to establish this concept as a reliable and practical method to identify potential endocrine disruptors and predict their modes or mechanisms of action. As an example, the chemical-responsive "fingerprint" of an androgen steroid biosynthesis inhibitor such as ketoconazole is a decrease in ASG weight and serum testosterone concentration, and an increase in serum LH concentration and, for an AR antagonist such as flutamide, there is a decrease in ASG weight and an increase in serum testosterone and LH concentrations.

The potential toxicity of a chemical substance is characterized by directional and temporal changes in final weight of target organs and corresponding serum hormone concentrations as well as the degree of histopathology when compared to the vehicle-control group and dose-response relationships. In the inter-laboratory validation study presented in Section 5, the observed results within laboratory (androgen-dependent organ weights, histology, and all but the thyroid hormones) for linuron (Table 11) corresponded to the expected results 46% (5/13), 46% (5/13), and 69% (9/13) of the time in the RTI, WIL, and Charles River laboratories, respectively (Table 11). Although the results were relatively weak within and across laboratories, they were considered supportive of the androgenic-like effects of linuron (Section 5.3.4). For phenobarbital, the observed results within and among laboratories (thyroid organ weight, histopathology, and thyroid hormone concentrations) corresponded to the expected results 80% (4/5), 100%, (5/5), and 100% (5/5) of the time in the RTI, WIL, and Charles River laboratories, respectively (Table 14). The results, both within and across laboratories, were more strongly supportive of the effects of phenobarbital on the thyroid system (Section 5.4.4).

6.1.3 A formal detailed protocol must be provided and must be available in the public domain. It should be sufficiently detailed to enable the user to adhere to it and should include data analysis and decision criteria.

An abbreviated version of the standardized protocol used in the inter-laboratory validation study is presented in Section 4 of this report for the purpose of emphasizing key aspects that were optimized and controlled to enhance and compare the protocol and results within and among laboratories. A detailed version of the protocol is presented in the Appendices of the individual inter-laboratory reports, included in the review package. Within each report, see Appendix IV for RTI, Appendix J for WIL and Appendix 3 for Charles River.

A final standardized protocol is presented in Appendix C of the ISR, which includes approaches to data analysis and interpretation.

6.1.4 Within-test, intra-laboratory, and inter-laboratory variability and how these parameters vary with time should have been evaluated.

The within- or intra-laboratory statistical analyses along with descriptive statistics including CVs describing the results using linuron and phenobarbital within the RTI, WIL, and Charles River laboratories are presented in detail in the individual laboratory reports which are included in the peer review package and summarized in Appendix B of the ISR. In addition, intra-laboratory results are summarized in Table 9 for the control group, Tables 10 and 11 for the linuron-treated group, and Tables 13 and 14 for the phenobarbital-treated group for purposes of making qualitative within- and among-laboratory comparisons for the multiple endpoints as discussed throughout Section 5.

For statistical comparisons across laboratories, the inter-laboratory analyses along with descriptive statistics including CVs describing the results using linuron and phenobarbital are presented in detail in Appendix B of the ISR and discussed throughout Section 5.

Historical results from the control and linuron- and phenobarbital-treated groups are also summarized from previous peer-reviewed publications in the respective tables listed above to qualitatively compare the current observed results with expected results within and among laboratories over time as discussed in Section 5.

6.1.5 The test method's performance must have been demonstrated using a series of reference chemicals preferably coded to exclude bias.

A total of 28 known positive chemicals with various MOAs and strengths have been run in the intact adult male assay during prevalidation (Table 4). Although most chemicals were run once in a single laboratory, some were run multiple times in the same or different laboratories involving four industrial laboratories and two CRO laboratories as discussed in Section 3. From these studies, several detailed examples using positive test chemicals with estrogenic, androgenic, anti-androgenic, and thyroidogenic activity are presented to illustrate the interrelationships between apical, histomorphological and hormonal endpoints. In addition, a known negative chemical was run in the bioassay that had a significant effect on the liver but not directly on any targeted endocrine organ or hormone.

Except for the prevalidation study with linuron and methoxychlor supported by the EPA (EPA, 2005), most test chemicals during prevalidation (Table 4) were run by industry and were not necessarily coded during development of the standardized protocol. However, after protocol standardization, the test chemicals linuron and phenobarbital were coded during the inter-laboratory validation process (summarized herein) to minimize bias as detailed in the laboratory reports of each of the three CRO laboratories and noted in Table 7 of this report.

To facilitate the pathologist's histomorphological evaluations of the testes, epididymides, and thyroid within each laboratory, the control and high-dose groups were identified, but the nature of each of the test chemicals were not revealed until after the assessments.

6.1.6 Sufficient data should be provided to permit a comparison of the performance of a proposed substitute test to that of the test it is designed to replace.

This criterion applies more to *in vitro* methods designed to replace animal tests as stated by ICCVAM (NIEHS, 1997). The 15-day intact adult male rat screening assay is considered a novel bioassay. Independent from the EDSP Tier-1 battery, the intact adult male assay does not have a predecessor or an *in vitro* successor assay at this time.

6.1.7 The limitations of the test method must be described (e.g., metabolic capability).

Technical and biological aspects pertaining to the strengths and weaknesses of the intact adult male assay are discussed in detail in Section 2.5. Technically, the assay of reproductive hormones may be one of the most challenging since these types of endpoints are not typically measured in most toxicology laboratories. Nonetheless, the assay kits are commercially available and are provided with technical support from the manufacturers to facilitate operation by individuals knowledgeable of the concept and trained in running immunoassays. Alternatively, endocrine laboratories that specialize in hormone assays are available for contracting. In addition, technical precautions have been incorporated into the protocol to minimize animal stress prior to termination and euthanasia to reduce the variation in serum hormone concentrations as discussed. In addition, animal numbers per group have been increased to an adequate but yet minimal number to compensate for the pulsatile nature and subsequent variation in hormone concentrations in a single blood sample/animal at necropsy. Despite those precautions, serum hormone concentrations serve a supporting role among the multiple apical and histological endpoints within the bioassay. Hormonal changes alone will not be sufficient to flag a chemical as a potential endocrine disruptor within the assay but may be considered among the results within the suite of *in vitro* and *in vivo* assays that comprise a Tier-1 screening battery.

Biologically, the adult male SD rat does not seem to be any more or less sensitive than the immature male SD rat according to comparable results obtained with several relatively weak positive test chemicals with different MOAs as discussed in Section 4.2. The adult rat was preferred over the immature rat for this 15-day bioassay, in part, to minimize the potential for confounding effects concerning relatively large changes in body weight due to accelerated growth from weaning to puberty and to provide a larger blood volume for analyzing the numerous serum hormones. In addition, a relatively larger data base is available for adult animals than for immature animals to provide historical reference for interpretation of results.

6.1.8 The data should be obtained in accordance with GLPs.

Except for the EPA prevalidation study (EPA, 2005), the majority of studies done during

prevalidation to standardize the intact adult male assay protocol were done in industry laboratories according to government GLP standards but were not formally audited by QA personnel. Regardless, the results of most studies were published in relevant peer-reviewed scientific journals and are cited accordingly throughout Sections 2 to 5. Subsequent to standardization of the protocol, the inter-laboratory study with linuron and phenobarbital was conducted in three different CRO laboratories according to government GLP and audited by QA personnel as detailed in the individual laboratory reports included in the peer review package along with any GLP deviations as noted in Section 5.

6.1.9 All data supporting the assessment of the validity of the test methods including the full data set collected during the validation studies must be publicly available and, preferably, published in an independent, peer-reviewed publication.

Most studies done during development of a standardized protocol for the intact adult male assay and those studies done using the standardized protocol are either published in relevant peer-reviewed scientific journals or are available at the EPA's EDSP web site (http://www.epa.gov/scipoly/oscpendo/) and referenced in Section 7 of this document. Prior to standardization of the protocol, the data collected by industry laboratories are primarily available as published papers in peer-reviewed scientific journals. For those studies sponsored by the EPA and conducted in CRO laboratories, especially during the inter-laboratory validation, the original data are available in the individual laboratory reports, included in the peer review package.

7.0 References

Ashby, J., Tinwell, H., Soames, A., and Foster, J. (1999). Induction of hyperplasia and increased DNA content in the uterus of immature rats exposed to coumestrol. *Environ. Health Perspect.* **107,** 819-822.

Biegel, L.B., Cook, J.C., O'Connor, J.C., Aschiero, M., Arduengo, A.J., Slone, T.W. (1995). Subchronic toxicity study in rats with 1-methyl-3-propylimidazole-2-thione (PTI): effects on the thyroid. Fundam. Appl. Toxicol. **27**(2),185-94.

Capen, C.C.(2001). Overview of structural and functional lesions in endocrine organs of animals. Toxicol. Pathol. **29**(1), 8-33.

Cook, J.C., Murray, S.M., Frame, S.R., and Hurtt, M.E. (1992). Induction of Leydig cell adenomas by ammonium perfluorooctanoate: a possible endocrine-related mechanism. Toxicol. Appl. Pharmacol. **113**, 209-217.

Cook, J.C., Mullin, L.S., Frame, S.R., and Biegel, L.B. (1993). Investigation of a mechanism for Leydig cell tumorigenesis by linuron in rats. *Toxicol. Appl. Pharmacol.* **119,** 195-204.

Danzo, B.J. (1997). Environmental xenobiotics may disrupt normal endocrine function by interfering with the binding of physiological ligands to steroid receptors and binding proteins. *Environ. Health Perspect.* **105,** 294-301.

Döhler, K.D., Wong, C.C., von zur Mühlen, A. (1979). The rat as model for the study of drug effects on thyroid function: consideration of methodological problems. Pharmacol. Ther. [B]. **5**(1-3), 305-18.

EPA (1997). Special report on environmental endocrine disruption: An effects assessment and analysis. EPA/630/R-96/012, February 1997.

EPA - EDSTAC (1998a). Endocrine Disruptor Screening and Testing Advisory Committee, Final Report, Volume I-II.
http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm

EPA - FRN (1998b). Federal Register Notice (FRN), volume 63, page 71542; Dec 28, 1998.

EPA-SAB/SAP (1999). Review of the Endocrine Disruptor Screening Program by a Joint Subcommittee of the Science Advisory Board and Scientific Advisory Panel. EPA-SAB-EC-99-013.

EPA - EDSP (2005). 15-Day Tier 1 screen of endocrine active compounds administered by gavage to adult male Sprague-Dawley CD® rats. Final report WA 2-28, pp.237.
http://www.epa.gov/scipoly/oscpendo/pubs/edmvac/wa2_28_final_report_042605.pdf

EPA (2006). U.S. Environmental Protection Agency Peer Review Handbook, 3rd Edition, EPA/100/B-06-002, Prepared by Members of the Peer Review Advisory Group, for EPA's Science Policy Council, available at http://www.epa.gov/peerreview/pdfs/Peer%20Review%20HandbookMay06.pdf.

EPA - FRN (2007a). Draft list of initial pesticide active ingredients and pesticide inerts to be considered for screening under the FFDCA. In: Federal Register Notice (FRN), Docket No. EPA-HQ-OPPT-2004-0109-0009, Volume 72, No. 116, pages 33486-33503, June 18, 2007.

EPA - EDSP (2007b) Validation of Screening and Testing Assays Proposed for the EDSP. Version 5.4. July, 2007.

Feron, V.J., De Groot, A.P., Spanjers, M.T., and Til, H.P. (1973). An evaluation of the criterion "organ weight" under conditions of growth retardation. *Fd. Cosmet. Toxicol.* **11**, 85-94.

ICCVAM (2000). Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) Authorization Act of 2000, 42 U.S.C. 285, Public Law 106-545, signed December 19, 2000.

Kang, I.H., Kim, H.S., Shin, J.H., Kim, T.S., Moon, H.J., Kim, I.Y., Choi, K.S., Kil, K.S., Park, Y.I., Dong, M.S., and Han, S.Y. (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.

Kang, H.G., Jeong, S.H., Cho, J.H., Kim, D.G., Park, J.M., and Cho, M.H. (2005). Evaluation of estrogenic and androgenic activity of butylated hydroxyanisole in immature female and castrated rats. *Toxicology* **213,** 147-156.

Kanno, J., Onyon, L., Peddada, S., Ashby, J., Jacob, E., and Owens, W. (2003a). The OECD program to validate the rat uterotrophic bioassay. Phase 2: dose-response studies. *Environ. Health Perspect.* **111,** 1530-1549.

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: coded single-dose studies. *Environ. Health Perspect.* **111,** 1550-1558.

Kelce, W.R., Stone, C.R., Laws, S.C., Gray, L.E., Kemppainen, J.A., and Wilson, E.M. (1995). Persistent DDT metabolite p,p'-DDE is a potent androgen receptor antagonist. *Nature* **375**, 581-585.

Maness, S.C., McDonnell, D.P., and Gaido, K.W. (1998). Inhibition of androgen receptor-dependent transcriptional activity by DDT isomers and methoxychlor in HepG2 human hepatoma cells. *Toxicol. Appl. Pharmacol.* **151,** 135-142.

Marty, M.S., Crissman, J.W., and Carney, E.W. (1999). Evaluation of the EDSTAC female pubertal assay in CD rats using 17β-estradiol, steroid biosynthesis inhibitors, and a thyroid inhibitor. *Toxicol. Sci.* **52,** 269-277.

Marty, M.S., Crissman, J.W., and Carney, E.W. (2001). Evaluation of the male pubertal onset assay to detect testosterone and steroid biosynthesis inhibitors in CD rats. *Toxicol. Sci.* **60,** 285-295.

Marty, M.S., Stebbins, K.E., and Carney, E.W. (2002). Evaluation of the 15-day intact male assay: capability to detect the anti-androgen flutamide. *The Toxicologist* **66** (1S), 275-276.

McClain, R. M., Levin, A. A., Posch, R., and Downing, J. C. (1989). Effect of phenobarbital on the metabolism and excretion of thyroxine in rats. *Toxicol. Appl. Pharmacol.* **99**, 216-228.

Mellert. (2003). P-Nonylphenol study with 15-day intact adult male rat assay by BASF. Abstract at: http://www.epa.gov/scipoly/oscpendo/pubs/edmvac/other_lab_results.pdf

Milburn. (2004). Genistein study with 15-day intact adult male rat assay by Syngenta. Abstract at http://www.epa.gov/scipoly/oscpendo/pubs/edmvac/other_lab_results.pdf

NIEHS (1997). National Institute of Environmental Health Sciences. "Validation and Regulatory Acceptance of Toxicological Test Methods, A Report of the ad hoc Interagency

O'Connor, J.C., Cook, J.C., Slone, T.W., Makovec, G.T., Frame, S.R., and Davis, L.G. (1998a). An on-going validation of a Tier I screening battery for detecting endocrine-active compounds (EACs). *Toxicol. Sci.* **46,** 45-60.

O'Connor, J.C., Frame, S.R., Biegel, L.B., Cook, J.C., and Davis, L.G. (1998b). Sensitivity of a Tier I screening battery compared to an *in utero* exposure for detecting the estrogen receptor agonist 17β-estradiol. *Toxicol. Sci.* **44,** 169-184.

O'Connor, J.C., Frame, S.R., Davis, L.G., and Cook, J.C. (1999a). Detection of the environmental anti-androgen p,p-DDE in CD and Long-Evans rats using a tier I screening battery and a Hershberger assay. *Toxicol. Sci.* **51,** 44-53.

O'Connor, J.C., Frame, S.R., Davis, L.G. and Cook, J.C. (1999b). Detection of thyroid toxicants in a Tier I screening battery and alterations in thyroid endpoints over 28 days of exposure. *Toxicol. Sci.* **51,** 54-70.

O'Connor, J.C., Davis, L.G., Frame, S.R., and Cook, J.C. (2000b). Detection of dopaminergic modulators in a Tier I screening battery for detecting endocrine-active compounds (EACs). *Reprod. Toxicol.* **14,** 193-205.

O'Connor, J.C., Frame, S.R., Smith, C. and Ladics, G.S. (2000c). Comparison of the oral (gavage) and intraperitoneal routes of administration for identifying endocrine-active compounds

(EAC) using an *in vivo* male battery, and evaluation of immune system endpoints. *Toxicol. Sci. Supplement – The Toxicologist* **54**, 261.

O'Connor, J.C., Frame, S.R., and Ladics, G.S. (2002a). Evaluation of a 15-day screening assay using intact male rats for identifying anti-androgens. *Toxicol. Sci.* **69,** 92-108.

O'Connor, J.C., Frame, S.R., and Ladics, G.S. (2002b). 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, J.C., Cook, J.C., Marty, M.S., Davis, L.G., Kaplan, A.M., and Carney, E.W. (2002c). Evaluation of Tier I screening approaches for detecting endocrine-active compounds (EACs). *Crit. Rev. Toxicol.* **32,** 521-549.

O'Connor, J.C., Davis, L.G., Frame, S.R., and Cook, J.C. (2002d). Evaluation of a Tier I screening battery for detecting endocrine-active compounds (EACs) using the positive controls testosterone, coumestrol, progesterone, and RU486. *Toxicol. Sci.* **54,** 338-354.

O'Connor, J.C. (2005). The 15-day intact adult male as an alternative Tier 1 screening assay for detecting endocrine active compounds. Slide presentation at: http://www.epa.gov/scipoly/oscpendo/presentations/intact_male_presentation_080205.ppt

O'Connor, J.C., Marty, M.S., Becker, R.A., and Kaplan, A.M.. (2007). Results of the negative control chemical allyl alcohol in the 15-day intact adult male rat screening assay for endocrine activity. [In preparation].

OECD (1996). Organisation for Economic Cooperation and Development. Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. August , 1996.

Physician's Desk Reference (2004). Medical Economics (ed), 58th Edition. Thomson Publishing; Hampshire.

Schmidt, W.N., and Katzenellenbogen, B.S. (1979). Androgen-uterine interactions: an assessment of androgen interaction with the testosterone- and estrogen-receptor systems and stimulation of uterine growth and progesterone-receptor synthesis. *Mol. Cell. Endocrinol.* **15,** 91-108.

Stump, DG. (2002). Methyltestosterone study with 15-day intact adult male rat assay by WIL laboratory. Abstract at: http://www.epa.gov/scipoly/oscpendo/pubs/edmvac/other_lab_results.pdf

Tinwell, H., Soames, A.R., Foster, J.R., and Ashby, J. (2000). Estradiol-type activity of coumestrol in mature and immature ovariectomized rat uterotrophic assays. *Environ. Health Perspect.* **108,** 631-634.

Yamasaki, K., Takeyoshi, M., Sawaki, M., Imatanaka, N., Shinoda, K., and Takatsuki, M. (2003a). Immature rat uterotrophic assay of 18 chemicals and Hershberger assay of 30 chemicals. *Toxicology* **183,** 93-115.

Yamasaki, K., Sawaki, M., Ohta, R., Okuda, H., Katayama, S., Yamada, T., Ohta, T., Kosaka, T., and Owens, W. (2003b). OECD validation of the Hershberger assay in Japan: phase 2 dose response of methyltestosterone, vinclozolin, and p,p'-DDE. *Environ. Health Perspect.* **111,** 1912-1919.

You, L., Casanova, M., Archibeque-Engle, S., Sar, M., Fan, L.-Q., and Heck, H.d'A. (1998). Impaired male sexual development in perinatal Sprague-Dawley and Long-Evans hooded rats exposed *in utero* and lactationally to *p,p'*-DDE. *Toxicol. Sci.* **45,** 162-173.

This page left blank intentionally.

APPENDIX A

Inter-Laboratory Validation of the 15-Day Intact Male Rat Assay: Hormonal Assay Quality Control Standards Data

October 3, 2006

This page left blank intentionally.

FINAL REPORT

REVISED STATISTICAL ANALYSIS

on

INTERLABORATORY VALIDATION OF THE 15-DAY INTACT MALE RAT ASSAY

HORMONAL ASSAY QUALITY CONTROL STANDARDS DATA

EPA CONTRACT NUMBER EP-W-06-032

WORK ASSIGNMENT 1-2
Technical Directive No. 1

October 3, 2006

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C.

ENDOCRINE DISRUPTOR SCREENING PROGRAM
TECHNICAL SUPPORT SERVICES

INTER-LABORATORY VALIDATION OF THE 15-DAY ADULT INTACT MALE RAT ASSAY: REVISED STATISTICAL ANALYSIS OF HORMONAL ASSAY QUALITY CONTROL STANDARDS DATA

EPA CONTRACT NUMBER 68-W-06-032
WORK ASSIGNMENT 1-2
Technical Directive No. 1

Zhenxu J. Ma, Author — Date: 7/10/2007

Paul I. Feder, Reviewer — Date: July 10, 2007

Table of Contents

Page

List of Tables

This page left blank intentionally.

Interlaboratory Validation of the 15-Day Intact Male Rat Assay
Hormonal Assay Quality Control Standards Data

Introduction and Background

Three laboratories − Charles River (Argus), Research Triangle Institute, and WIL − conducted the 15-day adult intact male rat assay according to the test method provided by the EPA. The test method specified 34 endpoints, divided among three categories:

- Growth - body weights and food consumption (7 endpoints)
- Organ weights (9 organs – 18 endpoints)
- Hormonal analysis (9 hormones)
  - Testosterone
  - LH
  - TSH
  - T4
  - T3
  - FSH
  - Estradiol
  - Prolactin
  - DHT

For each of the hormonal assays quality control (QC) standards were run along with the test samples. The data consist of quality control (QC) standards results from this 15-day intact adult male assay for the three test laboratories and multiple hormone assays. This report presents a summarization of the quality control standards results and for the RTI standards, comparison of the results across assays (i.e. across test days).

Test samples were collected from the three originating laboratories. The previous data analysis report (June 30, 2006) compared the QC standards results among the three originating laboratories. However, the Charles River samples were analyzed by RTI, in fact by the same chemist who analyzed the RTI samples. Thus for purposes of calculating variability and CVs across hormone assay QC standards in the revised analysis, the RTI standards data and the Charles River standards data were combined (and are referred to as RTI standards).

After the previous (June 30, 2006) QC standards statistical analysis had been submitted to EPA it was determined that the standards had been prepared differently at RTI and at WIL Laboratory. WIL Laboratory reported the results of the Kit QC assay standards provided by the manufacturers. For those assay kits that did not come with QC standards (e.g. the protein assays), WIL diluted the standard curve samples to make their own kit based QC standards. RTI's samples that were designated "Kit QC" standards were actually the results of validation samples, in which available lots of rat serum were spiked with known concentrations of hormones. This difference in procedures between

the test laboratories led to differences in the standard test concentrations, and therefore to differences in the recovered concentrations and to large among laboratory CVs.

Therefore in the present reanalysis the QC standards within assay and among assay variation was determined separately for the RTI standards and for the WIL Laboratory standards.

Data

For each hormonal assay, several types and groupings of quality control standards were run. As previously discussed, there were differences in the manner in which the QC standards were prepared at RTI and at WIL Laboratory.

The RTI samples are designated as "Rat Serum Calibration QC standards" or as "Zero Calibration QC standards." The "Rat Serum Calibration QC standards" were the results of validation samples, in which available lots of rat serum were spiked with known concentration of hormone. The "Zero Calibration QC standards" were the results of validation samples, in which media containing no hormone were spiked with known concentration of hormone. The RTI data include multiple true replicates, based on assays carried out on different assay dates (Table 2).

The WIL Laboratory samples were designated as "Kit QC standards" or "Non-Kit QC standards." When preparing the kit QC standards, WIL followed the manufacturers' instructions as closely as it could and it did not make any purposeful modifications. In those cases in which the assay kits did not come with QC standards (e.g. proteins) WIL made "Kit QC standards" by diluting standard curve samples. For the large majority of cases, WIL ran just one replicate of the QC standards, with two measurement duplicates within the replicate. On occasion, when WIL had to repeat the standard curve, it prepared a second true replicate set of QC standards.

Data Issues

The three test laboratories are referred to as RTI/RTI, RTI/Charles River, and WIL/WIL (analysis laboratory/originating laboratory) or simply as RTI (for the RTI and Charles River combined) and WIL).

Several unanticipated data issues arose, were discussed with EPA, and were treated in the manner specified by EPA. These issues and their treatment are discussed below.

1. For RTI/RTI – PRL, a handwritten notation on the data sheet indicated that the data collected on 11/10/2005 should not be used. These data, two samples per standard type, are included in Table 2, but were excluded from the data summaries.

2. For RTI/RTI and RTI/Charles Liver – DHT, four values for the "Serum Calibration QC standards-high" standards were recorded as "AAR" (above assay range). These values were excluded from the data summaries.
3. For RTI/RTI – T_4, there were two "Serum Calibration QC standards-mid" groups, mid 1 and mid 2, at concentrations 5µg/dl and 8µg/dL respectively. Both groups are included in Tables 1 and 2.
4. For WIL/WIL – FSH, a value in the "Non-kit QC low" group was reported as "off curve". This value was not included in the data summaries.
5. For testosterone, RTI reported concentrations in units "ng/mL". WIL reported concentrations in units "ng/dL". The WIL values were divided by 100 to transform them to the same units as those from RTI.
6. QC samples that were reported in the data as "NA" are shown as missing values (".") in the tables.

Data Summaries – Methods

Because of the differences in the manner in which the QC standards were prepared at RTI and at WIL Laboratory, separate comparisons were carried out for RTI results and for WIL Laboratory results.

RTI Results

For each hormone assay, QC test type, and test concentration, the combined sample results were categorized by test date regardless of originating laboratory (RTI or Charles River). A fixed effects one-way analysis of variance was carried out with test date as the group variable. The pooled mean (the unweighted average over the test date means) and the pooled within-day (within-assay) standard deviation was determined based on the analysis of variance and used to calculate the within-day (within-assay) CV for each hormone QC test type. The within assay CV was determined as the ratio of the pooled within-day (within-assay) standard deviation to the pooled mean. The among-day (among-assay) CV was determined as the ratio of the standard deviation among the test date means to the pooled mean.

WIL Laboratory Results

For the large majority of cases, WIL ran just one replicate of the QC standards with two measurement duplicates within the replicate. Just the within day CV was reported, based on the variation between measurement duplicates within each replicate. Note that these CVs, based on duplicates, do not reflect the total variation in the assay. They reflect primarily the measurement error component of variation. Thus they underestimate the total within assay variability.

WIL (telephone communication) indicated that its assay analysis instrument reported values of the counts (cpm) for each duplicate, values of percent bound for each duplicate, and then averaged the cpm or percent bound values and converted the percent bound average to a concentration (e.g. in units ng/mL) based on the standard curve. WIL

pointed out that there is no single conversion factor between cpm and ng/mL or between percent bound and ng/mL. This is because the standard curve does not in general go through (0, 0). Rather it is usually of the form percent bound = $b_0 + b_1$*(concentration). Thus the ratio of concentration to percent bound varies with the true concentration value with which one enters the standard curve. The assay analysis instrument output includes the duplicate values of cpm, the duplicate values of percent bound, the nominal test concentration, the mean calculated concentration, and the CV. WIL believes that the CV value reported by the assay instrument is either that between the duplicate cpm values or between the duplicate percent bound values.

The CVs between measurement duplicates, as determined directly by the assay instrument and included in WIL's report are included in the present report. In most cases this requires no calculation beyond the value reported by WIL. In the relatively small number of cases when WIL ran two true replicates, (i.e. when the standard curve determination needed to be repeated to fall in the correct concentration range) the average within-assay mean was determined as the simple average of the two reported individual QC standard values among duplicates within replicates. The reported CV is that which was originally reported by WIL.

Data Summaries – Results

Table 1 displays the descriptive statistics (mean standard deviation, sample size, minimum and maximum) of the RTI QC standards data, as determined on each test day (i.e., for each assay) for each QC standard, as well as the unweighted means of the test date means. This table also includes the pooled-within assay standard deviation, as determined by the one-way ANOVA, and the estimates of the within-assay coefficient of variation and the among-assay coefficient of variation. The reported individual QC standards values for RTI/RTI and RTI/Charles River are displayed in Table 2.

The individual WIL Laboratory data values and the within assay coefficients of variation based on variation between measurement duplicates are displayed in Table 3. As discussed previously, the within assay CVs reflect only the measurement component of variation and so underestimate the total within assay variation.

The ranges of CVs reported by RTI and by WIL for each hormone assay are summarized below. Note that these are rounded values, intended to provide a qualitative comparison of relative size. Tables 1 and 3 display the CV values in greater detail.

Hormone Assay	RTI		WIL
	Within Assay CV	Among Assay CV	Within Assay CV
DHT	5 - 30	6 - 35	1 - 6
Estradiol	9 - 19	1 - 17	1 - 27
FSH	3 - 10	0.3 - 14	0.1 - 3
LH	5 - 10	1 - 12	1.5 - 3
PRL	3 - 16	5 - 51	0.6 - 4
T_3	4 - 12	0.04 - 13	0.1 - 2
T_4	1 - 8	1 - 6	0.6 – 4.5
TSH	3 - 7	0.4 - 4	0.6 - 7
Testosterone	4 - 11	4 - 12	1 - 10

The summary table above indicates the following characteristics of the CVs.

For RTI:

- For each CV type the CVs vary considerably across the QC standards for each assay. The variation appears to be random.
- For each assay the range of the among-assay CVs is about the same as the within-assay CVs.
- PRL, DHT, and Estradiol have the largest CVs.

For WIL:

- The WIL within assay CVs are less than or equal to the RTI within assay CVs.
  The WIL within assay CVs are less than the RTI within assay CVs for six of the nine hormone assays (DHT, FSH, LH, PRL, T_3, T_4) and about the same for the other three (Estradiol, TSH, Testosterone).

It is to be expected that the WIL CVs would be less than the RTI CVs since the WIL CVs are based on variation between measurement duplicates whereas the RTI CVs are based on variation among true replicates.

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
DHT	**Rat Serum Calibration-Low**	01DEC05	4	302.03	101.67	184.620	406.280	225.09	67.60	30.03	34.86
		04DEC05	4	228.07	17.69	208.340	250.390				
		08DEC05	2	145.18	5.74	141.120	149.240				
DHT	**Rat Serum Calibration-Mid**	01DEC05	4	350.75	25.49	322.370	382.310	322.83	35.19	10.90	18.20
		04DEC05	4	362.44	45.53	317.020	422.640				
		08DEC05	2	255.31	22.34	239.520	271.110				
DHT	**Rat Serum Calibration-High**	01DEC05	1	748.40	.	748.400	748.400	681.50	38.31	5.62	19.86
		04DEC05	3	770.40	22.04	750.980	794.350				
		08DEC05	2	525.71	58.58	484.290	567.130				
DHT	**Zero Calibration-Low**	01DEC05	4	70.93	6.76	64.240	80.340	67.90	8.45	12.44	5.85
		04DEC05	4	69.37	10.88	56.680	81.920				
		08DEC05	2	63.40	2.69	61.500	65.310				
DHT	**Zero Calibration-High**	01DEC05	4	336.83	27.83	316.920	376.700	330.95	34.38	10.39	7.30
		04DEC05	4	351.64	44.51	297.110	401.020				
		08DEC05	2	304.38	2.26	302.780	305.980				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
Estradiol	**Rat Serum Calibration-Low**	03NOV05	4	9.72	2.00	7.880	11.650	10.65	1.60	15.05	12.32
		06DEC05	4	11.58	1.06	10.390	12.720				
Estradiol	**Rat Serum Calibration-Mid**	03NOV05	4	16.48	2.14	14.440	18.720	18.23	1.63	8.94	13.57
		06DEC05	4	19.97	0.85	19.020	20.720				
Estradiol	**Rat Serum Calibration-High**	03NOV05	4	31.55	4.74	26.880	35.990	35.97	3.54	9.83	17.40
		06DEC05	4	40.40	1.59	39.300	42.680				
Estradiol	**Zero Calibration-Low**	03NOV05	4	5.39	1.38	4.100	6.650	5.22	1.01	19.38	4.88
		06DEC05	4	5.04	0.38	4.520	5.320				
Estradiol	**Zero Calibration-High**	03NOV05	4	32.88	6.55	27.060	39.370	33.18	4.93	14.85	1.26
		06DEC05	4	33.47	2.36	30.990	35.860				
FSH	**Rat Serum Calibration-Low**	13OCT05	4	5.72	0.53	5.140	6.350	6.36	0.62	9.83	14.04
		02DEC05	4	6.99	0.70	6.170	7.850				
FSH	**Rat Serum Calibration-Mid**	13OCT05	4	11.72	0.38	11.150	11.980	11.58	0.43	3.69	1.69
		02DEC05	4	11.44	0.47	10.780	11.790				
FSH	**Rat Serum Calibration-High**	13OCT05	4	27.64	1.61	26.220	29.880	27.71	1.25	4.51	0.34
		02DEC05	4	27.77	0.73	26.780	28.520				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
FSH	**Zero Calibration-Low**	13OCT05	3	14.78	0.22	14.530	14.930	14.41	0.45	3.15	3.56
		02DEC05	4	14.05	0.56	13.430	14.770				
FSH	**Zero Calibration-High**	13OCT05	4	51.27	5.07	45.760	57.950	50.78	4.02	7.91	1.38
		02DEC05	4	50.28	2.56	46.740	52.500				
LH	**Rat Serum Calibration-Low**	01NOV05	4	1.06	0.12	0.970	1.230	1.16	0.11	9.78	11.61
		29NOV05	4	1.25	0.11	1.160	1.410				
LH	**Rat Serum Calibration-Mid**	01NOV05	4	3.81	0.29	3.430	4.040	3.63	0.24	6.72	6.96
		29NOV05	4	3.45	0.19	3.210	3.670				
LH	**Rat Serum Calibration-High**	01NOV05	4	12.59	0.63	11.890	13.180	12.66	0.60	4.76	0.77
		29NOV05	4	12.73	0.57	12.250	13.440				
LH	**Zero Calibration-Low**	01NOV05	4	5.28	0.36	4.800	5.560	5.08	0.29	5.73	5.43
		29NOV05	4	4.88	0.20	4.700	5.140				
LH	**Zero Calibration-High**	01NOV05	4	19.89	1.30	18.990	21.820	19.59	1.08	5.49	2.15
		29NOV05	4	19.30	0.79	18.560	20.260				
PRL	**Rat Serum Calibration-Low**	13OCT05	4	40.11	2.27	38.210	43.200	29.19	1.61	5.51	51.49
		08NOV05	2	44.15	1.34	43.200	45.090				
		10NOV05	0	.	.	.	.				
		07DEC05	4	15.86	0.54	15.070	16.210				
		09DEC05	2	16.65	1.60	15.520	17.780				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
PRL	**Rat Serum Calibration-Mid**	13OCT05	4	44.35	1.62	43.040	46.700	33.04	1.26	3.83	45.56
		08NOV05	2	47.70	0.65	47.240	48.160				
		10NOV05	0	.	.	.	.				
		07DEC05	4	19.99	0.97	19.090	21.180				
		09DEC05	2	20.13	1.30	19.210	21.050				
PRL	**Rat Serum Calibration-High**	13OCT05	4	57.08	1.31	55.630	58.680	43.44	1.34	3.09	36.23
		08NOV05	2	57.01	1.48	55.960	58.060				
		10NOV05	0	.	.	.	.				
		07DEC05	4	31.10	1.53	29.750	32.980				
		09DEC05	2	28.59	0.13	28.490	28.680				
PRL	**Zero Calibration-Low**	13OCT05	4	5.50	0.78	4.630	6.450	5.23	0.66	12.62	4.87
		08NOV05	2	4.89	0.06	4.850	4.930				
		10NOV05	0	.	.	.	.				
		07DEC05	4	5.24	0.69	4.530	6.110				
		09DEC05	2	5.29	0.48	4.950	5.630				
PRL	**Zero Calibration-High**	13OCT05	4	21.99	5.15	16.960	26.750	20.33	3.35	16.48	8.17
		08NOV05	2	18.02	1.21	17.170	18.880				
		10NOV05	0	.	.	.	.				
		07DEC05	4	20.72	1.57	18.850	22.210				
		09DEC05	2	20.57	1.17	19.750	21.400				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
T3	**Rat Serum Calibration-Low**	08NOV05	4	33.33	5.37	26.660	39.590	36.75	4.30	11.70	13.17
		28NOV05	4	40.17	2.86	37.080	43.740				
T3	**Rat Serum Calibration-Mid**	08NOV05	4	85.88	9.92	71.770	93.660	90.23	7.73	8.56	6.81
		28NOV05	4	94.57	4.58	87.780	97.580				
T3	**Rat Serum Calibration-High**	08NOV05	4	379.46	21.10	351.290	399.900	379.35	15.44	4.07	0.04
		28NOV05	4	379.23	5.63	371.710	385.370				
T3	**Zero Calibration-Low**	08NOV05	4	44.14	1.82	41.750	45.610	46.90	1.90	4.05	8.33
		28NOV05	4	49.67	1.97	47.420	51.650				
T3	**Zero Calibration-High**	08NOV05	4	325.91	18.08	307.880	342.750	311.62	16.53	5.30	6.49
		28NOV05	4	297.32	14.82	283.410	311.520				
T4	**Rat Serum Calibration-Low**	28OCT05	2	1.83	0.18	1.700	1.960	2.68	0.33	12.36	23.95
		02NOV05	2	2.56	0.40	2.280	2.850				
		18NOV05	4	3.02	0.28	2.690	3.310				
		22NOV05	2	3.30	0.48	2.960	3.640				
T4	**Rat Serum Calibration-Mid 1**	28OCT05	4	7.34	0.63	6.430	7.780	7.77	0.76	9.76	5.56
		02NOV05	2	7.45	0.75	6.920	7.980				
		18NOV05	4	8.18	0.82	7.350	9.000				
		22NOV05	2	8.10	0.92	7.450	8.750				
T4	**Rat Serum Calibration-Mid 2**	28OCT05	4	10.01	1.00	8.800	10.980	10.18	1.15	11.26	2.31
		02NOV05	2	10.34	1.51	9.280	11.410				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
T4	**Rat Serum Calibration-High**	28OCT05	4	13.75	1.00	12.880	14.700	14.41	1.20	8.34	4.30
		02NOV05	2	14.63	2.35	12.970	16.290				
		18NOV05	4	15.17	0.79	14.440	15.860				
		22NOV05	2	14.10	1.07	13.340	14.860				
T4	**Zero Calibration-Low**	28OCT05	4	2.87	0.21	2.650	3.150	2.92	0.23	7.80	4.62
		02NOV05	2	2.75	0.11	2.680	2.830				
		18NOV05	4	3.06	0.27	2.810	3.410				
		22NOV05	2	3.00	0.24	2.830	3.170				
T4	**Zero Calibration-High**	28OCT05	4	16.25	1.14	15.190	17.370	16.12	0.94	5.82	1.13
		02NOV05	2	16.13	1.32	15.200	17.060				
		18NOV05	4	16.24	0.55	15.780	17.040				
		22NOV05	2	15.86	0.72	15.350	16.370				
TSH	**Rat Serum Calibration-Low**	20OCT05	4	8.66	0.68	8.190	9.650	8.64	0.56	6.51	0.35
		05DEC05	4	8.62	0.41	8.100	8.960				
TSH	**Rat Serum Calibration-Mid**	20OCT05	4	12.59	0.92	11.240	13.300	13.00	0.80	6.17	4.42
		05DEC05	4	13.40	0.66	12.460	13.980				
TSH	**Rat Serum Calibration-High**	20OCT05	4	30.31	0.66	29.350	30.780	29.57	0.95	3.22	3.53
		05DEC05	4	28.83	1.17	27.730	30.380				
TSH	**Zero Calibration-Low**	20OCT05	4	6.02	0.11	5.910	6.160	5.97	0.18	2.94	1.07
		05DEC05	4	5.93	0.22	5.680	6.210				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
TSH	**Zero Calibration-High**	20OCT05	4	21.95	0.85	21.070	22.840	21.34	1.42	6.67	4.04
		05DEC05	4	20.73	1.82	19.680	23.460				
Testosterone	**Rat Serum Calibration-Low**	20OCT05	4	0.35	0.02	0.320	0.370	0.43	0.02	5.28	16.95
		24OCT05	2	0.39	0.01	0.390	0.400				
		17NOV05	4	0.48	0.02	0.460	0.500				
		22NOV05	2	0.50	0.04	0.480	0.530				
Testosterone	**Rat Serum Calibration-Mid**	20OCT05	4	3.89	0.38	3.600	4.430	3.95	0.33	8.39	3.67
		24OCT05	2	3.78	0.35	3.530	4.030				
		17NOV05	4	4.12	0.31	3.820	4.450				
		22NOV05	2	3.99	0.13	3.890	4.080				
Testosterone	**Rat Serum Calibration-High**	20OCT05	4	7.79	0.17	7.570	7.970	7.87	0.37	4.68	4.53
		24OCT05	2	8.37	0.44	8.060	8.680				
		17NOV05	4	7.79	0.39	7.230	8.100				
		22NOV05	2	7.52	0.60	7.100	7.950				
Testosterone	**Zero Calibration-Low**	20OCT05	4	0.44	0.04	0.400	0.490	0.53	0.06	11.28	11.78
		24OCT05	2	0.52	0.06	0.480	0.570				
		17NOV05	4	0.55	0.07	0.470	0.630				
		22NOV05	2	0.59	0.07	0.540	0.640				

Table 1. Quality Control Sample Summary Results and Within-Assay and Among-Assay Coefficients of Variation for the Hormonal Assays for RTI Data. By Hormone Type, QC Standard Type.

