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

08-Sep-2006

English - Or. English

ENV/JM/MONO(2006)26
Unclassified

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

Cancels & replaces the same document of 07 September 2006

SERIES ON TESTING AND ASSESSMENT
Number 59

REPORT OF THE VALIDATION OF THE UPDATED TEST GUIDELINE 407: REPEAT DOSE 28-DAY ORAL TOXICITY STUDY IN LABORATORY RATS

JT03213244

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

English - Or. English

OECD Environmental Health and Safety Publications

Series on Testing and Assessment

No. 59

REPORT OF THE VALIDATION OF THE UPDATED TEST GUIDELINE 407 REPEAT DOSE 28-DAY ORAL TOXICITY STUDY IN LABORATORY RATS

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris

September 2006

Also published in the Series on Testing and Assessment:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

© OECD 2006
Applications for permission to reproduce or translate all or part of this material should be made to: Head of Publications Service, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France

About the OECD

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

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

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

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

This publication is available electronically, at no charge.

For this and many other Environment, Health and Safety publications, consult the OECD's World Wide Web site (www.oecd.org/ehs/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

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

FOREWORD

This document is the validation report of the OCDE studies to update the TG 407 repeat dose 28-day oral toxicity study in laboratory rats. The aim of this work was to validate parameters suitable to potentially detect endocrine activity of test substance. This validation report provides a short description of Phase 1 feasibility and exploratory studies and a detailed description of Phase-2: it contains the background on how Phase-2 studies were organised and performed, a description of the model protocol used, detailed summaries of the data, and the conclusions drawn from the studies. Phase-2 consisted of twenty studies with ten test substances, six considered to be potent endocrine-active substances and four considered to be weak endocrine-active substances. The full TG 407 was performed to assess whether the updates might interfere with the current protocol, and most laboratories performed the studies under Good Laboratory Practices or in the spirit thereof.

This report then represents a significant body of experimental work performed, for phase 2, in a total of 13 laboratories from 7 countries between June 2000 and May 2001. The test substances were distributed from a central repository. The studies, the model protocol, the test substances and their doses, and other aspects of the studies have been approved by the Validation Management Group for Mammalian Effects Testing (VMG-mammalian), and subsequently endorsed by the Task Force on Endocrine Disrupters Testing and Assessment (EDTA).

Data and final reports for the 20 Phase-2 studies were requested from each laboratory to be provided to the Secretariat for archiving and for this report, which was written by Dr. William Owens (OECD Secretariat consultant). Comments and input were contributed by the Lead Laboratory (Dr. Alex Freyberger, Bayer HealthCare AG, Wuppertal, Germany), the various participating laboratories, and the members of the Mammalian Validation Management Group. Previous drafts of this report were submitted to the participating laboratories to verify the accuracy of the data and its extraction by the Secretariat, and any errors and inconsistencies both in the drafts and in the final report discovered in this process have been rectified.

The draft validation report was then submitted at the VMG-mammalian meeting (United States, 4-5 April 2006) and revised on the basis of the comments received. The changes were endorsed at the EDTA Task Force meeting that was held in Sweden on 26-27 April 2006. At its May 2006 meeting, the Working Group of the National Coordinators agreed to the submission of the revised draft report to the Joint Meeting with a view to its declassification.

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

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ANNEXES See document [ENV/JM/MONO(2006)27]

SUMMARY

i) This report summarises the results from an OECD inter-laboratory study conducted in 2000-2001 to assess whether potential updates added to the current TG 407 can reliably detect strong and weak endocrine-modulating substances without interference with or compromise of the current TG 407. These studies comprise Phase-2 of the updated TG 407 program, which was preceded by feasibility and exploratory studies in Phase-1. These studies were designed and directed by the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), which was established to develop and validate new and improved methods to identify and assess substances acting through endocrine mechanisms.

ii) The current 28-day repeat dose study (TG 407) provides information on the possible health hazards likely to arise from repeated exposure over limited period of time and includes observations on general, clinical, sensory, and neurological signs; haematology and clinical biochemistry; and gross pathology, weights, and histopathology for major organs and tissues (1). The TG 407 is intended to provide information on the major toxic effects of a test substance, indicate target organs; provide an estimate of a no-observed-adverse-effect level of exposure for certain toxic effects, and potentially provide a maximum tolerated dose.

iii) The proposed updates to the TG 407 are intended to provide the endpoints that can detect presumed (anti)oestrogenic, (anti)androgenic, and thyroid toxicant mechanisms of action. The current endpoints and updates added in Phase-2 include additional tissue weights and histopathology of target organs in the male and female reproductive tracts; tissue weight and histopathology of the pituitary; tissue weight and histopathology of the thyroid; circulating levels of T_3, T_4, and thyroid stimulating hormone (TSH); sperm counts and morphology; and staging of the oestrous cycle using vaginal smear cytology. Chemicals that potentially act as endocrine modulators may be identified by the TG 407 where they induce a statistically significant increase in the weights, pathological change of the target tissues, or affect the other endpoints. These updates had been chosen based upon the initial feasibility and exploratory studies in Phase-1. In addition, the power of the current and update endpoints was to be studied by conducting the study as two individual Subgroups of 5 animals per sex per dose as is done in the current TG 407. These individual Subgroups would then be combined in a statistical reanalysis to represent a combined Subgroup size of 10 animals per sex per dose.

iv) The TG 407 animals are sexually mature, young adults with the hypothalamic-pituitary-ovarian feedback loop established and intact. These mature and intact animals can respond to test substances with some degree of adaptation and compensation. This contrasts with the uterotrophic bioassay for (anti)oestrogens and the Hershberger bioassay for (anti)androgens where the animals are immature or ovariectomized/castrated. These animals with a sex hormonal feed-back system, that either is compromised or not yet in function, may then be more sensitive to active test substances. Further, the mature females are cycling in the TG 407, and the tissue weights of the female reproductive tract should then be more variable, increasing the coefficient of variations (CVs) of the ovarian and uterine weights. Thus, the degree of sensitivity of the TG 407 model needs to established relative to these other bioassays for (anti)oestrogens and (anti)androgens. For thyroid toxicants, the young adult rat appears to be very sensitive relative to humans, an intact animal is necessary to observe pituitary stimulation of the thyroid, and 28-days appears to be sufficient time for thyroid toxicants to induce effects at sufficient doses. Thus, the TG 407 is plausible as a sensitive model for thyroid toxicants, but this sensitivity needs to be demonstrated.

v) Phase-2 of the TG 407 studies involved 13 laboratories from France, Germany, Japan, Korea, Switzerland, the U.K., and the U.S. These laboratories were from both the public and private sectors. The lead laboratory was Bayer HealthCare AG, Wuppertal, Germany. In Phase-2, ten substances were tested in duplicate studies that were carried out in separate laboratories. These test substances were distributed from

a central repository. The laboratories were requested to conduct the full updated TG 407 so as to assess whether any interference or compromise would be encountered with the functional observation battery or any other current protocol requirements. The laboratories were allowed to utilize the rat strain, diet, bedding, vehicle, and other procedures normally employed in the conduct of current TG 407 studies. Data and final reports for the 20 Phase-2 studies were requested from each laboratory to be provided to the Secretariat for archiving and for this report.

vi) The 10 test substances included an array of potent and weak endocrine-active substances. Ethinyl oestradiol (EE), genistein (GN), and nonylphenol (NP) were selected as oestrogens. Tamoxifen (TAM) was selected as an antiestrogen. CGS 13820B (CGS) was selected as an aromatase inhibitor. Methyl testosterone (MT) was selected as an androgen. Flutamide (FLU) and *p,p'*-DDE (DDE) were selected as antiandrogens. Propylthiouracil (PTU) and l-thyroxine (THY) were selected as thyroid toxicants. For each of these test substances, three doses were selected with the intent to provide a no-effect dose, a dose causing moderate effects, and a maximum tolerated dose. In several cases, dose selection required preliminary range finding studies.

vii) Specific goals of Phase-2 were then to evaluate current and updated endpoints and procedures in order to increase the sensitivity of the TG 407 to toxic effects presumed to occur via endocrine modulation:

- evaluate both current and added hormone-sensitive organ weights and histopathology in the male and female reproductive tracts as well as the pituitary and thyroid;
- evaluate cauda epididymidal sperm counts and morphology;
- evaluated circulating thyroid hormone levels including T_3, T_4 and TSH;
- evaluate staging of the oestrous cycle using vaginal smear cytology to standardize female necropsy timing and thereby minimize oestrous cycle-induced variations in uterine and ovarian weights and histopathology; and
- evaluate the utility and benefit of any greater power (statistical probability to detect a positive finding) provided by increasing the number of animals from the number in the current TG 407 from five per sex per dose to ten per sex per dose.

viii) Overall, the TG 407 data and findings generally were in good agreement between the two laboratories conducting studies on the same substance at the same doses. A direct comparison has been conducted in the case of each test substance in the chapters of the main body of the report and in the Annexes. The statistical changes in body weights, haematology and clinical chemistry, and tissue weights; the changes in tissue histopathology; and other observations were generally consistent. The magnitude and shape of the dose response curve as well as the doses at which significant responses were observed were again similar between the two laboratories taking into account modest biological variability, strains, laboratory technique and other variables.

ix) The validation programme successfully achieved the goal of demonstrating the ability of the updated TG 407 protocol to detect potent (anti)oestrogenic and (anti)androgenic substances. Both of the laboratories employing a test substance observed effects in target organs with EE, TAM, CGS, MT, and FLU. Moreover, both laboratories were able to detect the activity of the potent test substances with group sizes of only five animals per sex per dose. The findings were detected using the tissue weights and histopathology of the current and updated tissues of the male and female reproductive tracts. Although changes were observed in the tissue weight and histopathology of the pituitary, the pituitary weight findings were variable and often inconsistent with the expected mechanism of action of a test substance. Sperm counts and morphology were found to be insensitive to many potent test substances, and, when findings were present in sperm counts and morphology, these findings occurred at higher doses than

observations in tissue weights and histopathology.

x) In contrast to the potent (anti)oestrogenic and (anti)androgenic substances, findings were limited in the cases of the weak oestrogens GN and NP, and no antiandrogenic findings were present in either laboratory in the case of DDE. For GN, the evidence of estrogen activity was equivocal. One study detected a statistically significant increase in the uterine weight. Both studies did observe a possible desynchronisation of the female reproductive cycle among the tissue, but only when the data were reinterpreted by an expert pathologist. These changes started at 400 mg/kg/d. The subtle and atypical nature of the observations (asynchrony of several tissues and cell types in the female reproductive tract to the oestrous cycle) necessary to arrive at this conclusion must be noted. The detection of very weak estrogens such as GN may then require increased vigilance and changes in the current practice of histopathological examination of the female reproductive tract. For NP, the updated TG 407 did not detect estrogenic responses with NP at doses below a maximum tolerated dose. Equivocal evidence in uterine histopathology and decreases in the absolute weights of male accessory reproductive tissues were observed at doses that caused mortalities and other clinical signs in the animals. The effects on the synchronisation of the female reproductive tract tissues seen with EE and GN may have been seen in one NP study at a low frequency, but confirmation was absent at the higher NP dose in the second study. Therefore, this evidence is judged to be equivocal. The absence of clear effects of weak oestrogens is not surprising in view of the use of adult animals with an apparently well-functioning homeostasis. It indicates that the test discriminates among hazardous strong oestrogens, which are positive in the test, and weak oestrogens which are relatively innocuous in the adults, which appear negative. For DDE, neither study provided evidence of antiandrogenic activity in the male reproductive tract, despite mortalities in both sexes indicating that the maximum tolerated dose was exceeded in one study. Both DDE studies used Sprague-Dawley rats, which have been shown to be less sensitive to DDE at the administered doses than Long Evans rats, and this may have contributed to the lack of observed antiandrogenic effects.