Hormone Parameter	QC Standard Type	Assay Date	Sample Size	Mean	Std. Deviation	Minimum	Maximum	Among Assay Mean[1]	Pooled Within Assay Std. Deviation[2]	Pooled Within Assay Coefficient of Variation (%)[3]	Among Assay Coefficient of Variation (%)[4]
Testosterone	**Zero Calibration-High**	20OCT05	4	3.85	0.22	3.630	4.040	4.14	0.18	4.37	5.17
		24OCT05	2	4.29	0.17	4.170	4.410				
		17NOV05	4	4.10	0.17	3.900	4.270				
		22NOV05	2	4.32	0.04	4.290	4.340				

1. The among assay mean is the unweighted average of the assay date means for each QC standard type.
2. The pooled within assay standard deviation is the square root of the pooled within assay variance as determined by the one-way ANOVA.
3. The pooled within-assay CV is the pooled within assay standard deviation divided by the among assay mean.
4. The among-assay CV is the CV among the assay date means.

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
DHT	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	pg/mL	04DEC05	250.39
						221.86
						231.69
						208.34
DHT	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	pg/mL	08DEC05	149.24
						141.12
DHT	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	pg/mL	01DEC05	364.60
						184.62
						252.60
						406.28
DHT	**Rat Serum Calibration-Mid**	**RTI/Charles River**	100.00	pg/mL	04DEC05	317.02
						340.51
						422.64
						369.57
DHT	**Rat Serum Calibration-Mid**	**RTI/Charles River**	100.00	pg/mL	08DEC05	271.11
						239.52
DHT	**Rat Serum Calibration-Mid**	**RTI/RTI**	100.00	pg/mL	01DEC05	340.70
						382.31
						322.37
						357.61
DHT	**Rat Serum Calibration-High**	**RTI/Charles River**	400.00	pg/mL	04DEC05	750.98
						794.35
						765.88
						AAR
DHT	**Rat Serum Calibration-High**	**RTI/Charles River**	400.00	pg/mL	08DEC05	567.13
						484.29
DHT	**Rat Serum Calibration-High**	**RTI/RTI**	400.00	pg/mL	01DEC05	748.40
						AAR
						AAR
						AAR
DHT	**Zero Calibration-Low**	**RTI/Charles River**	64.00	pg/mL	04DEC05	65.18
						56.68
						73.71
						81.92

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
DHT	**Zero Calibration-Low**	**RTI/Charles River**	64.00	pg/mL	08DEC05	65.31
						61.50
DHT	**Zero Calibration-Low**	**RTI/RTI**	64.00	pg/mL	01DEC05	69.41
						69.72
						80.34
						64.24
DHT	**Zero Calibration-High**	**RTI/Charles River**	320.00	pg/mL	04DEC05	338.12
						370.33
						401.02
						297.11
DHT	**Zero Calibration-High**	**RTI/Charles River**	320.00	pg/mL	08DEC05	302.78
						305.98
DHT	**Zero Calibration-High**	**RTI/RTI**	320.00	pg/mL	01DEC05	318.56
						376.70
						335.13
						316.92
Estradiol	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	pg/mL	06DEC05	12.18
						11.01
						10.39
						12.72
Estradiol	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	pg/mL	03NOV05	11.24
						11.65
						8.11
						7.88
Estradiol	**Rat Serum Calibration-Mid**	**RTI/Charles River**	7.50	pg/mL	06DEC05	19.02
						20.67
						19.49
						20.72
Estradiol	**Rat Serum Calibration-Mid**	**RTI/RTI**	7.50	pg/mL	03NOV05	17.88
						14.87
						18.72
						14.44

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
Estradiol	**Rat Serum Calibration-High**	**RTI/Charles River**	25.00	pg/mL	06DEC05	42.68
						39.30
						40.30
						39.32
Estradiol	**Rat Serum Calibration-High**	**RTI/RTI**	25.00	pg/mL	03NOV05	35.99
						28.07
						35.26
						26.88
Estradiol	**Zero Calibration-Low**	**RTI/Charles River**	5.00	pg/mL	06DEC05	4.98
						5.32
						4.52
						5.32
Estradiol	**Zero Calibration-Low**	**RTI/RTI**	5.00	pg/mL	03NOV05	6.52
						6.65
						4.31
						4.10
Estradiol	**Zero Calibration-High**	**RTI/Charles River**	33.33	pg/mL	06DEC05	30.99
						35.09
						35.86
						31.96
Estradiol	**Zero Calibration-High**	**RTI/RTI**	33.33	pg/mL	03NOV05	27.42
						39.37
						37.68
						27.06
FSH	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ng/mL	02DEC05	7.85
						6.77
						6.17
						7.16
FSH	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/mL	13OCT05	5.14
						6.35
						5.46
						5.95
FSH	**Rat Serum Calibration-Mid**	**RTI/Charles River**	6.25	ng/mL	02DEC05	11.46
						10.78
						11.79
						11.75
FSH	**Rat Serum Calibration-Mid**	**RTI/RTI**	6.25	ng/mL	13OCT05	11.86

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						11.90
						11.98
						11.15
FSH	**Rat Serum Calibration-High**	**RTI/Charles River**	25.00	ng/mL	02DEC05	27.76
						26.78
						28.52
						28.03
FSH	**Rat Serum Calibration-High**	**RTI/RTI**	25.00	ng/mL	13OCT05	26.22
						29.88
						27.67
						26.78
FSH	**Zero Calibration-Low**	**RTI/Charles River**	7.50	ng/mL	02DEC05	14.12
						14.77
						13.43
						13.88
FSH	**Zero Calibration-Low**	**RTI/RTI**	7.50	ng/mL	13OCT05	14.87
						14.93
						.
						14.53
FSH	**Zero Calibration-High**	**RTI/Charles River**	30.00	ng/mL	02DEC05	46.74
						51.77
						50.12
						52.50
FSH	**Zero Calibration-High**	**RTI/RTI**	30.00	ng/mL	13OCT05	49.86
						51.51
						45.76
						57.95
LH	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ng/mL	29NOV05	1.41
						1.16
						1.20
						1.24
LH	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/mL	01NOV05	1.05
						1.23
						0.97
						1.00
LH	**Rat Serum Calibration-Mid**	**RTI/Charles River**	3.13	ng/mL	29NOV05	3.46
						3.21

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						3.67
						3.47
LH	**Rat Serum Calibration-Mid**	**RTI/RTI**	3.13	ng/mL	01NOV05	4.03
						3.74
						4.04
						3.43
LH	**Rat Serum Calibration-High**	**RTI/Charles River**	12.50	ng/mL	29NOV05	12.95
						13.44
						12.28
						12.25
LH	**Rat Serum Calibration-High**	**RTI/RTI**	12.50	ng/mL	01NOV05	13.18
						11.89
						13.07
						12.23
LH	**Zero Calibration-Low**	**RTI/Charles River**	3.75	ng/mL	29NOV05	5.14
						4.76
						4.70
						4.94
LH	**Zero Calibration-Low**	**RTI/RTI**	3.75	ng/mL	01NOV05	5.55
						5.56
						4.80
						5.19
LH	**Zero Calibration-High**	**RTI/Charles River**	15.00	ng/mL	29NOV05	18.56
						18.76
						20.26
						19.60
LH	**Zero Calibration-High**	**RTI/RTI**	15.00	ng/mL	01NOV05	21.82
						19.24
						19.51
						18.99
PRL	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ng/mL	07DEC05	15.98
						16.21
						16.19
						15.07
PRL	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ng/mL	09DEC05	17.78
						15.52
PRL	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/mL	13OCT05	43.20

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						38.62
						40.40
						38.21
PRL	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/mL	08NOV05	45.09
						43.20
PRL	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/mL	10NOV05	23.53[2]
						27.20[2]
PRL	**Rat Serum Calibration-Mid**	**RTI/Charles River**	3.10	ng/mL	07DEC05	21.18
						19.09
						19.31
						20.36
PRL	**Rat Serum Calibration-Mid**	**RTI/Charles River**	3.10	ng/mL	09DEC05	21.05
						19.21
PRL	**Rat Serum Calibration-Mid**	**RTI/RTI**	3.10	ng/mL	13OCT05	43.04
						46.70
						43.64
						44.02
PRL	**Rat Serum Calibration-Mid**	**RTI/RTI**	3.10	ng/mL	08NOV05	47.24
						48.16
PRL	**Rat Serum Calibration-Mid**	**RTI/RTI**	3.10	ng/mL	10NOV05	26.67[2]
						27.30[2]
PRL	**Rat Serum Calibration-High**	**RTI/Charles River**	12.50	ng/mL	07DEC05	29.97
						29.75
						31.71
						32.98
PRL	**Rat Serum Calibration-High**	**RTI/Charles River**	12.50	ng/mL	09DEC05	28.68
						28.49
PRL	**Rat Serum Calibration-High**	**RTI/RTI**	12.50	ng/mL	13OCT05	55.63
						57.50
						56.51
						58.68
PRL	**Rat Serum Calibration-High**	**RTI/RTI**	12.50	ng/mL	08NOV05	55.96
						58.06
PRL	**Rat Serum Calibration-High**	**RTI/RTI**	12.50	ng/mL	10NOV05	34.75[2]
						34.37[2]
PRL	**Zero Calibration-Low**	**RTI/Charles River**	2.50	ng/mL	07DEC05	5.41
						4.53

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						4.89
						6.11
PRL	**Zero Calibration-Low**	**RTI/Charles River**	2.50	ng/mL	09DEC05	4.95
						5.63
PRL	**Zero Calibration-Low**	**RTI/RTI**	2.50	ng/mL	13OCT05	6.45
						4.63
						5.17
						5.77
PRL	**Zero Calibration-Low**	**RTI/RTI**	2.50	ng/mL	08NOV05	4.93
						4.85
PRL	**Zero Calibration-Low**	**RTI/RTI**	2.50	ng/mL	10NOV05	5.17[2]
						5.03[2]
PRL	**Zero Calibration-High**	**RTI/Charles River**	10.00	ng/mL	07DEC05	18.85
						21.81
						20.01
						22.21
PRL	**Zero Calibration-High**	**RTI/Charles River**	10.00	ng/mL	09DEC05	19.75
						21.40
PRL	**Zero Calibration-High**	**RTI/RTI**	10.00	ng/mL	13OCT05	26.09
						26.75
						16.96
						18.15
PRL	**Zero Calibration-High**	**RTI/RTI**	10.00	ng/mL	08NOV05	17.17
						18.88
PRL	**Zero Calibration-High**	**RTI/RTI**	10.00	ng/mL	10NOV05	18.30[2]
						19.39[2]
T3	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ng/dL	28NOV05	40.95
						38.91
						43.74
						37.08

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
T3	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/dL	08NOV05	26.66
						32.36
						39.59
						34.70
T3	**Rat Serum Calibration-Mid**	**RTI/Charles River**	50.00	ng/dL	28NOV05	95.96
						97.58
						87.78
						96.97
T3	**Rat Serum Calibration-Mid**	**RTI/RTI**	50.00	ng/dL	08NOV05	86.24
						71.77
						91.85
						93.66
T3	**Rat Serum Calibration-High**	**RTI/Charles River**	300.00	ng/dL	28NOV05	385.37
						380.00
						379.84
						371.71
T3	**Rat Serum Calibration-High**	**RTI/RTI**	300.00	ng/dL	08NOV05	399.90
						351.29
						376.49
						390.18
T3	**Zero Calibration-Low**	**RTI/Charles River**	50.00	ng/dL	28NOV05	48.65
						47.42
						50.95
						51.65
T3	**Zero Calibration-Low**	**RTI/RTI**	50.00	ng/dL	08NOV05	45.52
						45.61
						41.75
						43.68
T3	**Zero Calibration-High**	**RTI/Charles River**	300.00	ng/dL	28NOV05	283.41
						308.66
						311.52
						285.70
T3	**Zero Calibration-High**	**RTI/RTI**	300.00	ng/dL	08NOV05	307.88
						312.88
						340.13
						342.75
T4	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ug/dL	18NOV05	3.31

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						2.90
						3.17
						2.69
T4	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ug/dL	22NOV05	2.96
						3.64
T4	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ug/dL	28OCT05	.
						1.96
						.
						1.70
T4	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ug/dL	02NOV05	2.28
						2.85
T4	**Rat Serum Calibration-Mid 1**	**RTI/Charles River**	5.00	ug/dL	18NOV05	9.00
						7.62
						8.75
						7.35
T4	**Rat Serum Calibration-Mid 1**	**RTI/Charles River**	5.00	ug/dL	22NOV05	7.45
						8.75
T4	**Rat Serum Calibration-Mid 1**	**RTI/RTI**	5.00	ug/dL	28OCT05	6.43
						7.74
						7.43
						7.78
T4	**Rat Serum Calibration-Mid 1**	**RTI/RTI**	5.00	ug/dL	02NOV05	7.98
						6.92
T4	**Rat Serum Calibration-Mid 2**	**RTI/RTI**	8.00	ug/dL	28OCT05	10.98
						9.61
						10.66
						8.80
T4	**Rat Serum Calibration-Mid 2**	**RTI/RTI**	8.00	ug/dL	02NOV05	9.28
						11.41
T4	**Rat Serum Calibration-High**	**RTI/Charles River**	12.00	ug/dL	18NOV05	15.86
						14.44
						15.85
						14.52

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
T4	**Rat Serum Calibration-High**	**RTI/Charles River**	12.00	ug/dL	22NOV05	13.34
						14.86
T4	**Rat Serum Calibration-High**	**RTI/RTI**	12.00	ug/dL	28OCT05	12.88
						12.89
						14.70
						14.53
T4	**Rat Serum Calibration-High**	**RTI/RTI**	12.00	ug/dL	02NOV05	12.97
						16.29
T4	**Zero Calibration-Low**	**RTI/Charles River**	2.67	ug/dL	18NOV05	3.41
						3.10
						2.81
						2.90
T4	**Zero Calibration-Low**	**RTI/Charles River**	2.67	ug/dL	22NOV05	3.17
						2.83
T4	**Zero Calibration-Low**	**RTI/RTI**	2.67	ug/dL	28OCT05	2.65
						3.15
						2.88
						2.79
T4	**Zero Calibration-Low**	**RTI/RTI**	2.67	ug/dL	02NOV05	2.83
						2.68
T4	**Zero Calibration-High**	**RTI/Charles River**	16.00	ug/dL	18NOV05	17.04
						16.01
						15.78
						16.14
T4	**Zero Calibration-High**	**RTI/Charles River**	16.00	ug/dL	22NOV05	16.37
						15.35
T4	**Zero Calibration-High**	**RTI/RTI**	16.00	ug/dL	28OCT05	17.08
						17.37
						15.35
						15.19
T4	**Zero Calibration-High**	**RTI/RTI**	16.00	ug/dL	02NOV05	15.20
						17.06
TSH	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.00	ng/mL	05DEC05	8.95
						8.96
						8.10
						8.48
TSH	**Rat Serum Calibration-Low**	**RTI/RTI**	0.00	ng/mL	20OCT05	8.24

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						8.19
						8.58
						9.65
TSH	**Rat Serum Calibration-Mid**	**RTI/Charles River**	4.00	ng/mL	05DEC05	13.50
						12.46
						13.98
						13.67
TSH	**Rat Serum Calibration-Mid**	**RTI/RTI**	4.00	ng/mL	20OCT05	13.03
						12.79
						13.30
						11.24
TSH	**Rat Serum Calibration-High**	**RTI/Charles River**	16.00	ng/mL	05DEC05	27.73
						29.06
						30.38
						28.14
TSH	**Rat Serum Calibration-High**	**RTI/RTI**	16.00	ng/mL	20OCT05	29.35
						30.41
						30.78
						30.68
TSH	**Zero Calibration-Low**	**RTI/Charles River**	2.50	ng/mL	05DEC05	5.98
						6.21
						5.84
						5.68
TSH	**Zero Calibration-Low**	**RTI/RTI**	2.50	ng/mL	20OCT05	5.98
						6.02
						6.16
						5.91
TSH	**Zero Calibration-High**	**RTI/Charles River**	10.00	ng/mL	05DEC05	19.68
						23.46
						19.78
						20.00
TSH	**Zero Calibration-High**	**RTI/RTI**	10.00	ng/mL	20OCT05	21.07
						21.40
						22.49
						22.84
Testosterone	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.50	ng/mL	17NOV05	0.50
						0.50

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						0.48
						0.46
Testosterone	**Rat Serum Calibration-Low**	**RTI/Charles River**	0.50	ng/mL	22NOV05	0.53
						0.48
Testosterone	**Rat Serum Calibration-Low**	**RTI/RTI**	0.50	ng/mL	20OCT05	0.32
						0.34
						0.37
						0.37
Testosterone	**Rat Serum Calibration-Low**	**RTI/RTI**	0.50	ng/mL	24OCT05	0.40
						0.39
Testosterone	**Rat Serum Calibration-Mid**	**RTI/Charles River**	4.00	ng/mL	17NOV05	3.82
						3.89
						4.45
						4.33
Testosterone	**Rat Serum Calibration-Mid**	**RTI/Charles River**	4.00	ng/mL	22NOV05	3.89
						4.08
Testosterone	**Rat Serum Calibration-Mid**	**RTI/RTI**	4.00	ng/mL	20OCT05	4.43
						3.64
						3.91
						3.60
Testosterone	**Rat Serum Calibration-Mid**	**RTI/RTI**	4.00	ng/mL	24OCT05	4.03
						3.53
Testosterone	**Rat Serum Calibration-High**	**RTI/Charles River**	8.00	ng/mL	17NOV05	7.96
						7.88
						8.10
						7.23
Testosterone	**Rat Serum Calibration-High**	**RTI/Charles River**	8.00	ng/mL	22NOV05	7.10
						7.95
Testosterone	**Rat Serum Calibration-High**	**RTI/RTI**	8.00	ng/mL	20OCT05	7.85
						7.78
						7.97
						7.57
Testosterone	**Rat Serum Calibration-High**	**RTI/RTI**	8.00	ng/mL	24OCT05	8.06
						8.68
Testosterone	**Zero Calibration-Low**	**RTI/Charles River**	0.50	ng/mL	17NOV05	0.47
						0.63
						0.53

Table 2 Individual RTI QC Assay Data Values by Hormone, Originating Laboratory, QC Standard Type, Spike Concentration, Unit, and Assay Date.

Hormone Parameter	QC Standard Type	Analysis Laboratory /Test Laboratory	Spike Concentration	Unit	Assay Date	Individual QC data[1]
						0.56
Testosterone	**Zero Calibration-Low**	**RTI/Charles River**	0.50	ng/mL	22NOV05	0.64
						0.54
Testosterone	**Zero Calibration-Low**	**RTI/RTI**	0.50	ng/mL	20OCT05	0.49
						0.40
						0.47
						0.41
Testosterone	**Zero Calibration-Low**	**RTI/RTI**	0.50	ng/mL	24OCT05	0.57
						0.48
Testosterone	**Zero Calibration-High**	**RTI/Charles River**	4.00	ng/mL	17NOV05	4.01
						4.27
						4.21
						3.90
Testosterone	**Zero Calibration-High**	**RTI/Charles River**	4.00	ng/mL	22NOV05	4.34
						4.29
Testosterone	**Zero Calibration-High**	**RTI/RTI**	4.00	ng/mL	20OCT05	4.04
						4.04
						3.63
						3.70
Testosterone	**Zero Calibration-High**	**RTI/RTI**	4.00	ng/mL	24OCT05	4.17
						4.41

1. Data with values above assay range ("AAR") were not included in the summary calculation.
2. PRL QC standards concentration data with assay date 11/10/2005 were excluded from the summary calculations, as per instructions on the RTI data form.

Table 3 Individual WIL Laboratory Assay QC Standards Data Values and Within-Assay Coefficient of Variation for the Hormonal Asssays. By Hormone Type, QC Standard Type, Test Concentration, Assay Date, and Unit.

Hormone Parameter	Dose Group	Test Concentration	Unit	Assay Date	Individual QC Standard	Mean	Within Assay Coefficient of Variation (%)[1]
DHT	**Kit QC Low**	100.00	pg/mL	01JAN05	92.30	92.30	2.00
DHT	**Kit QC Mid**	500.00	pg/mL	01JAN05	596.70	596.70	6.30
DHT	**Kit QC High**	.		.		-	-
DHT	**Non-Kit QC Low**	50.00	pg/mL	01JAN05	242.20	242.20	1.10
DHT	**Non-Kit QC High**	500.00	pg/mL	01JAN05	322.70	322.70	4.10
Estradiol	**Kit QC Low**	250.00	pg/mL	01JAN05	251.80	251.80	1.10
Estradiol	**Kit QC Mid**	.		.		-	-
Estradiol	**Kit QC High**	1000.00	pg/mL	01JAN05	988.70	988.70	2.30
Estradiol	**Non-Kit QC Low**	50.00	pg/mL	01JAN05	87.40	87.40	6.90
Estradiol	**Non-Kit QC High**	750.00	pg/mL	01JAN05	662.90	662.90	27.40
FSH	**Kit QC Low**	4.80	ng/mL	01JAN05	2.00	2.00	0.10
FSH	**Kit QC Mid**	.		.		-	-
FSH	**Kit QC High**	61.00	ng/mL	01JAN05	71.20	71.20	0.90
FSH	**Non-Kit QC Low**	0.60	ng/mL	01JAN05	Off Curve	-	0.90
FSH	**Non-Kit QC High**	5.00	ng/mL	01JAN05	8.60	8.60	2.60
LH	**Kit QC Low**	1.20	ng/mL	01JAN05	1.00	1.00	3.20
LH	**Kit QC Mid**	.		.		-	-
LH	**Kit QC High**	33.30	ng/mL	01JAN05	30.40	30.40	3.20
LH	**Non-Kit QC Low**	0.60	ng/mL	01JAN05	0.60	0.60	1.80
LH	**Non-Kit QC High**	2.50	ng/mL	01JAN05	1.80	1.80	1.50
PRL	**Kit QC Low**	1.10	ng/mL	01JAN05	1.10	1.05	0.60
		1.10	ng/mL	01JAN05	1.00		
PRL	**Kit QC Mid**	.		.		-	-
PRL	**Kit QC High**	31.20	ng/mL	01JAN05	33.20	31.65	1.70
		31.20	ng/mL	01JAN05	30.10		
PRL	**Non-Kit QC Low**	0.60	ng/mL	01JAN05	0.70	0.40	2.80
		0.60	ng/mL	01JAN05	0.10		
PRL	**Non-Kit QC High**	2.50	ng/mL	01JAN05	3.60	4.25	3.80
		2.50	ng/mL	01JAN05	4.90		
T3	**Kit QC Low**	74.00	ng/dL	01JAN05	71.50	71.50	1.50
T3	**Kit QC Mid**	153.00	ng/dL	01JAN05	140.10	140.10	2.20
T3	**Kit QC High**	235.00	ng/dL	01JAN05	220.60	220.60	0.10
T3	**Non-Kit QC Low**	50.00	ng/dL	01JAN05	30.90	30.90	6.40
T3	**Non-Kit QC High**	200.00	ng/dL	01JAN05	42.60	42.60	0.20

Table 3 Individual WIL Laboratory Assay QC Standards Data Values and Within-Assay Coefficient of Variation for the Hormonal Asssays. By Hormone Type, QC Standard Type, Test Concentration, Assay Date, and Unit.

Hormone Parameter	Dose Group	Test Concentration	Unit	Assay Date	Individual QC Standard	Mean	Within Assay Coefficient of Variation (%)[1]
T4	**Kit QC Low**	2.60	ug/dL	01JAN05	1.40	1.40	1.60
T4	**Kit QC Mid 1**	7.70	ug/dL	01JAN05	8.20	8.20	0.60
T4	**Kit QC Mid 2**	.		.		-	-
T4	**Kit QC High**	11.90	ug/dL	01JAN05	12.60	12.60	4.30
T4	**Non-Kit QC Low**	4.00	ug/dL	01JAN05	0.90	0.90	4.50
T4	**Non-Kit QC High**	16.00	ug/dL	01JAN05	15.90	15.90	3.50
TSH	**Kit QC Low**	3.20	ng/mL	01JAN05	2.50	2.50	0.60
TSH	**Kit QC Mid**	.		.		-	-
TSH	**Kit QC High**	41.80	ng/mL	01JAN05	41.80	41.80	6.60
TSH	**Non-Kit QC Low**	0.40	ng/mL	01JAN05	1.00	1.00	2.30
TSH	**Non-Kit QC High**	1.60	ng/mL	01JAN05	1.90	1.90	2.10
Testosterone[2]	**Kit QC Low**	1.10	ng/mL	01JAN05	1.00	1.04	1.20
		1.10	ng/mL	01JAN05	1.08		
Testosterone[2]	**Kit QC Mid**	4.17	ng/mL	01JAN05	4.57	4.91	2.20
		4.17	ng/mL	01JAN05	5.26		
Testosterone[2]	**Kit QC High**	7.33	ng/mL	01JAN05	7.89	7.56	3.50
		7.33	ng/mL	01JAN05	7.23		
Testosterone[2]	**Non-Kit QC Low**	1.00	ng/mL	01JAN05	1.32	3.31	10.30
		1.00	ng/mL	01JAN05	5.31		
Testosterone[2]	**Non-Kit QC High**	8.00	ng/mL	01JAN05	6.50	8.39	3.40
		8.00	ng/mL	01JAN05	10.29		

1. Based on duplicate measurement cpm values.
2. WIL Laboratory reported testosterone concentrations in units "ng/dL". RTI reported testosterone concentrations in units "ng/mL". The WIL "Test Concentration" and "Individual QC Standard" values were divided by 100 to transform them to the same units as the RTI concentrations.

APPENDIX B

Inter-Laboratory Validation of the 15-Day Intact Adult Male Rat Assay: Statistical Analysis of Among-Laboratory Results

This page left blank intentionally.

AMENDED FINAL REPORT

on

INTER-LABORATORY VALIDATION OF THE 15-DAY ADULT INTACT MALE RAT ASSAY: STATISTICAL ANALYSIS OF AMONG LABORATORY RESULTS

COMPARISON OF LIKELIHOOD RATIO TESTS AND WALD Z TESTS OF HOMOGENEITY OF LABORATORIES

EPA CONTRACT NUMBER 68-W-01-023
WORK ASSIGNMENT 5-15

DECEMBER 20, 2006

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
ENDOCRINE DISRUPTOR SCREENING PROGRAM
WASHINGTON, D.C.

Prepared by

BATTELLE
505 King Avenue
Columbus, Ohio 43201

INTER-LABORATORY VALIDATION OF THE 15-DAY ADULT INTACT MALE RAT ASSAY: STATISTICAL ANALYSIS OF AMONG LABORATORY RESULTS

EPA CONTRACT NUMBER 68-W-01-023
WORK ASSIGNMENT 5-15

Zhenxu J. Ma, Author — Date: 12/21/06

Paul I. Feder, Reviewer — Date: 12/21/2006

Quality Assurance Statement

Study Number: WA 5-15

This study was inspected by the Quality Assurance Unit and reports were submitted to the Study Director and Management as follows:

Phase Inspected	Inspection Date	Date Reported to Battelle Task Leader/ Battelle Management
Audit study file	6/23/2006	6/23/2006
Audit inter-laboratory statistics draft report	6/23/2006	6/23/2006
Audit inter-laboratory statistics revised draft report	9/1/2006	9/1/2006
Audit inter-laboratory statistics revised final report	9/20/2006	9/20/2006
Audit inter-laboratory statistics draft amended final report	12/14/2006	12/14/2006
Audit inter-laboratory statistics amended final report	12/20/2006	12/20/2006

Quality Assurance Unit **Date**

Terri L Pollock 12-21-06

Terri L. Pollock

This page left blank intentionally.

Table of Contents

Page

List of Tables

This page left blank intentionally.

Executive Summary

Three contract research organization (CRO) laboratories—RTI International, Charles River Laboratories, Inc., and WIL Research Laboratories, LLC—conducted a 15-day adult intact male rat assay according to a test protocol provided by the EPA. Within each laboratory, two chemicals, linuron and phenobarbital, were tested, each at three dose levels plus a common vehicle control as specified by the EPA. The sample size was n=15 adult male rats per dose level (low, intermediate, and high) and a control for a total of seven groups and 105 animals per laboratory.

Separate intra-laboratory statistical analyses were carried out with each laboratory's data to summarize the within laboratory results, which are presented in detail in individual laboratory reports. An inter-laboratory statistical analysis was carried out to compare the intra-laboratory analysis results among the laboratories. The principal objective of the inter-laboratory analysis was to assess the extent of variation of results across laboratories as a measure of the reliability of this bioassay. The results of the inter-laboratory analysis were presented in the final report dated September 20, 2006 (1).

The September 2006 inter-laboratory analysis final report presented estimates of the intra-laboratory and the inter-laboratory variability, the among laboratory coefficient of variation, and the ratio of the among laboratory variability to the average within laboratory variability. It also presented probability values for the null hypothesis that the among laboratories component of variability is zero.

The probability values presented in the September, 2006 report were those given by the SAS statistical computing system (2) based on the ratio of the estimated variance component to its estimated standard error. These are referred to as "Wald Z statistic" p-values. This is the default procedure. The SAS system documentation for PROC MIXED, paragraph titled "Inference and Test Statistics", states "…the Wald Z is valid for large samples …A better alternative is the likelihood ratio χ^2…". This amended report presents the probability values of the likelihood ratio test of the significance of the among laboratory variation side-by-side with the probability values of the Wald Z test. All other results provided in the September, 2006 report stay the same and are repeated in this report.

The September 2006 report stated that the among laboratory variation did not significantly differ from zero (based on the Wald Z test) for any endpoint and within any chemical-dose group combination. By contrast the likelihood ratio test results indicated multiple endpoint-chemical-dose group combinations for which the among laboratories variance component was significantly ($p<0.05$) greater than zero. These are summarized below.

Body Weight and Food Consumption

Linuron 50 mg/kg/day

Body weight change (TD 15 – TD 8) ($p<0.05$)
Body weight change (TD 15 – TD 1) ($p<0.05$)

Linuron 100 mg/kg/day

Body weight change (TD 8 – TD 1) ($p<0.05$)
Body weight change (TD 15 – TD 1) ($p<0.05$)

Phenobarbital 100 mg/kg/day

Body weight change (TD 8 – TD 1) ($p<0.05$)
Body weight change (TD 15 – TD 1) ($p<0.05$)
Food consumption (TD 8 – TD 1) ($p<0.05$)

Organ Weights

Linuron (laboratory main effect averaged over dose groups)

Adjusted liver ($p<0.05$)

Phenobarbital 100 mg/kg/day

Liver ($p<0.00833$)
Adjusted liver ($p<0.00833$)

Phenobarbital (laboratory main effect averaged over dose groups)

Adjusted liver ($p<0.05$)

Hormonal Analysis

Linuron 100 mg/kg/day

T4 ($p<0.00833$)

Linuron 150 mg/kg/day

T4 ($p<0.00833$)

Linuron (laboratory main effect averaged over dose groups)

Testosterone ($p<0.01$)
LH ($p<0.05$)
T4 ($p<0.01$)
T3 ($P<0.05$)
FSH ($p<0.05$)

Phenobarbital 50 mg/kg/day

LH ($p<0.05$)
T4 ($p<0.00833$)

Phenobarbital 100 mg/kg/day

T4 ($p<0.00833$)

Phenobarbital (laboratory main effect averaged over dose groups)

Testosterone ($p<0.05$)
LH ($p<0.05$)
T4 ($p<0.05$)

Prolactin ($p<0.01$)

For T4 the significant among laboratories variances were due to the ratios of dose group to control group mean being smaller for WIL than for the other laboratories.

Reanalysis Excluding Outliers

The analysis was repeated after excluding outliers. The endpoint-chemical-dose group combinations for which the among laboratories variance was statistically significantly greater than zero are summarized below.

<u>Linuron 50 mg/kg/day</u>
Body weight change (TD15 – TD8) ($p<0.05$)
Prolactin ($p<0.00833$)

<u>Linuron 100 mg/kg/day</u>
Prolactin ($p<0.00833$)

<u>Linuron (laboratory main effect averaged over dose groups)</u>
Food consumption (TD 15 – TD 8) ($p<0.05$)
Testosterone ($p<0.01$)
FSH ($p<0.05$)
Prolactin ($p<0.01$)

<u>Phenobarbital 50 mg/kg/day</u>
Prolactin ($p<0.05$)

<u>Phenobarbital 100 mg/kg/day</u>
Prolactin ($p<0.00833$)

<u>Phenobarbital (laboratory main effect averaged over dose groups)</u>
Prolactin ($p<0.01$)

Conclusion

The likelihood ratio test, included in the amended report, had consistently lower probability values than the Wald Z test for the same endpoint-chemical-dose group. Based on the likelihood ratio test there were significant differences among laboratories for a number of the endpoints. About two thirds of the significant differences among laboratories were for hormonal analysis endpoints, particularly T4 and prolactin. Testosterone and prolactin also showed some significant differences among groups. These were the endpoints that were identified previously as having relatively large among laboratory CVs.

In relation to the principal objective, the 15-day adult intact male rat assay is relatively homogeneous across laboratories for the food consumption, body weight, organ weight, and some of the hormonal endpoints. However for several of the hormonal endpoints (particularly

testosterone, T4, and prolactin) it exhibits relatively high CVs and significant heterogeneity among laboratories.

Introduction and Background

Three contract research organization (CRO) laboratories—RTI International, Charles River Laboratories, Inc., and WIL Research Laboratories, Inc.—conducted a 15-day adult intact male rat assay according to a test protocol provided by the EPA. Within each laboratory, a common vehicle control and two chemicals, linuron and phenobarbital, were tested, each at three dose levels (low, intermediate, and high) specified by the EPA. The sample size was n=15 adult male rats per dose level and a control, for a total of seven groups and 105 animals per laboratory.

Separate intra-laboratory statistical analyses were carried out for each laboratory to summarize the within laboratory results. The summaries, displays, and inferences that were included in the intra-laboratory statistical analyses were based on a uniform intra-laboratory statistical analysis plan. The results are detailed in individual laboratory reports. The results of the inter-laboratory analyses are discussed in detail in the September 20, 2006 report (1).

Objective

The inter-laboratory statistical analysis was carried out to compare the intra-laboratory analysis results among the laboratories. The principal objective of the inter-laboratory analysis is to assess the extent of variation of results across laboratories as a measure of the reliability of this bioassay. This was accomplished by using the results of the ratios of treated group means to the control group mean from within laboratories for comparison among laboratories and assessing the variation among laboratories based on the among laboratory coefficients of variation (CV) and the among laboratory component of variability. The treated to control group ratios from within laboratories were used, in part, to standardize systematic and operational differences inherent in different laboratories so that the focus of the comparisons among laboratories is on the relative effects of treatment and not necessarily on whether actual organ weights or hormone concentrations are the same within each laboratory. The inter-laboratory statistical analysis studies the variation in the relative responses to the treatments among laboratories. The overall hypothesis is that the variation in mean relative responses among laboratories does not differ from zero, or equivalently that the mean relative results are equal across laboratories. This is hypothesized to be true for three main categories of data: final body and organ weights and serum hormone concentrations.

The September 2006 report included tests of significance for the hypothesis that the among laboratories component of variation was zero. It presented probability values based on the Wald Z test, as reported by the SAS statistical analysis system (2). The SAS system documentation discusses the use of an alternative test, the likelihood ratio chi square test. This amended report compares the results of the likelihood ratio chi square test side-by-side with the previously reported Wald Z test results.

Data

This amended report is based on the same responses that were used in the September 20, 2006 inter-laboratory analysis report. The test method specifies 34 endpoints, divided among three categories:

1. Body weights and food consumption – (7 endpoints)
    - Body weight change (TD8 – TD1)
    - Body weight change (TD15 – TD8)
    - Body weight change (TD15 – TD1)
    - Final body weight (TD15)
    - Food consumption (TD8 - TD1)
    - Food consumption (TD15 - TD8)
    - Food consumption (TD15 - TD1)

The Test Day (TD) 15 final body weight is the live weight before sacrifice.
Body weights were reported in grams. Body weight changes were reported in g/day. Food consumption was reported in g/kg/day.

2. Organ weights – (9 organs)
    - Liver
    - Right testis
    - Left testis
    - Testes paired (sum of left and right testis weights)
    - Epididymides (paired weight)
    - Entire prostate
    - Seminal vesicles with fluid and coagulating gland
    - Accessory sex gland (ASG) (sum of entire prostate and seminal vesicles with fluid and coagulating gland weights)
    - Thyroid

Organ weights were reported in grams. Organ weights were analyzed in two ways:
- Unadjusted
- Adjusted − Organ to final body weight ratio (expressed as percent)

There are 18 organ weight endpoints (nine organs × unadjusted and adjusted).

3. Hormonal analysis - (9 serum hormone concentrations)
    - Testosterone (ng/ml)
    - LH (ng/ml)
    - TSH (ng/ml)
    - T_4 (μg/dl)
    - T_3 (ng/dl)
    - FSH (ng/ml)
    - Estradiol (pg/ml)

Prolactin (ng/ml)
DHT (pg/ml)

The test protocol specified that all rats were to be sacrificed on TD 15. If animals died or were euthanized prior to necropsy, their body weights were included in summaries and displays up to the time of death, but were not imputed beyond date of death nor included in the final body weight data summaries. The numbers of deaths per group prior to necropsy were reported in the intra-laboratory statistical analysis reports.

Statistical Analysis Methods

The test results reported in the September 20, 2006 report were based on the Wald Z test. This is the default procedure. Since there are just three test laboratories there were just 3 -1≡ 2 degrees of freedom with which to estimate the laboratory-to-laboratory variance component. This is a small sample of laboratories.

The SAS PROC MIXED system documentation, "Inference and Test Statistics" section, states "…For inferences concerning the covariance parameters in your model, you can use likelihood-based statistics. One common likelihood-based statistic is the Wald Z, which is computed as the parameter estimate divided by its asymptotic standard error….the Wald Z is valid for large samples but it can be unreliable for small data sets and for parameters such as variance components, which are known to have a skewed or bounded sampling distribution….A better alternative is the likelihood ratio χ^2…as long as the reduced model does not occur on the boundary of the covariance parameter space, the χ^2 statistic…has a large-sample sampling distribution that is χ^2 with degrees of freedom equal to the difference in the number of covariance parameters between the two models. If the reduced model does occur on the boundary of the covariance parameter space, the asymptotic distribution becomes a mixture of χ^2 distributions (3). A common example of this is when you are testing that a variance component equals its lower boundary constraint of 0." The large-sample distribution of the likelihood ratio χ^2 when the reduced model occurs on the boundary of the covariance parameter space is also discussed in (4).

In the application discussed in this report, the results of likelihood ratio tests that the among laboratories variance component is zero are reported. Under the null hypothesis the distribution of the likelihood ratio test statistic is 0 with probability ½ and χ^2 with one degree of freedom, with probability ½.

The probability values of the likelihood ratio test are reported side-by-side with those of the previously reported Wald Z test.

Appendix A contains an example output from SAS PROC MIXED that illustrates where the Wald Z test probability values were reported and how the likelihood ratio test probability values were calculated. The example corresponds to Phenobarbital 100 mg/kg/day, adjusted liver weight. The Wald Z test and the likelihood ratio test probability values are reported in Table 4c.

Statistical Analysis Results

The September 20, 2006 inter-laboratory statistical analysis report included a one-way mixed effects analysis of variance and a two-way mixed effects analysis of variance.

The tables in this amended report are numbered in correspondence with those in the September 20, 2006 report. They include the summaries and Wald Z test probability values presented in the previous report as well as the probability values of the likelihood ratio test for side-by-side comparison with the Wald Z test.

The tables are numbered as follows:

- Tables 1a through 2c display results for the seven body weight and food consumption endpoints
- Tables 3a through 4c display results for the 18 unadjusted organ weight and adjusted organ weight (organ weight to final body weight ratio) endpoints
- Tables 5a through 6c display results for the nine serum hormonal analysis endpoints.

Tables 7 through 12 display the components of variance associated with laboratory main effect, laboratory-by-dose interaction, their probability values and indications of their significance for each chemical and endpoint category. Tables 7, 9, and 11 display results for linuron. Tables 8, 10, and 12 display results for phenobarbital.

The significant likelihood ratio test results are summarized below. Refer to the September 20, 2006 for discussion of the remaining summaries in the tables.