xi) The validation programme successfully achieved the goal of demonstrating the ability of the updated TG 407 protocol to detect thyroid toxicants. Both of the laboratories employing a test substance observed effects in the thyroid with PTU and THY. Both laboratories were able to detect the activity of these test substances with group sizes of only five animals per sex per dose. In addition, thyroid modulation by MT and DDE were detected in both laboratories. The effects of MT were mild, and the thyroid effects of DDE including histopathology and hormonal changes were preceded in dose by significant hepatic enlargement and histopathological suggesting these thyroid effects may be the result of increased elimination of circulating thyroid hormones resulting in increased pituitary stimulation of thyroid TSH secretion. Thus, the ability of the TG 407 to detect thyroid modulation appears to be robust and extend beyond classical potent compounds. Thyroid histopathology was consistently the most reliable and most sensitive endpoint for the detection of thyroid modulation. Thyroid weight was reliable, but was somewhat less sensitive when compared to thyroid histopathology. Circulating thyroid hormone levels (T_3, T_4, and TSH) were not always reliable and sensitive, but the standard operating procedures for blood sampling and for thyroid hormone analyses were not standardised to reduce stress induced variability and to reduce analytical variability, respectively. Circulating T_4 levels were the most promising of the three thyroid hormonal values.

xii) The ability of laboratories to perform the entire TG 407 protocol was not negatively impacted by the updates. Laboratories were able to conduct the functional observation and motor activity batteries without interference. Study time for some females to reach diestrous was extended beyond 28-days in the studies, sometimes extending the work of the necropsy staff into weekends.

xiii) An important factor affecting sensitivity of the endpoint is its variability as represented by the coefficient of variation (CV). A lower CV improves the ability to statistically detect significance, and a higher CV diminishes that ability. The CV values of the proposed enhancement tissue weights were

slightly higher than a number of current tissues. For example, current tissues, that are larger and easier to dissect, such as liver and testes, had lower mean CVs ranging from about 8 to 11. As tissues became smaller and more difficult to dissect, CVs increased. Mean CVs were 12-14 for paired adrenals, 18-19 for pituitary, and 20-21 for the thyroid. Fluid-filled male accessory tissues that are difficult to dissect, such as the prostate lobes, had higher mean CVs approaching 24. The CVs for the uterus and for the male accessory tissues were consistent with those observed in the uterotrophic and Hershberger validation programs. In addition, as also with the uterotrophic and Hershberger studies, CV values varied and appeared to be related with the individual performing laboratories. This suggests that laboratory technique is a possible variable and could be connected to the ability to detect weakly acting substances.

xiv) The utility of increasing the group size or the number of animals per group for detecting endocrine mediated effects were assessed between individual Subgroups of 5 animals per sex and combined Subgroups of 10 animals per sex. The assessment concludes that, for potent test substances, 5 animals per sex are sufficient to detect endocrine modulation with (anti)oestrogenic, (anti)androgenic, and thyroid toxicants. First, power calculations were performed and showed that the same degree of benefit cannot be generalized. The increased power will depend upon the variability of a particular endpoint, or its inherent CV, as well as the magnitude of change induced by a test substance. Moreover, while the animal number may be doubled (increased by 100%), the increased power provided is usually 80% or less. Second, given that CV is a major contributor to power, improvements in laboratory skill and the practice of the tissues' histopathological examination should also be entertained and may have higher priority than increasing animal numbers. The sensitivity was investigated here in terms of statistically significant differences from the control group. An alternative approach, e.g. the benchmark calculation, which uses the complete dose response assessment to calculate the critical dose level for each endpoint, may be applied to evaluate the data more efficiently. This may also help to overcome the need for doubling the animal numbers in the protocol, which may only marginally increase power.

xv) Ultimately, the decision to increase or not increase the group size may be even more complicated. First, it is a value judgment that weighs the benefit of greater power achieved with increased group size against the animal welfare concerns of doubling the animal numbers. Second, the benefit of detection would depend upon the overall chemical testing strategy employed by regulatory jurisdictions.

xvi) A comparison has been made between the results of the updated TG 407 and other toxicological assays with the same or similar endocrine active test substances in most cases. At least some data were found for comparison with for all substances except methyl testosterone and l-thyroxine. These comparisons support the ability of the TG 407 to detect systemic, target organ and endocrine related toxicities for potent chemicals. Where chronic studies or reproductive and development studies easily detected effects, concordant results were typically found in the current TG 407 studies for systemic, target organ, and endocrine toxicities. These comparisons are summarised at the end of each test substance chapter in the body of the report.

xvii) The Phase-2 results support the inclusion of several current endpoints and related updates in a revised guideline as useful for the detection of endocrine modulating substances. Tissue weights and histopathology of the male and female reproductive tracts as well as the male and female mammary glands clearly contributed to the detection of (anti)oestrogenic and (anti)androgenic substances with potent compounds in a reliable manner and with less potent compound in substance-specific cases. While there was no consistent relationship between pituitary weight change and an expected mode of action, histopathology of the pituitary was often relevant in providing support for findings in other tissues. Likewise, adrenal weights and histopathology also provided supplemental and supporting information for findings in other tissues. However, neither pituitary histopathology, adrenal weights, nor adrenal histopathology should by themselves be considered diagnostic of potential endocrine modulation in the case of (anti)oestrogens, (anti)androgen, and thyroid toxicants.

xviii) The Phase-2 results clearly support the inclusion of thyroid weights and histopathology in a revised guideline as able to detect thyroid toxicants of various potencies. As the thyroid weight is somewhat less sensitive, care should be exercised not to compromise the histopathology of the tissue. Thyroid trimming should then continue to be performed after fixation of the tissue and by carefully trained technicians.

xix) The Phase-2 results do not support the inclusion of the spermatology endpoints. The spermatology endpoints provided no substantive benefit for detecting possible endocrine modulation. Sperm effects were found only with potent compounds at doses above those where other endpoints had already identified clear effects, and the CVs indicate that substantial work might be needed to achieve reliable methodological practice among laboratories.