Body Weight and Food Consumption

Linuron 50 mg/kg/day

Body weight change (TD 15 – TD 8) ($p<0.05$)
Body weight change (TD 15 – TD 1) ($p<0.05$)

Linuron 100 mg/kg/day

Body weight change (TD 8 – TD 1) ($p<0.05$)
Body weight change (TD 15 – TD 1) ($p<0.05$)

Phenobarbital 100 mg/kg/day

Body weight change (TD 8 – TD 1) ($p<0.05$)
Body weight change (TD 15 – TD 1) ($p<0.05$)
Food consumption (TD 8 – TD 1) ($p<0.05$)

Organ Weights

Linuron (laboratory main effect averaged over dose groups)

Adjusted liver ($p<0.05$)

Phenobarbital 100 mg/kg/day
Liver ($p<0.00833$)
Adjusted liver ($p<0.00833$)

Phenobarbital (laboratory main effect averaged over dose groups)
Adjusted liver ($p<0.05$)

Hormonal Analysis

Linuron 100 mg/kg/day
T4 ($p<0.00833$)

Linuron 150 mg/kg/day
T4 ($p<0.00833$)

Linuron (laboratory main effect averaged over dose groups)
Testosterone ($p<0.01$)
LH ($p<0.05$)
T4 ($p<0.01$)
T3 ($P<0.05$)
FSH ($p<0.05$)

Phenobarbital 50 mg/kg/day
LH ($p<0.05$)
T4 ($p<0.00833$)

Phenobarbital 100 mg/kg/day
T4 ($p<0.00833$)

Phenobarbital (laboratory main effect averaged over dose groups)
Testosterone ($p<0.05$)
LH ($p<0.05$)
T4 ($p<0.05$)
Prolactin ($p<0.01$)

For T4 the significant among laboratories variances were due to the ratios of treatment to control group mean being smaller for WIL than for the other laboratories.

Reanalysis Excluding Outliers

Additional analyses of variance were applied to the data excluding outliers for those responses that had at least one outlying value in one of the seven control or chemical-dose group combinations. There were relatively few outliers. The results are summarized in Tables 13a through 14c and in Tables 15 and 16. The results were nearly the same as those including the outliers.

The endpoint-chemical-dose group combinations for which the among laboratories variance was statistically significantly greater than zero after excluding outliers are summarized below.

Linuron 50 mg/kg/day
- Body weight change (TD 15 – TD 8) ($p<0.05$)
- Prolactin ($p<0.00833$)

Linuron 100 mg/kg/day
- Prolactin ($p<0.00833$)

Linuron (laboratory main effect averaged over dose groups)
- Food consumption (TD 15 – TD 8) ($p<0.05$)
- Testosterone ($p<0.01$)
- FSH ($p<0.05$)
- Prolactin ($p<0.01$)

Phenobarbital 50 mg/kg/day
- Prolactin ($p<0.05$)

Phenobarbital 100 mg/kg/day
- Prolactin ($p<0.00833$)

Phenobarbital (laboratory main effect averaged over dose groups)
- Prolactin ($p<0.01$)

Conclusion

The principal objective of the amended inter-laboratory analysis report is to present comparisons of the results of likelihood ratio tests of homogeneity among laboratories with those of the previously reported Wald Z test results. In all cases the probability values of the likelihood ratio tests were less than or equal to those of the corresponding Wald Z tests.

A number of the likelihood ratio test results were significant ($p\leq 0.05$). For the most part (22 of 35) the significant likelihood ratio test results occurred for the hormonal analysis endpoints. Of the 22 significant hormonal analysis endpoints, 19 corresponded to prolactin and T4 and to a lesser extent testosterone and LH. This is in agreement with the previously reported large among laboratory variance component and CVs for these parameters.

References

15-DAY ADULT INTACT MALE RAT ASSAY: STATISTICAL ANALYSIS OF AMONG LABORATORY RESULTS

1. Battelle, 2006. Interlaboratory Validation of the 15-Day Intact Adult Male Rat Assay: Statistical Analysis of Among Laboratory Results. Final Report to U. S. Environmental Protection Agency, EPA Contract Number 68-W-01-023.

2. SAS Institute Inc., 2003. SAS Statistical Analysis System. Version 9.1.

3. Self, S.G. and K. Y. Liang, 1987. "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions." *Journal of the American Statistical Association* **82**: 605-610.

4. Chernoff, H,, 1954. "On the Distribution of the Likelihood Ratio." *Ann. Math. Statist*. **25**: 573-578.

APPENDIX A

Example SAS Output Illustrating Wald Z Test and Likelihood Ratio Test Significance Level Determinations

Adjusted Liver
Phenobarbital 100 mg/kg/day
(Table 4c)

Mixed Procedure for Adj Liver with Phenobarbital 100

22

The Mixed Procedure

Estimated V Matrix for Subject 1

Row	Col1	Col2	Col3
1	84.0836		
2		84.0304	
3			83.2449

Covariance Parameter Estimates

Cov Parm	Group	Estimate	Standard Error	Z Value	Pr Z
lab		68.7288	68.2507	1.01	**0.1570**
Residual	lab Charles River	15.3548	0	.	.
Residual	lab RTI	14.5161	0	.	.
Residual	lab WIL	15.3016	0	.	.

(Wald Z test probability value)

Fit Statistics

-2 Log Likelihood	21.8
AIC (smaller is better)	25.8
AICC (smaller is better)	37.8
BIC (smaller is better)	24.0

(full model)

Mixed Procedure for Adj Liver with Phenobarbital 100 - Without random lab

24

The Mixed Procedure

Fit Statistics

-2 Log Likelihood	30.5
AIC (smaller is better)	32.5
AICC (smaller is better)	36.5
BIC (smaller is better)	31.6

(reduced model)

Likelihood Ratio Test: [-2log likelihood (reduced model)] – [-2log likelihood (full model)] = 30.5 – 21.8 = 8.7

½ [1 - Prob ($\chi_1^2 > 8.7$)] = **0.0016** (Likelihood ratio test probability value)

.

Table 1a. Comparisons Among Laboratories of Ratio of Linuron 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption

	Linuron (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Body Weight Change (TD8-TD1)	2.149 (9.239)	28.100 (11.800)	-21.617 (14.105)	11.715	4.007 (10.748)	464.538	14.542	0.2485	124.134	0.1246
Body Weight Change (TD15-TD8)	58.606 (10.796)	105.800 (9.700)	95.871 (13.429)	11.309	86.715 (12.127)	24.223	17.714	0.1841	156.639	0.0177*
Body Weight Change (TD15-TD1)	26.504 (7.636)	60.500 (8.100)	29.569 (8.533)	8.090	38.818 (8.888)	39.660	13.099	0.1870	161.922	0.0217*
Final Body Weight	87.523 (2.330)	93.600 (1.900)	89.888 (2.080)	2.103	90.618 (1.455)	2.781	1.408	0.3497	66.924	0.3096
Food Consumption (TD8-TD1)	75.992 (3.483)	85.100 (3.600)	82.540 (3.475)	3.519	81.141 (2.210)	4.718	1.510	0.4246	42.908	0.4158
Food Consumption (TD15-TD8)	89.016 (3.885)	94.700 (6.200)	100.521 (3.683)	4.589	94.887 (3.226)	5.888	3.454	0.2964	75.258	0.2160
Food Consumption (TD15-TD1)	81.735 (3.358)	88.400 (4.100)	90.419 (2.659)	3.372	87.131 (2.278)	4.529	2.192	0.3446	65.008	0.3006

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and ½. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 1b. Comparisons Among Laboratories of Ratio of Linuron 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption

	Linuron (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Body Weight Change (TD8-TD1)	-18.843 (9.399)	-27.400 (11.800)	-82.246 (15.268)	12.156	-40.621 (15.383)	-65.591	23.690	0.1799	194.891	0.0135*
Body Weight Change (TD15-TD8)	52.941 (10.539)	68.100 (11.400)	64.807 (11.552)	11.164	61.439 (6.429)	18.126	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD1)	12.124 (7.435)	12.400 (6.900)	-18.180 (8.317)	7.551	2.720 (8.027)	511.089	11.676	0.2017	154.636	0.0414*
Final Body Weight	85.407 (2.305)	86.900 (1.800)	82.867 (2.009)	2.038	85.181 (1.159)	2.356	0.000	1.0000	0.000	1.0000
Food Consumption (TD8-TD1)	68.017 (3.201)	75.600 (4.100)	63.478 (3.446)	3.582	68.553 (2.637)	6.662	2.867	0.3311	80.036	0.2795
Food Consumption (TD15-TD8)	81.451 (3.814)	100.600 (5.700)	91.192 (3.499)	4.338	90.160 (4.182)	8.034	5.837	0.2300	134.563	0.0800
Food Consumption (TD15-TD1)	74.261 (3.112)	87.200 (5.100)	75.941 (2.517)	3.576	76.807 (1.827)	4.120	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$.
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 1c. Comparisons Among Laboratories of Ratio of Linuron 150 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption

	Linuron (150 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Body Weight Change (TD8-TD1)	-80.826 (11.877)	-81.200 (13.600)	-91.106 (15.983)	13.820	-83.402 (7.806)	-16.212	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD8)	55.773 (10.665)	65.100 (17.200)	40.988 (10.792)	12.886	51.177 (6.941)	23.491	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD1)	-21.898 (7.556)	-20.200 (7.300)	-33.557 (8.904)	7.920	-24.254 (4.522)	-32.295	0.000	1.0000	0.000	1.0000
Final Body Weight	79.623 (2.241)	81.000 (1.800)	79.571 (2.020)	2.020	80.170 (1.152)	2.490	0.000	1.0000	0.000	1.0000
Food Consumption (TD8-TD1)	52.689 (3.074)	63.300 (4.500)	60.271 (3.569)	3.715	57.940 (2.649)	7.918	2.787	0.3200	75.026	0.2583
Food Consumption (TD15-TD8)	78.821 (3.900)	89.900 (5.500)	80.396 (3.494)	4.298	81.561 (2.352)	4.995	0.000	1.0000	0.000	1.0000
Food Consumption (TD15-TD1)	63.641 (3.227)	72.200 (5.700)	69.414 (2.578)	3.835	67.724 (1.899)	4.857	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$.
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [$\sqrt{3}$×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 2a. Comparisons Among Laboratories of Ratio of Phenobarbital 25 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption

	Phenobarbital (25 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Body Weight Change (TD8-TD1)	103.636 (13.302)	104.300 (11.200)	98.311 (12.517)	12.340	102.202 (7.070)	11.982	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD8)	109.804 (13.834)	108.500 (13.200)	104.430 (14.016)	13.683	107.634 (7.892)	12.700	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD1)	106.297 (10.772)	106.100 (10.100)	100.977 (11.629)	10.834	104.698 (6.224)	10.296	0.000	1.0000	0.000	1.0000
Final Body Weight	102.132 (2.506)	100.500 (1.900)	99.742 (2.185)	2.197	100.657 (1.244)	2.141	0.000	1.0000	0.000	1.0000
Food Consumption (TD8-TD1)	100.425 (3.751)	99.700 (4.300)	103.681 (2.589)	3.547	102.053 (1.909)	3.241	0.000	1.0000	0.000	1.0000
Food Consumption (TD15-TD8)	97.672 (2.909)	99.300 (3.100)	99.757 (2.655)	2.888	98.950 (1.657)	2.901	0.000	1.0000	0.000	1.0000
Food Consumption (TD15-TD1)	98.647 (2.372)	98.400 (3.400)	101.650 (2.020)	2.597	100.050 (1.401)	2.426	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$.
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [$\sqrt{3}$×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 2b. Comparisons Among Laboratories of Ratio of Phenobarbital 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption

	Phenobarbital (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Body Weight Change (TD8-TD1)	91.736 (12.535)	89.700 (10.400)	89.001 (12.133)	11.689	90.067 (6.681)	12.848	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD8)	109.586 (13.819)	109.300 (10.900)	92.261 (13.190)	12.636	104.329 (7.179)	11.919	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD1)	99.436 (10.409)	97.900 (9.600)	90.422 (11.032)	10.347	96.229 (5.945)	10.700	0.000	1.0000	0.000	1.0000
Final Body Weight	101.008 (2.492)	99.900 (1.900)	98.026 (2.166)	2.186	99.561 (1.239)	2.156	0.000	1.0000	0.000	1.0000
Food Consumption (TD8-TD1)	100.204 (3.747)	106.000 (4.100)	105.077 (2.674)	3.507	103.997 (1.922)	3.202	0.000	1.0000	0.000	1.0000
Food Consumption (TD15-TD8)	100.990 (2.906)	100.900 (4.600)	98.957 (2.689)	3.398	100.051 (1.814)	3.140	0.000	1.0000	0.000	1.0000
Food Consumption (TD15-TD1)	100.569 (2.364)	103.200 (3.300)	100.548 (2.117)	2.594	101.049 (1.423)	2.439	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$.
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 2c. Comparisons Among Laboratories of Ratio of Phenobarbital 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption

	Phenobarbital (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Body Weight Change (TD8-TD1)	36.860 (9.844)	-4.400 (8.200)	29.630 (10.791)	9.612	19.787 (10.790)	94.450	16.029	0.1752	166.761	0.0105*
Body Weight Change (TD15-TD8)	102.708 (13.583)	87.100 (12.800)	110.776 (14.697)	13.693	99.123 (7.868)	13.748	0.000	1.0000	0.000	1.0000
Body Weight Change (TD15-TD1)	65.857 (9.055)	33.800 (7.800)	64.983 (10.001)	8.952	53.984 (8.956)	28.736	12.679	0.1979	141.636	0.0357*
Final Body Weight	94.536 (2.458)	89.600 (1.900)	95.077 (2.174)	2.177	92.784 (1.537)	2.869	1.558	0.3266	71.572	0.2720
Food Consumption (TD8-TD1)	85.803 (3.488)	79.800 (3.800)	93.527 (2.643)	3.310	86.918 (3.311)	6.598	4.690	0.2014	141.694	0.0390*
Food Consumption (TD15-TD8)	101.699 (2.967)	110.000 (4.400)	105.446 (2.829)	3.399	104.790 (1.856)	3.068	0.000	1.0000	0.000	1.0000
Food Consumption (TD15-TD1)	93.057 (2.296)	94.500 (3.200)	99.269 (2.109)	2.535	95.907 (1.746)	3.153	1.751	0.3236	69.063	0.2659

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$.
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 3a. Comparisons Among Laboratories of Ratio of Linuron 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[8]

Endpoint	Linuron (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Liver	85.014 (4.279)	96.890 (3.290)	98.455 (4.489)	4.019	93.666 (3.289)	6.083	4.070	0.2767	101.271	0.1791
Right Testis	104.172 (2.892)	101.590 (3.010)	101.162 (2.335)	2.746	102.147 (1.555)	2.637	0.000	1.0000	0.000	1.0000
Left Testis	103.484 (2.730)	101.440 (2.900)	101.868 (2.809)	2.813	102.305 (1.623)	2.747	0.000	1.0000	0.000	1.0000
Paired Testes	103.825 (2.743)	101.510 (2.910)	101.514 (2.332)	2.662	102.219 (1.516)	2.569	0.000	1.0000	0.000	1.0000
Paired Epididymides	94.545 (3.269)	98.980 (3.070)	96.029 (3.035)	3.125	96.594 (1.801)	3.230	0.000	1.0000	0.000	1.0000
Entire Prostate	90.559 (6.933)	93.180 (6.070)	82.958 (5.979)	6.327	88.696 (3.629)	7.087	0.000	1.0000	0.000	1.0000
SVCGF	86.625 (7.230)	103.340 (8.080)	82.787 (5.201)	6.837	89.077 (4.632)	9.006	4.501	0.3651	65.838	0.3325
Accessory Sex Gland	88.407 (5.511)	98.850 (6.030)	82.842 (4.687)	5.409	89.129 (3.715)	7.219	3.563	0.3574	65.867	0.3215
Thyroid Glands	114.156 (11.471)	105.760 (5.390)	99.472 (5.830)	7.563	104.064 (3.741)	6.227	0.000	1.0000	0.000	1.0000
Adj Liver	97.305 (2.950)	103.430 (3.220)	109.488 (3.243)	3.138	103.259 (2.900)	4.864	3.923	0.2253	125.030	0.0814
Adj Right Testis	119.060 (3.693)	108.630 (3.760)	112.578 (3.266)	3.573	113.410 (2.321)	3.544	1.870	0.4028	52.334	0.3876
Adj Left Testis	118.353 (3.803)	108.470 (3.650)	113.241 (3.082)	3.512	113.224 (2.024)	3.096	0.502	0.4916	14.286	0.4915

Endpoint	Linuron (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Adj Paired Testes	118.704 (3.676)	108.550 (3.660)	112.908 (3.266)	3.534	113.345 (2.247)	3.433	1.654	0.4185	46.810	0.4082
Adj Paired Epididymides	107.947 (4.449)	105.980 (4.210)	107.051 (4.012)	4.224	106.961 (2.432)	3.938	0.000	1.0000	0.000	1.0000
Adj Entire Prostate	102.704 (8.436)	100.070 (6.780)	92.772 (6.649)	7.288	97.878 (4.137)	7.321	0.000	1.0000	0.000	1.0000
Adj SVCGF	98.105 (8.474)	110.200 (8.000)	92.129 (5.975)	7.483	98.704 (4.455)	7.817	2.604	0.4436	34.800	0.4387
Adj Accessory Sex Gland	100.195 (6.820)	105.720 (5.210)	92.334 (5.323)	5.785	99.400 (3.584)	6.246	2.497	0.4131	43.172	0.4011
Adj Thyroid Glands	129.691 (13.612)	112.590 (6.070)	110.553 (6.611)	8.765	113.414 (4.248)	6.488	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [$\sqrt{3}$×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 3b. Comparisons Among Laboratories of Ratio of Linuron 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[8]

Endpoint	Linuron (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Liver	87.784 (4.338)	86.760 (3.150)	93.083 (4.370)	3.953	88.629 (2.202)	4.303	0.000	1.0000	0.000	1.0000
Right Testis	100.590 (2.840)	99.390 (2.980)	95.276 (2.283)	2.701	97.894 (1.528)	2.703	0.000	1.0000	0.000	1.0000
Left Testis	99.995 (2.683)	98.750 (2.860)	94.745 (2.333)	2.625	97.485 (1.499)	2.664	0.000	1.0000	0.000	1.0000
Paired Testes	100.290 (2.695)	99.070 (2.870)	95.011 (2.275)	2.613	97.707 (1.487)	2.636	0.000	1.0000	0.000	1.0000
Paired Epididymides	92.895 (3.242)	91.400 (2.960)	89.661 (2.940)	3.048	91.219 (1.754)	3.331	0.000	1.0000	0.000	1.0000
Entire Prostate	83.941 (6.709)	86.020 (5.850)	70.167 (5.621)	6.060	79.636 (4.287)	9.325	4.333	0.3305	71.493	0.2786
SVCGF	86.441 (7.224)	87.990 (7.230)	71.731 (4.930)	6.461	80.433 (4.678)	10.074	5.065	0.3028	78.395	0.2300
Accessory Sex Gland	85.309 (5.428)	87.370 (5.810)	71.233 (4.431)	5.223	80.542 (4.374)	9.406	5.514	0.2476	105.581	0.1239
Thyroid Glands	105.459 (11.212)	99.760 (5.400)	98.516 (5.802)	7.471	99.876 (3.728)	6.465	0.000	1.0000	0.000	1.0000
Adj Liver	102.832 (3.033)	99.940 (3.170)	112.244 (3.287)	3.163	104.915 (2.979)	4.919	4.078	0.2264	128.898	0.0838
Adj Right Testis	117.446 (3.663)	114.530 (3.870)	115.522 (3.290)	3.608	115.852 (2.069)	3.093	0.000	1.0000	0.000	1.0000
Adj Left Testis	116.859 (3.775)	113.810 (3.750)	114.869 (4.101)	3.875	115.190 (2.232)	3.356	0.000	1.0000	0.000	1.0000

Endpoint	Linuron (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Adj Paired Testes	117.151 (3.648)	114.170 (3.760)	115.196 (3.285)	3.564	115.508 (2.047)	3.070	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides	108.374 (4.458)	105.310 (4.200)	108.590 (4.043)	4.234	107.420 (2.438)	3.932	0.000	1.0000	0.000	1.0000
Adj Entire Prostate	96.795 (8.191)	99.430 (6.760)	84.542 (6.383)	7.111	92.874 (4.160)	7.758	1.703	0.4701	23.948	0.4688
Adj SVCGF	99.975 (8.553)	101.230 (8.150)	86.055 (5.797)	7.500	93.584 (4.471)	8.274	2.781	0.4327	37.079	0.4258
Adj Accessory Sex Gland	98.530 (6.764)	100.750 (5.270)	85.574 (5.148)	5.727	94.503 (4.167)	7.638	4.479	0.3066	78.202	0.2361
Adj Thyroid Glands	122.470 (13.373)	115.110 (6.250)	118.374 (6.873)	8.832	117.216 (4.370)	6.457	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 3c. Comparisons Among Laboratories of Ratio of Linuron 150 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[8]

Endpoint	Linuron (150 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Liver	87.101 (4.323)	84.180 (3.180)	89.117 (4.369)	3.957	86.206 (2.210)	4.440	0.000	1.0000	0.000	1.0000
Right Testis	99.310 (2.822)	97.820 (3.010)	98.139 (2.358)	2.730	98.408 (1.551)	2.730	0.000	1.0000	0.000	1.0000
Left Testis	99.183 (2.672)	97.600 (2.890)	96.795 (1.944)	2.502	97.617 (1.381)	2.450	0.000	1.0000	0.000	1.0000
Paired Testes	99.246 (2.681)	97.710 (2.900)	97.469 (2.347)	2.643	98.097 (1.508)	2.663	0.000	1.0000	0.000	1.0000
Paired Epididymides	86.479 (3.140)	90.780 (3.010)	94.063 (3.062)	3.071	90.510 (1.772)	3.391	0.000	1.0000	0.000	1.0000
Entire Prostate	73.991 (6.393)	77.750 (5.740)	76.330 (5.918)	6.017	76.160 (3.463)	7.876	0.000	1.0000	0.000	1.0000
SVCGF	70.685 (6.692)	77.030 (5.840)	76.494 (5.156)	5.896	75.217 (3.347)	7.708	0.000	1.0000	0.000	1.0000
Accessory Sex Gland	72.183 (5.093)	77.350 (5.540)	76.442 (4.644)	5.092	75.296 (2.917)	6.710	0.000	1.0000	0.000	1.0000
Thyroid Glands	90.170 (10.796)	102.850 (5.450)	96.577 (5.851)	7.366	98.763 (3.741)	6.561	0.000	1.0000	0.000	1.0000
Adj Liver	109.276 (3.132)	103.860 (3.280)	112.162 (3.338)	3.250	108.418 (1.928)	3.080	0.782	0.4741	24.047	0.4731
Adj Right Testis	124.393 (3.790)	121.080 (4.060)	123.251 (3.435)	3.762	123.008 (2.157)	3.037	0.000	1.0000	0.000	1.0000
Adj Left Testis	124.278 (3.915)	120.860 (3.940)	121.547 (2.863)	3.573	122.079 (1.994)	2.828	0.000	1.0000	0.000	1.0000
Adj Paired Testes	124.335 (3.779)	120.970 (3.940)	122.402 (3.427)	3.716	122.598 (2.134)	3.015	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides	108.077 (4.452)	112.460 (4.420)	118.141 (4.302)	4.391	113.011 (2.534)	3.884	0.000	1.0000	0.000	1.0000
Adj Entire Prostate	91.078 (7.960)	95.790 (6.760)	95.662 (6.870)	7.197	94.480 (4.122)	7.557	0.000	1.0000	0.000	1.0000
Adj SVCGF	88.013 (8.058)	94.710 (6.370)	95.383 (6.185)	6.871	93.417 (3.887)	7.207	0.000	1.0000	0.000	1.0000

Endpoint	Linuron (150 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Adj Accessory Sex Gland	89.406 (6.463)	95.190 (5.200)	95.472 (5.508)	5.723	93.814 (3.264)	6.025	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands	113.372 (13.085)	126.710 (6.530)	120.993 (7.062)	8.892	122.808 (4.502)	6.349	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 4a. Comparisons Among Laboratories of Ratio of Phenobarbital 25 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[8]

Endpoint	Phenobarbital (25 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Liver	129.396 (5.331)	116.000 (4.270)	121.730 (5.040)	4.880	121.615 (3.168)	4.511	2.593	0.3920	53.131	0.3726
Right Testis	101.227 (2.849)	99.920 (2.990)	98.628 (2.559)	2.799	99.826 (1.606)	2.786	0.000	1.0000	0.000	1.0000
Left Testis	101.637 (2.705)	100.350 (2.880)	99.213 (2.271)	2.618	100.251 (1.489)	2.572	0.000	1.0000	0.000	1.0000
Paired Testes	101.434 (2.710)	100.130 (2.890)	98.920 (2.528)	2.709	100.101 (1.557)	2.694	0.000	1.0000	0.000	1.0000
Paired Epididymides	97.303 (3.314)	102.060 (3.120)	99.457 (3.088)	3.174	99.695 (1.830)	3.179	0.000	1.0000	0.000	1.0000
Entire Prostate	102.961 (7.376)	102.520 (6.350)	104.195 (6.646)	6.790	103.219 (3.898)	6.540	0.000	1.0000	0.000	1.0000
SVCGF	102.275 (7.817)	108.910 (5.890)	95.302 (5.534)	6.414	101.844 (3.755)	6.386	1.878	0.4540	29.274	0.4509
Accessory Sex Gland	102.586 (5.916)	106.080 (6.250)	98.136 (5.057)	5.741	101.679 (3.274)	5.577	0.000	1.0000	0.000	1.0000
Thyroid Glands	131.212 (10.124)	119.490 (7.760)	129.716 (6.769)	8.218	126.495 (4.556)	6.238	0.000	1.0000	0.000	1.0000
Adj Liver	126.684 (3.413)	115.360 (3.420)	122.031 (3.450)	3.428	121.360 (2.692)	3.842	3.161	0.2861	92.231	0.1993
Adj Right Testis	99.131 (3.344)	99.670 (3.590)	98.984 (2.583)	3.172	99.193 (1.776)	3.102	0.000	1.0000	0.000	1.0000
Adj Left Testis	99.552 (3.463)	100.140 (3.500)	99.525 (2.348)	3.104	99.676 (1.699)	2.952	0.000	1.0000	0.000	1.0000
Adj Paired Testes	99.343 (3.339)	99.910 (3.500)	99.253 (2.534)	3.124	99.442 (1.749)	3.045	0.000	1.0000	0.000	1.0000

Endpoint	Phenobarbital (25 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Adj Paired Epididymides	95.362 (4.178)	101.980 (4.130)	100.041 (3.874)	4.061	99.195 (2.340)	4.087	0.000	1.0000	0.000	1.0000
Adj Entire Prostate	100.376 (8.338)	102.150 (6.850)	104.398 (7.047)	7.412	102.504 (4.232)	7.151	0.000	1.0000	0.000	1.0000
Adj SVCGF	100.722 (8.585)	108.730 (5.680)	95.695 (6.082)	6.782	102.288 (3.747)	6.346	0.457	0.4972	6.734	0.4972
Adj Accessory Sex Gland	100.565 (6.833)	105.820 (6.110)	98.464 (5.489)	6.144	101.437 (3.505)	5.985	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands	128.195 (10.286)	118.310 (8.210)	130.920 (7.307)	8.601	125.972 (4.821)	6.629	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [$\sqrt{3}$×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 4b. Comparisons Among Laboratories of Ratio of Phenobarbital 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[8]

Endpoint	Phenobarbital (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Liver	134.548 (5.465)	127.470 (4.430)	127.826 (5.192)	5.029	129.529 (2.868)	3.836	0.000	1.0000	0.000	1.0000
Right Testis	103.277 (2.879)	101.120 (3.010)	99.939 (2.570)	2.819	101.333 (1.617)	2.764	0.000	1.0000	0.000	1.0000
Left Testis	100.582 (2.691)	100.800 (2.890)	92.664 (5.777)	3.786	99.848 (1.864)	3.233	0.000	1.0000	0.000	1.0000
Paired Testes	101.920 (2.717)	100.960 (2.900)	96.312 (2.507)	2.708	99.486 (1.555)	2.707	0.000	1.0000	0.000	1.0000
Paired Epididymides	101.001 (3.376)	102.450 (3.130)	101.626 (3.121)	3.209	101.726 (1.849)	3.149	0.000	1.0000	0.000	1.0000
Entire Prostate	109.046 (7.603)	105.310 (6.440)	112.466 (6.925)	6.989	108.744 (4.008)	6.383	0.000	1.0000	0.000	1.0000
SVCGF	105.491 (7.944)	110.070 (6.880)	102.301 (5.731)	6.852	105.485 (3.851)	6.324	0.000	1.0000	0.000	1.0000
Accessory Sex Gland	107.101 (6.050)	108.520 (6.430)	105.539 (5.247)	5.909	106.846 (3.374)	5.470	0.000	1.0000	0.000	1.0000
Thyroid Glands	123.194 (9.810)	132.990 (8.180)	131.003 (6.812)	8.267	129.905 (4.618)	6.157	0.000	1.0000	0.000	1.0000
Adj Liver	133.164 (3.521)	127.470 (3.630)	130.429 (3.594)	3.582	130.412 (2.067)	2.746	0.000	1.0000	0.000	1.0000
Adj Right Testis	102.064 (3.393)	101.430 (3.630)	101.826 (2.608)	3.211	101.796 (1.797)	3.057	0.000	1.0000	0.000	1.0000
Adj Left Testis	99.417 (3.461)	101.140 (3.520)	94.315 (5.796)	4.259	99.351 (2.271)	3.958	0.000	1.0000	0.000	1.0000

Endpoint	Phenobarbital (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Adj Paired Testes	100.731 (3.362)	101.280 (3.530)	98.081 (2.523)	3.138	99.588 (1.752)	3.047	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides	99.892 (4.273)	102.700 (4.150)	103.762 (3.947)	4.123	102.216 (2.377)	4.027	0.000	1.0000	0.000	1.0000
Adj Entire Prostate	107.269 (8.631)	105.800 (6.980)	115.042 (7.430)	7.680	109.394 (4.383)	6.939	0.000	1.0000	0.000	1.0000
Adj SVCGF	104.673 (8.756)	111.640 (7.220)	104.185 (6.346)	7.441	106.803 (4.186)	6.789	0.000	1.0000	0.000	1.0000
Adj Accessory Sex Gland	105.853 (7.016)	109.900 (6.330)	107.639 (5.746)	6.364	107.906 (3.638)	5.840	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands	121.560 (9.996)	134.260 (8.720)	133.451 (7.396)	8.704	130.836 (4.912)	6.503	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [$\sqrt{3}$×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 4c. Comparisons Among Laboratories of Ratio of Phenobarbital 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[8]

Endpoint	Phenobarbital (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Liver	145.395 (5.818)	118.070 (4.500)	139.074 (5.546)	5.288	133.692 (6.891)	8.928	10.697	0.1575	202.297	0.0017**
Right Testis	103.396 (2.930)	101.860 (3.130)	96.039 (2.601)	2.887	100.130 (1.956)	3.384	1.799	0.3570	62.323	0.3217
Left Testis	103.204 (2.773)	102.160 (3.020)	96.899 (2.275)	2.689	100.291 (1.755)	3.031	1.487	0.3757	55.282	0.3497
Paired Testes	103.299 (2.782)	102.010 (3.020)	96.468 (2.570)	2.791	100.359 (1.822)	3.145	1.499	0.3850	53.705	0.3632
Paired Epididymides	102.591 (3.462)	102.410 (3.240)	102.006 (3.182)	3.294	102.321 (1.898)	3.213	0.000	1.0000	0.000	1.0000
Entire Prostate	94.225 (7.193)	107.250 (6.740)	104.957 (6.783)	6.905	102.469 (3.982)	6.731	0.000	1.0000	0.000	1.0000
SVCGF	99.966 (7.864)	119.500 (11.230)	105.493 (5.921)	8.338	105.905 (4.359)	7.129	0.000	1.0000	0.000	1.0000
Accessory Sex Gland	97.365 (5.868)	114.080 (6.720)	105.322 (5.330)	5.972	104.916 (3.483)	5.750	1.273	0.4811	21.317	0.4805
Thyroid Glands	124.519 (10.039)	133.660 (8.670)	129.229 (6.843)	8.518	129.503 (4.736)	6.335	0.000	1.0000	0.000	1.0000
Adj Liver	153.712 (3.919)	131.650 (3.810)	145.885 (3.912)	3.880	143.709 (5.285)	6.369	8.290	0.1570	213.662	0.0016**
Adj Right Testis	109.061 (3.571)	114.050 (3.990)	101.011 (2.661)	3.407	107.380 (3.193)	5.151	4.373	0.2153	128.333	0.0609
Adj Left Testis	109.040 (3.690)	114.360 (3.890)	101.899 (2.469)	3.350	107.705 (3.068)	4.933	4.148	0.2192	123.840	0.0677

Endpoint	Phenobarbital (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Adj Paired Testes	109.050 (3.561)	114.210 (3.880)	101.454 (2.611)	3.351	107.598 (3.119)	5.021	4.255	0.2169	126.984	0.0639
Adj Paired Epididymides	108.355 (4.530)	114.870 (4.550)	107.526 (4.088)	4.389	110.045 (2.525)	3.974	0.000	1.0000	0.000	1.0000
Adj Entire Prostate	98.998 (8.429)	120.180 (7.730)	110.247 (7.371)	7.843	110.369 (4.683)	7.349	2.186	0.4669	27.866	0.4653
Adj SVCGF	104.635 (8.903)	134.150 (12.990)	110.574 (6.656)	9.516	112.150 (4.932)	7.616	0.000	1.0000	0.000	1.0000
Adj Accessory Sex Gland	102.073 (7.004)	127.970 (6.750)	110.470 (5.921)	6.558	113.514 (6.029)	9.199	8.133	0.2363	124.018	0.1022
Adj Thyroid Glands	130.770 (10.562)	150.730 (9.540)	135.448 (7.561)	9.221	138.812 (5.168)	6.448	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 5a. Comparisons Among Laboratories of Ratio of Linuron 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Hormonal Analysis Endpoints

Endpoint	Linuron (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err[2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Testosterone	48.618 (13.330)	120.490 (29.650)	102.929 (26.081)	23.020	81.903 (19.207)	40.619	24.740	0.2365	107.470	0.1004
LH	80.502 (6.420)	103.720 (8.440)	126.923 (22.858)	12.573	95.358 (8.715)	15.829	10.836	0.2654	86.184	0.1051
TSH	75.172 (10.686)	99.880 (19.720)	90.602 (13.321)	14.576	84.044 (7.678)	15.824	0.000	1.0000	0.000	1.0000
T4	65.520 (3.816)	69.730 (3.570)	54.021 (5.181)	4.189	64.119 (3.427)	9.257	4.261	0.2881	101.725	0.2002
T3	83.520 (4.138)	97.870 (4.670)	98.292 (5.012)	4.607	92.806 (4.117)	7.683	5.450	0.2292	118.311	0.0888
FSH	96.366 (5.338)	104.460 (7.470)	119.797 (6.468)	6.426	106.458 (5.860)	9.535	7.885	0.2205	122.719	0.0714
Estradiol	130.686 (9.911)	135.310 (12.420)	115.044 (53.161)	25.164	132.122 (7.666)	10.049	0.000	1.0000	0.000	1.0000
Prolactin	12.981 (3.874)	81.510 (37.750)	106.097 (30.636)	24.087	15.135 (3.824)	43.761	0.000	1.0000	0.000	1.0000
DHT	70.935 (12.045)	120.850 (20.280)	71.384 (23.108)	18.478	84.471 (12.324)	25.270	12.366	0.3551	66.923	0.3169

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 5b. Comparisons Among Laboratories of Ratio of Linuron 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Hormonal Analysis Endpoints

Endpoint	Linuron (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err[2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Testosterone	40.099 (13.124)	75.920 (21.010)	101.432 (25.940)	20.025	64.691 (14.725)	39.426	16.865	0.2938	84.223	0.2086
LH	81.818 (7.882)	102.160 (8.830)	96.154 (19.625)	12.112	91.643 (6.463)	12.215	4.865	0.3914	40.170	0.3699
TSH	93.259 (16.670)	64.240 (12.010)	79.558 (12.435)	13.705	76.211 (7.670)	17.431	0.000	1.0000	0.000	1.0000
T4	38.469 (3.420)	46.950 (3.240)	25.871 (4.267)	3.642	37.479 (4.902)	22.656	7.667	0.1638	210.506	0.0034**
T3	80.366 (4.074)	87.860 (4.440)	87.349 (4.746)	4.420	84.808 (2.537)	5.181	0.000	1.0000	0.000	1.0000
FSH	105.507 (5.487)	118.410 (7.720)	126.193 (10.029)	7.745	113.689 (4.957)	7.551	4.491	0.3672	57.988	0.3350
Estradiol	161.184 (11.424)	138.340 (12.510)	148.618 (60.501)	28.145	150.754 (8.355)	9.599	0.000	1.0000	0.000	1.0000
Prolactin	15.259 (4.383)	58.520 (24.590)	68.421 (30.135)	19.703	17.633 (4.271)	41.957	0.000	1.0000	0.000	1.0000
DHT	73.353 (17.149)	86.880 (15.930)	93.721 (25.454)	19.511	82.891 (10.609)	22.169	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 5c. Comparisons Among Laboratories of Ratio of Linuron 150 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Hormonal Analysis Endpoints

	Linuron (150 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err[2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Testosterone	33.004 (8.970)	73.680 (20.870)	66.845 (23.747)	17.863	48.147 (11.195)	40.275	11.373	0.3476	63.672	0.3169
LH	85.614 (9.600)	110.020 (9.170)	90.659 (19.465)	12.745	97.174 (7.553)	13.464	6.627	0.3556	51.993	0.3160
TSH	79.905 (12.208)	74.730 (15.260)	73.965 (12.237)	13.235	76.405 (7.520)	17.048	0.000	1.0000	0.000	1.0000
T4	32.506 (3.356)	30.370 (3.160)	10.484 (1.364)	2.627	24.045 (5.875)	42.319	9.799	0.1221	373.036	<0.0001**
T3	78.955 (4.046)	89.090 (4.560)	83.686 (4.758)	4.455	83.496 (2.554)	5.297	0.000	1.0000	0.000	1.0000
FSH	107.443 (5.540)	114.640 (7.850)	115.736 (7.675)	7.022	111.358 (3.899)	6.064	0.000	1.0000	0.000	1.0000
Estradiol	148.563 (10.905)	187.410 (14.370)	137.005 (58.801)	28.025	164.280 (12.888)	13.589	13.812	0.2904	49.285	0.1838
Prolactin	5.536 (1.367)	117.720 (51.830)	54.868 (32.352)	28.517	5.702 (1.366)	41.485	0.000	1.0000	0.000	1.0000
DHT	61.481 (12.043)	92.340 (16.780)	75.691 (24.093)	17.639	72.499 (9.065)	21.657	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 6a. Comparisons Among Laboratories of Ratio of Phenobarbital 25 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Hormonal Analysis Endpoints

Endpoint	Phenobarbital (25 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err[2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Testosterone	61.164 (16.798)	77.700 (23.200)	72.911 (15.747)	18.582	69.442 (10.295)	25.679	0.000	1.0000	0.000	1.0000
LH	83.134 (6.621)	99.820 (8.400)	93.269 (19.344)	11.455	89.959 (5.265)	10.136	2.334	0.4635	20.375	0.4614
TSH	178.307 (31.733)	117.050 (22.300)	155.998 (25.109)	26.381	143.762 (14.760)	17.783	0.000	1.0000	0.000	1.0000
T4	79.306 (4.073)	83.880 (3.820)	83.110 (6.916)	4.936	81.931 (2.585)	5.464	0.000	1.0000	0.000	1.0000
T3	79.419 (4.055)	89.380 (4.480)	90.058 (4.811)	4.449	85.807 (2.963)	5.982	2.597	0.3709	58.377	0.3427
FSH	93.480 (5.167)	78.030 (3.880)	84.010 (5.073)	4.707	84.451 (3.712)	7.613	4.424	0.2787	93.993	0.1831
Estradiol	132.694 (10.007)	114.730 (10.980)	108.343 (49.798)	23.595	124.193 (7.316)	10.203	0.000	1.0000	0.000	1.0000
Prolactin	38.518 (13.417)	59.510 (24.600)	78.019 (24.203)	20.740	49.974 (10.591)	36.709	0.000	1.0000	0.000	1.0000
DHT	79.885 (15.327)	87.290 (16.410)	62.776 (15.238)	15.658	76.123 (9.025)	20.535	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 6b. Comparisons Among Laboratories of Ratio of Phenobarbital 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Hormonal Analysis Endpoints

	Phenobarbital (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err[2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Testosterone	35.202 (9.854)	78.230 (23.290)	57.053 (14.092)	15.746	47.354 (8.414)	30.776	5.330	0.4403	33.849	0.4339
LH	65.871 (4.572)	96.520 (8.420)	80.769 (18.184)	10.392	79.475 (8.910)	19.419	12.154	0.1907	116.952	0.0180*
TSH	196.579 (29.597)	143.550 (27.910)	167.822 (26.019)	27.842	168.249 (16.008)	16.479	0.000	1.0000	0.000	1.0000
T4	77.009 (4.028)	77.890 (3.710)	58.177 (3.909)	3.882	71.045 (5.247)	12.791	8.216	0.1585	211.630	0.0020**
T3	80.105 (4.069)	83.720 (4.350)	78.473 (4.544)	4.321	80.798 (2.487)	5.331	0.000	1.0000	0.000	1.0000
FSH	84.142 (4.933)	81.730 (3.970)	85.178 (6.396)	5.100	83.152 (2.784)	5.800	0.000	1.0000	0.000	1.0000
Estradiol	143.782 (10.548)	133.200 (11.520)	127.331 (55.424)	25.831	138.731 (7.704)	9.619	0.000	1.0000	0.000	1.0000
Prolactin	22.263 (5.993)	43.700 (18.060)	82.663 (24.560)	16.204	27.356 (5.541)	35.085	0.000	1.0000	0.000	1.0000
DHT	61.793 (13.392)	89.060 (16.620)	56.287 (14.434)	14.815	66.957 (8.453)	21.865	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 6c. Comparisons Among Laboratories of Ratio of Phenobarbital 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Hormonal Analysis Endpoints

	Phenobarbital (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
Endpoint	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[8])
Testosterone	22.076 (5.716)	78.750 (24.140)	46.763 (13.477)	14.444	38.819 (11.059)	49.342	14.167	0.2762	98.077	0.1560
LH	71.396 (5.215)	92.280 (8.430)	112.294 (21.604)	11.750	83.660 (7.758)	16.063	9.452	0.2716	80.445	0.1270
TSH	227.017 (33.909)	142.370 (30.900)	206.954 (29.649)	31.486	190.820 (20.540)	18.644	16.758	0.3942	53.224	0.3757
T4	55.459 (3.748)	53.540 (3.520)	34.326 (4.360)	3.876	48.061 (5.381)	19.391	8.473	0.1616	218.605	0.0029**
T3	68.763 (3.946)	66.430 (4.220)	71.363 (4.494)	4.220	68.749 (2.426)	6.113	0.000	1.0000	0.000	1.0000
FSH	82.921 (5.006)	76.710 (4.000)	89.358 (6.116)	5.041	81.285 (2.819)	6.006	0.734	0.4894	14.560	0.4892
Estradiol	151.645 (11.058)	142.450 (12.360)	233.617 (86.303)	36.574	148.335 (8.204)	9.579	0.000	1.0000	0.000	1.0000
Prolactin	11.557 (4.598)	40.600 (18.630)	75.301 (25.975)	16.401	15.004 (4.399)	50.782	0.000	1.0000	0.000	1.0000
DHT	51.021 (9.188)	94.010 (17.660)	44.135 (13.362)	13.403	58.076 (9.473)	28.253	10.313	0.3806	76.947	0.3715

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).