xx) The Phase-2 results indicated that further work may be useful with T_4 and TSH hormone levels as optional and supplementary endpoints. Where a compound is suspected of thyroid activity or changes are noted in thyroid histopathology, T_4 and TSH analyses of the retained serum samples may be warranted. The CV of the T_4 measurement was modest, averaging 20-21 across laboratories, but the TSH CV was nearly double that value. The primary concerns at this time are better understanding and possible standardisation of necropsy sampling conditions and analytical methods. The T_3 hormone levels did not appear sufficiently sensitive and yielded several false positives. Therefore, T_3 hormone analyses are not recommended for an initial general toxicology screen such as the TG 407.

xxi) The Phase-2 results suggest that a review of the histopathology procedures for the male and female reproductive tracts and for the mammary glands of both sexes has merit. The central problem confronting the histopathologists with (anti)oestrogens and (anti)androgens is that the pattern of effects is typically not one of frank pathological changes and lesions in any tissue. In the male, the pattern may be one of modest atrophy or hypertrophy or changes in particular target cells, such as the Leydig cells. In the female, the overall synchronization of a temporal sequence in several tissues during the oestrous cycle may be altered so that 'normal' histology is observed for individual parameters, but the overall coordination of the oestrous cycle is impaired. Thus, the effects are subtle and not necessarily atypical, and one must carefully correlate the pattern in several cell types of the ovary, uterus, cervix and vagina with that of the oestrous cycle in order to evaluate whether there are dissimilarities with the normal, expected sequence and its orchestrated pattern.

xxii) Such a careful staging of the female reproductive tract tissues is not the current practice in TG 407 studies and was not conducted as part of these update studies. As this appears to be the most promising possibility at this time to reliably identify weak oestrogens with the TG 407 protocol, it may be worthwhile to conduct pilot studies with specific compounds. Due to the costs and animal use, major studies should await positive results from such feasibility studies.

xxiii) This discussion also raises a question about the degree of blinding of the pathologist in TG 407 studies. As noted, the histopathological changes in the male and female reproductive tracts are often subtle, making the recognition of changes in the context of normal patterns technically challenging. Together with improved guidance on histopathological practice for these tissues, identification of substances is likely to be improved where the pathologist is informed of 1) structural, *in vitro*, or other *in vivo* alerts as indicators of a test substance's possible endocrine activity and 2) changes in both the significance and trends in the absolute and relative weights of particular tissues.

xxiv) In conclusion, the following endpoints should be included in an updated TG 407 guideline for the detection of toxicants with possible (anti)oestrogenic, (anti)androgenic, and antithyroid modes of action for young animals:

- Absolute and relative tissue weights of the male reproductive tract, including testes, epididymides, ventral and dorsolateral prostate and seminal vesicles with coagulating glands.
- Absolute and relative tissue weights of the female reproductive tract, including the paired ovaries and uterus.
- Histopathology of the pituitary.
- Absolute and relative thyroid weights and histopathology of the thyroid. The recommendation for the addition of the thyroid weights is qualified in that the dissection and trimming procedures must not compromise the more valuable thyroid histopathology.
- Histopathology of the male and female reproductive tract tissues (adding cervix and vagina and mandatory staging assessment of the oestrous cycle can be realised).
- Histopathology of the male and female mammary glands.
- If there are indications that the test substance is a thyroid toxicant or thyroid histopathology indicates effects, it could be appropriate to consider performing T_4 and TSH analyses on retained serum samples if the necropsy sampling and analytical questions noted in this are resolved.

xxv) The decision whether or not to increase the group size from 5 animals per sex per dose to 10 animals per sex per dose requires wider input and consideration. Benchmark calculations using the whole data set may be a further adjunct to increase the sensitivity of the test based on 5 animals per sex per dose. In particular, the need for sensitivity and increased power should be evaluated in the context of an overall testing hierarchy. Finally, such a decision should also take into account alternatives for improving the power of current endpoints by improving technique and practice as well as possibilities noted above such as the staging of the female reproductive tract cycle.

INTRODUCTION

1. The original Test Guideline (TG) 407 for the 28-day repeated dose oral toxicity was adopted in 1981. The OECD Guidelines for the Testing of Chemicals are periodically reviewed, and TG 407 was revised in 1995 (1). That revision resulted from a Consultation Meeting of an ad hoc Working Group of Experts on Systemic Short-term and (Delayed) Neurotoxicity, held in Paris in February 1992 (2), which was based on an earlier proposal by the United Kingdom, dated February 1991. The primary objective of those revisions was to obtain additional information from the animals used in the study and to incorporate a series of observations that would indicate the potential of a test substance to cause neurotoxic effects.

2. In the assessment and evaluation of the toxic characteristics of a chemical, the determination of oral toxicity using repeated doses is central to hazard identification and characterization. The 28-day repeat dose study (TG 407) provides information on the possible health hazards likely to arise from repeated exposure over a relatively limited period of time. The current updated version of TG 407 requires observations on general and clinical signs; sensory and neurological signs; haematology and clinical biochemistry; and gross pathology, weights, and histopathology for major organs and tissues (1). As a result, the study provides information on a wide range of toxic effects and indicates target organs, but is not designed to identify particular mechanisms or modes of action. The study also serves as a dose range finder for follow-up studies and provides an estimate of a no-observed-adverse-effect level (NOAEL) of exposure for certain toxic effects, and potentially a maximum tolerated dose. This information can be used in selecting dose levels for chronic studies, if needed. As a result of its broad spectrum of endpoints and observations, TG 407 is frequently part of regulatory data requirements a pivotal study for the hazard assessment of new chemicals in many member countries.