Table 7. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Linuron Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption[5]

Endpoint	Among Laboratories Main Effect Variance Component[1]	P-value[2]	P-value, Based On Likelihood Ratio[6]	Among Laboratories by Dose Interaction Variance Component[3]	P-value[4]	P-value, Based On Likelihood Ratio[7]
Body Weight Change (TD8-TD1)	212.43	0.1846	0.0665	21.4663	0.4125	0.4032
Body Weight Change (TD15-TD8)	47.4756	0.3030	0.2677	33.0512	0.3563	0.3328
Body Weight Change (TD15-TD1)	71.5580	0.1946	0.0911	21.4862	0.3259	0.2908
Final Body Weight	0.8390	0.3128	0.2479	0	1.0000	1.0000
Food Consumption (TD8-TD1)	8.8671	0.2109	0.0873	0	1.0000	1.0000
Food Consumption (TD15-TD8)	17.8476	0.1918	0.0527	0	1.0000	1.0000
Food Consumption (TD15-TD1)	6.7857	0.2388	0.0926	0	1.0000	1.0000

1. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
2. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
3. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories by dose interaction variance component. Significance is indicated by "*" for the 0.05 level.
5. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
6. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.
7. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 8. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Phenobarbital Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Body Weight Changes, Final Body Weight, and Food Consumption[5]

Endpoint	Among Laboratories Main Effect Variance Component[1]	P-value[2]	P-value Based On Likelihood Ratio[6]	Among Laboratories by Dose Interaction Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[7]
Body Weight Change (TD8-TD1)	0	1.0000	1.0000	62.6647	0.1846	0.1678
Body Weight Change (TD15-TD8)	0	1.0000	1.0000	0	1.0000	1.0000
Body Weight Change (TD15-TD1)	0	1.0000	1.0000	28.7604	0.2783	0.2527
Final Body Weight	0	1.0000	1.0000	0	1.0000	1.0000
Food Consumption (TD8-TD1)	4.3526	0.2441	0.1690	0	1.0000	1.0000
Food Consumption (TD15-TD8)	0	1.0000	1.0000	0	1.0000	1.0000
Food Consumption (TD15-TD1)	0.3111	0.4258	0.4171	0	1.0000	1.0000

1. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
2. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
3. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories by dose interaction variance component. Significance is indicated by "*" for the 0.05 level.
5. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
6. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
7. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 9. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Linuron Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[5,6]

Endpoint	Among Laboratories Main Effect Variance Component[1]	P-value[2]	P-value Based On Likelihood Ratio[7]	Among Laboratories by Dose Interaction Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[8]
Liver	0.2211	0.4852	0.4849	0	1.0000	1.0000
Paired Epididymides	0	1.0000	1.0000	0	1.0000	1.0000
Left Testis	0	1.0000	1.0000	0	1.0000	1.0000
Right Testis	0	1.0000	1.0000	0	1.0000	1.0000
Paired Testes	0	1.0000	1.0000	0	1.0000	1.0000
Entire Prostate	4.3652	0.3687	0.3394	0	1.0000	1.0000
SVCGF	9.1906	0.3032	0.2320	0	1.0000	1.0000
Accessory Sex Gland	10.8089	0.2523	0.1708	0	1.0000	1.0000
Thyroid Glands	0	1.0000	1.0000	0	1.0000	1.0000
Adj Liver	12.9270	0.1677	0.0361*	0	1.0000	1.0000
Adj Paired Epididymides	0	1.0000	1.0000	0	1.0000	1.0000
Adj Left Testis	0	1.0000	1.0000	0	1.0000	1.0000
Adj Right Testis	0.1002	0.4901	0.4900	0	1.0000	1.0000
Adj Paired Testes	0.2150	0.4786	0.4779	0	1.0000	1.0000
Adj Entire Prostate	0	1.0000	1.0000	0	1.0000	1.0000
Adj SVCGF	3.6835	0.4075	0.3936	0	1.0000	1.0000
Adj Accessory Sex Gland	7.6296	0.2922	0.2139	0	1.0000	1.0000
Adj Thyroid Glands	0	1.0000	1.0000	0	1.0000	1.0000

1. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
2. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
3. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories by dose interaction variance component indicated by "*" for the 0.05 level.
5. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
6. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
7. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
8. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 10. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Phenobarbital Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Unadjusted and Adjusted Organ Weights[5,6]

Endpoint	Among Laboratories Main Effect Variance Component[1]	P-value[2]	P-value Based On Likelihood Ratio[7]	Among Laboratories by Dose Interaction Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[8]
Liver	32.4932	0.1696	0.0510	2.1279	0.4438	0.4406
Paired Epididymides	0	1.0000	1.0000	0	1.0000	1.0000
Left Testis	1.2307	0.3505	0.3115	0	1.0000	1.0000
Right Testis	1.0146	0.3608	0.3274	0	1.0000	1.0000
Paired Testes	2.3652	0.2675	0.1634	0	1.0000	1.0000
Entire Prostate	0	1.0000	1.0000	0	1.0000	1.0000
SVCGF	6.6638	0.3590	0.3236	0	1.0000	1.0000
Accessory Sex Gland	0	1.0000	1.0000	0	1.0000	1.0000
Thyroid Glands	0	1.0000	1.0000	0	1.0000	1.0000
Adj Liver	22.3784	0.1524	0.0177*	0	1.0000	1.0000
Adj Paired Epididymides	0	1.0000	1.0000	0	1.0000	1.0000
Adj Left Testis	2.4189	0.3151	0.2561	0	1.0000	1.0000
Adj Right Testis	0	1.0000	1.0000	0	1.0000	1.0000
Adj Paired Testes	1.8660	0.3211	0.2618	0	1.0000	1.0000
Adj Entire Prostate	0	1.0000	1.0000	0	1.0000	1.0000
Adj SVCGF	12.0648	0.3155	0.2508	0	1.0000	1.0000
Adj Accessory Sex Gland	8.4783	0.3224	0.2643	0	1.0000	1.0000
Adj Thyroid Glands	0	1.0000	1.0000	0	1.0000	1.0000

1. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
2. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
3. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories by dose interaction variance component indicated by "*" for the 0.05 level.
5. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
6. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
7. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
8. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 11. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Linuron Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Hormonal Analysis Endpoints[5]

Endpoint	Among Laboratories Main Effect Variance Component[1]	P-value[2]	P-value Based On Likelihood Ratio[6]	Among Laboratories by Dose Interaction Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[7]
Testosterone	423.07	0.1610	0.0085**	0	1.0000	1.0000
LH	89.5965	0.1809	0.0219*	0	1.0000	1.0000
TSH	0	1.0000	1.0000	0	1.0000	1.0000
T4	79.1721	0.1194	0.0013**	0	1.0000	1.0000
T3	15.7074	0.1892	0.0344*	0	1.0000	1.0000
FSH	40.0078	0.1852	0.0345*	0	1.0000	1.0000
Estradiol	0	1.0000	1.0000	0	1.0000	1.0000
Prolactin	0	1.0000	1.0000	0	1.0000	1.0000
DHT	86.6113	0.2728	0.1709	0	1.0000	1.0000

1. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
2. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
3. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories by dose interaction variance component. Significance is indicated by "*" for the 0.05 level.
5. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
6. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
7. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 12. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Phenobarbital Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Hormonal Analysis Endpoints[5]

Endpoint	Among Laboratories Main Effect Variance Component[1]	P-value[2]	P-value Based On Likelihood Ratio[6]	Among Laboratories by Dose Interaction Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[7]
Testosterone	170.05	0.2110	0.0426*	0	1.0000	1.0000
LH	102.21	0.1609	0.0102*	0	1.0000	1.0000
TSH	432.96	0.2262	0.0818	0	1.0000	1.0000
T4	48.8079	0.1403	0.0366*	0	1.0000	1.0000
T3	0	1.0000	1.0000	0	1.0000	1.0000
FSH	7.0468	0.2662	0.1616	0	1.0000	1.0000
Estradiol	0	1.0000	1.0000	0	1.0000	1.0000
Prolactin	389.30	0.1700	0.0037**	0	1.0000	1.0000
DHT	123.19	0.2364	0.1039	0	1.0000	1.0000

1. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
2. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
3. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories by dose interaction variance component. Significance is indicated by "*" for the 0.05 level.
5. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
6. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
7. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 13a. Comparisons Among Laboratories of Ratio of Linuron 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Selected Endpoints.[10] Outliers Excluded.[11]

Endpoint	Linuron (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Body Weight Change (TD15-TD8)	58.606 (10.796)	105.800 (9.700)	95.871 (13.429)	11.309	86.715 (12.127)	24.223	17.714	0.1841	156.639	0.0177*
Food Consumption TD15-TD8	89.016 (3.885)	100.200 (3.200)	100.521 (3.683)	3.589	96.876 (2.951)	5.276	3.649	0.2759	101.681	0.1793
Paired Epididymides	94.545 (3.269)	98.980 (2.990)	96.029 (3.035)	3.098	96.638 (1.785)	3.199	0.000	1.0000	0.000	1.0000
Entire Prostate	90.559 (6.933)	93.180 (5.930)	82.958 (5.979)	6.281	88.771 (3.599)	7.022	0.000	1.0000	0.000	1.0000
Thyroid Glands	114.156 (11.471)	105.760 (5.090)	99.472 (5.830)	7.463	104.157 (3.636)	6.047	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides[8]	107.947 (4.449)	105.980 (4.110)	107.051 (4.012)	4.190	106.945 (2.412)	3.907	0.000	1.0000	0.000	1.0000
Adj Entire Prostate[8]	102.704 (8.436)	100.070 (6.600)	92.772 (6.649)	7.228	97.922 (4.095)	7.244	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands[8]	129.691 (13.612)	112.590 (5.850)	110.553 (6.611)	8.691	113.385 (4.170)	6.371	0.000	1.0000	0.000	1.0000
Testosterone	48.618 (13.330)	120.490 (29.650)	102.929 (26.081)	23.020	81.903 (19.207)	40.619	24.740	0.2365	107.470	0.1004
TSH	75.172 (10.686)	118.560 (15.500)	90.602 (13.321)	13.169	91.775 (9.882)	18.651	11.162	0.3090	84.755	0.2383

Endpoint	Linuron (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
FSH	96.366 (5.338)	104.460 (7.470)	119.797 (6.468)	6.426	106.458 (5.860)	9.535	7.885	0.2205	122.719	0.0714
Estradiol	130.686 (9.911)	135.310 (12.420)	131.598 (17.066)	13.132	132.334 (7.054)	9.233	0.000	1.0000	0.000	1.0000
Prolactin	12.981 (3.874)	129.420 (39.550)	68.679 (15.492)	19.639	57.012 (24.273)	73.741	36.838	0.1921	187.579	0.0008**
DHT	70.935 (12.045)	112.260 (18.630)	71.384 (23.108)	17.928	82.160 (10.366)	21.854	7.173	0.4291	40.012	0.4211

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
10. Endpoints for which flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
11. Outliers were excluded from the analysis.

Table 13b. Comparisons Among Laboratories of Ratio of Linuron 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Selected Endpoints.[10] Outliers Excluded.[11]

Endpoint	Linuron (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Body Weight Change (TD15-TD8)	52.941 (10.539)	68.100 (11.400)	64.807 (11.552)	11.164	61.439 (6.429)	18.126	0.000	1.0000	0.000	1.0000
Food Consumption TD15-TD8	81.451 (3.814)	100.600 (5.700)	91.192 (3.499)	4.338	90.160 (4.182)	8.034	5.837	0.2300	134.563	0.0800
Paired Epididymides	92.895 (3.242)	91.400 (2.880)	89.661 (2.940)	3.021	91.222 (1.737)	3.298	0.000	1.0000	0.000	1.0000
Entire Prostate	83.941 (6.709)	83.320 (5.760)	70.167 (5.621)	6.030	78.637 (3.831)	8.438	2.858	0.4062	47.387	0.3920
Thyroid Glands	105.459 (11.212)	99.760 (5.080)	98.516 (5.802)	7.365	99.869 (3.618)	6.274	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides[8]	108.374 (4.458)	105.310 (4.100)	108.590 (4.043)	4.200	107.385 (2.418)	3.901	0.000	1.0000	0.000	1.0000
Adj Entire Prostate[8]	96.795 (8.191)	96.300 (6.600)	84.542 (6.383)	7.058	91.794 (4.003)	7.553	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands[8]	122.470 (13.373)	115.110 (6.020)	118.374 (6.873)	8.755	117.138 (4.289)	6.342	0.000	1.0000	0.000	1.0000
Testosterone	40.099 (13.124)	75.920 (21.010)	101.432 (25.940)	20.025	64.691 (14.725)	39.426	16.865	0.2938	84.223	0.2086
TSH	93.259 (16.670)	76.250 (8.720)	79.558 (12.435)	12.608	79.808 (6.563)	14.243	0.000	1.0000	0.000	1.0000

Endpoint	Linuron (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
FSH	105.507 (5.487)	118.410 (7.720)	126.193 (10.029)	7.745	113.689 (4.957)	7.551	4.491	0.3672	57.988	0.3350
Estradiol	161.184 (11.424)	138.340 (12.510)	140.371 (17.550)	13.828	148.839 (7.603)	8.848	0.000	1.0000	0.000	1.0000
Prolactin	15.259 (4.383)	93.390 (21.660)	68.421 (15.842)	13.962	53.805 (19.513)	62.814	30.416	0.1518	217.855	0.0002**
DHT	73.353 (17.149)	86.880 (15.170)	93.721 (25.454)	19.258	83.065 (10.376)	21.635	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
10. Endpoints for which flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
11. Outliers were excluded from the analysis.

Table 13c. Comparisons Among Laboratories of Ratio of Linuron 150 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Selected Endpoints.[10] Outliers Excluded.[11]

Endpoint	Linuron (150 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Body Weight Change (TD15-TD8)	55.773 (10.665)	65.100 (17.200)	40.988 (10.792)	12.886	51.177 (6.941)	23.491	0.000	1.0000	0.000	1.0000
Food Consumption TD15-TD8	78.821 (3.900)	89.900 (5.500)	80.396 (3.494)	4.298	81.561 (2.352)	4.995	0.000	1.0000	0.000	1.0000
Paired Epididymides	86.479 (3.140)	90.780 (2.920)	94.063 (3.062)	3.041	90.516 (1.753)	3.355	0.000	1.0000	0.000	1.0000
Entire Prostate	73.991 (6.393)	77.750 (5.620)	76.330 (5.918)	5.977	76.185 (3.436)	7.813	0.000	1.0000	0.000	1.0000
Thyroid Glands	90.170 (10.796)	99.450 (5.190)	96.577 (5.851)	7.279	97.267 (3.654)	6.506	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides[8]	108.077 (4.452)	112.460 (4.310)	118.141 (4.302)	4.355	113.002 (2.513)	3.852	0.000	1.0000	0.000	1.0000
Adj Entire Prostate[8]	91.078 (7.960)	95.790 (6.580)	95.662 (6.870)	7.137	94.507 (4.080)	7.478	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands[8]	113.372 (13.085)	123.490 (6.340)	120.993 (7.062)	8.829	121.340 (4.438)	6.335	0.000	1.0000	0.000	1.0000
Testosterone	33.004 (8.970)	73.680 (20.870)	66.845 (23.747)	17.863	48.147 (11.195)	40.275	11.373	0.3476	63.672	0.3169

Endpoint	Linuron (150 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
TSH	79.905 (12.208)	80.560 (9.160)	73.965 (12.237)	11.202	78.646 (6.286)	13.844	0.000	1.0000	0.000	1.0000
FSH	107.443 (5.540)	114.640 (7.850)	115.736 (7.675)	7.022	111.358 (3.899)	6.064	0.000	1.0000	0.000	1.0000
Estradiol	148.563 (10.905)	187.410 (14.370)	134.462 (17.562)	14.279	157.511 (12.105)	13.311	15.542	0.2637	108.843	0.1496
Prolactin	5.536 (1.367)	186.910 (50.430)	54.868 (15.605)	22.467	6.044 (1.362)	39.022	0.000	1.0000	0.000	1.0000
DHT	61.481 (12.043)	92.340 (16.010)	75.691 (24.093)	17.382	73.054 (8.938)	21.190	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
10. Endpoints for which flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
11. Outliers were excluded from the analysis.

Table 14a. Comparisons Among Laboratories of Ratio of Phenobarbital 25 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Selected Endpoints.[10] Outliers Excluded.[11]

Endpoint	Phenobarbital (25 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Body Weight Change (TD15-TD8)	109.804 (13.834)	100.300 (10.200)	104.430 (14.016)	12.683	103.847 (7.084)	11.815	0.000	1.0000	0.000	1.0000
Food Consumption TD15-TD8	97.672 (2.909)	99.300 (3.100)	99.757 (2.655)	2.888	98.950 (1.657)	2.901	0.000	1.0000	0.000	1.0000
Paired Epididymides	97.303 (3.314)	102.060 (3.040)	99.457 (3.088)	3.147	99.738 (1.813)	3.149	0.000	1.0000	0.000	1.0000
Entire Prostate	102.961 (7.376)	102.520 (6.210)	104.195 (6.646)	6.744	103.208 (3.865)	6.486	0.000	1.0000	0.000	1.0000
Thyroid Glands	131.212 (10.124)	119.490 (7.760)	129.716 (6.769)	8.218	126.495 (4.556)	6.238	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides[8]	95.362 (4.178)	101.980 (4.030)	100.041 (3.874)	4.027	99.239 (2.322)	4.052	0.000	1.0000	0.000	1.0000
Adj Entire Prostate[8]	100.376 (8.338)	102.150 (6.670)	104.398 (7.047)	7.352	102.497 (4.189)	7.078	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands[8]	128.195 (10.286)	118.310 (8.210)	130.920 (7.307)	8.601	125.972 (4.821)	6.629	0.000	1.0000	0.000	1.0000
Testosterone	61.164 (16.798)	77.700 (21.590)	72.911 (15.747)	18.045	69.686 (10.142)	25.208	0.000	1.0000	0.000	1.0000

Endpoint	Phenobarbital (25 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
TSH	178.307 (31.733)	138.950 (16.670)	155.998 (25.109)	24.504	149.654 (12.723)	14.725	0.000	1.0000	0.000	1.0000
FSH	93.480 (5.167)	75.430 (3.570)	84.010 (5.073)	4.603	83.539 (4.366)	9.051	6.017	0.2132	130.717	0.0576
Estradiol	132.694 (10.007)	114.730 (10.980)	123.934 (16.001)	12.329	124.436 (6.714)	9.345	0.000	1.0000	0.000	1.0000
Prolactin	38.518 (13.417)	94.490 (21.090)	78.019 (24.203)	19.570	65.092 (15.188)	40.414	18.200	0.2623	92.997	0.1523
DHT	79.885 (15.327)	87.290 (16.410)	62.776 (15.238)	15.658	76.123 (9.025)	20.535	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
10. Endpoints for which flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
11. Outliers were excluded from the analysis.

Table 14b. Comparisons Among Laboratories of Ratio of Phenobarbital 50 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Selected Endpoints.[10] Outliers Excluded.[11]

Endpoint	Phenobarbital (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Body Weight Change (TD15-TD8)	109.586 (13.819)	109.300 (10.900)	92.261 (13.190)	12.636	104.329 (7.179)	11.919	0.000	1.0000	0.000	1.0000
Food Consumption TD15-TD8	100.990 (2.906)	104.400 (3.200)	98.957 (2.689)	2.932	101.137 (1.680)	2.877	0.000	1.0000	0.000	1.0000
Paired Epididymides	101.001 (3.376)	101.020 (3.070)	101.626 (3.121)	3.189	101.224 (1.837)	3.143	0.000	1.0000	0.000	1.0000
Entire Prostate	109.046 (7.603)	105.310 (6.300)	112.466 (6.925)	6.943	108.686 (3.973)	6.332	0.000	1.0000	0.000	1.0000
Thyroid Glands	123.194 (9.810)	132.990 (8.180)	131.003 (6.812)	8.267	129.905 (4.618)	6.157	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides[8]	99.892 (4.273)	100.870 (4.080)	103.762 (3.947)	4.100	101.608 (2.363)	4.029	0.000	1.0000	0.000	1.0000
Adj Entire Prostate[8]	107.269 (8.631)	105.800 (6.790)	115.042 (7.430)	7.617	109.316 (4.334)	6.868	0.000	1.0000	0.000	1.0000
Adj Thyroid Glands[8]	121.560 (9.996)	134.260 (8.720)	133.451 (7.396)	8.704	130.836 (4.912)	6.503	0.000	1.0000	0.000	1.0000
Testosterone	35.202 (9.854)	63.480 (19.430)	57.053 (14.092)	14.459	45.487 (7.457)	28.396	0.000	1.0000	0.000	1.0000

Endpoint	Phenobarbital (50 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
TSH	196.579 (29.597)	170.400 (21.500)	167.822 (26.019)	25.705	175.853 (14.461)	14.243	0.000	1.0000	0.000	1.0000
FSH	84.142 (4.933)	81.730 (3.680)	85.178 (6.396)	5.003	83.046 (2.679)	5.587	0.000	1.0000	0.000	1.0000
Estradiol	143.782 (10.548)	133.200 (11.520)	129.579 (16.905)	12.991	137.317 (7.067)	8.914	0.000	1.0000	0.000	1.0000
Prolactin	22.263 (5.993)	69.390 (15.460)	82.663 (24.560)	15.338	51.272 (15.911)	53.751	23.022	0.1823	150.098	0.0102*
DHT	61.793 (13.392)	89.060 (16.620)	56.287 (14.434)	14.815	66.957 (8.453)	21.865	0.000	1.0000	0.000	1.0000

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
10. Endpoints for which flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
11. Outliers were excluded from the analysis.

Table 14c. Comparisons Among Laboratories of Ratio of Phenobarbital 100 mg/kg/day Group Response to Vehicle Group Response in Adult Intact Male Assay for Selected Endpoints.[10] Outliers Excluded.[11]

Endpoint	Phenobarbital (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
Body Weight Change (TD15-TD8)	102.708 (13.583)	87.100 (12.800)	110.776 (14.697)	13.693	99.123 (7.868)	13.748	0.000	1.0000	0.000	1.0000
Food Consumption TD15-TD8	101.699 (2.967)	110.000 (4.400)	105.446 (2.829)	3.399	104.790 (1.856)	3.068	0.000	1.0000	0.000	1.0000
Paired Epididymides	102.591 (3.462)	102.410 (3.150)	102.006 (3.182)	3.264	102.322 (1.880)	3.182	0.000	1.0000	0.000	1.0000
Entire Prostate	94.225 (7.193)	107.250 (6.580)	104.957 (6.783)	6.852	102.549 (3.948)	6.668	0.000	1.0000	0.000	1.0000
Thyroid Glands	124.519 (10.039)	133.660 (8.670)	129.229 (6.843)	8.518	129.503 (4.736)	6.335	0.000	1.0000	0.000	1.0000
Adj Paired Epididymides[8]	108.355 (4.530)	114.870 (4.430)	107.526 (4.088)	4.349	110.125 (2.504)	3.938	0.000	1.0000	0.000	1.0000
Adj Entire Prostate[8]	98.998 (8.429)	120.180 (7.520)	110.247 (7.371)	7.773	110.517 (4.726)	7.407	2.667	0.4511	34.311	0.4475
Adj Thyroid Glands[8]	130.770 (10.562)	150.730 (9.540)	135.448 (7.561)	9.221	138.812 (5.168)	6.448	0.000	1.0000	0.000	1.0000
Testosterone	22.076 (5.716)	78.750 (22.350)	46.763 (13.477)	13.848	40.406 (11.667)	50.010	15.561	0.2546	112.370	0.1118

Endpoint	Phenobarbital (100 mg/kg/day)									
	Charles River Laboratories	Research Triangle Institute	WIL Laboratories	Among Laboratories						
	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Ratio to Vehicle (%) (Within Lab Std Err)[1]	Pooled Average Within Laboratory Std Err [2]	Among Laboratories Overall Average (Among Lab Std Error)[3]	Among Laboratories CV (%)[4]	Among Laboratories Standard Deviation[5]	P-value[6]	Ratio of Among Labs Std Devn to Average Within Labs Std Err (%)[7]	Homogeneity of Among Labs Results (P-value, Based On Likelihood Ratio[9])
TSH	227.017 (33.909)	169.000 (26.830)	206.954 (29.649)	30.130	196.568 (17.159)	15.120	0.000	1.0000	0.000	1.0000
FSH	82.921 (5.006)	76.710 (3.670)	89.358 (6.116)	4.931	81.060 (2.819)	6.023	1.477	0.4575	29.949	0.4547
Estradiol	151.645 (11.058)	142.450 (12.360)	174.019 (20.632)	14.683	151.198 (7.653)	8.767	0.000	1.0000	0.000	1.0000
Prolactin	11.557 (4.598)	49.910 (11.420)	75.301 (25.975)	13.998	37.316 (14.334)	66.533	20.843	0.1931	148.906	0.0052**
DHT	51.021 (9.188)	94.010 (17.660)	44.135 (13.362)	13.403	58.076 (9.473)	28.253	10.313	0.3806	76.947	0.3715

1. Ratio to the vehicle was calculated as percent of vehicle group mean. The standard error for the ratio was approximated as $Se[R(X, Y)] \approx |1/X| [(Y/X)^2 S_X^2 + S_Y^2]^{1/2} \times 100\%$
2. Unweighted average of the standard errors within each laboratory.
3. Weighted average across laboratories of within laboratory ratio to vehicle and associated standard error based on one-way heterogeneous variance random effects analysis of variance.
4. CV was calculated as [√3×standard error of among laboratories mean]/[overall average] x 100%.
5. Square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance.
6. P-value of among laboratories variation. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
7. Ratio of square root of among laboratories standard deviation based on one-way heterogeneous variance random effects analysis of variance to average of three within laboratories standard errors x 100%.
8. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.05/6 = 0.00833 significance level (Bonferroni adjustment).
10. Endpoints for which flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
11. Outliers were excluded from the analysis.

Table 15. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Linuron Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Selected Endpoints.[1] Outliers Excluded.[2,8]

Endpoint[1]	Among Laboratories Main Effect Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[9]	Among Laboratories by Dose Interaction Variance Component[5]	P-value[6]	P-value Based On Likelihood Ratio[10]
Body Weight Change (TD15-TD8)	47.4756	0.3030	0.2677	33.0512	0.3563	0.3328
Food Consumption (TD15-TD8)	19.8882	0.1705	0.0185*	0	1.0000	1.0000
Paired Epididymides	0	1.0000	1.0000	0	1.0000	1.0000
Entire Prostate	2.0065	0.4276	0.4196	0	1.0000	1.0000
Thyroid Glands	0	1.0000	1.0000	0	1.0000	1.0000
Adj Paired Epididymides[7]	0	1.0000	1.0000	0	1.0000	1.0000
Adj Entire Prostate[7]	0	1.0000	1.0000	0	1.0000	1.0000
Adj Thyroid Glands[7]	0	1.0000	1.0000	0	1.0000	1.0000
Testosterone	423.07	0.1610	0.0085**	0	1.0000	1.0000
TSH	0	1.0000	1.0000	0	1.0000	1.0000
FSH	40.0078	0.1852	0.0345*	0	1.0000	1.0000
Estradiol	0	1.0000	1.0000	0	1.0000	1.0000
Prolactin	1467.89	0.1324	<0.0001**	0	1.0000	1.0000
DHT	72.8970	0.2833	0.1910	0	1.0000	1.0000

1. Endpoints for which a flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
2. Outliers were excluded from the analysis.
3. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
5. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
6. P-value of among laboratories by dose interaction variance component. Significance is indicated by "*" for the 0.05 level.
7. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
8. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
10. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

Table 16. Among Laboratories Main Effects and Laboratory by Dose Group Interactions for Ratios for Phenobarbital Dose Group Responses to Vehicle Control in Adult Intact Male Assay for Selected Endpoints.[1] Outliers Excluded.[2,8]

Endpoint[1]	Among Laboratories Main Effect Variance Component[3]	P-value[4]	P-value Based On Likelihood Ratio[9]	Among Laboratories by Dose Interaction Variance Component[5]	P-value[6]	P-value Based On Likelihood Ratio[10]
Body Weight Change (TD15-TD8)	0	1.0000	1.0000	0	1.0000	1.0000
Food Consumption (TD15-TD8)	0	1.0000	1.0000	0	1.0000	1.0000
Paired Epididymides	0	1.0000	1.0000	0	1.0000	1.0000
Entire Prostate	0	1.0000	1.0000	0	1.0000	1.0000
Thyroid Glands	0	1.0000	1.0000	0	1.0000	1.0000
Adj Paired Epididymides[7]	0	1.0000	1.0000	0	1.0000	1.0000
Adj Entire Prostate[7]	0	1.0000	1.0000	0	1.0000	1.0000
Adj Thyroid Glands[7]	0	1.0000	1.0000	0	1.0000	1.0000
Testosterone	138.70	0.2097	0.0515	0	1.0000	1.0000
TSH	0	1.0000	1.0000	0	1.0000	1.0000
FSH	11.0593	0.2177	0.1013	0	1.0000	1.0000
Estradiol	0	1.0000	1.0000	0	1.0000	1.0000
Prolactin	496.05	0.1402	0.0012**	0	1.0000	1.0000
DHT	123.19	0.2364	0.1039	0	1.0000	1.0000

1. Endpoints for which a flagged "potential outlier" was to be treated as an outlier for at least one chemical-dose group.
2. Outliers were excluded from the analysis.
3. Among laboratories main effect variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
4. P-value of among laboratories main effect variance component. Significance is indicated by "*" for the 0.05 level.
5. Laboratory by dose interaction variance component based on two-way heterogeneous variance mixed effects analysis of variance across doses and laboratories.
6. P-value of among laboratories by dose interaction variance component. Significance is indicated by "*" for the 0.05 level.
7. Adjusted organ weights are defined as organ weight to final body weight ratios x 100%.
8. Note that entries in this table are variance components rather than standard deviations. Entries in Tables 1a to 6c are standard deviations or standard errors.
9. P-value of among laboratories variation based on likelihood ratio test for homogeneous of variance among laboratories. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01.
10. P-value of among laboratories by dose interaction variation based on likelihood ratio test for homogeneous of variance among laboratories by dose combination. -2 log(likelihood ratio) follows a mixture of χ^2_1 and 0 with probability 1/2 and 1/2. Significance is indicated by "*" for the 0.05 significance level and "**" for the 0.01 significance level.

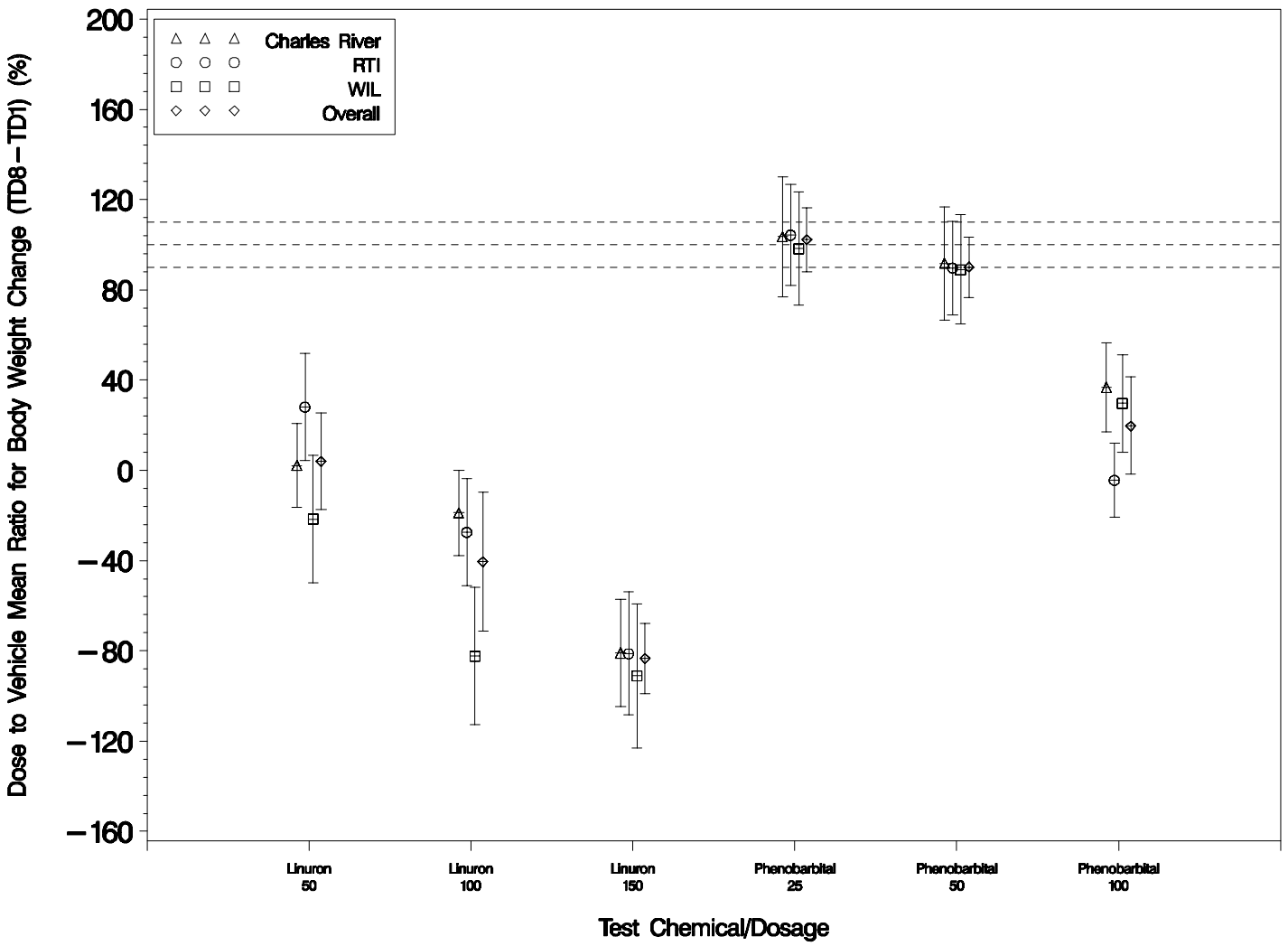
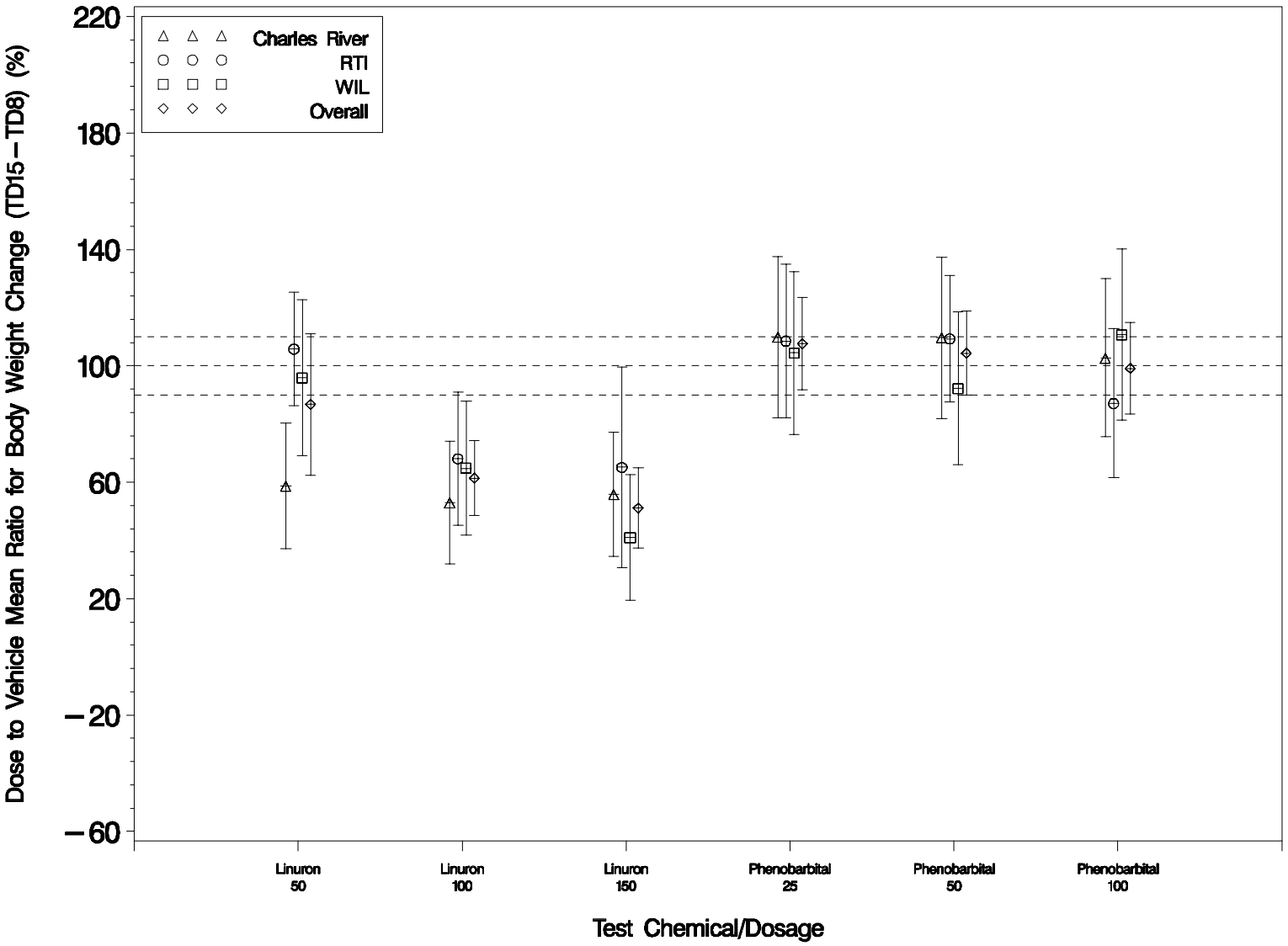
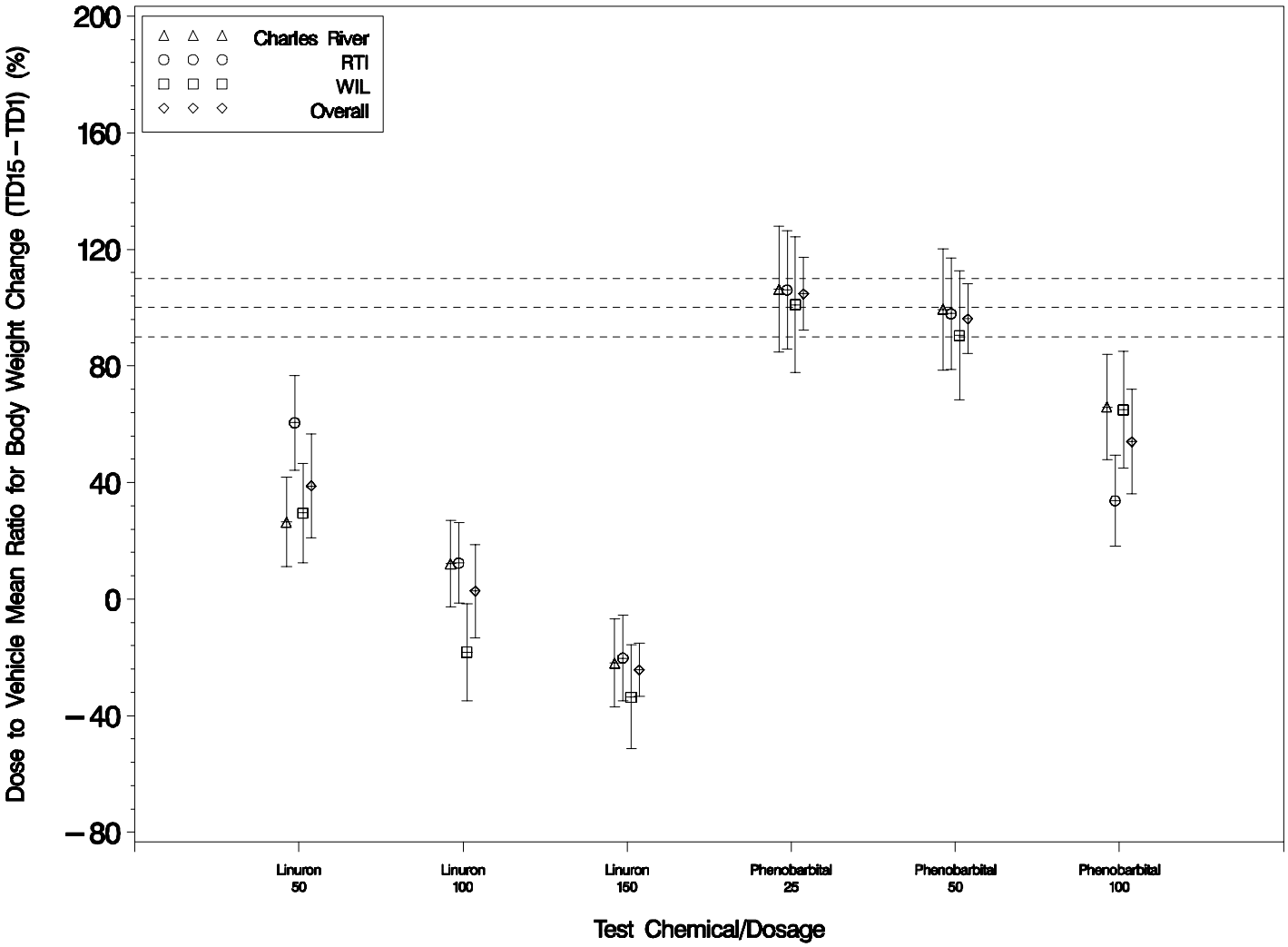