3. Endocrine disruption or modulation is not in itself a toxic endpoint, but is an underlying mechanism that leads to classical toxicities such as reproductive and developmental toxicity (3)(4). The need to address key endocrine mechanisms has arisen from the concerns that ambient levels of natural and industrial chemicals may interact with the endocrine system and as a consequence possibly elicit reproductive, developmental, and other adverse effects in humans and wildlife. The supporting evidence for these concerns have expressed or articulated in number of expert reviews (3)(4)(5)(6)(7)(8). This leads then to the need to ensure that endpoints sensitive to these mechanisms are evaluated for the usefulness and that appropriate endpoints are incorporated into toxicity tests such as the TG 407.

4. In response to these concerns, the OECD member countries concluded in 1997 that existing Test Guidelines were generally insufficient to identify certain endocrine mechanisms (oestrogen, androgen, and thyroid) and might not be adequate to fully characterize the hazards of these mechanisms. Therefore, a *Special Activity on the Testing and Assessment of Endocrine Disrupters* was initiated as part of the OECD Test Guidelines Programme. The purpose of this activity was to revise existing Guidelines and to develop new OECD Test Guidelines for test methods that are considered useful tools and to identify and characterise substances that could interact with the endocrine system (Further information concerning the OECD Endocrine Disruptor Testing and Assessment Programme can be found at http://www.oecd.org/EN/document/0,,EN-document-524-nodirectorate-no-24-6685-8,00.html). An OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) was established to manage the work and identify and recommend priorities for the development and validation of new and improved methods to identify and to assess substances acting through endocrine mechanisms (9). Member countries were all invited to nominate national representatives to the Task Force. Other stakeholders (e.g., BIAC, NGOs, ICAPO) were invited to nominate observers.

5. The EDTA Task Force considered various national and stake holder initiatives to develop and validate *in vitro* and *in vivo* assays for the detection of chemicals that may interfere with endocrine responses; took note of the recommendations of a number of national, regional and international

workshops (10)(11)(12)(13)(14); and took into account the recommendations from a detailed OECD review of the status of both existing standardised assays and available research methodologies (15). The Task Force agreed that several possible mechanisms of actions needed to be addressed: estrogens, antiestrogens, androgens, antiandrogens, and thyroid toxicants. The EDTA Task Force chose as a priority work on two *in vivo* bioassays that were specific for particular mechanisms: the uterotrophic bioassay for estrogens and antiestrogens and the Hershberger bioassay for androgens and antiandrogens. In addition, the update of TG 407 was selected to potentially address both (anti)estrogen and (anti)androgen as well as thyroid toxicants (9). As TG 407 is commonly performed for most high production volume chemicals, update of this Test Guideline suggested greater efficiency and lower animal use than constructing a battery of individual assays to address different mechanisms.

PHASE-1: INITIAL ASSESSMENT OF THE UPDATE OF TEST GUIDELINE 407

6. In order to identify possible estrogens, antiestrogens, androgens, antiandrogens, and thyroid toxicants during the course of the 28-day repeat dose toxicity study, the principle for TG 407 updates is that the protocol should incorporate the necessary array of endpoints sensitive to these mechanisms. This array would in principle first include the primary endocrine organs that produce to these hormones and that respond to their homeostatic feedback systems (pituitary, thyroid, gonads, and adrenals). This array would also potentially include other primary target organs (the male and female reproductive tracts such as the prostate, epididymides, uterus, and mammary glands). These evaluations would utilize traditional toxicological techniques such as tissue weights and full microscopic histopathology, and the use of other potential endpoints in the evaluation such as circulating hormone levels, sperm parameters, and oestrous cyclicity would be entertained and assessed.

7. Consistent with current approaches to the use of the TG 407 data, TG 407 is not as a definitive test, but is designed to identify possible chemical hazards from the toxic effects observed in the study. That is, the results from the TG 407 are designed for initial hazard identification and "to raise a red flag" about potential toxicity concerns such as endocrine modulation. These data would need to be incorporated with other hazard and exposure data using a weight of the evidence approach. Any concerns raised by the weight of the evidence may then need to be specifically addressed in subsequent studies by further defining the effects seen for hazard characterization or for risk assessment purposes. Therefore, any indications of toxicity mediated by endocrine mechanisms that were seen in the updated TG407 protocol may need to be followed up and investigated in greater depth in tests designed specifically to examine adverse reproductive or developmental effects and to more fully characterize their hazard.

8. After a series of consultations, it was originally proposed to add the following updates to the current TG 407 for Phase-1 feasibility investigations:

- Organ and tissue weights: The current TG 407 includes weights of the adrenals, ovaries, testes, and epididymides. The initial updated proposal was to weigh the testes separately and to add weights of the following tissues: seminal vesicles with coagulating glands, whole prostate with the possibility to measure the ventral and dorsolateral prostate separately, uterus, and thyroid.
- Histopathology: In addition to the current TG 407 requirements, the initial proposal to update TG 407 was to perform additional histopathology on the pituitary, vagina, one epididymis, seminal vesicles with coagulation glands, and mammary gland.
- Serum hormone levels: In order to possibly assess changes in the feedback control of estrogens, androgens, and thyroid through the pituitary, the initial proposal to update TG 407 was to add the analysis of both the pituitary hormone and the target organ hormone in the serum. These hormones were: leutinising hormone (LH), follicle stimulating hormone (FSH), prolactin, testosterone, 17β-estradiol, corticosterone, T_3, T_4, and thyroid stimulating hormone (TSH).

- Sperm production and quality parameters: As an essential function of the male under androgen and pituitary control is sperm production, the initial proposal to update TG 407 was to add sperm number, morphology, and motility endpoints.
- Oestrous cyclicity: As an essential feature of the female under estrogen and pituitary control is the oestrous cycle, the initial proposal to update TG 407 was to add daily vaginal cytological smears to cover at least two full cycles in order to evaluate the stage of the oestrous cycle each day and, therefore, assess the number of days per cycle and its regularity.

9. The VMG at their first meeting in Tokyo in February 1999, concurred with the proposed TG 407 updates, an initial feasibility phase (Phase-1), and development of a draft protocol by experts in collaboration with the Secretariat with the aim towards undertaking the Phase-1 studies in the summer and early fall of 1999 (16). The 1st VMG submitted this proposal to the EDTA. The EDTA agreed to initiate the first phase of the validation work (Phase-1) by testing the feasibility and the practicability of the updates, their lack of interference with the established protocol endpoints, and the sensitivity and utility of endpoints with at least one reference substance (17). Flutamide was selected for the initial prevalidation test based on the following arguments: Flutamide is a known, potent antiandrogen, its effects in toxicity studies have been well characterized, and a laboratory with experience in Flutamide toxicity tests volunteered to perform the study. This laboratory was designated the lead laboratory, and this laboratory agreed to follow the updated TG 407 protocol including all current endpoints. In addition, five laboratories in Japan and one in Korea volunteered to perform the updated TG 407 in parallel using flutamide and other endocrine-active chemicals. These additional chemicals included ethinyl oestradiol, propylthiouracil, methoxychlor, tamoxifen, and methyl testosterone.

10. In Phase-1, only the lead laboratory study with flutamide actually followed the updated TG 407 protocol exactly, performing all of the current and updated endpoints. In addition, only this study and a study with propylthiouracil were performed under the standard TG 407 dosing recommendations. That is, using three dose levels where the highest dose level was chosen with the aim of either 1) clearly inducing toxic effects at a maximum tolerated dose, but not mortality or suffering or 2) achieving a limit dose of 1000 mg/kg body weight/day. The lower dose levels were selected to demonstrate any dose-related responses with the lowest dose intended to be a no-observed-adverse effect level (NOAEL). Dosage intervals, while chemical and study specific, were recommended not to exceed a factor of 10.

11. The other studies, however, did not include all of the current TG 407 endpoints (e.g., the functional observation battery was often not included) and were performed at test chemical dose levels considerably below the doses that would have been used in a standard TG 407 study. For example, some of these studies used the range of doses that would be used in the uterotrophic or Hershberger bioassays. These two bioassays are performed with immature or gonadectomised animals, so that the target organs are regressed and offer sensitive response to administered compounds. In contrast, the TG 407 animals are sexually mature, intact, and near or at full capacity for reproductive function. The TG 407 animals, therefore, have intact hormone feedback systems (hypothalamic-pituitary-target organ axis), which may compensate for some degree of endocrine insult. With the unknown degree of difference in the dose responses between the uterotrophic and Hershberger bioassays and the updated TG 407, these tests could only be used to examine the extent of agreement between different procedures and their dose responses.

12. A meeting was arranged with expert representatives of all participating laboratories and other expert toxicologists at Wuppertal, Germany, in December, 1999 to present and to discuss the Phase-1 results (18). Several laboratories have subsequently published their results in the scientific literature (19)(20)(21)(22). The primary conclusions from the Phase-1 work at the Wuppertal meeting were:

- Histopathology: The histopathological examination of the additional tissues was judged to be the most sensitive and promising endpoint for the detection of the hormonal effects. In many of the

studies, histopathological effects were seen at test chemical doses that did not reveal changes in circulating hormones, tissue weights, oestrous cycling or other effects.

- Spermatology: There was general agreement that, with the possible exception of sperm morphology, the sperm counts and motility measurements were not useful. Further, sperm motility and morphology measurements were the most time- and resource-consuming new endpoints. It was the judgment of the meeting that histopathology of the testes provides sufficient information and is more sensitive for effects on sperm maturity and production.

  With one exception, the experts concluded that spermatology was unnecessary, if histopathology of the testes was performed. The sole dissent recommended that sperm morphology studies continue to be performed as a TG 407 update because detailed mature sperm morphology would not be discernible through histopathologic evaluation.

- Functional Observational Battery (FOB): There were concerns regarding the timing of the FOB tests and the additional numbers of animals making the FOB more cumbersome and difficult to perform logistically. To prevent interference with the FOB, it was recommended to begin the vaginal cytology measurements later and to sacrifice the females in diestrous, which is the longest period and presumably most stable portion of the oestrous cycle, in order to make the weights and histopathology of the female reproductive tract more homogenous.

- Hormonal measurements: As few of the toxicology laboratories were set up to perform the hormonal measurements (e.g., 17β-estradiol and testosterone) in-house, this task was often contracted out, adding significant cost to the study. A comparison of these hormonal data showed significant inter-animal variability within dose groups in the measurements in all laboratories. There were also differences in sample timing among and even within laboratories, and it was pointed out that the effects of stress or circadian rhythms on the various hormonal levels had been published in the scientific literature. Few statistically significant hormonal differences were seen in the studies, and these events were never the most sensitive endpoint. Therefore, other than the thyroid hormones, the other hormonal analyses were not recommended for inclusion in Phase-2.

- Estrous cycling: Vaginal smears were considered time consuming to perform, there was considerable inter-animal variation, and several parties noted that, with the rather young adults used in the TG 407, the oestrous cycle may not have entirely stabilised. There were also a number of examples where the histopathology was inconsistent with the smears and suggested a different cycle stage of the oestrous cycle. However, continued work with the oestrous cycle measurements was recommended for Phase-2.

- Fresh or fixed organ/tissue weights; dissection of accessory sex glands (ASG's): The Wuppertal meeting agreed that fixation before weighing was necessary only for the pituitary and thyroid because of the difficulties in dissecting these small tissues from extensive fascia and surrounding tissue, e.g., prostate. It was agreed that fixation before weighing was unnecessary to assist the trimming and dissection of the male reproductive tract accessory sex tissues. While the data indicated that, with flutamide, the weight of the combined (unseparated) accessory sex glands was just as sensitive a parameter as after they were separated and weighed separately, there was also evidence to suggest that other mechanisms might be specific to individual tissues. In such cases, these mechanisms might otherwise not be detected. Therefore, it was agreed that individual weights of the male accessory reproductive tissues would be continued, including separate weights for the ventral and dorsolateral prostate lobes.

- Trimming of thyroid: Trimming of the thyroid increases the possibility to damage the outer layers (from cutting, trimming and compression of the tissue while it is being dissected), but that trimming is necessary for an accurate weight. It was recommended that one lobe be trimmed and weighed. The other lobe would be fixed untrimmed and this lobe would be used for histopathology.