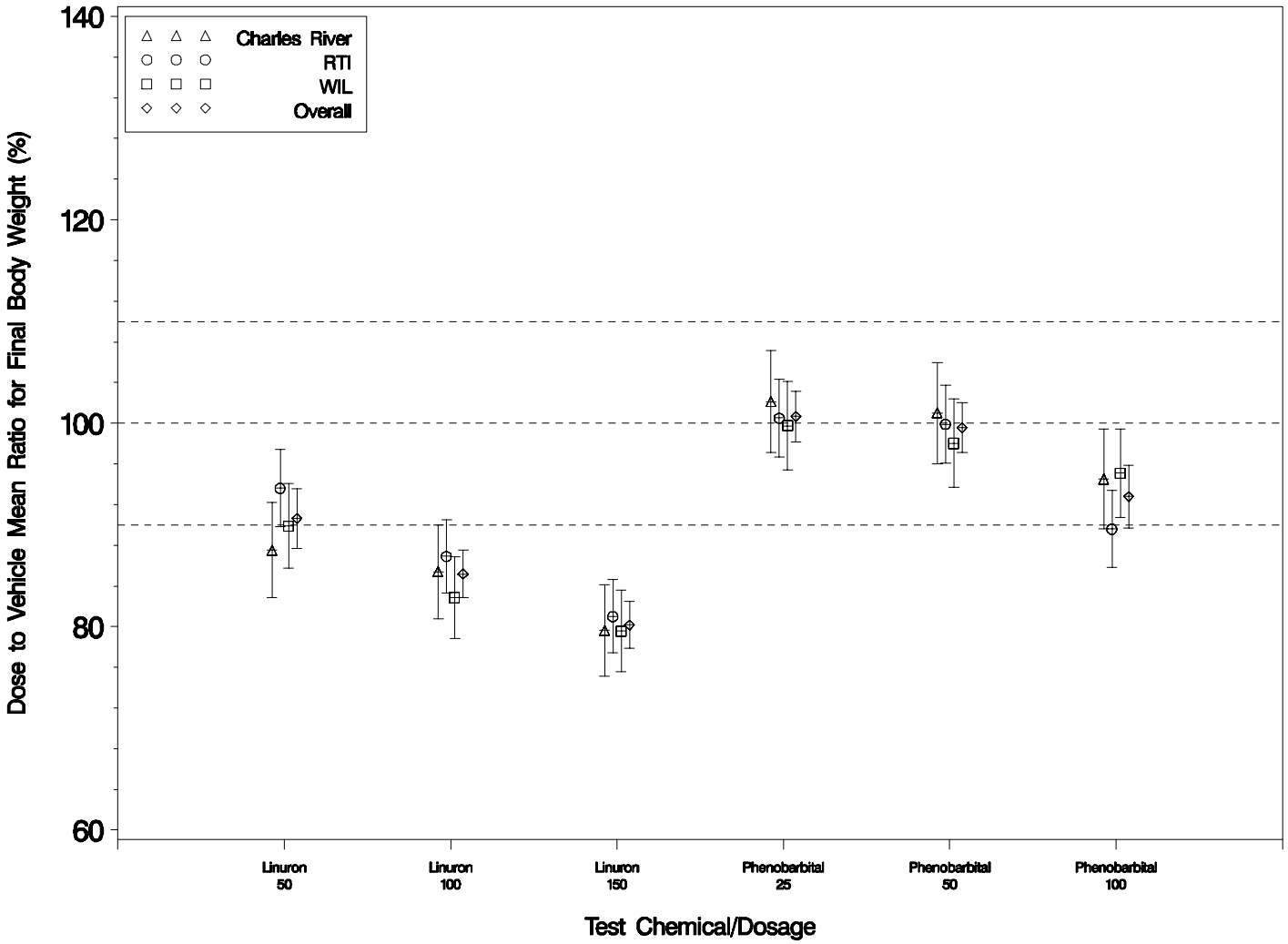
Figure 1. **Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Response in Adult Intact Male Assay for Average Daily Body Weight Changes from Day 1 to Day 8, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.**



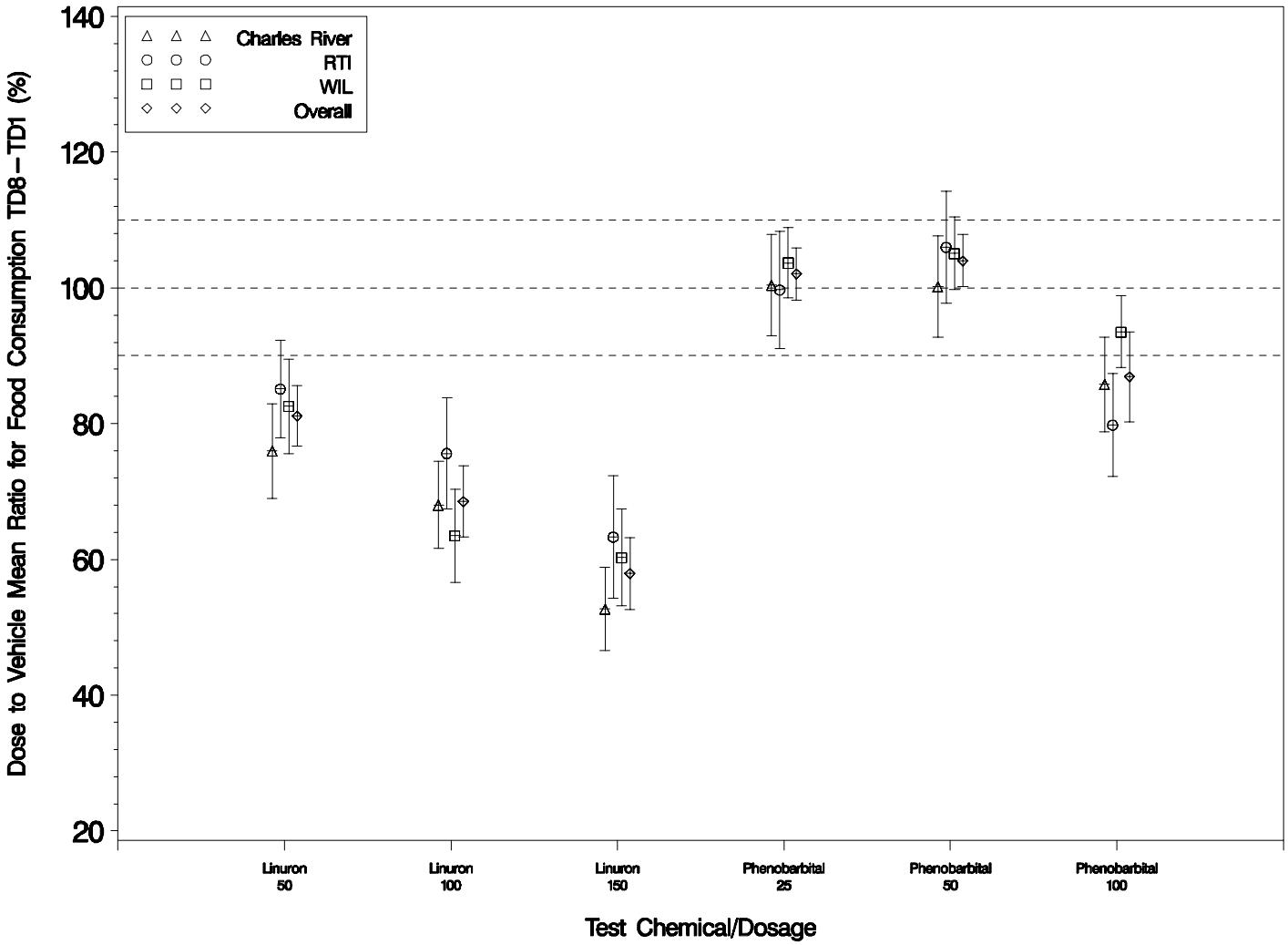
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Average Daily Body Weight Changes from Day 8 to Day 15, by Dose Group within Each La
Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose G
Vehicle Group Ratio.



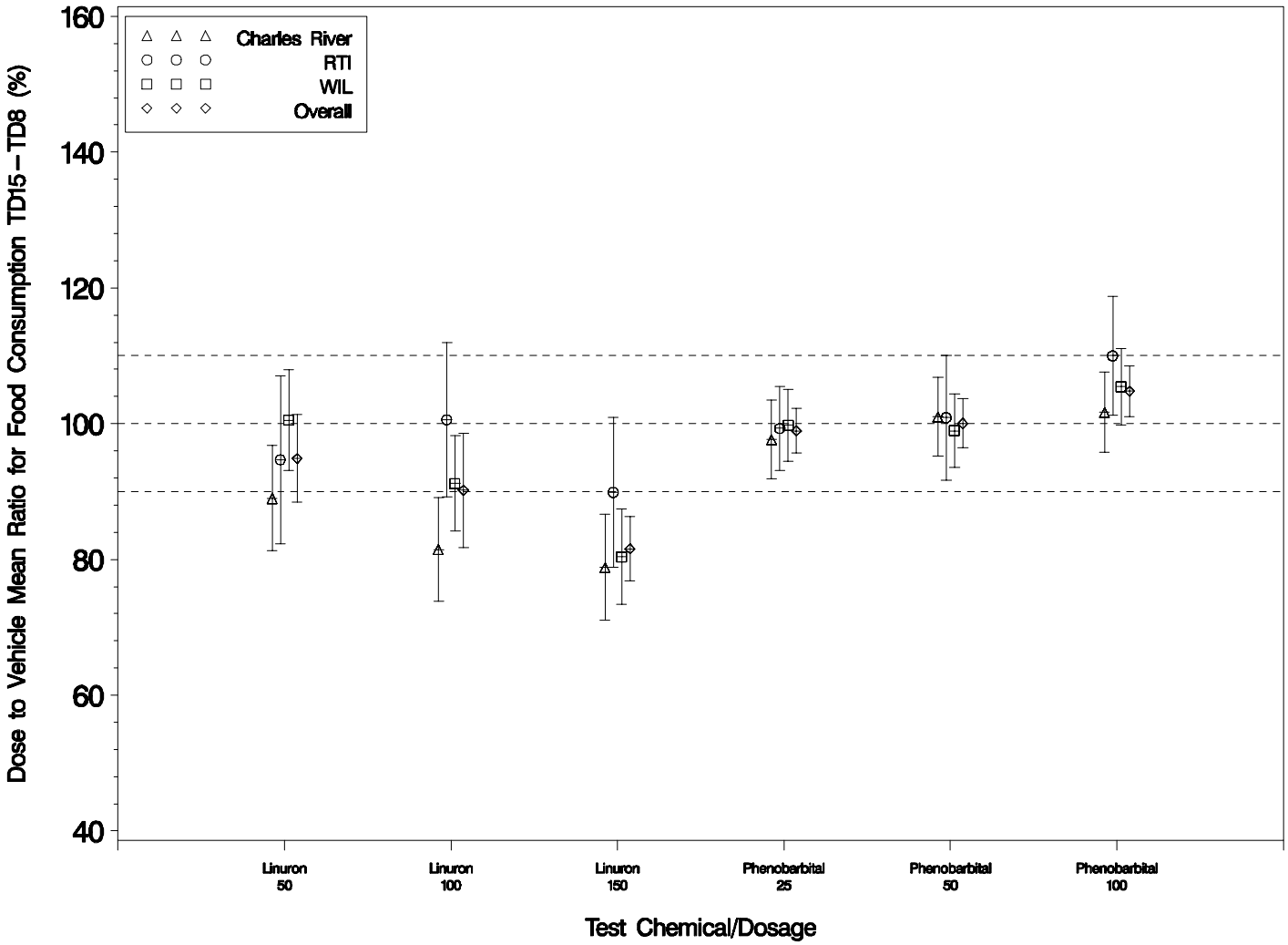
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp Adult Intact Male Assay for Average Daily Body Weight Changes from Day 1 to Day 15, by Dose Group within Each La Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose G Vehicle Group Ratio.



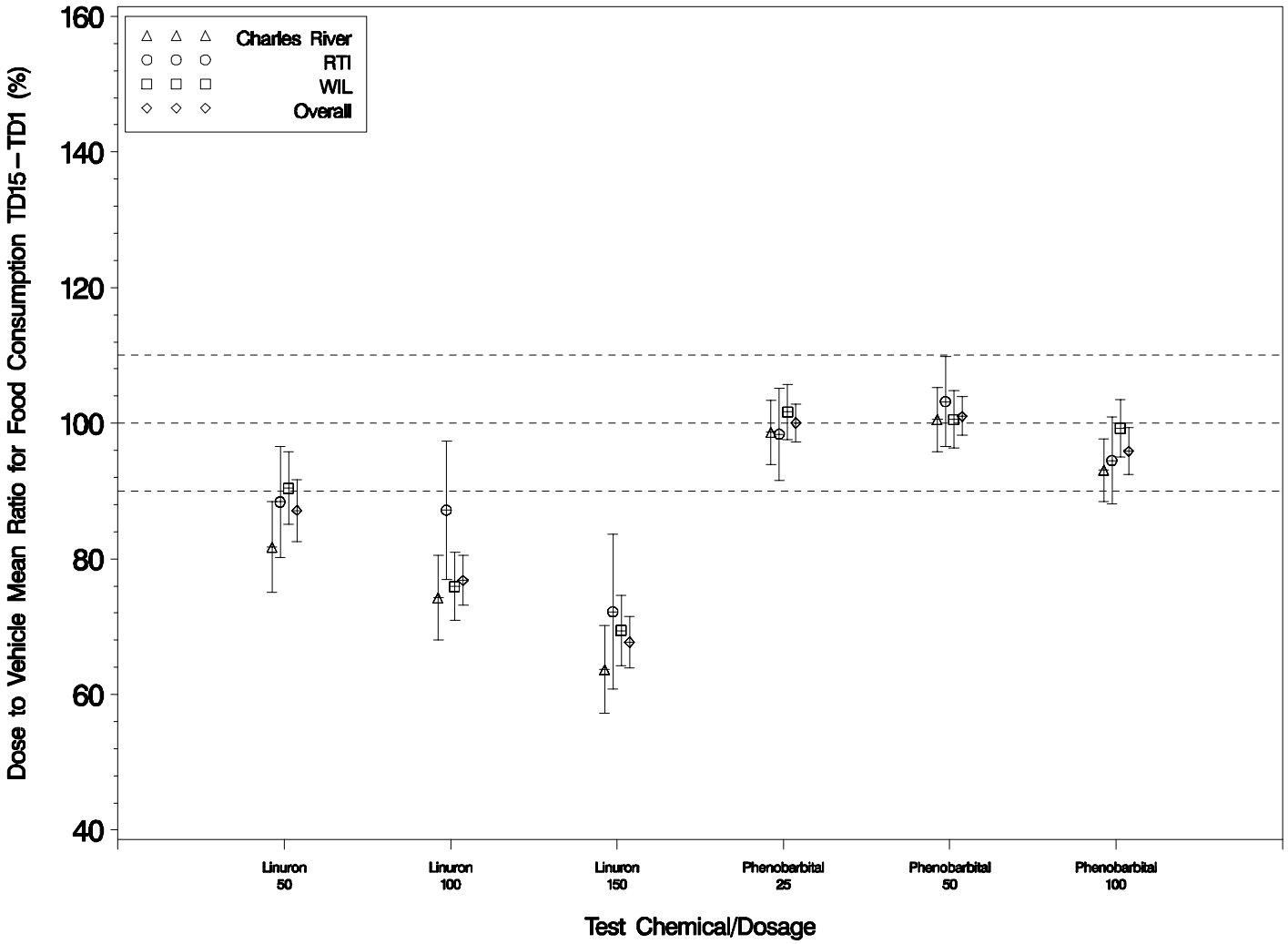
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Res
Adult Intact Male Assay for Final Body Weight, by Dose Group within Each Laboratory and Across Laboratories. The
Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



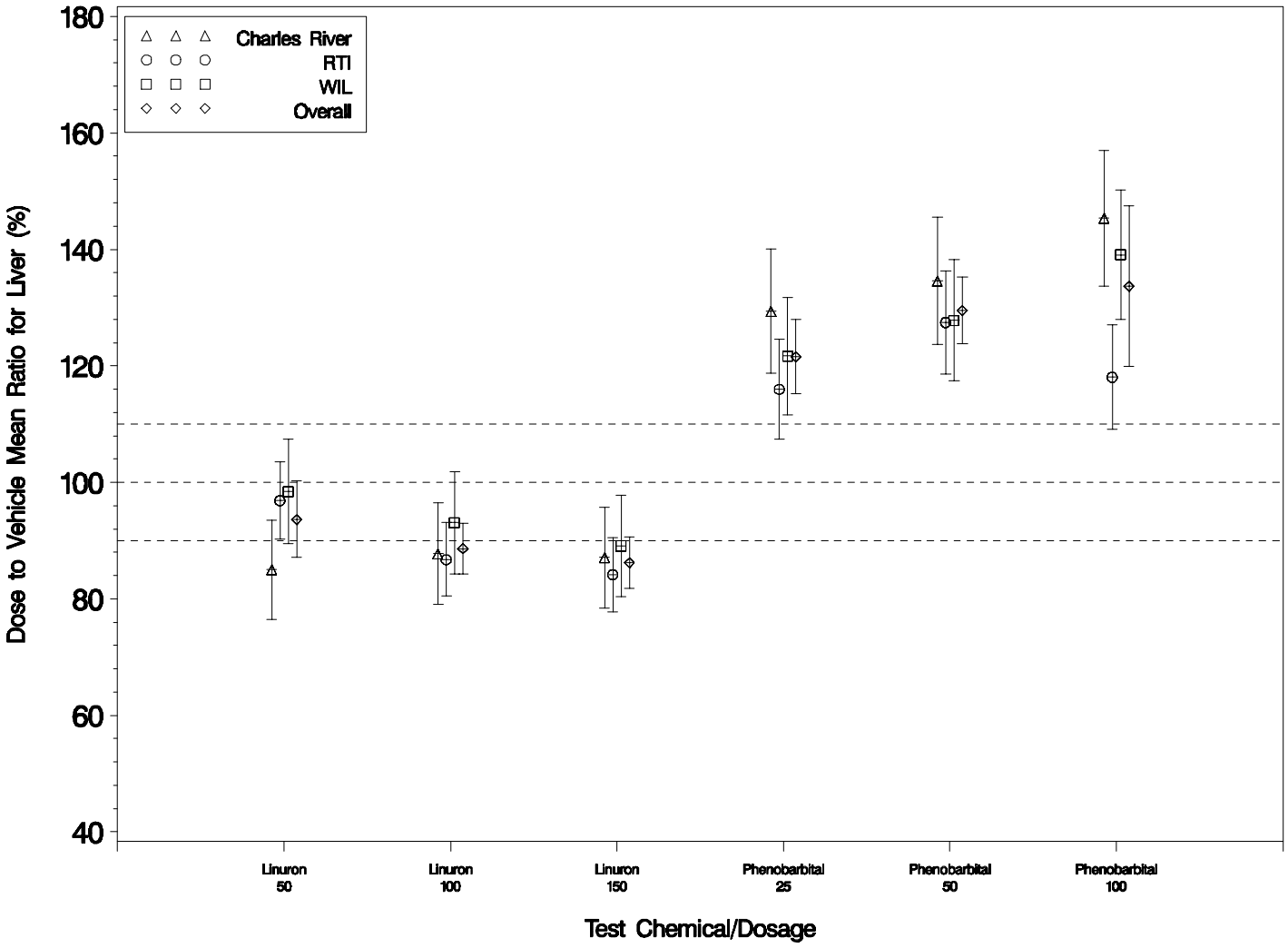
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Average Daily Food Consumption from Day 1 to Day 8, by Dose Group within Each Labora
Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose G
Vehicle Group Ratio.**



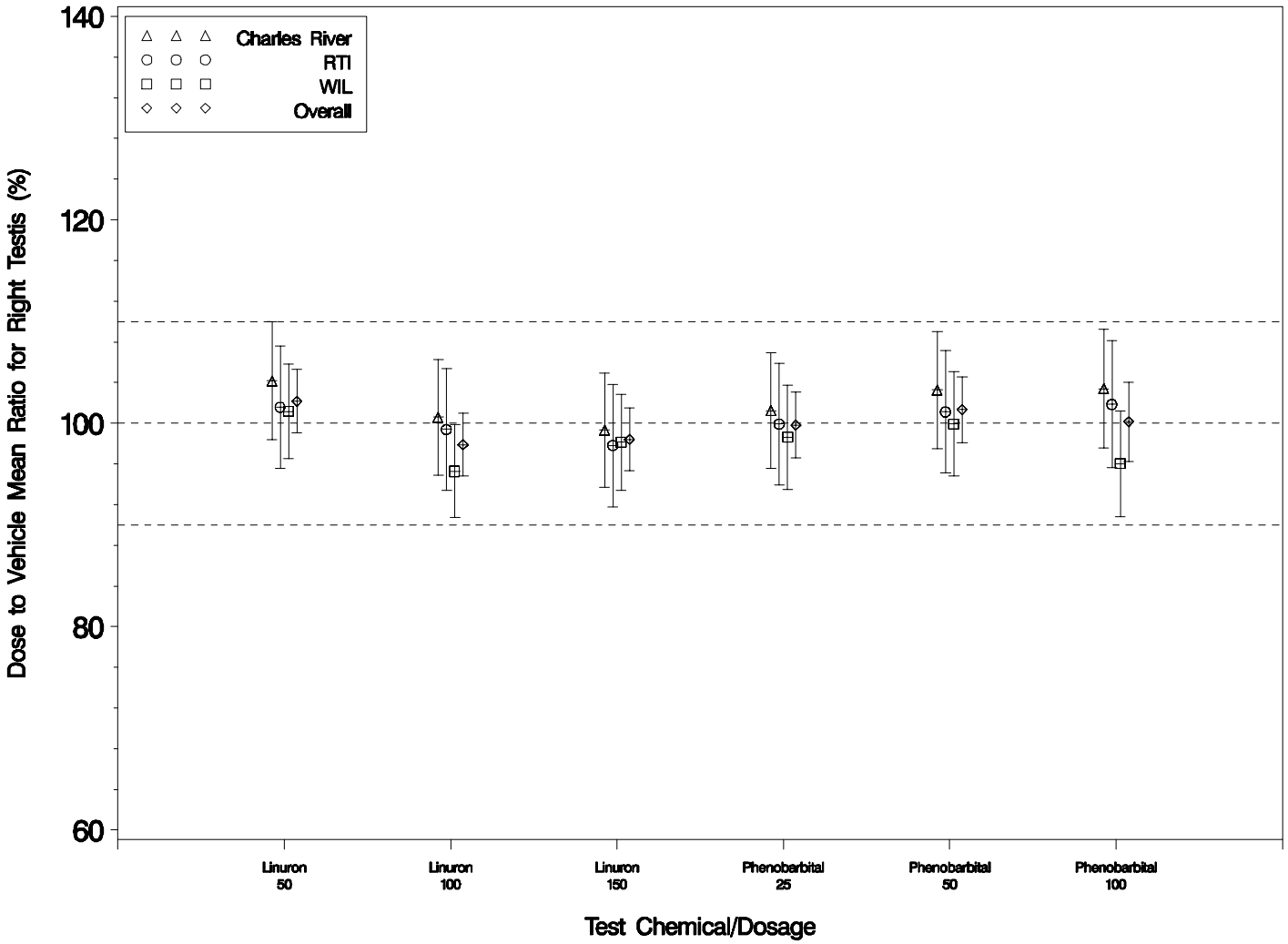
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Average Daily Food Consumption from Day 8 to Day 15, by Dose Group within Each Labo
Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose G
Vehicle Group Ratio.**



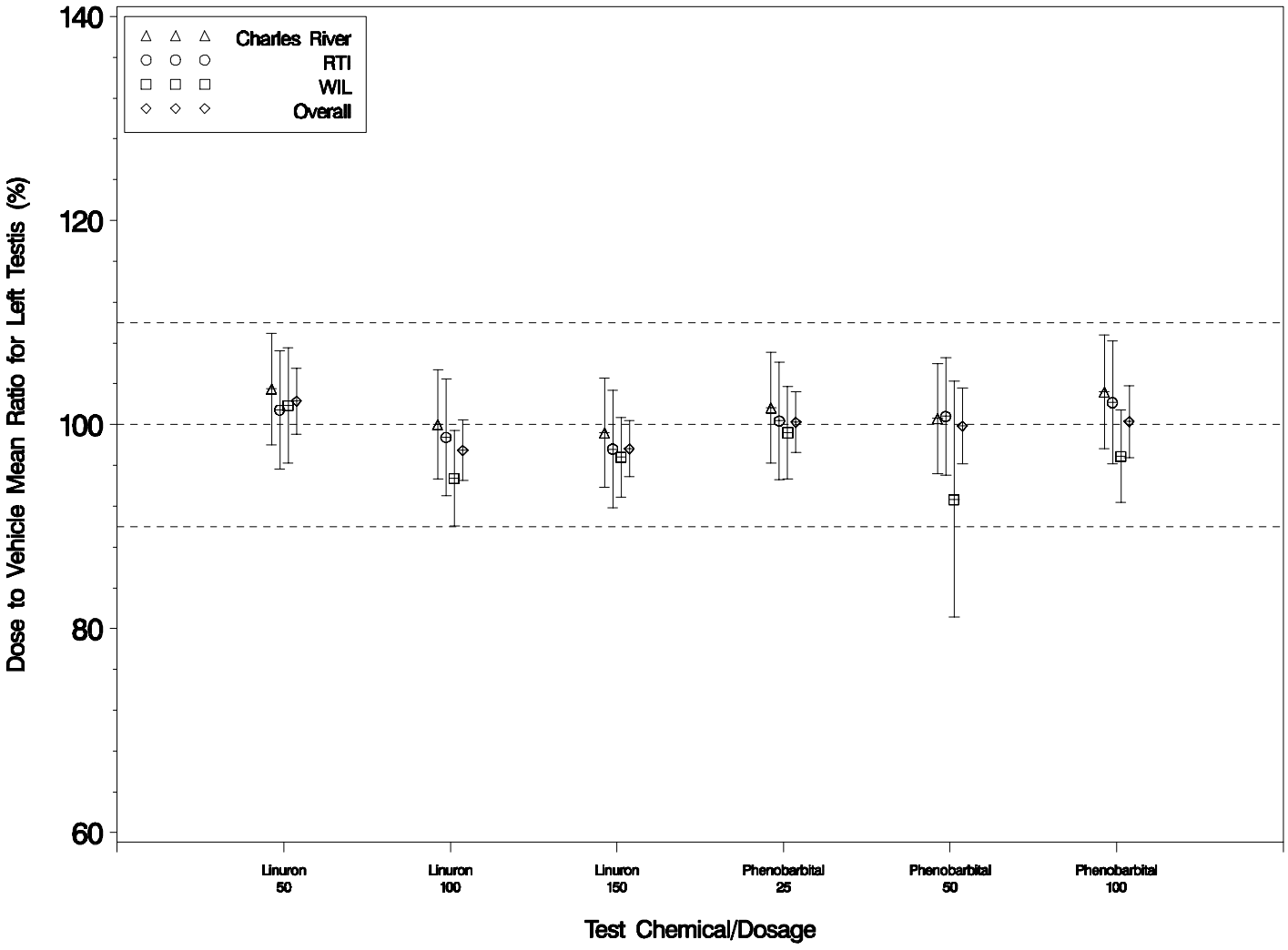
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Average Daily Food Consumption from Day 1 to Day 15, by Dose Group within Each Labo
Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose G
Vehicle Group Ratio.**



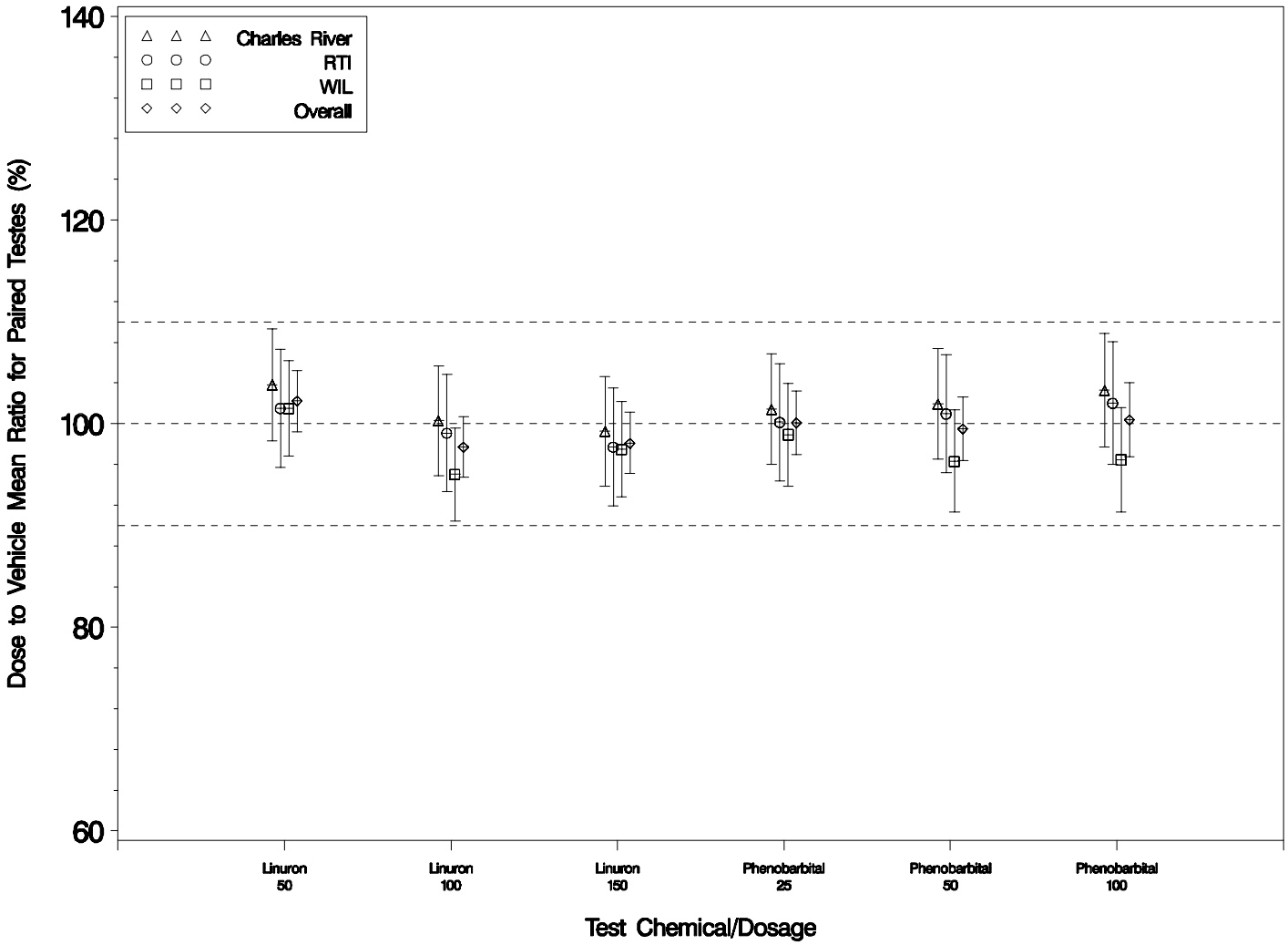
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Liver, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizo
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



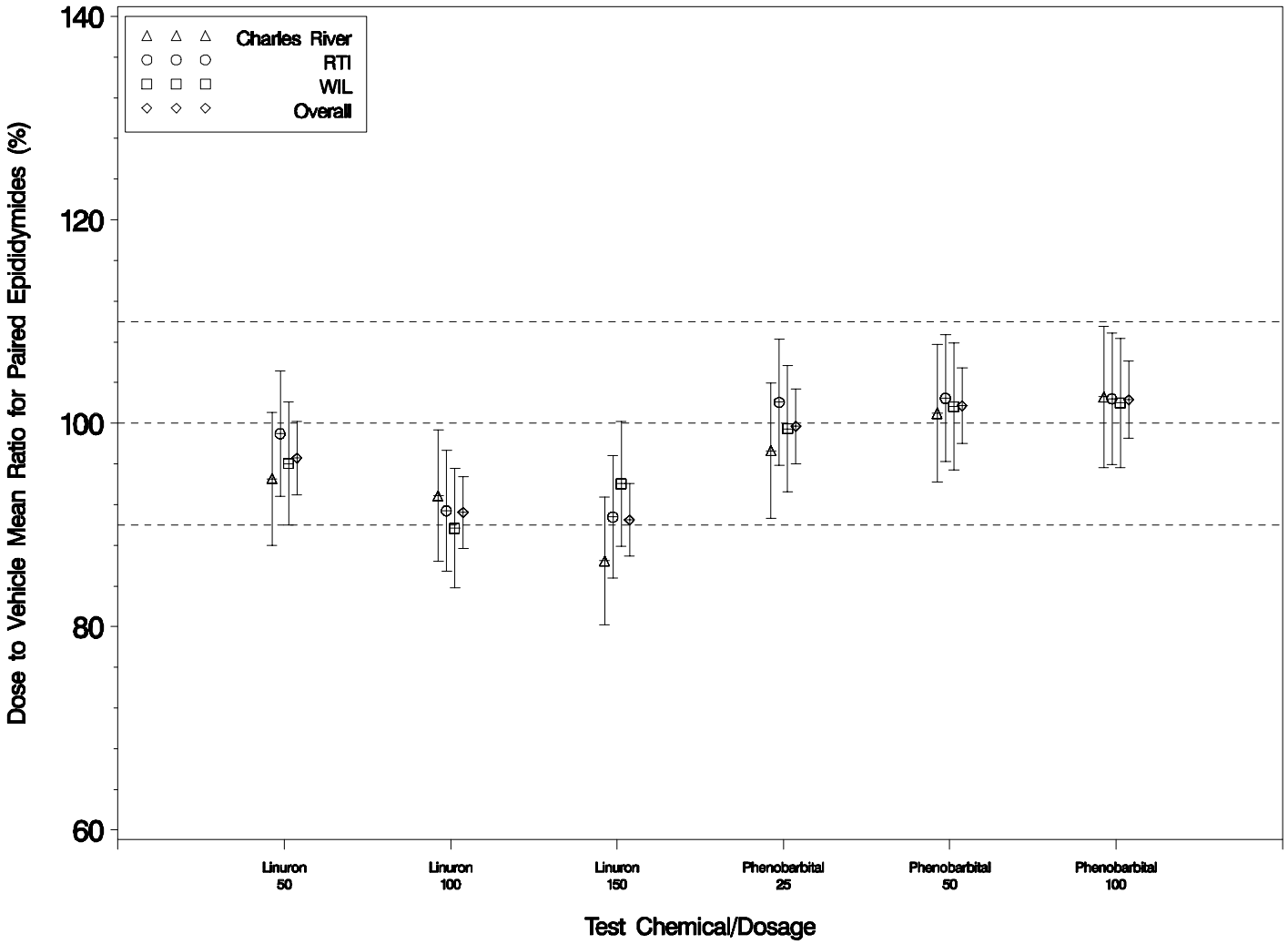
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Right Testis, by Dose Group within Each Laboratory and Across Laboratories. The Three
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



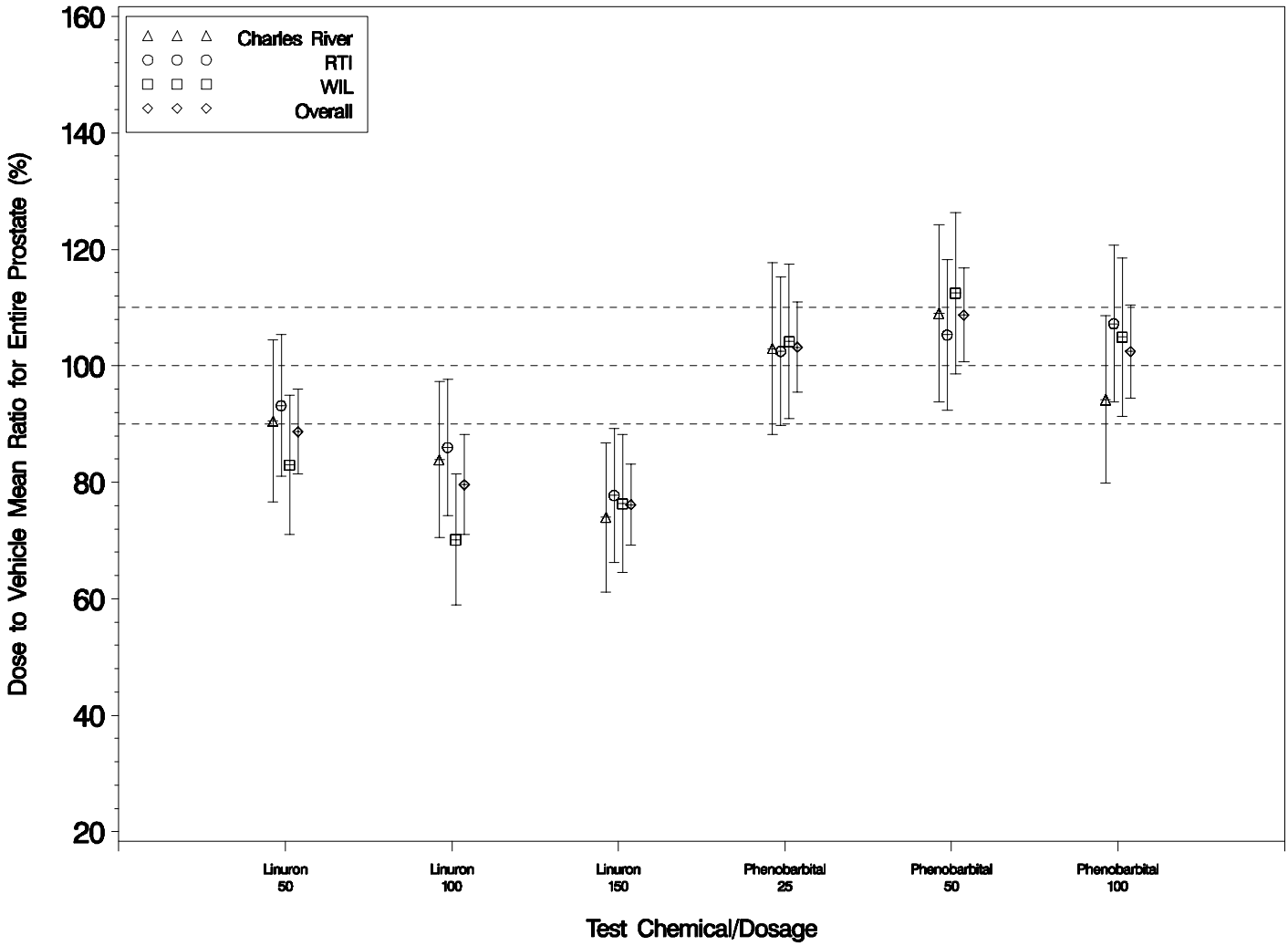
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Left Testis, by Dose Group within Each Laboratory and Across Laboratories. The Three H
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