- Animal numbers: Several studies used more than the 5 animals/dose group as specified by the current TG 407 protocol to provide data on intra-laboratory variability and statistical power. The results suggested that ten animals could reduce variability and increase statistical power. It was recommended to conduct Phase-2 using two individual Subgroups of 5 animals per sex per dose and, after statistically analyzing the individual Subgroups, to combine the animals into groups of 10 per sex per dose and then to repeat the statistical analyses in order to assess the impact of increasing the group size.

- Dosing: The TG 407 animals generally did not respond to estrogen/antiestrogens or androgens/antiandrogens as do the uterotrophic and Hershberger bioassays where the animals are either immature or gonadectomised. Higher doses were expected to be necessary to elicit responses in the intact and young adult animals used in TG 407.

- Resources and Costs: The use of the complete, proposed updated OECD protocol nearly doubles the cost of the 407 procedure. The primary cost-intensive additions were the hormone determinations, particularly with external contracting. The most labour-intensive (and next most expensive) additions to the protocol were the spermatology parameters, particularly sperm motility and morphology. Proper timing of the sperm collection immediately requires 2-3 additional people at the time of necropsy, and limits the numbers of males that can be necropsied in one day to 20.

13. The results of the Phase-1 studies and the conclusions and recommendations from the Wuppertal meeting were presented to the VMG at their second meeting in Paris in January, 2000, along with the proposal to undertake Phase-2 studies with a battery of test substances having different endocrine mechanisms (23). Further, it was proposed that two independent, duplicate studies would be performed with each chemical, that dose selection should be based upon standard TG 407 considerations, and that the full TG 407 should be performed including current endpoints such as the FOB and all updated endpoints. The 2nd VMG endorsed the recommendations and forwarded them for consideration to the EDTA Task Force. The VMG requested that urgent expert consultations should begin to consider and recommend the battery of test substances and that discussions should begin to solicit participation by a sufficient number of laboratories to perform two studies on each test substance.

14. Consultation by the Lead Laboratory with the participating laboratories and other experts led to the proposal for a battery of ten test substances as summarised in Table 1 (24). These include 6 strongly acting substances and 4 substances that are considered to be weakly acting. The six strongly acting substances were ethinyl estradiol, tamoxifen, methyl testosterone, flutamide, L-thyroxine, and propylthiouracil, and the four additional weak substances were CGS 18320B, *p,p'*-DDE, genistein, and nonylphenol. Their modes of action are indicated in Table 1. Further consultations were recommended to consider the test substance doses to be used based on available literature and information. If the information for certain substances were to prove insufficient, then range finding studies were recommended for those substances.

Table 1. Chemicals used in Phase-2 of the updated TG 407 guideline studies.

	Chemical	CASR-Number	Mode of Action
1	Ethinyl Oestradiol	57-63-6	Oestrogen agonist
2	Genistein	446-72-0	Possible Oestrogen agonist
3	Nonylphenol	25154-52-3	Possible Oestrogen agonist
4	Tamoxifen		Antioestrogen – Oestrogen antagonist
5	CGS 18320B[a]		Antioestrogen - Aromatase inhibitor
6	Methyl Testosterone	58-18-4	Androgen agonist
7	Flutamide	1311-84-7	Antiandrogen – Androgen antagonist
8	*p,p'*-DDE[b]	72-55-9	Possible Antiandrogen– Androgen antagonist
9	Propylthiouracil	51-52-5	Thyroid Toxicant – Inhibitor of enzymes necessary for thyroid hormone synthesis
10	l-Thyroxine	6106-07-6	Thyroid Hormone (Natural Agonist)

[a] CGS 18320 replaces the original aromatase inhibitor Fadrozole recommended in May, 2000, as it was not made available by the supplier.

[b] 1,1-dichloro-2,2-bis-(*p*-chlorophenyl)ethylene

15. At the 4th Meeting of the EDTA in Paris in May, 2000, the VMG recommendations and those of the expert consultations, including a revised model protocol for the updated TG 407 studies, were endorsed (25). EDTA-4 agreed to undertake Phase-2 of the update of the TG 407 taking into account the following specific goals:

- to evaluate an increase in the group size per sex per dose from the current 5 animals to 10 animals in order to improve the reproducibility, statistical power and sensitivity of the TG 407 protocol, particularly for endocrine endpoints;
- to evaluate the inclusion of current and additional hormone-sensitive tissue weights (testes, epididymides, whole prostate, ventral prostate, dorsolateral prostate, seminal vesicles with coagulating glands, ovaries, uterus, pituitary, and thyroid gland) on the capability to detect endocrine-active substances;
- to evaluate the inclusion of current and additional hormone-sensitive tissue histopathology (testes, epididymides, whole prostate with ventral and dorsolateral portions, seminal vesicles with coagulating glands, ovaries, uterus, vagina, pituitary, thyroid gland, and mammary gland) on the capability to detect endocrine-active substances;
- to evaluate the inclusion of epididymal sperm counts and sperm morphology evaluations on the capability to detect endocrine-active substances;
- to evaluate the inclusion of circulating T_3, T_4, and TSH measurements on the capability to detect thyroid active substances; and
- to evaluate the inclusion of vaginal cytology smears on the capability to detect changes in the oestrous cycle of the young adult female rat.

16. With this EDTA endorsement, the further work to plan, prepare, and conduct Phase-2 of the studies to update TG 407 was undertaken. These steps included:

- The recruitment of participating laboratories so that duplicate studies could be conducted with each of the agreed ten test substances. Including the Lead Laboratory, a total of thirteen laboratories in seven countries agreed to participate.