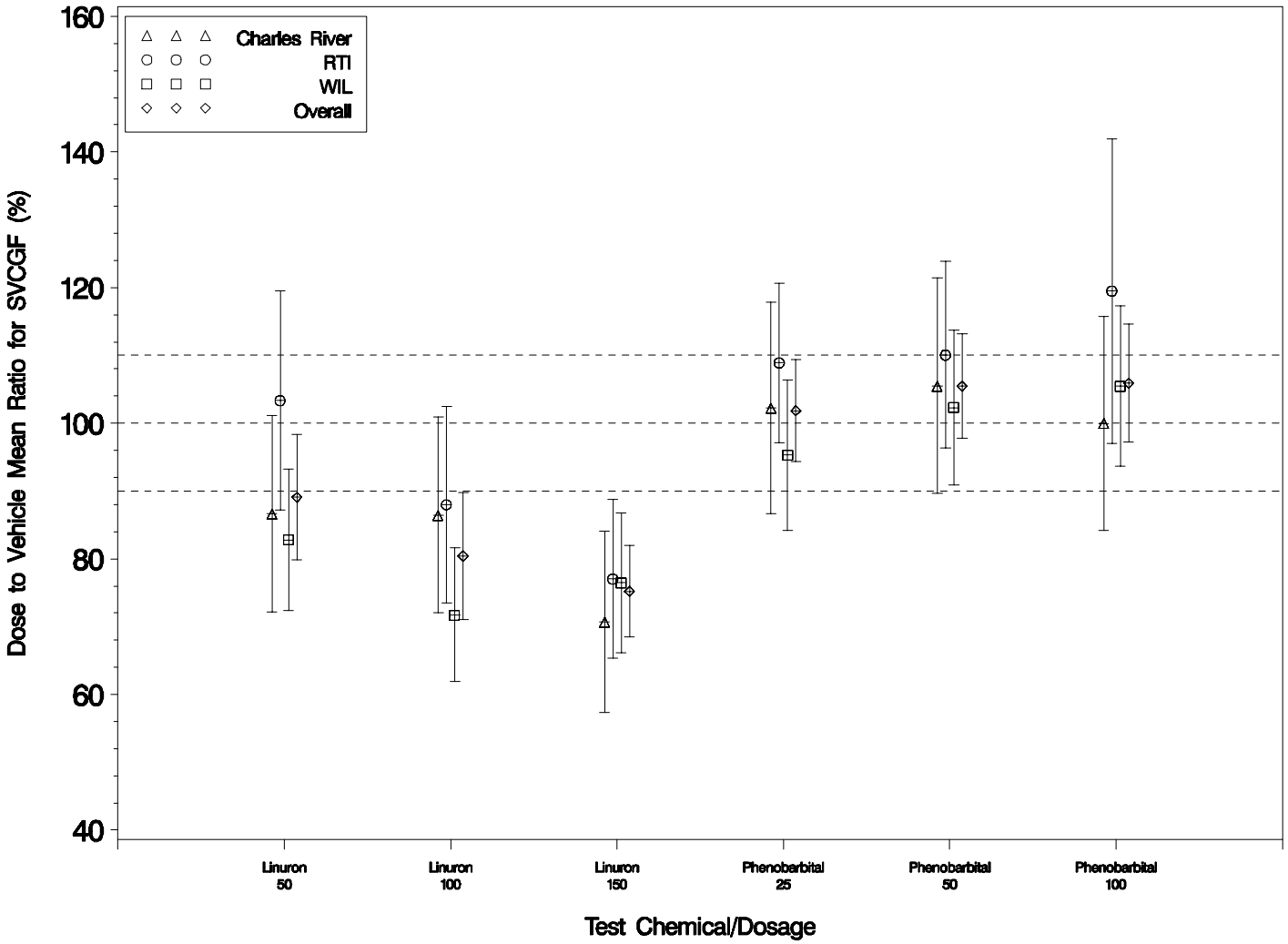
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Paired Testes, by Dose Group within Each Laboratory and Across Laboratories. The Thre
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.**



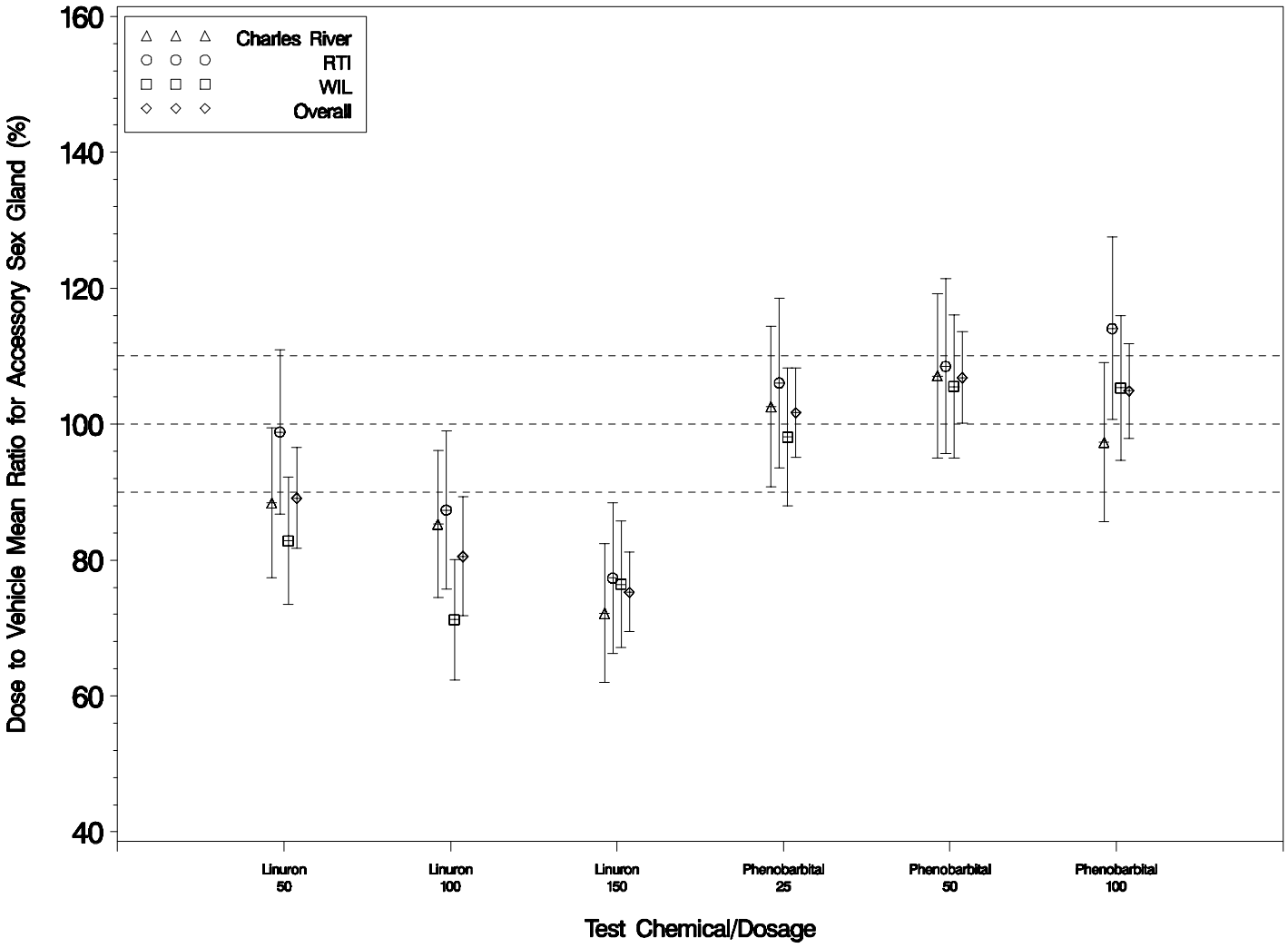
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Paired Epididymides, by Dose Group within Each Laboratory and Across Laboratories. Th
Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



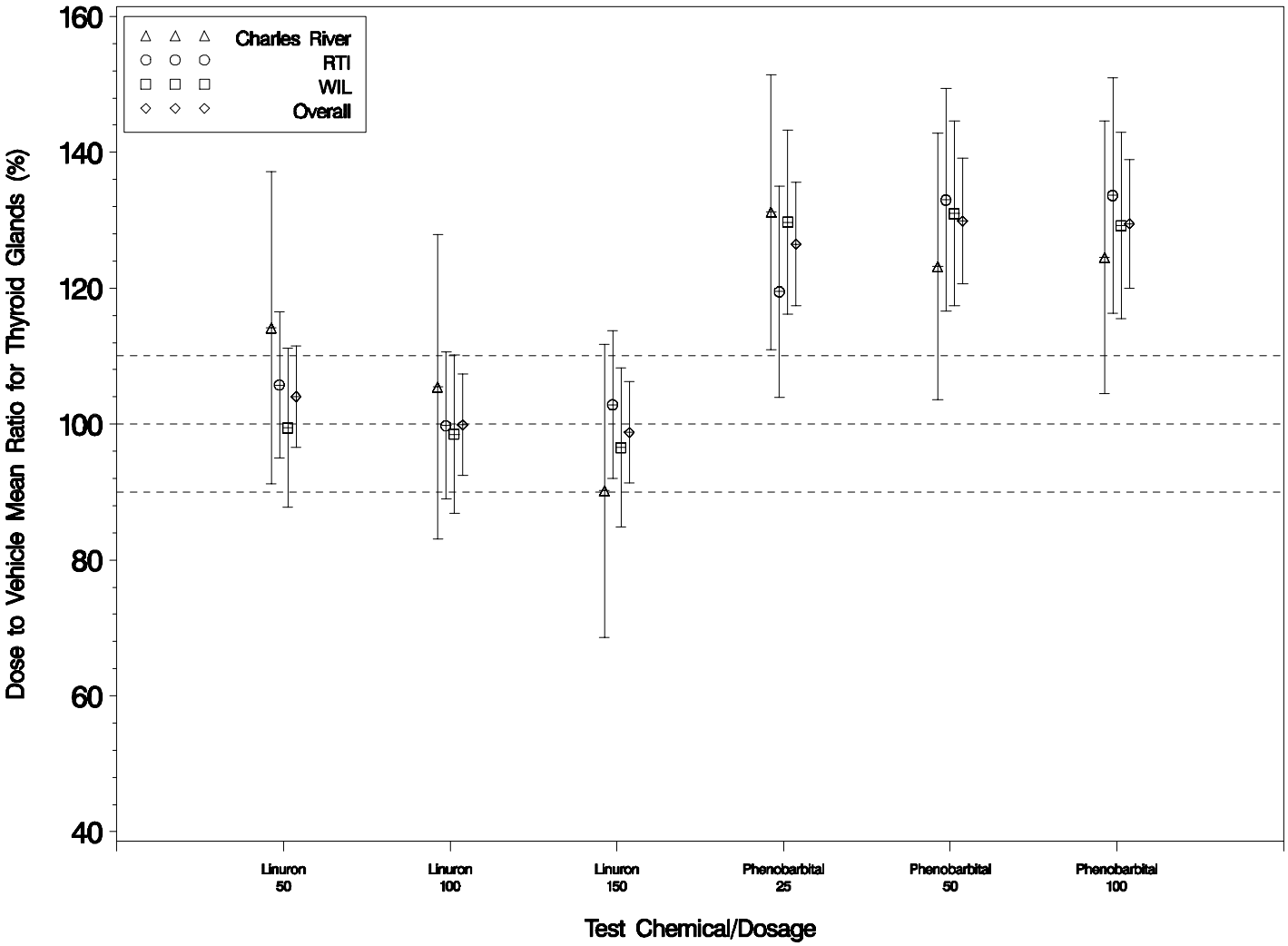
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Entire Prostate, by Dose Group within Each Laboratory and Across Laboratories. The Th
Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



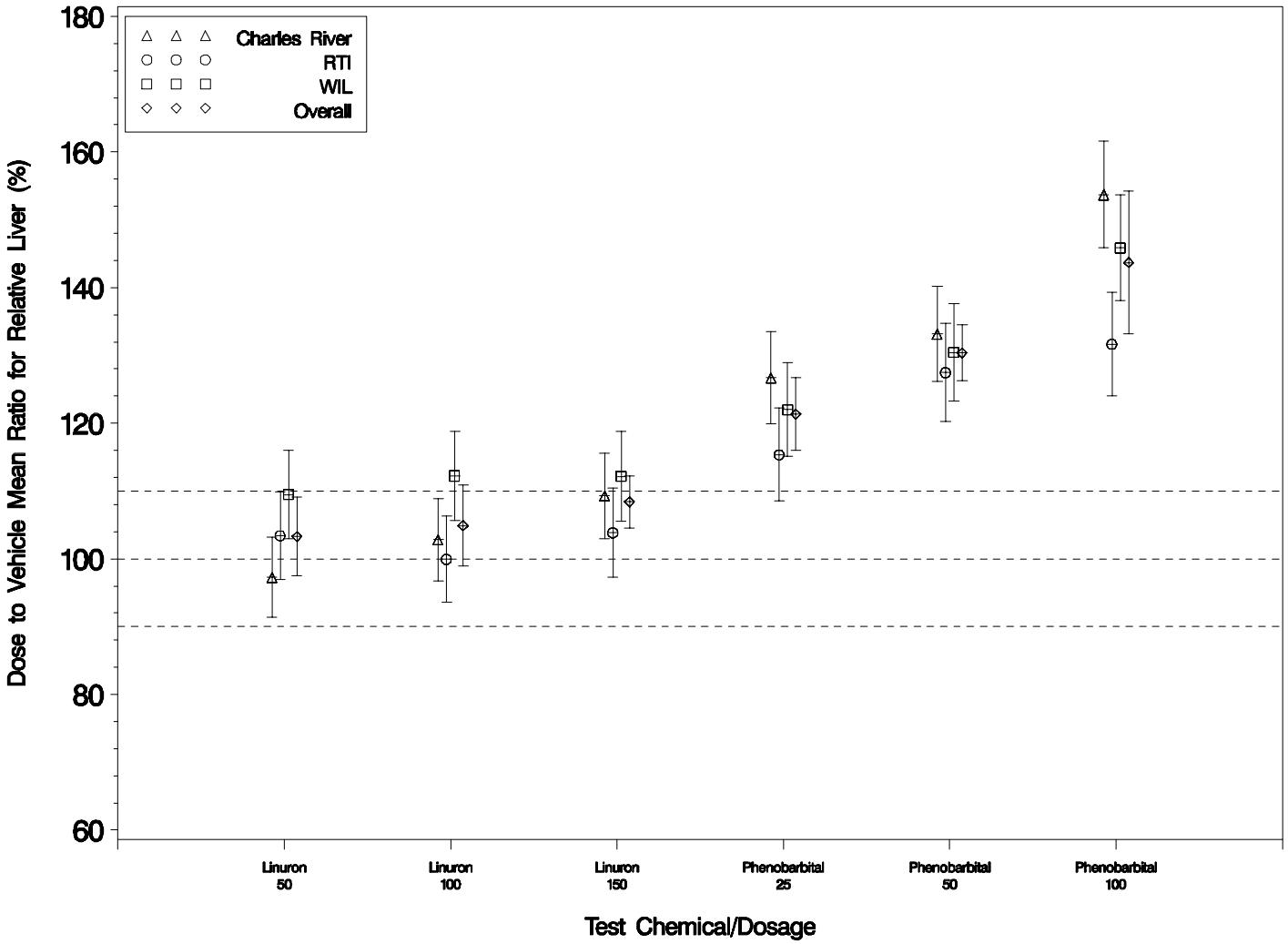
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Seminal Vesicle Plus Coagulating Gland with Fluid (SVCGF) Weight, by Dose Group withi
Laboratory and Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of C
Dose Group to Vehicle Group Ratio.**



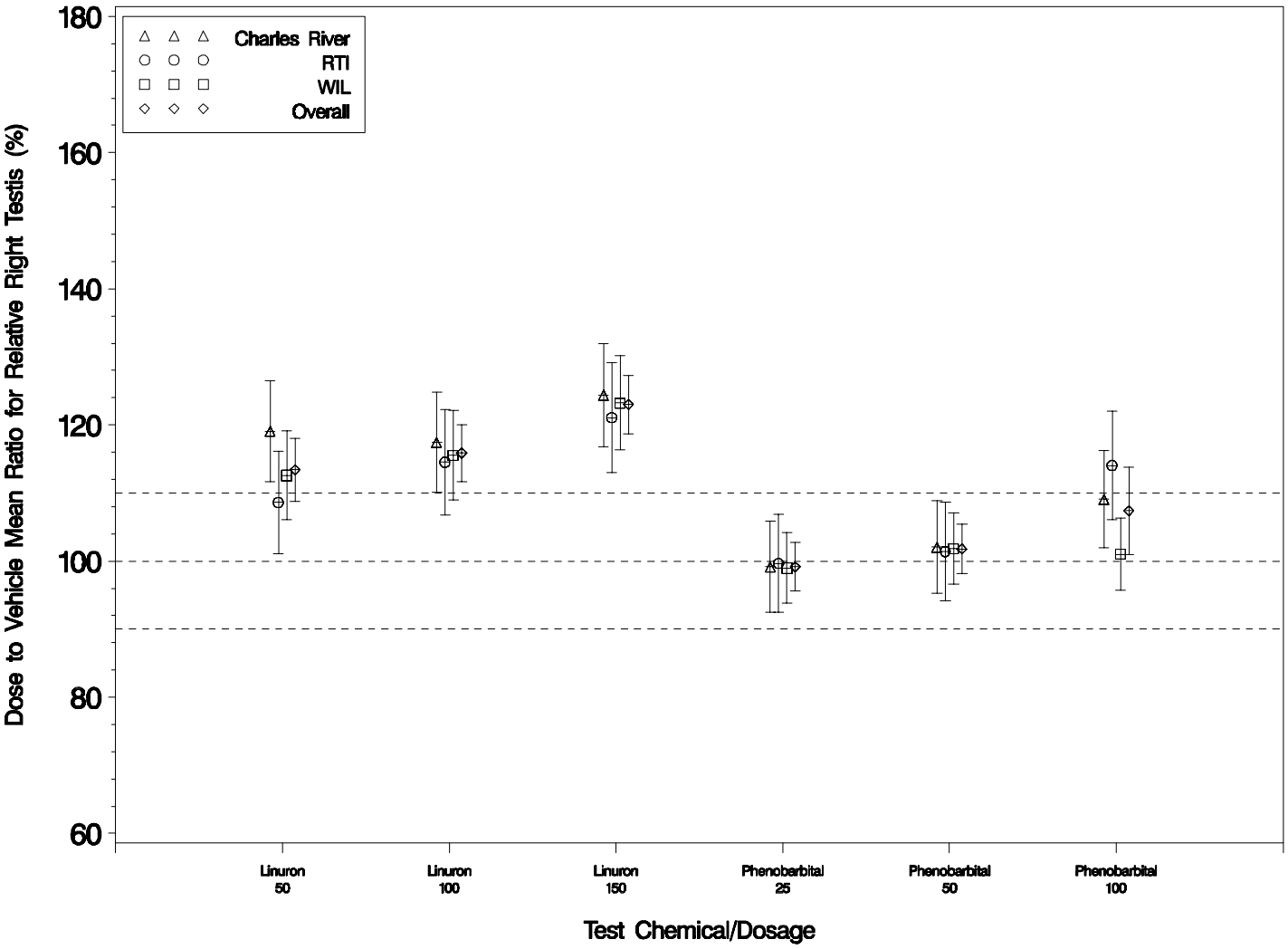
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Accessory Sex Gland Weight, by Dose Group within Each Laboratory and Across Laborato
Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ra



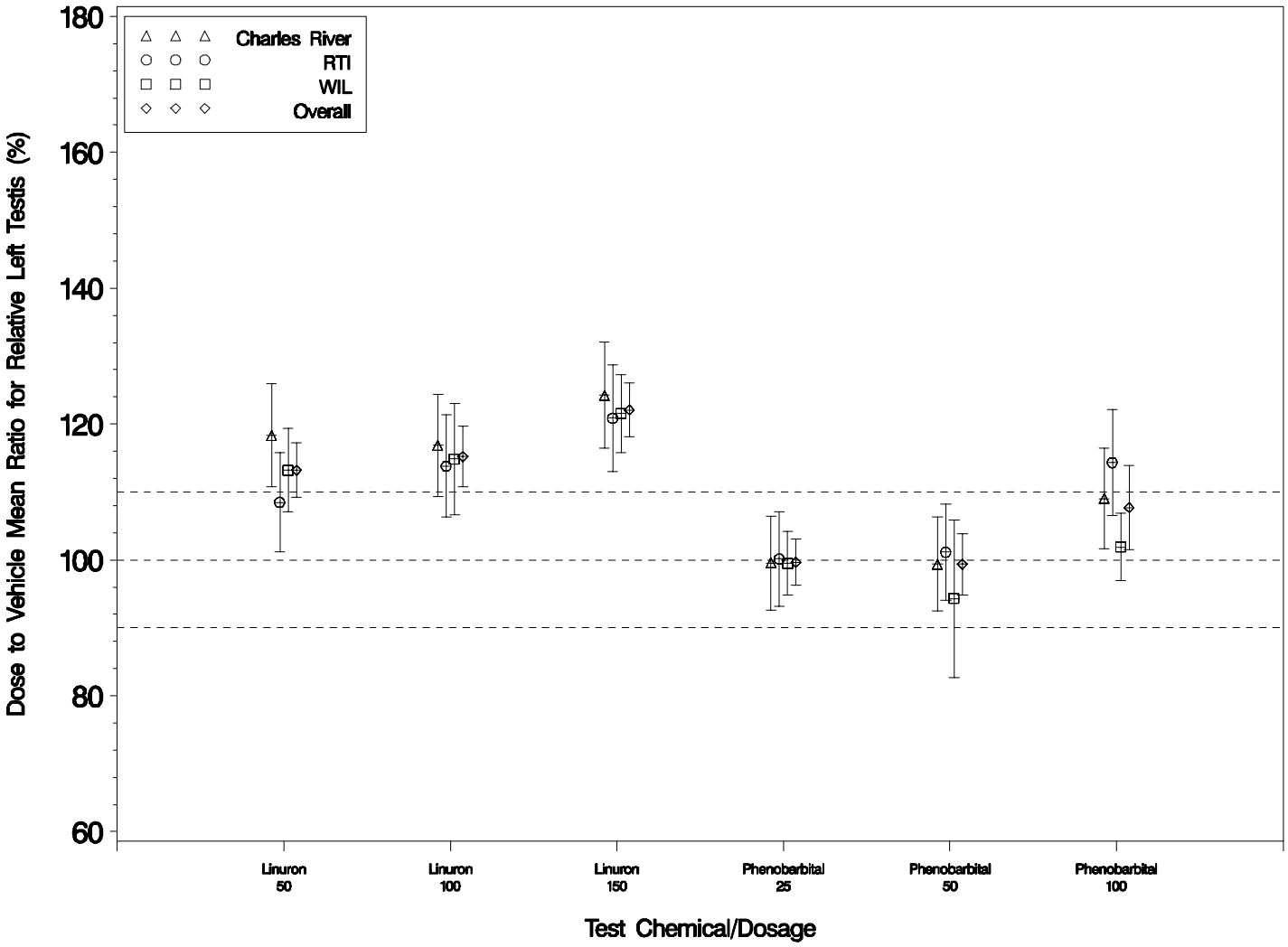
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Res
Adult Intact Male Assay for Thyroid Glands Weight, by Dose Group within Each Laboratory and Across Laboratories.
Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



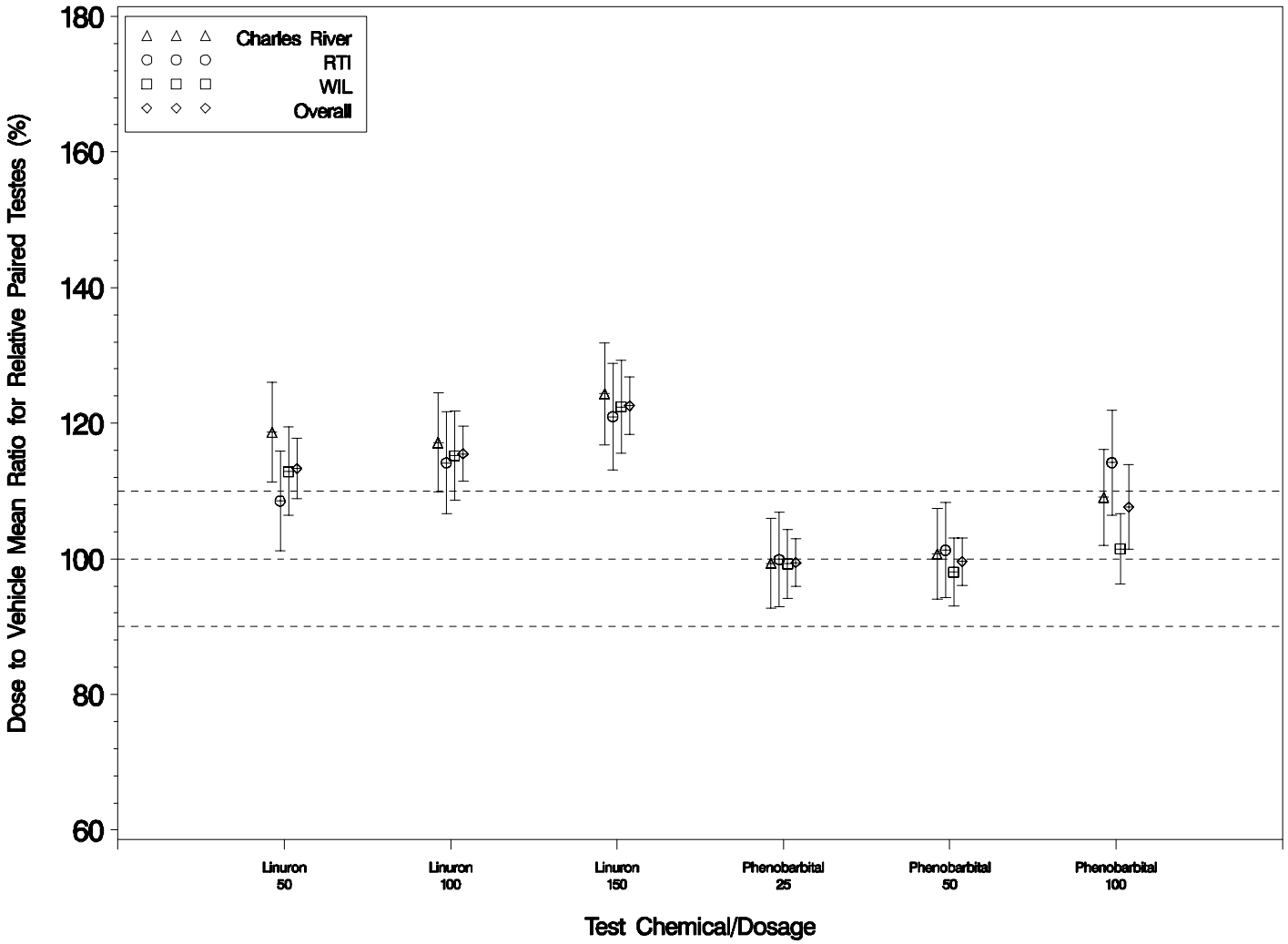
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Liver to Body Weight Ratio, by Dose Group within Each Laboratory and Across Laborator
Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ra



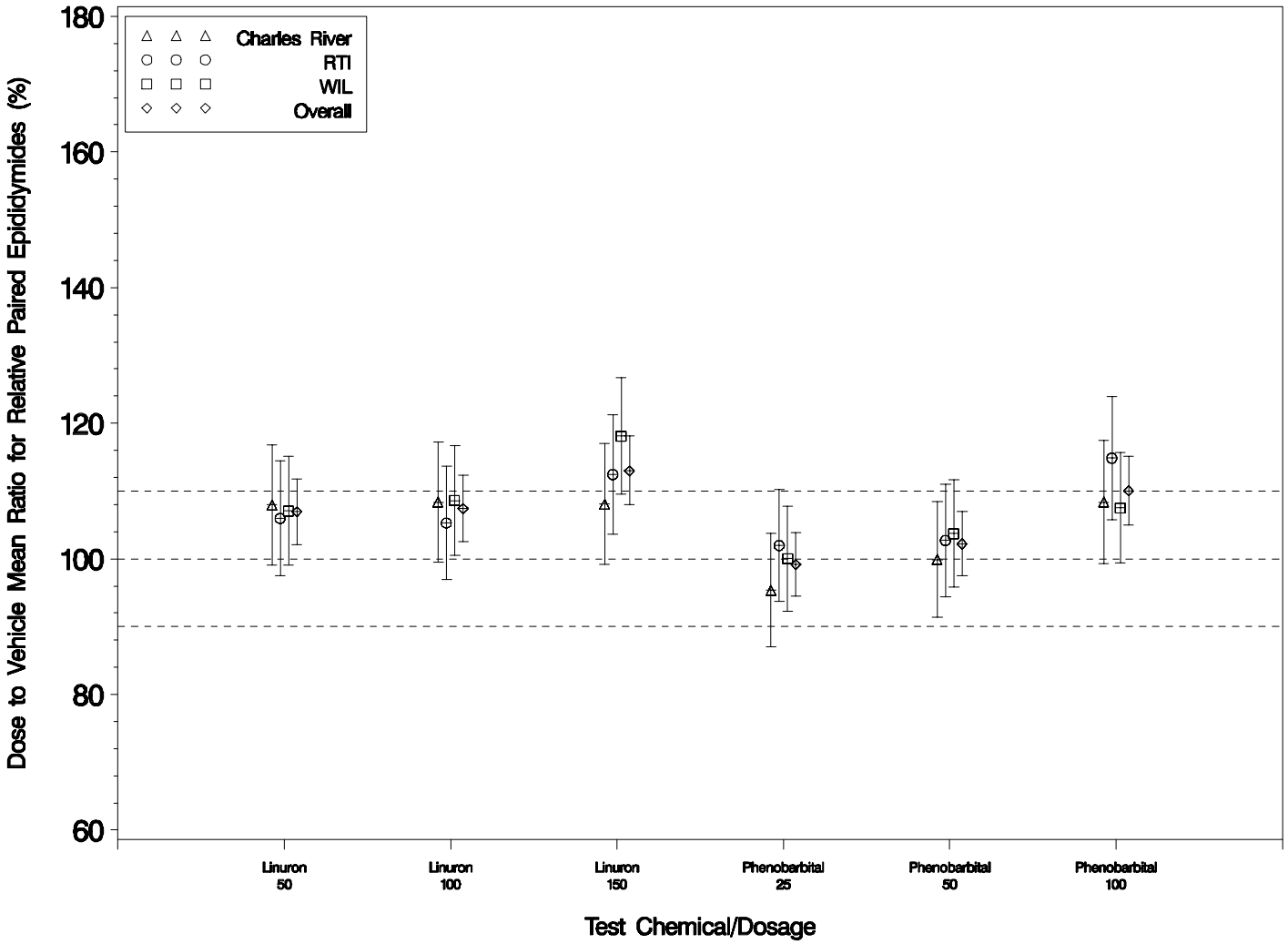
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Right Testis to Body Weight Ratio, by Dose Group within Each Laboratory and Across Lab
The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group



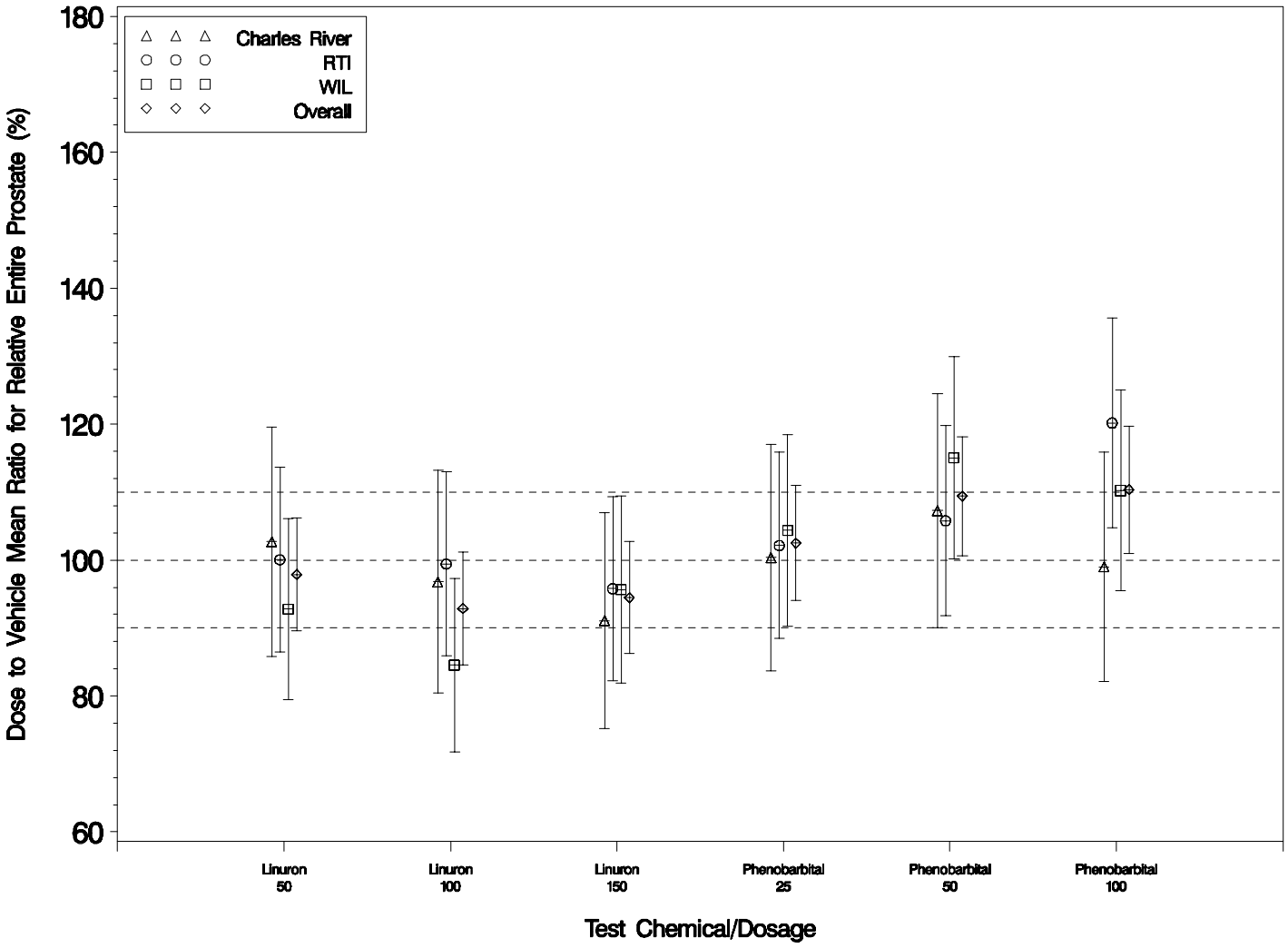
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Left Testis to Body Weight Ratio, by Dose Group within Each Laboratory and Across Labo
Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ra



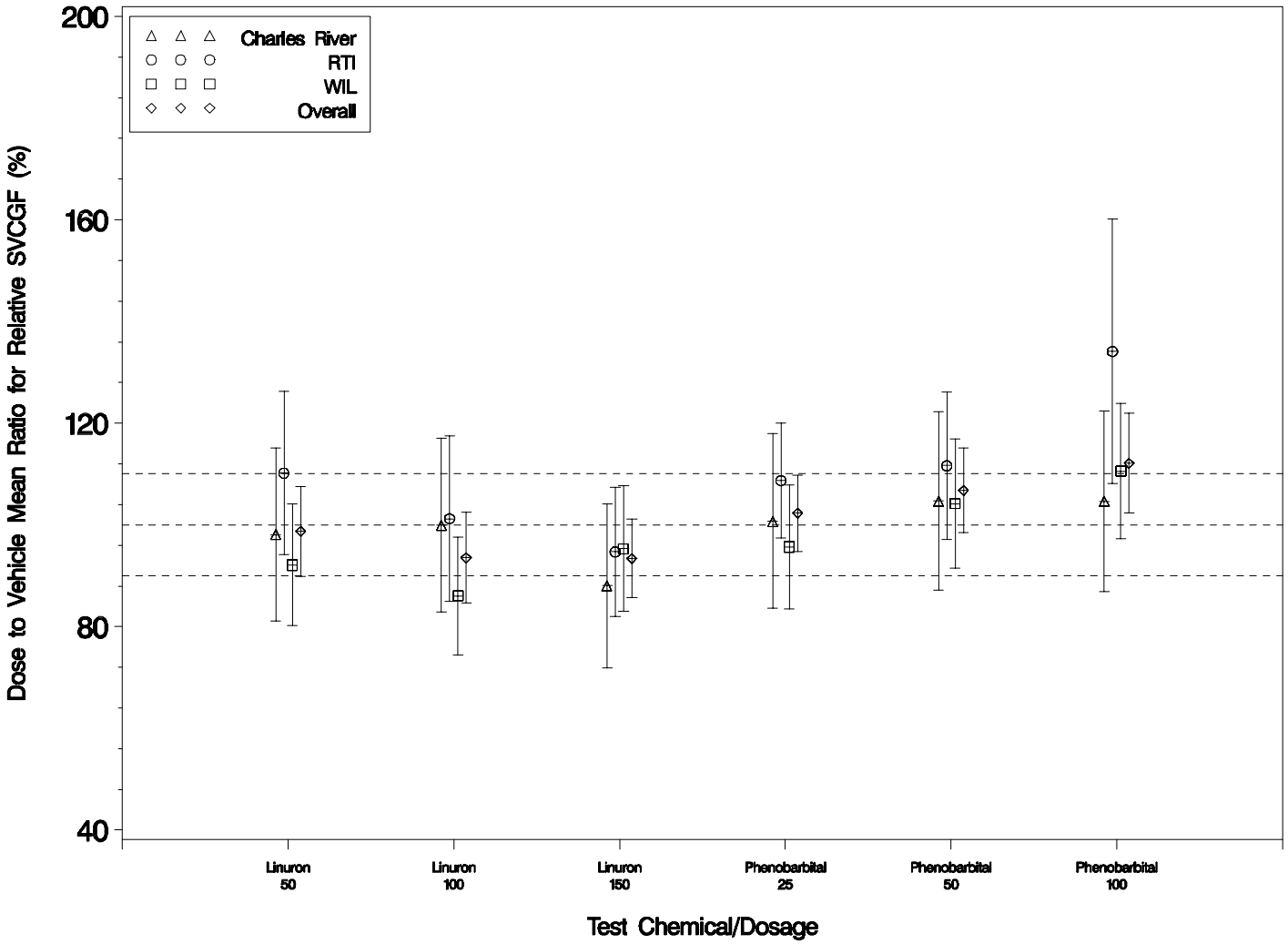
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Paired Testes to Body Weight Ratio, by Dose Group within Each Laboratory and Across La
The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group



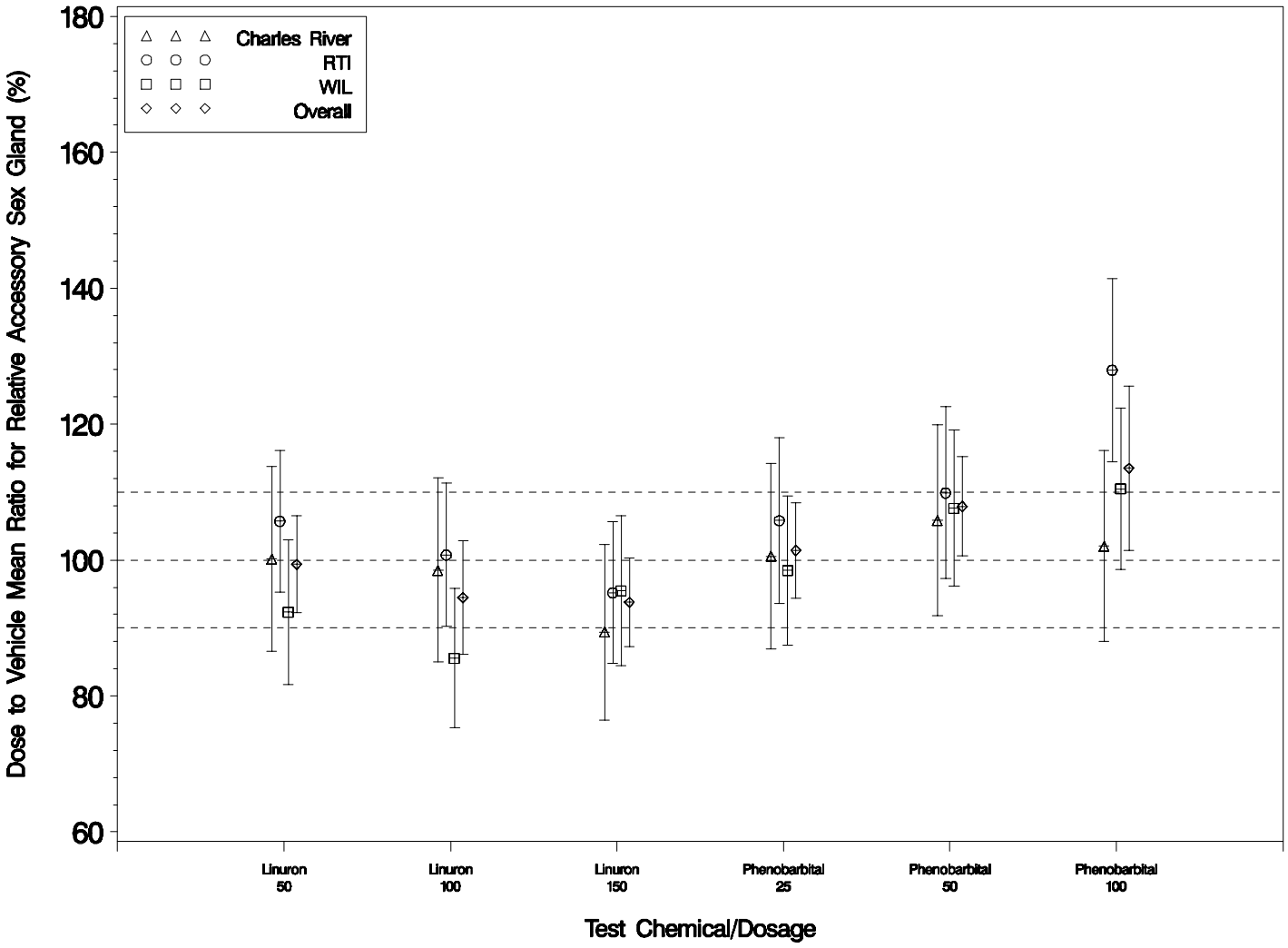
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Res
Adult Intact Male Assay for Paired Epididymides to Body Weight Ratio, by Dose Group within Each Laboratory and A
Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to
Group Ratio.



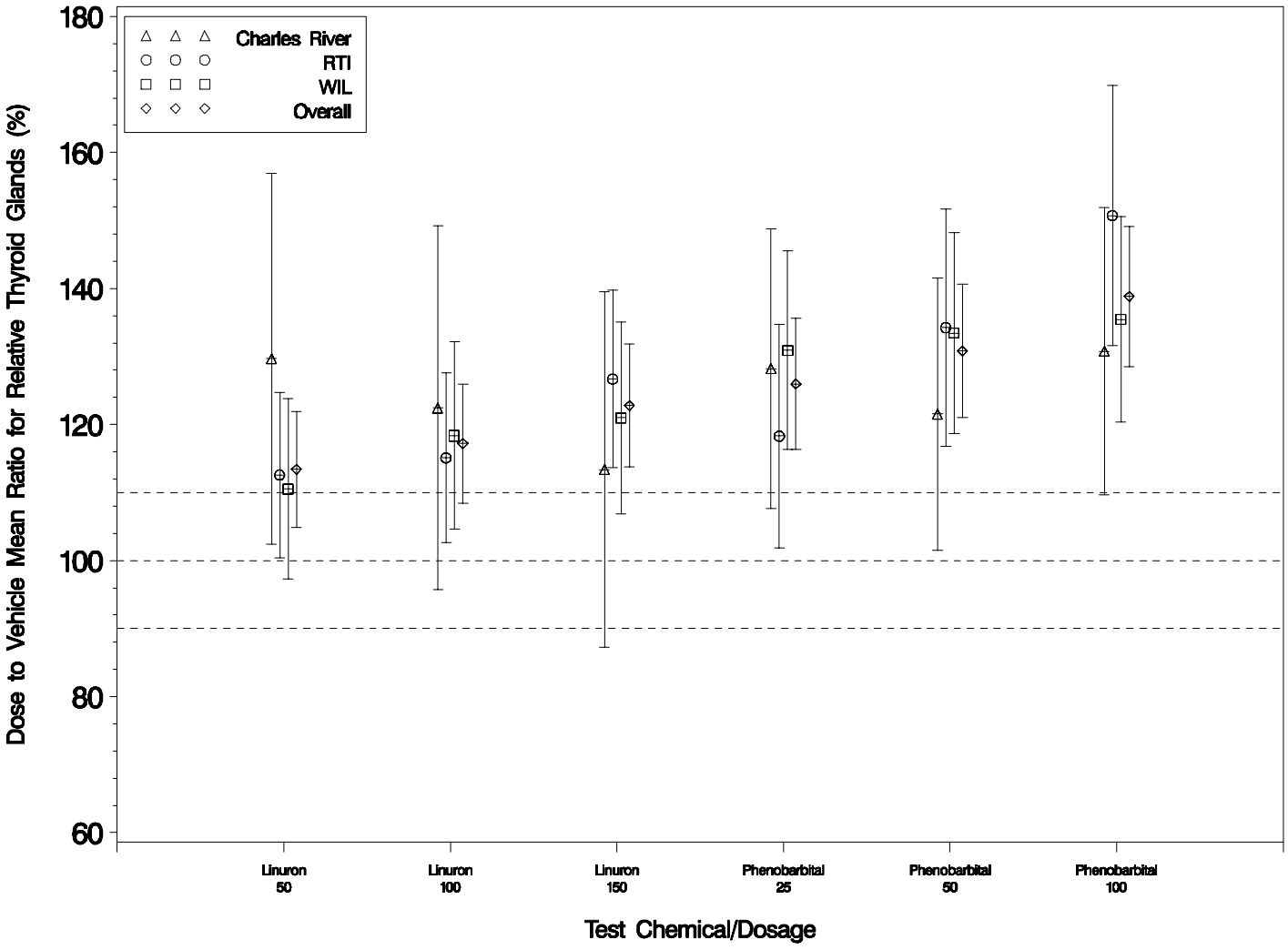
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Res
Adult Intact Male Assay for Entire Prostate to Body Weight Ratio, by Dose Group within Each Laboratory and Across
The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Grou



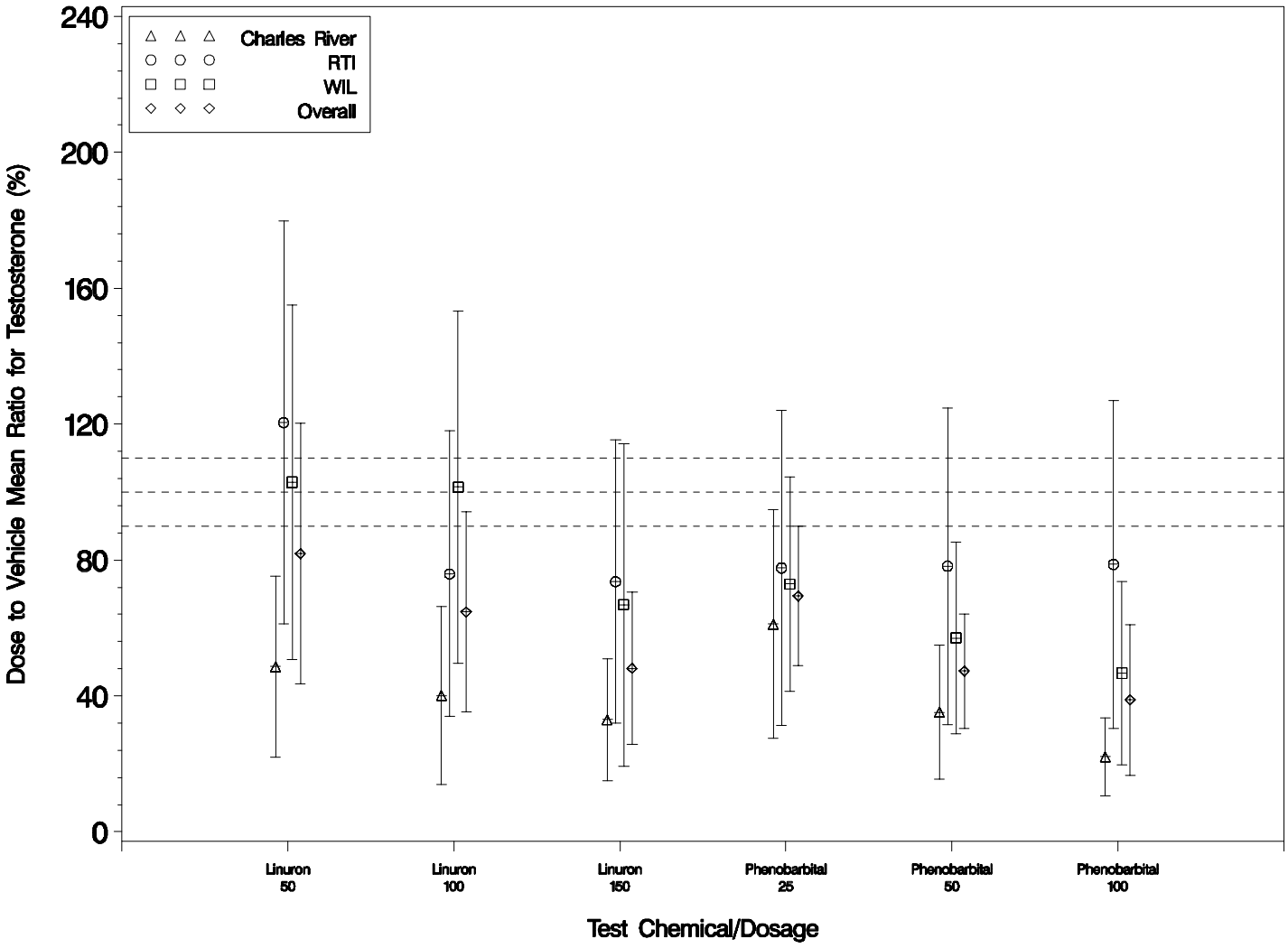
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Seminal Vesicle Plus Coagulating Gland with Fluid (SVCGF) to Body Weight Ratio, by Dos
within Each Laboratory and Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, a
Chemical-Dose Group to Vehicle Group Ratio.**



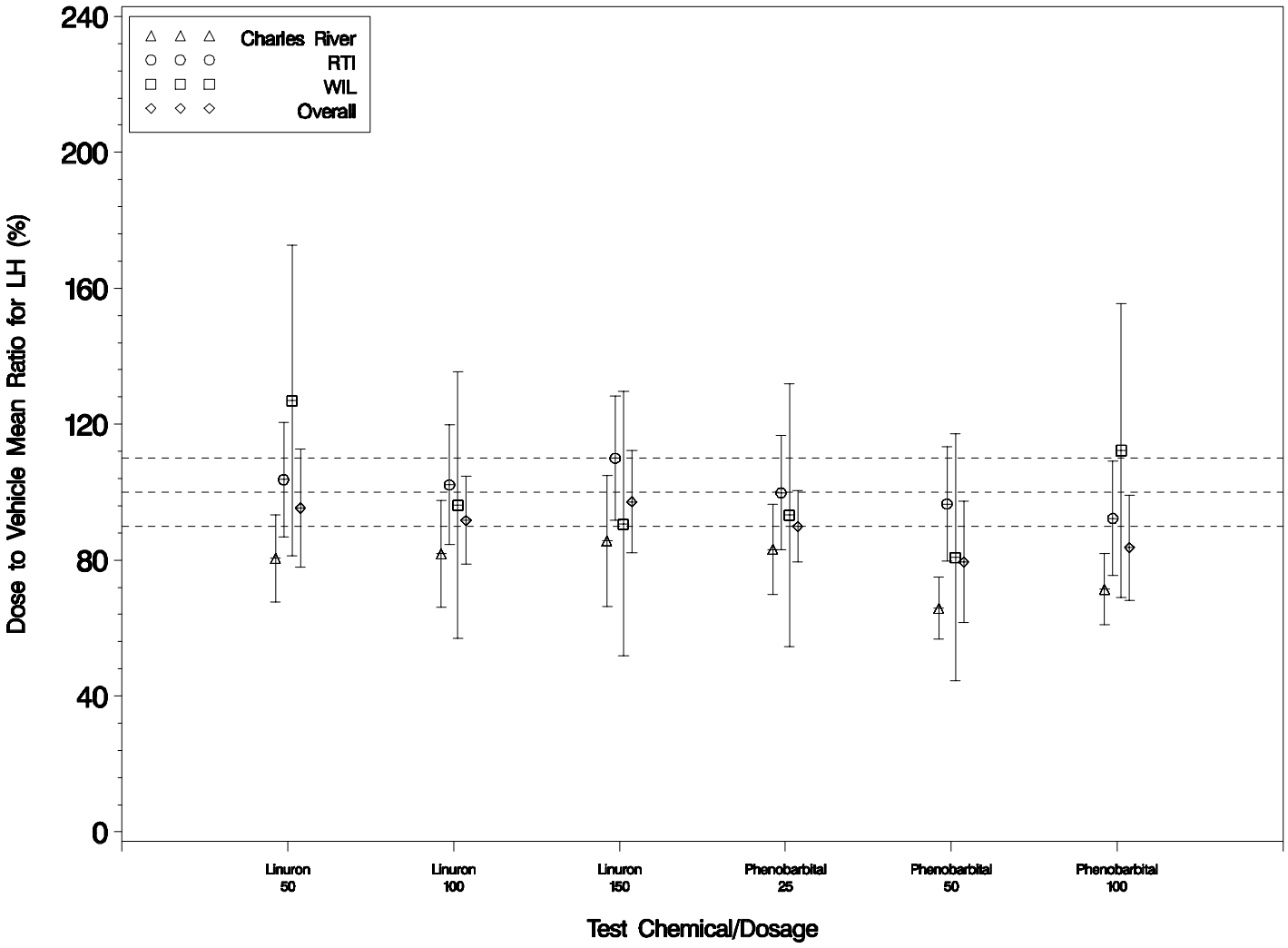
**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Accessory Sex Glands to Body Weight Ratio, by Dose Group within Each Laboratory and A
Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to
Group Ratio.**



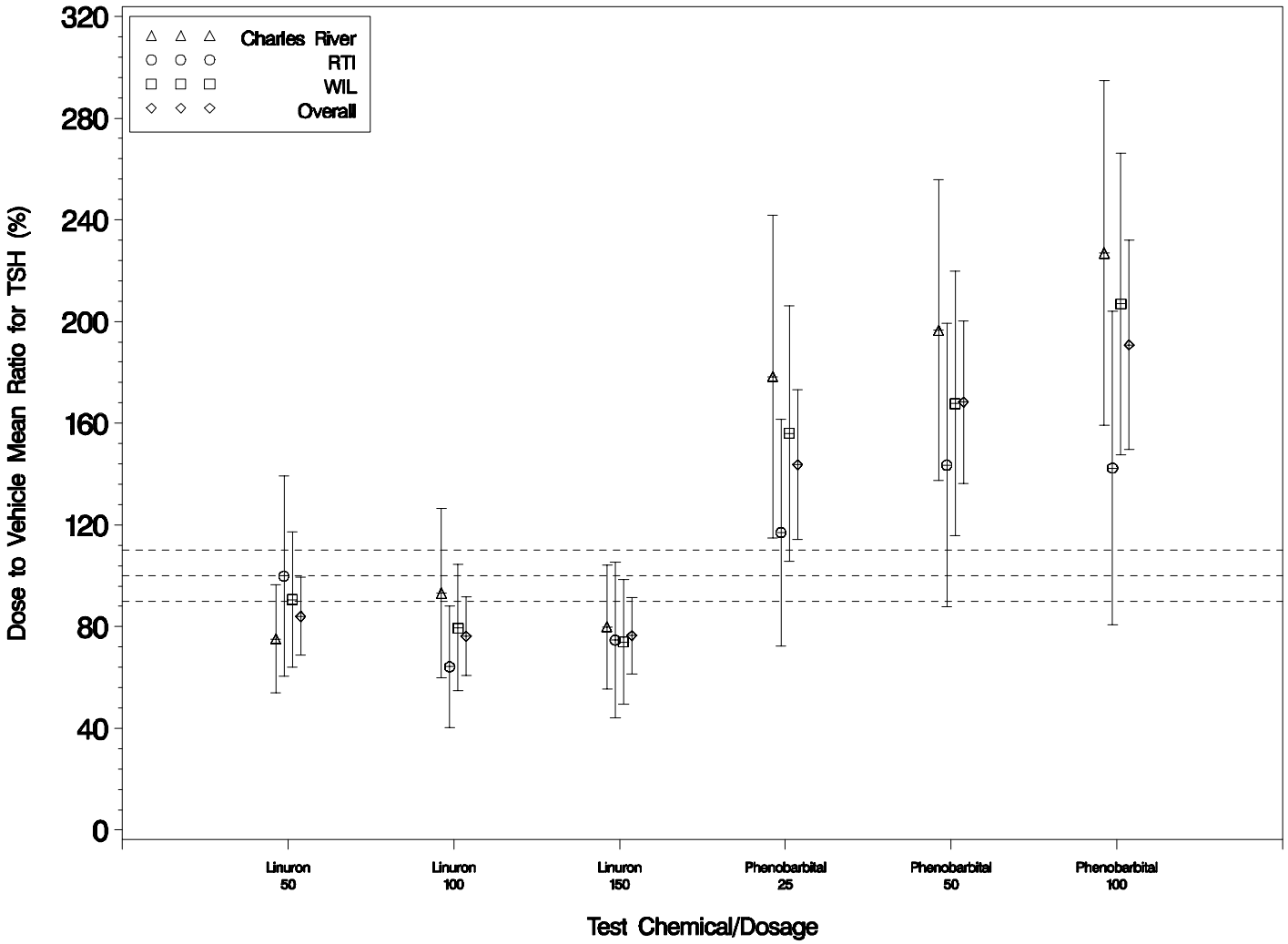
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Thyroid Gland to Body Weight Ratio, by Dose Group within Each Laboratory and Across
The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Grou



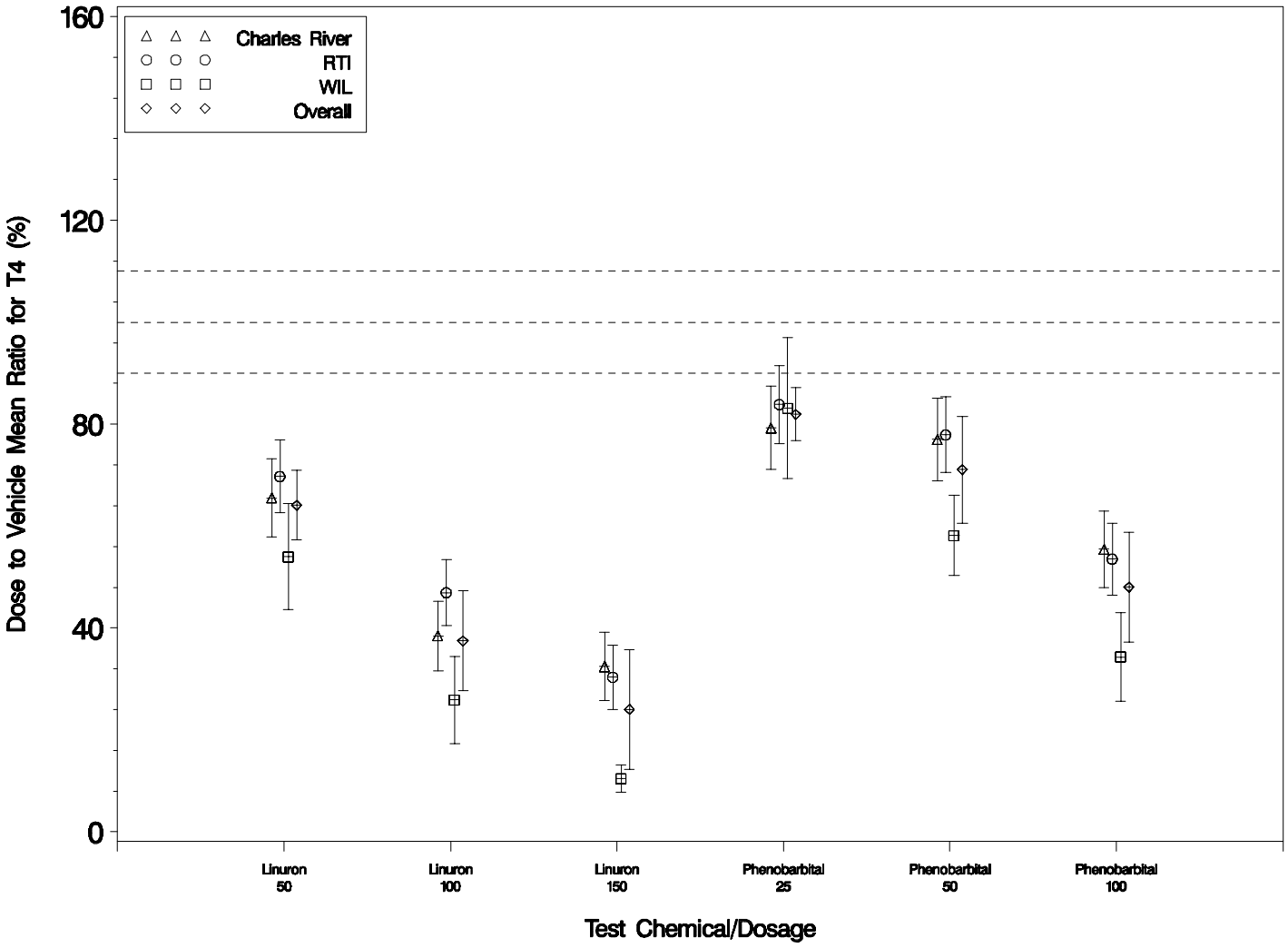
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Testosterone, by Dose Group within Each Laboratory and Across Laboratories. The Three
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



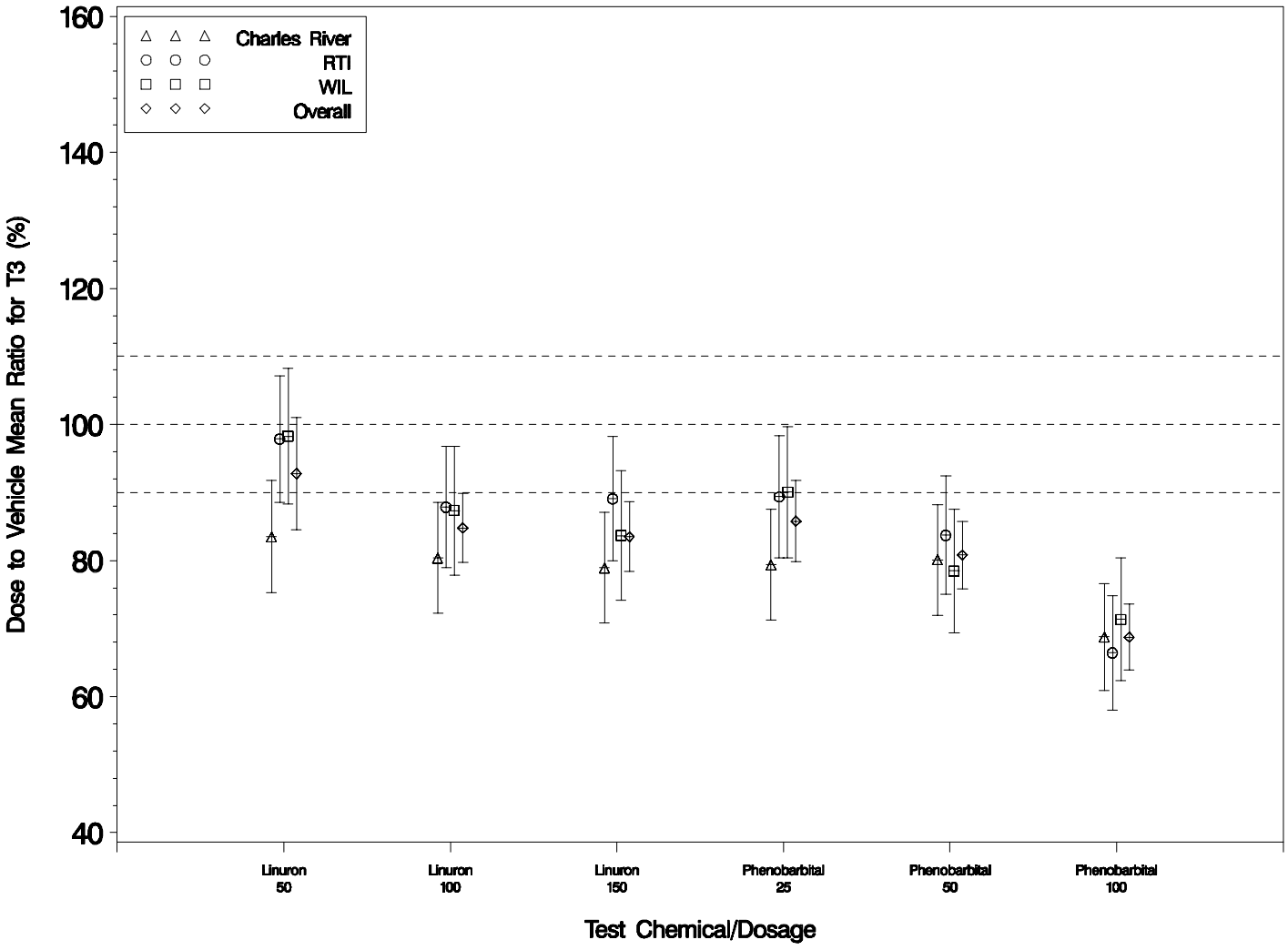
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for LH, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizont
Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



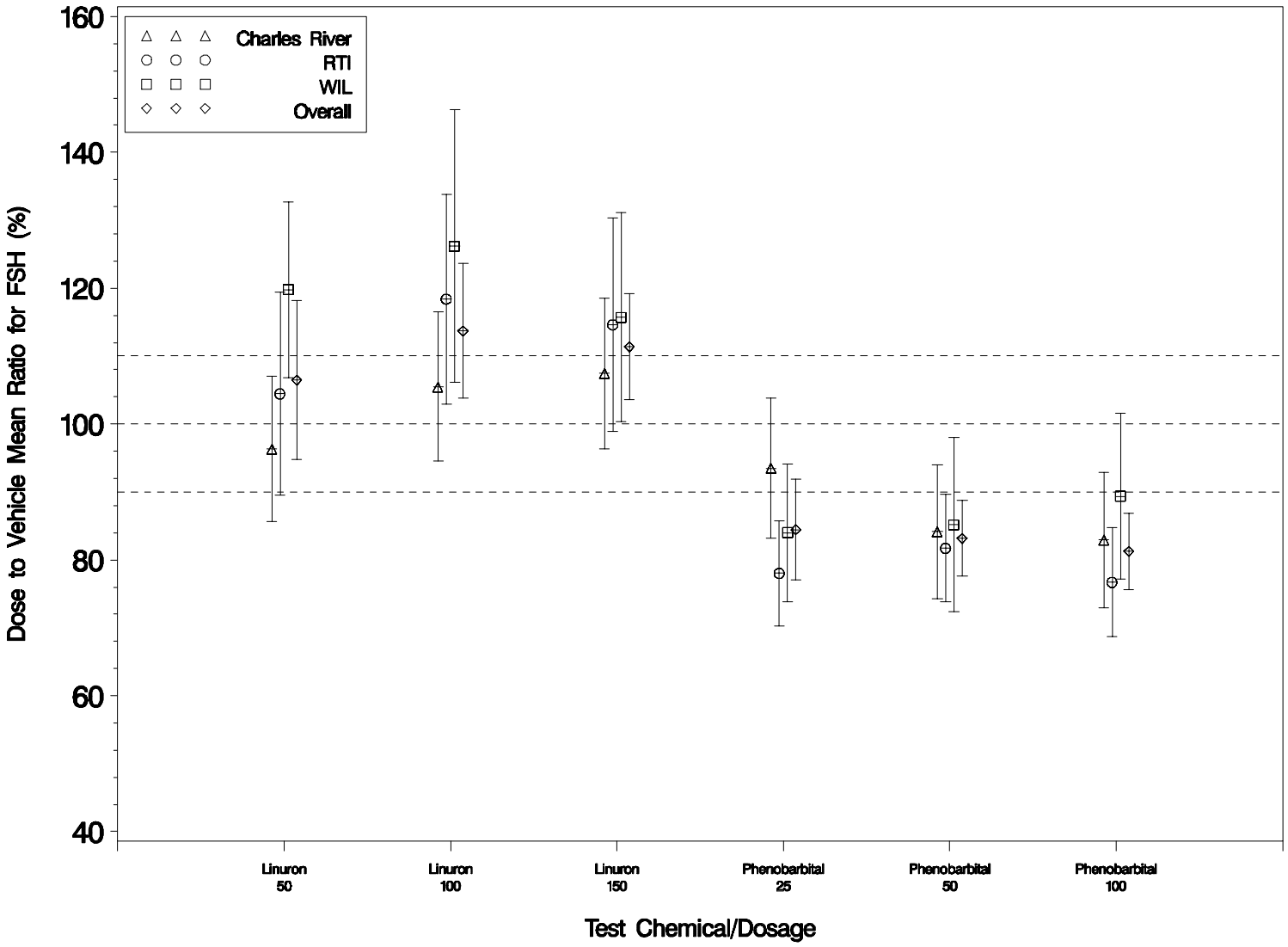
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for TSH, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizo
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



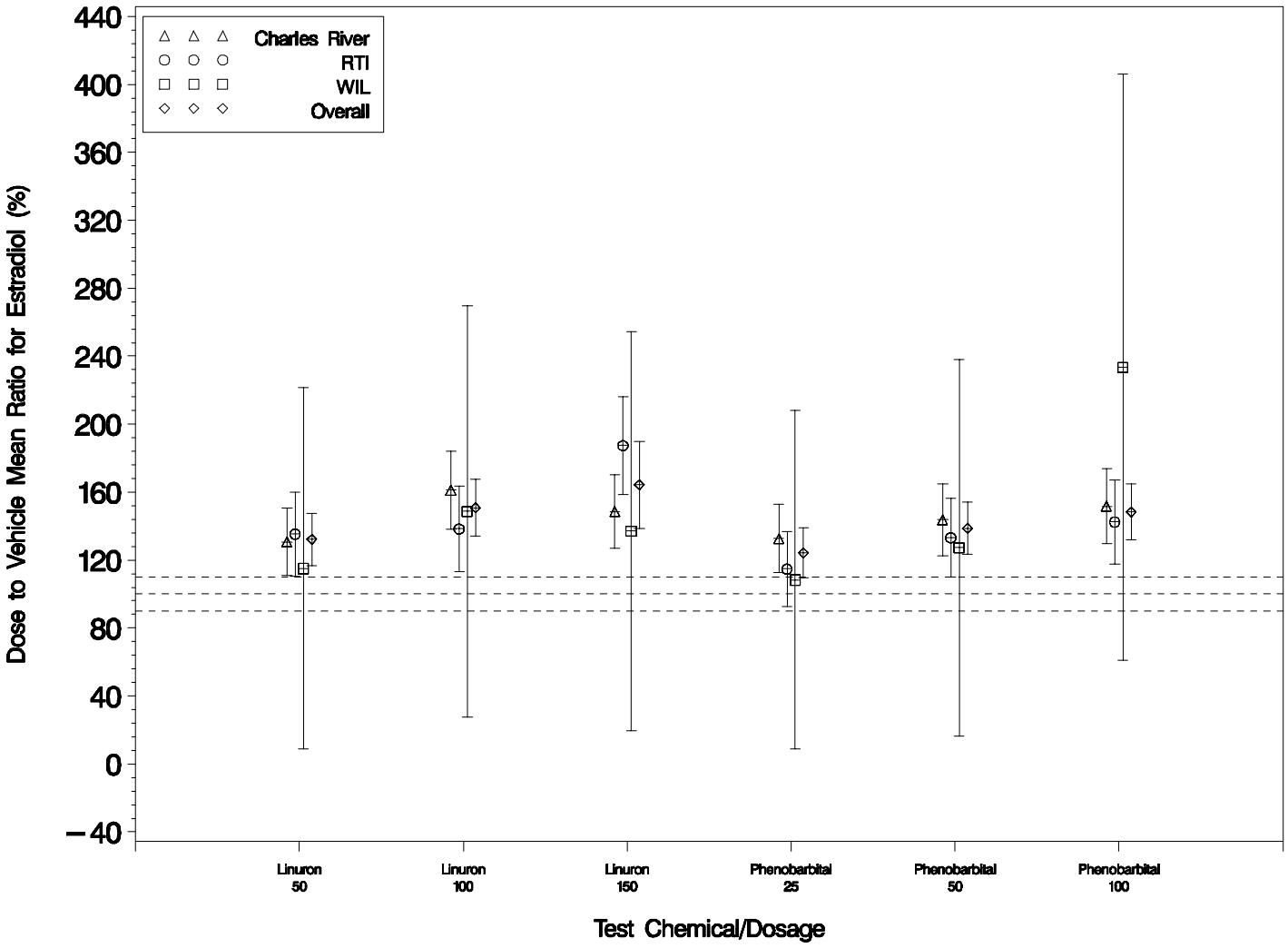
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp Adult Intact Male Assay for T4, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizont Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



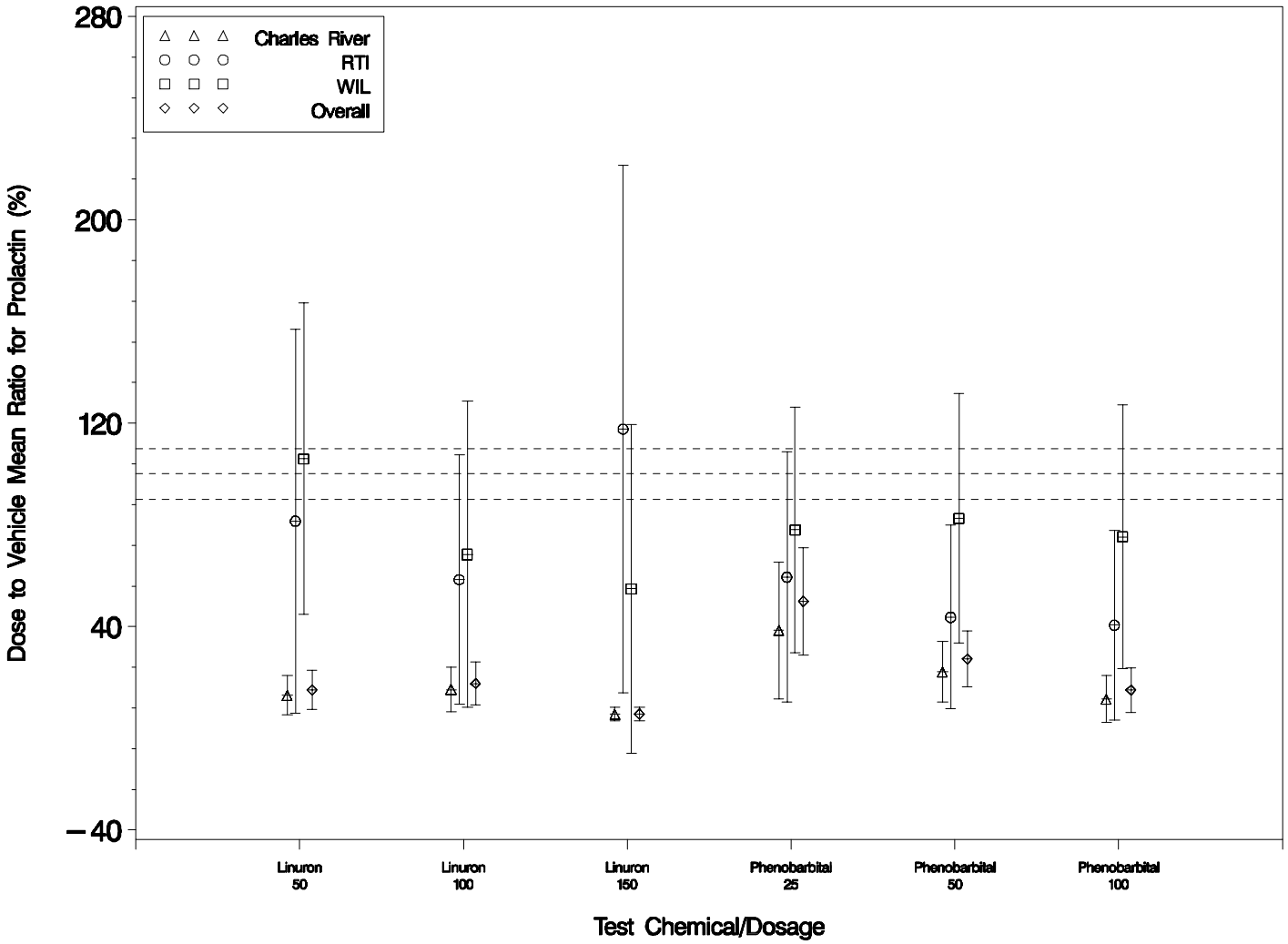
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp Adult Intact Male Assay for T3, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizont Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



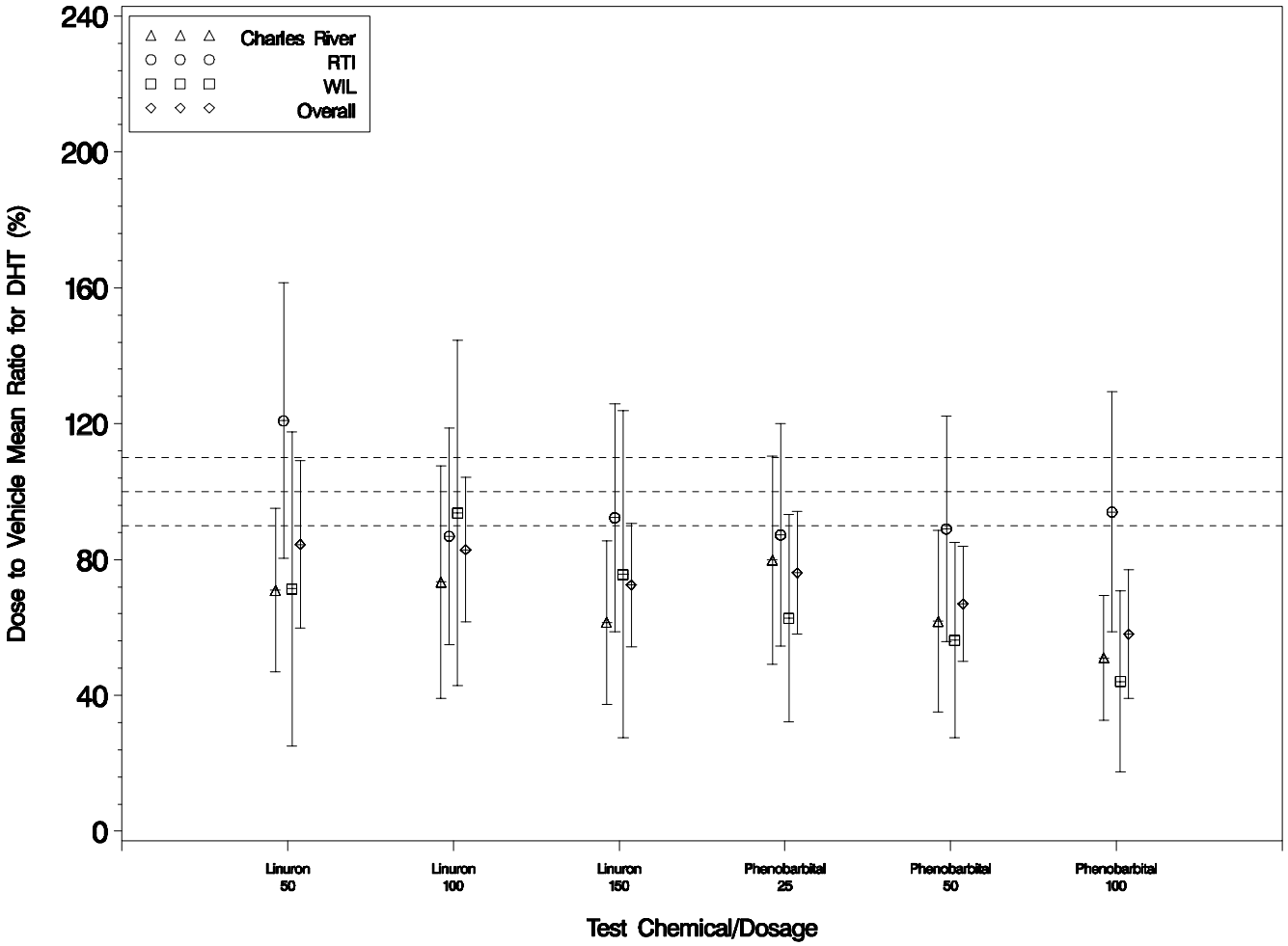
Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for FSH, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizon
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Estradiol, by Dose Group within Each Laboratory and Across Laboratories. The Three Ho
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.



**Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Resp
Adult Intact Male Assay for Prolactin, by Dose Group within Each Laboratory and Across Laboratories. The Three Ho
Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dose Group to Vehicle Group Ratio.**



ure 34. Least Squares Means (with ± 2 Standard Error Bars) of Ratio of Chemical-Dose Group Response to Vehicle Group Response in Adult Intact Male Assay for DHT, by Dose Group within Each Laboratory and Across Laboratories. The Three Horizontal Reference Lines Correspond to 90%, 100%, and 110% of Chemical-Dos Group to Vehicle Group Ratio.

This page left blank intentionally.

APPENDIX C

Final Standardized Protocol for the 15-Day Intact Adult Male Rat Assay

This page left blank intentionally.

FINAL STANDARDIZED PROTOCOL

for

15-DAY INTACT ADULT MALE RAT ASSAY

August 29, 2007

U.S. Environmental Protection Agency
Office of Science Coordination and Policy
Washington, D.C.

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviation	Definition
AR	Androgen receptor
ASG	Accessory sex gland
AWA	Animal Welfare Act
CFR	Code of Federal Regulations
CO_2	Carbon dioxide
CRO	Contract Research Organization
CV	Coefficient of variation
DHT	Dihydrotestosterone
DSL	Diagnostic Systems Laboratory
EAC	Endocrine-active compound
EAT	Estrogen, androgen, and thyroid
EDSP	Endocrine Disruptor Screening Program
ER	Estrogen receptor
FIFRA	Federal Insecticide, Fungicide, Rodenticide Act
FSH	Follicle-stimulating hormone
G	Gram(s)
GLP	Good Laboratory Practices
H&E	Hematoxylin and eosin
kg	Kilogram(s)
LH	Luteinizing hormone
MAFF	Ministry of Agriculture, Forestry and Fisheries (Japan)
MOA	Mode Of Action
MSDS	Material Safety Data Sheet
ND	Not determined
PDF	Portable Document Format (Adobe Systems®)
PRL	Prolactin
QA	Quality assurance
QC	Quality control
RIA	Radioimmunoassay
SD	Sprague-Dawley (rat strain)
SE	Standard error
SOP	Standard Operating Procedure
SVCG	Seminal vesicles and coagulating glands (with fluid)
T_3	Triiodothyronine
T_4	Thyroxine
TD	Test day
TSCA	Toxic Substances Control Act
U.S.C.	United States Code

INTRODUCTION

According to numerous publications as reviewed by O'Connor *et al.*, (2002c), the 15-day intact adult male rat assay has been developed to detect ER agonists/antagonists, AR agonists/antagonists, progesterone agonists/antagonists, steroid biosynthesis inhibitors, gonadotropin and thyroid modulators either directly or indirectly by altering the HPG or -hypothalamus-pituitary-thyroidal (HPT) axes, and prolactin (PRL) modulators through neuroendocrine pathways.

Briefly, the design of the intact adult male rat assay consists of multiple endpoints, principally, terminal weights of primary and secondary sex organs and thyroid gland; histomorphology of the testes, epididymides, and thyroid; and serum concentrations of reproductive steroids, gonadotropins, and thyroid hormones. Results of the comparisons of these endpoints between control and treated groups at three dose levels (n=15 rats/group) administered by oral gavage are evaluated using a weight-of-evidence approach within the bioassay to determine whether a chemical has a positive effect on the EAT hormonal systems. Criteria for interpretation of endocrine-mediated effects within the bioassay are presented in Section 3.3.

The extent of the diversity of this assay to detect effects on the EAT hormonal system using a variety of endocrine-active compounds (EACs) has been hypothesized, tested, and reported in published peer-reviewed scientific journals (O'Connor *et al.*, 1998a,b; 1999a,b; 2000a,b; 2002a,b,c). Within the EDSP, the purpose of the intact adult male screening assay would be to provide overlap and sufficient breadth to cover MOAs, especially AR agonists/antagonists, steroid biosynthesis inhibitors, gonadotropin and thyroid modulators either directly or indirectly through intact HPG or HPT axes, that are not wholly covered or by other assays in the Tier-1 battery. Thus, the results from within the intact adult male rat assay are expected to contribute to the results among assays within a Tier-1 battery and be considered using a weight-of-evidence approach within the battery for detecting estrogenic, androgenic and thyroidogenic effects of potential endocrine disruptors.

<enter additional background information relevant to the purpose of this study>

OBJECTIVE

<enter specific purpose of the study>

PERSONNEL INVOLVED IN THE STUDY

<enter sponsor and contract laboratory information>

STUDY SCHEDULE

Proposed Animal Receipt Date:

Proposed Experimental Start Date:
(First Day of Dose Administration)

Proposed Experimental Termination Date:

Proposed Audited Report Date:

TEST SUBSTANCE

Reserve samples of the test substance will be taken in accordance with the contract research organization (CRO) laboratory's Standard Operating Procedures (SOPs) and stored in the Archives at the CRO laboratory indefinitely, unless otherwise specified.

Personnel safety data are to be provided by the Sponsor. It is the responsibility of the Sponsor to notify the testing facility of any special handling requirements of the test material. A material safety data sheet (MSDS) will accompany the test material upon arrival at the laboratory.

Neat test substances will be provided by the Sponsor.

Identification:

Lot Number:

Purity:

Stability:

Physical Description:

Storage Conditions:

Target Dosages:

Target Dose Concentrations:

Type of Formulations

TEST SYSTEM

Species: Rat

Strain: Sprague-Dawley Crl:CD®(SD)

Source:

Number on Study: 60 (15 rats for each of the 4 dose levels at start of dosing).

Body Weight Range: Approximately 225 to 350 grams (g) at randomization.

Age: Approximately 10 weeks of age at the start of dose administration.

Identification System:

Justification for Selection: The basis for selecting the Crl:CD®(SD) rat is that it is readily available and has often been the animal model of choice for determining general toxicological and, to a lesser extent, endocrinological effects. More recently, SD rats have been used to examine specific endocrine-mediated effects of natural and synthetic compounds on reproduction and thyroid function in intact rodent models. Many laboratories use SD rats for multigeneration studies, including the two-generation reproduction toxicity test currently proposed for the EDSP Tier-2 battery and, therefore, this model will allow for an examination of reproducibility of endpoints common to Tiers 1 and 2 in the same strain of rats. Furthermore, relatively large historical data bases are available for reference.

SPECIFIC MAINTENANCE SCHEDULE

Animal Housing:

Animals will be housed individually in solid-bottom, polycarbonate cages fitted with stainless steel wire lids with appropriate cage bedding or wire-mesh cages.

Environmental Conditions:

Animal rooms will be maintained on a 12:12 hours light:dark cycle. Target conditions for temperature and relative humidity in the animal rooms will be between 64 and 79°F and between 30 and 70%, respectively.

Drinking Water:

Water will be available *ad libitum* through plastic bottles with stainless steel sipper tubes or an automatic watering system.

Basal Diet:

Certified animal feed will be used, guaranteed by the manufacturer to meet specified nutritional requirements. Analysis will include ensuring that heavy metals, pesticides, and phytoestrogens (e.g., genistein, daidzein, and glycitein) are not present at concentrations that would be expected to affect the outcome of the study.

In addition, the following procedures are to be performed periodically to ensure that contaminant levels are below those that would be expected to impact the scientific integrity of the study:

- Water samples are to be analyzed for total bacterial counts, and the presence of coliforms, lead, and other contaminants.
- Feed samples are to be analyzed for the presence of bacteria and fungi.
- Samples from freshly washed cages and cage racks are to be analyzed to ensure adequate sanitation by the cagewashers.

The animal health monitoring program is to be administered by the laboratory animal veterinarian. Data are to be maintained separately from study records and may be included in the Final Report at the discretion of the study director.

EXPERIMENTAL DESIGN

Animal Receipt and Quarantine:

Each rat will be inspected by a qualified technician upon receipt. Rats judged to be in good health and suitable as test animals will be immediately placed in quarantine for a minimum of seven days. All rats will be initially weighed and permanently identified. During the quarantine period, each rat will be observed twice daily for changes in general appearance and behavior and weighed two more times. Prior to the start of the in-life phase, those rats judged to be suitable test subjects will be identified.

Randomization:

At the conclusion of the quarantine period, rats will be released by the laboratory veterinarian as suitable test subjects and meeting acceptable body weight requirements for assignment to the study. Animals will be divided by computerized, stratified randomization based on pre-study body weights into at least four groups (vehicle control group and low-, intermediate- and high-dose) of 15 rats each so that there are no statistically significant differences among group body weight means.

Route and Rationale of Test Substance Administration:

The route of administration will be oral (gavage). Historically, this route has been used extensively for studies of this nature. Appropriate-sized flexible, Teflon®-shafted, stainless steel ball-tipped dosing cannulae will be used for the oral administration by gavage.

Organization of Test Groups, Dosage Levels and Treatment Regimen:

Test Substance and Dose Level Rationale:

All dose levels should be selected by taking into account any existing toxicological data available for the test substance. The highest dose level should take into account all relevant information (e.g., LD_{50}, acute toxicity, and range-finding studies) in order to avoid death, severe suffering, or distress in the animals as well as available information on the MTD in previous subchronic or chronic toxicological studies. In general, the MTD should not cause a reduction in final body weight of the animals greater than 10% of control weight but may need to be

exaggerated to induce an effect on the EAT hormonal system without inducing acute toxicity as discussed in the data interpretation section (O'Connor *et al.*, 1999b; 2000b).

If a single limit dose of at least 1000 mg/kg/day is used and fails to produce statistically significant changes in target organ weights and serum hormone concentrations; and if histopathology of the testes, epididymides, and thyroid gland is not detected; and if the results are interpreted to be unequivocal that there was no effect on the EAT hormonal systems, then additional dose levels may not be necessary.

<enter additional information accordingly>

Organization of Test Groups:

The dosage levels will be provided by the Sponsor.

Group	No. Males	Treatments	Dosage Level (mg/kg/d)[a]	Dosage Volume (ml/kg)
1	15	Vehicle control[b]	0	5
2	15	Low	TBD	5
3	15	Intermediate	TBD	5
4	15	High	TBD	5

TBD=to be determined

[a]Test compounds administered once daily by oral gavage on Test Days (TD) 1 through 15.

[b]0.25% aqueous methylcellulose, vehicle only

Vehicle Control Substance:

0.25% methylcellulose in water.

Treatment Regimen:

The suspensions of formulated test and vehicle control substances will be administered once daily by oral gavage for 15 consecutive days (TD 1 through TD 15). Prior to dose administration, all formulations to be used for dosing that day will be removed from refrigeration and placed on a stir plate to vortex for at least 45 minutes to equilibrate to room temperature. The formulations will continue to be stirred throughout dose administration.

Animals will be dosed beginning early in the morning so that at termination blood collection and necropsy can be completed within a 2- to 3- hour window after the last dose on TD 15 in the morning hours. Typical necropsy times used in previous experiments were from 0700 to 1000 hours and, in some instances, the laboratories staggered the start of the study in a manner across

dose groups to accommodate the number of animals scheduled for necropsy within a defined time (2 to 3 hours) after administration of last dose on TD 15.

Adjustment of Dosages Levels:

Individual doses will be calculated based on each daily body weight to provide the proper dosage except on TD 15, which will use the previous day's weight (TD 14). Individual animal body weights and individual animal dosages will be recorded.

Dosage Preparation and Analysis

Method and Frequency of Preparation:

Formulation of dosage levels will be done in general accordance with Sponsor-provided formulation instructions. The dosages may be corrected for purity according to the sponsor. Dosages of each test substance will be prepared according to stability, aliquoted into amber bottles per group, sampled as described below, and stored refrigerated. The Study Director or designee will visually inspect all formulations. This visual inspection will be performed to ensure that the formulations are visibly homogeneous.

Homogeneity, Stability, and Concentration of Test Substance Formulation:

The development of stability data of test substance formulations (bracketing dose levels used on study) is the responsibility of the Sponsor and appropriate documentation of stability will be provided to the Study Director and included as an appendix to the Final Report.

From the dosage preparations, samples will be collected from the top, middle and bottom of each dosage level (extra samples should be made for backup) prior to TD 1 and again on TD 15 from the TD 15 dosing aliquot. The first set of samples will be analyzed to confirm homogeneity and concentration of the test substance formulations. The concentrations will be within 10% of target and homogeneity will have a percentage difference between top and bottom concentrations of 5% or less before they will be approved for dosage administration. The TD 15 analysis will be analyzed for resuspension homogeneity.

The contract laboratory will conduct an analytical method transfer validation for each test substance based on validated methods provided by the Sponsor. The Sponsor will provide detailed formulation, sampling, and analytical method instructions for guidance. Any samples or backup samples will be discarded after issuance of the Final Report.

General Observations during the Experimental Period

Clinical Signs:

The rats will be observed twice daily for appearance, behavior, moribundity, and mortality. A detailed physical examination will be conducted on the day of randomization and daily prior to

dosage administration. Observations shall include, but are not limited to, evaluations for changes in appearance of the skin and fur, eyes and mucous membranes, respiratory, circulatory, autonomic and central nervous system functions, somatomotor activity and behavior patterns. Observations will be recorded. During the treatment period, the rats will also be observed according to a designated time provided by the Sponsor following dosing and the observations will be recorded.

Body Weights:

Body weights will be recorded individually on a daily basis from TD 1 to TD 14 (day prior to necropsy), inclusively. Note, final body weight on TD 15 will be the live weight before euthanasia.

Food Consumption:

Food consumption data will be recorded individually on a weekly basis from TD 1 to TD 14 (day prior to necropsy), inclusively. Food intake will be reported as g/animal/day and g/kg/day for each corresponding body weight interval.

***Deaths and Animals Euthanized* in Extremis:**

Animals not surviving until the scheduled euthanasia will be necropsied and cause of death recorded, if possible. Rats not expected to survive to the next observation period (moribund) will be euthanized by carbon dioxide (CO_2) inhalation and subjected to gross necropsy. Tissues with unusual gross findings will be preserved in 10% neutral buffered formalin. All carcasses will be discarded.

POST-MORTEM EXAMINATION

Blood Collection and Macroscopic Examination:

On the morning of TD 15 following dosing, all surviving study animals will be moved to the necropsy holding room at least 1 hour before euthanasia of study animals begins (to minimize stress-induced changes in hormone levels related to cage transport). Animals will not be fasted prior to euthanasia. Rats will be euthanized by decapitation with prior anesthesia using CO_2 for approximately 60 seconds; time of euthanasia will be recorded. Blood will be collected via the site of decapitation as described below. Rapid euthanasia is necessary because of the likelihood that prolonged anesthesia or stress associated with the administration of anesthesia will interfere with the accurate measurement of the various hormones that are essential endpoints with this assay (Holson, 1992).

The order in which animals will be necropsied for blood and tissue collection will be stratified across all groups. Euthanasia for all animals should occur between 0700 and 1000 hour (2 to 3 hours after final dose) in order to minimize variability associated with serum hormone measurements.

Immediately following euthanasia, trunk blood will be collected (target volume of 3 mL) into a serum separator tube. Tubes containing blood will be kept cold (e.g., on ice) until serum is prepared. Blood samples will then be centrifuged for isolation of serum. Aliquots of serum will be made based on the number of different assays that will be run in a day to minimize the potential freeze and thaw effect on hormone concentrations. Serum will be stored in a freezer set to maintain ≤ -65°C for subsequent hormone analyses. Extra serum will be stored at ≤ -65°C. Remaining serum samples will be discarded after the final report has been issued to the Sponsor.

The necropsy examinations will include the external surface, all orifices, the external surface of the brain and the thoracic, abdominal, and pelvic cavities including viscera. Organs/tissues to be weighed and preserved are described below. Tissues with gross findings will be preserved in 10% neutral buffered formalin, if possible (unless a different fixative is specified below) for possible histological examination.

Organ Weights:

The following tissues will be weighed (to the nearest mg) from all animals:

Accessory sex gland (ASG)[1]	Testes[3]
Entire prostate[2]	Epididymides[3]
Seminal vesicles with coagulating gland containing fluid (SVCG)[3]	Liver
Thyroid[4]	

[1] Entire prostate, seminal vesicles, and coagulating gland with fluid
[2] Dorsolateral and ventral prostate
[3] Weighed as paired organs
[4] Weighed following fixation and dissection

With the exception of the thyroid trimming described below, organ harvesting and weighing procedures will be divided as equally as possible among the prosecting and weighing technicians, such that all animals from a group are not processed by a single individual (operator number will be recorded) in order to minimize systematic bias in the weighing procedures.

Tissue Fixation and Processing:

The testes will be placed in Bouin's fixative for approximately 24 hours, after which they will be rinsed and stored in 70% alcohol until histological processing. The epididymides and liver from each rat will be placed in 10% neutral buffered formalin. The thyroid, with attached trachea, is fixed in 10% neutral buffered formalin for at least 48 hours. Afterwards, the thyroid is dissected under a dissecting microscope from the trachea, blotted, weighed, and placed in 10% neutral buffered formalin until histological processing. The fixed thyroid dissection will be performed by one individual in order to reduce the variability of the dissection procedure with the expectation to minimize the variability of thyroid weights. The testes, epididymides, and the thyroid from the control and high-dose animals are then embedded in paraffin, sectioned, and stained with hematoxylin and eosin (H&E) for subsequent histological evaluations (slide

preparation of respective tissues from the lower dose groups is at the discretion of the Sponsor). Sections of 2 to 5 microns will be made for the testis (transverse), epididymis (longitudinal), and thyroid (positioning according to the laboratory's governing SOP).

Microscopic Evaluation:

Testes, epididymides, and thyroid gland histomorphology from the control and high-dose groups will be evaluated for pathologic abnormalities and potential treatment-related effects. A minimum of two sections for the thyroid and a sufficient number of sections for each testis and epididymis shall be examined. The type, incidence and degree of severity of histomorphologic changes will be recorded, especially the height of the follicular epithelium and colloidal area of the thyroid gland.

Microscopic evaluations on tissues from lower dose groups will be done by protocol amendment if necessary. Liver will be evaluated microscopically at the discretion of the study pathologist or Study Director and Sponsor.

Histological interpretation will be done by a board-certified veterinary pathologist knowledgeable of the control and high-dose groups but not the nature of the chemicals until after the evaluation has been completed. Peer review by a second board-certified veterinary pathologist is highly recommended since this bioassay has assigned histopathological effects associated with the testes, epididymides, and thyroid gland a high priority in the weight-of-evidence approach within the bioassay to determine test-article related effects on the EAT hormonal systems.

HORMONE ASSAYS

The following hormones will be analyzed from serum samples from all animals:

Testosterone (ng/ml)	Luteinizing Hormone (LH, ng/ml)
Estradiol (pg/ml)	Prolactin (PRL, ng/ml)
Dihydrotestosterone (DHT, pg/ml))	Thyroid-Stimulating Hormone (TSH, ng/ml)
Follicle-Stimulating Hormone (FSH, ng/ml)	Thyroxine (T_4, ug/dl)
	Triiodothyronine (T_3, ng/dl)

All hormones will be measured using commercially available radioimmunoassay (RIA) kits (Biotrak™, Amersham Biosciences and Diagnostic Systems Laboratory). Model numbers and, perhaps, other sources of assay kits will be specified by the Sponsor. The sequence in which the hormones should be assayed is testosterone, LH, TSH, T_4, T_3, FSH, estradiol, PRL, and DHT but this sequence may be rearranged by the Sponsor according to the nature of the test material and availability of serum. If serum is limiting, the Study Director should contact the Sponsor to establish a priority list of hormones to be measured.

Each assay will be run according to the manufactures instructions and include all samples from the control group and each treatment group, except for reanalysis of specific samples that may be

out of range of the reference standard curve. Each reference standard and serum sample will be run in duplicate. Each assay will also include high and low quality control (QC) samples run in duplicate and replicate at the beginning, middle, and end of each assay. The QC standards for rat FSH, LH, TSH, and PRL can be obtained from the National Hormone and Pituitary Program, and QC standards for testosterone, estradiol, DHT, T_4, and T_3 can be obtained from a commercial supplier and designated "non-kit QC standards". For the non-kit QC samples, the buffer/medium in which the reference standards are prepared (e.g., zero control standard) will be spiked with respective hormones at concentrations that are expected to be within 70% B/B_0 (±10%) and 30% B/B_0 (±10%) of the reference standard curve (i.e., linear portion of the standard curve). The assay kits may contain "kit QC standards" that will be run according to the manufacturer's instructions. Coefficients of variation (CV) for within- and between-assays will be calculated from the non-kit and kit QC standards and reported. Assay sensitivity will be calculated according to the manufacturer's instructions and reported.

DURATION OF STUDY

The duration of the in-life phase of this study will require approximately 3 weeks for acclimation, dosing and necropsy.

STATISTICAL METHODS

Endpoints for the statistical analysis described below include the following:

TD 15 body weight
Body weight change, TD 1 to 8, 8 to 15, 1 to 15
Food consumption (g/kg/day only), TD 1 to 8, 8 to 15, 1 to 15
Absolute organ weights (7 total)
Organ weights relative to final body weight (7 total)
Hormones concentrations (9 total).

Note, ASG weights (entire prostate plus the SVCG) will be calculated per animal and analyzed.

Based on 7 absolute organ weight values, 7 relative organ weight values, 9 possible hormones, and the 7 body weight and food values, 30 possible endpoints are to be evaluated statistically. A test for extreme or outlying values (Grubbs, 1969) and an evaluation of normality will be carried out prior to statistical analyses. Tests for homogeneity of variance will be carried out on the data (excluding values identified as potential outliers) using a one-way analysis variance (ANOVA) model fitted to the data including the fixed factor of treatment and the residual replicate per treatment. Following the homogeneity of variance evaluation, transformation of the data may be performed as appropriate to minimize the degree of heterogeneity of the data. Subsequent analyses will be carried out based on transformed data.

A one-way ANOVA model will also be fitted to the data to estimate treatment effects for each endpoint described above. Probability values will be indicated for each endpoint where the level

of significance will be two-tailed (two-sided) at 0.05 and 0.01. The data used for the above-described analyses will exclude potential outliers and may be performed on transformed versions of the variables. The factors in the ANOVA models will include treatment and residual replicate (treatment). Linear trend statistics will also be evaluated for each endpoint using the means of two-sample t-tests at 0.05 and 0.01. Summary statistics will be transformed back to the original scale for the purposes of data presentation.

DATA SUMMARY

The following tables and figures for each test substance along with the respective control will be provided:

Tables

The first set of tables will display summary values for the final live body weight (TD 15), body weight change intervals (TD 1 to 8, 8 to 15 and 1 to 15), and food consumption (g/kg/day) intervals (TD 1 to 8, 8 to 15, 1 to 15). For each endpoint and each dose and control group the following will be reported:

- Number of animals per group
- Mean +/- standard error (SE)
- Coefficient of variation (CV)
- Mean as a percent of control group mean +/- SE
- P-value.

In addition, the linear trend slope contrast will be estimated for each test substance based on the control group and the three graded dose groups. The estimated treatment slope and its SE will be reported.

The second set of tables will display summary statistics described above for the nine absolute organ weights.

The third set of tables will display summary statistics described above for the nine relative organ weights (ratio of organ weight to final body weight).

The fourth set of tables will display summary statistics described above for the nine hormones.

Figures

The first set of figures will display, in a line graph, the mean body weight for each TD from TD 1 through TD 15 for the control group and each of the three dose levels per test substance.

The second set of figures will display, in a scatter plot, the TD 15 mean absolute body weight, the three mean body weight change intervals, and the three mean food consumption intervals listed above, +/- 2 SE.

The third set of figures will display, in a scatter plot, the mean absolute organ weight for each organ, +/- 2 SE.

The fourth set of figures will display, in a scatter plot, the mean relative organ weight for each organ, +/- 2 SE.

The fifth set of figures will display, in a scatter plot, the mean hormone concentration for each hormone (as applicable), +/- 2 SE.

INTERPRETATION OF ENDOCRINE-MEDIATED EFFECTS DATA

Effect of final body weight on target organ weight and hormone concentrations

Interpretation of changes in target organ weights and histomorphology as well as serum hormone concentrations are expected to be interpreted in the context of final body weight decrements according to results obtained in dietary restriction experiments conducted during prevalidation of the intact adult male rat (O'Connor *et al.*, 1999b; 2000b).

The first consideration in this series of studies was to determine the dependency of target organ weight on final body weight. As shown in Table C-1, relative (organ-to-body weight ratio) testis and epididymal weights significantly increased in association with a ≥10% decrease in final body weight in the feed-restricted animals compared to the *ad libitum*-fed controls, whereas absolute testis and epididymal weights were not significantly different between the feed-restricted animals and the *ad libitum*-fed control animals until a body weight decrement of 26% was reached. In contrast, the thyroid, ASG (total prostate plus SVCG), SVCG and prostate were considered body-weight dependent since relative organ weights did not change significantly between feed-restricted animals and the *ad libitum*-fed control animals throughout a 26% decrement in final body weight. While both absolute and relative liver weights were affected by dietary restriction, relative liver weight corrected for most of the body weight decrement. This was in keeping with the generally accepted theory that liver weight is body weight dependent and that expression on a relative to body weight basis will correct for body weight decrements (Feron *et al.*, 1973). Thus, when evaluating target organ weight data following chemical exposure using the 15-day intact adult male rat assay, weights of the testes and epididymides should be evaluated on an absolute organ weight basis, and weights of the liver, thyroid, ASG, SVCG, prostate glands should be evaluated on a relative to final body weight basis in order to optimize interpretation of endocrine-related effects.

.

Table C-1. Mean (±SE) effect of dietary restriction on final body and target organ weights in the intact adult male rat assay (O'Connor *et al.*, 1999b; 2000a).

Feed/day (grams)	Final body (grams)	Final body weight (% control)	Liver	Thyroid	Testes	Epididymides	Accessory sex gland	Seminal vesicles	Prostate
Absolute organ weights (g)									
ad libitum[a]	414 ± 6[b]	100	16.0 ± 0.4	0.025 ± 0.001	3.3 ± 0.1	1.14 ± 0.02	2.3 ± 0.1	1.6 ± 0.1	0.617 ± 0.021
22	373 ± 4*	90	13.4 ± 0.1*	0.021 ± 0.001*	3.2 ± 0.0	1.11 ± 0.03	2.0 ± 0.1	1.5 ± 0.1	0.555 ± 0.031
19	351 ± 3*	85	12.0 ± 0.2*	0.019 ± 0.001*	3.3 ± 0.1	1.08 ± 0.02	1.8 ± 0.1*	1.2 ± 0.1*	0.529 ± 0.034
16	328 ± 3*	79	10.5 ± 0.2*	0.019 ± 0.001*	3.2 ± 0.1	1.11 ± 0.01	1.8 ± 0.1*	1.3 ± 0.1*	0.524 ± 0.039
13	307 ± 2*	74	9.8 ± 0.1*	0.019 ± 0.001*	3.2 ± 0.1	1.06 ± 0.02*	1.6 ± 0.1*	1.1 ± 0.1*	0.454 ± 0.029*
Relative organ weights (% body weight)									
ad libitum[a]	414 ± 6[b]	100	3.9 ± 0.1	0.006 ± 0.0003	0.79 ± 0.02	0.276 ± 0.006	0.552 ± 0.018	0.396 ± 0.017	0.149 ± 0.005
22	373 ± 4*	90	3.6 ± 0.1*	0.006 ± 0.0003	0.86 ± 0.01*	0.296 ± 0.007*	0.548 ± 0.020	0.394 ± 0.016	0.149 ± 0.008
19	351 ± 3*	85	3.4 ± 0.1*	0.006 ± 0.0003	0.94 ± 0.02*	0.308 ± 0.005*	0.504 ± 0.020	0.350 ± 0.020	0.150 ± 0.009
16	328 ± 3*	79	3.2 ± 0.0*	0.005 ± 0.0003	0.97 ± 0.02*	0.338 ± 0.004*	0.561 ± 0.036	0.411 ± 0.026	0.160 ± 0.012
13	307 ± 2*	74	3.2 ± 0.0*	0.006 ± 0.0003	1.04 ± 0.02*	0.344 ± 0.006*	0.516 ± 0.022	0.364 ± 0.018	0.148 ± 0.010

[a] *Ad libitum* control rats consumed 25.8 g/day.
[b] Mean ± standard error.
* Significantly different ($p < 0.05$) from control by Dunnett's Test. (n=15 animals/feed group)

A second consideration in this series of studies was to determine the degree of body weight loss that can occur before target organ weights and serum hormone concentrations are secondarily affected by an extreme decrease in final body weight that may be indicative of acute toxicity or overexposure to chemical treatment (O'Connor *et al.*, 1999b; 2000b). As shown in Table C-1, absolute weight of the testes and epididymides and relative weights of the liver, thyroid, ASG, SVCG, and prostate were not significantly different between feed-restricted and *ad libitum*-fed control animals until a decrement in final body weight of ≥26% was reached. As shown in Table C-2, serum hormone concentrations were not significantly different between feed-restricted and *ad libitum*-fed control animals until a final body weight decrement of 15% was reached for T_3 and T_4, 21% for estradiol and DHT, and ≥26% for PRL, FSH, LH and TSH. Although targeting a final body weight decrement in the high-dose group in the intact adult male rat assay of around 10% of control at the time of euthanasia minimizes the potential for confounding secondary effects due to acute toxicity or overexposure of treatment, final body weight decrements from 15 to 20% relative to controls may be acceptable for interpretation of endocrine-mediated effects on some target organs, histomorphology and serum hormones.

Table C-2. Mean (±SE) effect of dietary restriction on serum hormone concentrations in the intact adult male rat assay (O'Connor *et al.*, 1999b; 2000a).

Feed/ day (grams)	Final body (% of control)	Estradiol (pg/ml)	Testos-terone (ng/ml)	Dihydro-testosterone (pg/ml)	Prolactin (ng/ml)	Follicle stimulating hormone (ng/ml)	Luteiniz-ing hormone (ng/ml)	Thyroid stimulating hormone (ng/ml)a	T_3 (ng/dl)[a]	T_4 (µg/dl)[a]
ad libitum[b]	100	3.5 ± 0.5[c]	11.1 ± 1.4	162.3 ± 25.4	17.9 ± 2.9	13.1 ± 0.7	4.4 ± 0.3	17.3 ± 1.3	80.7 ± 4.0	4.3 ± 0.2
22	90	3.9 ± 0.6	11.9 ± 1.2	175.3 ± 19.6	11.8 ± 1.5	14.9 ± 0.9	5.2 ± 0.4	17.0 ± 1.8	79.9 ± 3.6	4.0 ± 0.2
19	85	3.7 ± 0.8	12.9 ± 1.2	176.3 ± 32.6	16.5 ± 2.3	13.4 ± 0.6	4.8 ± 0.3	16.7 ± 1.5	68.1 ± 3.7#	3.6 ± 0.2#
16	79	1.6 ± 0.4#	12.8 ± 1.6	81.3 ± 14.0#	9.9 ± 1.4	13.9 ± 0.6	5.1 ± 0.3	14.1 ± 1.1	70.5 ± 3.7#	3.2 ± 0.2#
13	74	0.9 ± 0.3#	ND	60.6 ± 12.9#	10.1 ± 2.1#	12.8 ± 0.7	5.1 ± 0.3	10.8 ± 1.5#	60.8 ± 2.8#	3.1 ± 0.2#

[a] Data from O'Connor *et al.* (1999b).
[b] *Ad libitum* control rats consumed 25.8 g/day.
[c] Mean ± standard error.
ND – not determined due to a lack of serum for analysis.
Significantly different ($p < 0.05$) from control by Jonckheere's test for trend. (n=15 animals/feed group).

Thus, interpretation of whether the results of chemical exposure are endocrine-related involves consideration of whether weight changes of target organs are affected on an absolute or relative basis and whether the final body weight decrement is within the limits of interpretation of an endocrine-related effect rather than an acute toxic effect secondary to an extreme decrease in final body weight during treatment.

Priority of endpoints for interpretation of results

Weight changes and histopathology of target organs are expected to carry a heavier weight of evidence within the intact adult male assay than changes in serum hormone concentrations alone to indicate whether a substance affects the EAT hormonal system. That is, hormonal changes alone are of insufficient weight within the bioassay to make a conclusion. An increased incidence of histopathologic alterations of the testes, epididymides or thyroid gland in treated animals compared to controls would be an indication of a compound-induced effect independent of effects on target organ weights or serum hormone concentrations. However, statistically significant changes in respective organ weights and related hormones between treated and control groups would add weight-of-evidence within the assay to the histopathological results, and also allow differentiation of MOA based on the pattern of the effects. Statistically significant target organ weight changes alone would also be considered compound-related with a relatively high degree of confidence if the results correspond to a significant linear trend indicating that the results are dose-dependent. If the linear trend analysis is not significant, it is possible that a significant difference between treated and control groups at any dose level is spurious and not compound-related; however, a weight-of-evidence approach among the multiple endpoints within the assay combined with biological plausibility can help distinguish compound-related from spurious alterations of an endpoint result.

Statistically significant changes in serum hormone concentrations are expected to support target organ weight and histopathological changes as well as provide additional information to differentiate between various MOAs for unknown chemicals. Instances when only serum

hormone concentrations are significantly altered will not be considered sufficient evidence alone within the assay to identify a positive endocrine test result but, perhaps, may be considered relevant in a weight-of-evidence approach between or among assays when interpreting the entire EDSP Tier-1 screening battery. In addition, if the results among the endpoints for organ weights and histomorphology are equivocal with respect to an effect on the endocrine system within the bioassay, they too, perhaps, may be considered relevant in a weight-of-evidence approach between or among assays in the Tier-1 screening battery.

QUALITY ASSURANCE

The study will be audited by the CRO laboratory Quality Assurance (QA) Unit with in-phase inspections to ensure compliance with the study protocol and protocol amendments, CRO laboratory SOPs and the appropriate provisions of the EPA Toxic Substances Control Act (TSCA) and Federal Insecticide, Fungicide, Rodenticide Act (FIFRA) Good Laboratory Practice (GLP) Standards published in the Federal Register (40 CFR Part 792 and 40 CFR Part 160), Ministry of Agriculture, Forestry and Fisheries (MAFF) Japan Good Laboratory Standards (59 NohSan No. 3850, 10-Aug-1984) and OECD Principles of GLP [C(97)187/Final November 27, 1997] standards. The raw data and draft report will be audited by the CRO laboratory QA Unit prior to submission to the Sponsor to ensure that the Final Report accurately describes the conduct and the findings of the study.

RECORDS TO BE MAINTAINED

All original raw data records, as defined by the CRO laboratory's SOPs and the applicable GLPs, will be stored as described in the next section in the Archives at the CRO laboratory.

WORK PRODUCT

The Sponsor will have title to all documentation records, raw data, slides, specimens, and other work products generated during the performance of the study. All work products, including raw paper data, pertinent electronic storage media, and specimens will be returned to the Sponsor after a period of six months following issuance of the final report. All work products will be stored in compliance with regulatory requirements.

Any work product, including documents, specimens, and samples, that are required by this protocol, its amendments, or other written instructions of the Sponsor, to be shipped by the CRO laboratory to another location will be appropriately packaged and labeled as defined by the CRO laboratory SOPs and delivered to a common carrier for shipment.

REPORTS

The final report will contain a summary, methods and procedures, summary tables and figures, animal data, and an interpretation and discussion of the study results.

Draft final and final reports will be written. The draft final report, audited summary tables, and electronic copy of individual data spreadsheets will be submitted to the Sponsor. It is expected that the Sponsor will review the draft report and provide comments to the CRO laboratory within a prescribed time following submission. The CRO laboratory will submit the final report in a timely manner following receipt of comments. One electronic copy (Adobe® Portable Document File, or PDF file) will be provided; requests for additional copies of the final report may result in additional charges.

ANIMAL WELFARE ACT COMPLIANCE

The study will comply with all applicable sections of the Final Rules of the Animal Welfare Act (AWA) regulations (9 CFR Parts 1, 2 and 3). The Sponsor should make particular note of the following:

- The certification of the Sponsor Representative's approval as noted on this protocol documents for the Study Director the Sponsor's assurance that the study described in this protocol does not unnecessarily duplicate previous experiments.

- Whenever possible, procedures used in this study have been designed to avoid or minimize discomfort, distress, or pain to animals. All methods are described in this study protocol or in written laboratory SOPs.

- Animals that experience severe or chronic pain or distress that cannot be relieved will be painlessly euthanized as deemed appropriate by the veterinary staff and Study Director. The Sponsor will be advised by the Study Director of all circumstances which could lead to this action in as timely a manner as possible.

- Methods of euthanasia used during this study are in conformance with the above-referenced regulation.

- The Sponsor/Study Director has considered alternatives to procedures that may cause more than momentary or slight pain or distress to the animals and has provided a written narrative description (AWA covered species only) of the methods and sources used to determine that alternatives are not available.

PROTOCOL MODIFICATION

Modification of the protocol may be accomplished during the course of this investigation. However, no changes will be made in the study design without the written permission of the Sponsor. In the event that the Sponsor requests or approves a change in the protocol, such changes will be made by appropriate documentation in the form of protocol amendment. All alterations of the protocol and reasons for the modification(s) will be signed by the Study Director and the Sponsor Representative.

REFERENCES

Feron, V.J., De Groot, A.P., Spanjers, M.T., and Til, H.P. (1973). An evaluation of the criterion "organ weight" under conditions of growth retardation. *Fd. Cosmet. Toxicol.* **11**, 85-94.

Grubbs, F. Procedures for detecting outlying observations in samples. *Technometrics* **1969**, *11*, 1-21.

Holson, R.R. Euthanasia by decapitation: evidence that this technique produces prompt, painless unconsciousness in laboratory rodents. *Neurotoxicology and Teratology* **1992**, *14*, 253-257.

O'Connor, J.C., Cook, J.C., Slone, T.W., Makovec, G.T., Frame, S.R., and Davis, L.G. (1998a). An on-going validation of a Tier I screening battery for detecting endocrine-active compounds (EACs). *Toxicol. Sci.* **46,** 45-60.

O'Connor, J.C., Frame, S.R., Biegel, L.B., Cook, J.C., and Davis, L.G. (1998b). Sensitivity of a Tier I screening battery compared to an *in utero* exposure for detecting the estrogen receptor agonist 17β-estradiol. *Toxicol. Sci.* **44,** 169-184.

O'Connor, J.C., Frame, S.R., Davis, L.G., and Cook, J.C. (1999a). Detection of the environmental anti-androgen p,p-DDE in CD and Long-Evans rats using a tier I screening battery and a Hershberger assay. *Toxicol. Sci.* **51,** 44-53.

O'Connor, J.C., Cook, J.C., Slone, T.W., Makovec, G.T., Frame, S.R., and Davis, L.G. (1999b). Detection of thyroid toxicants in a Tier I screening battery and alterations in thyroid endpoints over 28 days of exposure. *Toxicol. Sci.* **51,** 54-70.

O'Connor, J.C., Davis, L.G., Frame, S.R., and Cook, J.C. (2000a). Evaluation of a Tier I screening battery for detecting endocrine-active compounds (EACs) using the positive controls testosterone, coumestrol, progesterone, and RU486. *Toxicol. Sci.* **54,** 338-354.

O'Connor, J.C., Davis, L.G., Frame, S.R., and Cook, J.C. (2000b). Detection of dopaminergic modulators in a Tier I screening battery for detecting endocrine-active compounds (EACs). *Reprod. Toxicol.* **14,** 193-205.

O'Connor, J.C., Frame, S.R., Smith, C. and Ladics, G.S. (2000c). Comparison of the oral (gavage) and intraperitoneal routes of administration for identifying endocrine-active compounds (EAC) using an *in vivo* male battery, and evaluation of immune system endpoints. *Toxicol. Sci. Supplement – The Toxicologist* **54**, 261.

O'Connor, J.C., Frame, S.R., and Ladics, G.S. (2002a). Evaluation of a 15-day screening assay using intact male rats for identifying anti-androgens. *Toxicol. Sci.* **69,** 92-108.

O'Connor, J.C., Frame, S.R., and Ladics, G.S. (2002b). 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, J.C., Cook, J.C., Marty, M.S., Davis, L.G., Kaplan, A.M., and Carney, E.W. (2002c). Evaluation of Tier I screening approaches for detecting endocrine-active compounds (EACs). *Crit. Rev. Toxicol.* **32,** 521-549.

PROTOCOL APPROVAL

Sponsor approval received via _____ on _____.
Date

_____	_____
Date	Date
Study Director	Director of Toxicology
CRO laboratory	CRO laboratory