- The organisation of a central chemical repository to acquire and to distribute the test substances with the aim that the laboratories would use test substances from the same lot. The European chemical industry agreed to manage and to fund the central chemical repository.
- Consultations began on the available information for each test substance in order to select common doses for each test substances and that these doses would be used in both laboratories. Where information was judged to insufficient, range finding studies were to be conducted by one of the participating laboratories.
- Based on the dose selection, sufficient quantities of the test substances were acquired and deposited in the central chemical repository.
- The general protocol for the updated TG 407 was finalised in the summer of 2000 and distributed to the participating laboratories, and this protocol was adapted to the test substance and the particulars of the participating laboratory, e.g., rat strain, vehicle, and so on.

PHASE-2: MULTI-CHEMICAL, MULTI-LABORATORY VALIDATION OF THE UPDATED TG 407

METHODS

Introduction

17. This part of the report describes in detail the multi-chemical, multi-laboratory study design for the validation of the updated TG 407 protocol. It includes an overview of the protocol and the participating laboratories, the doses of the 10 endocrine-active test substances used, briefly reviews how these doses were chosen, summarises the conditions used in the laboratories during the updated TG 407 studies (e.g., strain of rat, vehicle, diet), summarises the various laboratory measurements, and summarises the statistical methods used by the laboratories.

Phase-2 Protocol and Participating Laboratories

18. The model protocol used in Phase-2 and the changes from the current TG 407 are outlined in Table 2. The full model protocol is in Annex 1. Allowances were also made for the individual laboratories to investigate and to report the results for additional measures and alternative procedures in their individual protocols, so long as the specific objectives of the program would not be compromised.

Table 2. A comparison of the current TG407 endpoints with the proposed update endpoints investigated during the Phase-2 studies

Endpoint/effect	In Current TG 407	Proposed Updates to TG 407
Organ/tissue weights	liver, kidney, adrenals, testes, epididymides, thymus, spleen, brain, heart	1. testes (each weighed separately) 2. seminal vesicles + coagulating glands 3. prostate (possible dissection and separate weights for ventral and dorsolateral prostate), ovaries 4. thyroid 5. uterus
Histopathology	brain, spinal cord, stomach, small and large intestines, liver; kidneys, adrenals, spleen; heart, thymus, trachea, lungs, gonads, accessory sex organs (i.e., uterus, prostate), urinary bladder, lymph nodes, peripheral nerve, bone marrow, all gross lesions	1. pituitary 2. vagina 3. one epididymidis, seminal vesicles + coagulation glands 4. mammary gland,
Thyroid Hormones	none	1. circulating levels of T_3 and T_4 2. circulating levels of TSH
Spermatology	none	1. epididymal sperm number 2. sperm morphology
Oestrous cycle	none	Daily vaginal smears to assess oestrous cycling via epithelial cytology for at least five days to ensure necropsy during diestrus

19. A total of 13 laboratories participated in the Phase-2 studies. Laboratories were geographically diverse with participants from France, Germany, Japan, Korea, Switzerland, the UK, and the US. The laboratories were from both the public and private sectors as noted in Table 3. The participating laboratories and their lead investigators are identified in Annex 2.

Table 3. Laboratories participating in the OECD Phase-2 updated TG 407 studies

Country	Laboratory	Number of Laboratories
France	Industry	1
Germany	Industry	2
Japan	Industry Private Contract	2 4
Korea	Academic	1
Switzerland	Industry	1
UK	Industry	1
United States	Industry	1
	Total laboratories performing studies	**13**

20. The intended purpose of the TG 407 is a comprehensive *in vivo* assessment after a relatively short period of exposure. It is intended to be used as an initial screening assessment of the possible hazard and could be applied to a potentially large number of chemicals. The TG 407 has been in use since May 1981 and was substantially revised in 1995. Although the TG 407 has not undergone what would now be considered a formal validation process, regulatory authorities consider the TG 407 to be acceptable based upon its history of practice and demonstrated usefulness in measuring systemic toxicities. Therefore, no further standardisation of the updated protocol variables, such as animal strain or diet were undertaken. Consequently, the protocol allows variations in a number of study conditions, such as the choice of rat strain and laboratory diet. The specific supplies and conditions used in each laboratory of the participating laboratories during the updated TG 407 studies are summarised in Table 4.

Measurements and Data Reporting

21. The most common measurements performed in each laboratory during Phase-2 and consistent with the current TG 407 guideline have been compiled and summarised. The haematology and clinical chemistry study measurements are summarised in Table 5, the organ and tissue weights measured are summarised in Table 6, and the histopathology observations are summarised in Table 7. The Tables are not intended to be exhaustive, as some laboratories may have performed certain additional measures that were not performed in the other laboratories. Examples are serum magnesium levels, serum cholinesterase levels, and histopathology of the aorta or the Zymbal's glands. As there is no robust data set for making comparisons among the laboratories as to the value or reproducibility of such individual and sporadic measurements, these details have not been included in this report.

22. Upon completion of the studies, the data analyses, and the data quality audit, the Secretariat requested that each participating laboratory submit a final report in both hard copy and in electronic form. Twelve of the 13 laboratories provided complete final reports. One laboratory provided a summary report that did not include individual animal data and, originally, did not include histopathological data numbers and grades. Upon the specific request of the Secretariat, the histopathological data from this laboratory were provided for inclusion in the Annexes of this report.

Table 4. Laboratory parameters and conditions in Phase-2 updated TG 407 studies.

Study	Lab	Chemical	Rat strain/ Supplier	Administration/ Vehicle	Diet	Animals per cage
1	1	Nonylphenol	Wistar rats CrlGlxBrlHan:WI/ Charles River, GR	Gavage/Corn oil	Basic maintenance diet, Provimi KLIBA, GR	1
2	1	Propylthiouracil	Wistar rats CrlGlxBrlHan:WI/ Charles River, GR	Gavage/Distilled water	Basic maintenance diet, Provimi KLIBA, GR	1
3	2	Ethinyl Oestradiol	Wistar rats Hsd Cpb:WU/ Harlan Winkelmann, GR	Gavage/Corn oil	Nafag 9439 25 W10 pellets, Eberle Nafag, SW	1
4	2	Flutamide	Wistar rats Hsd Cpb:WU/ Harlan Winkelmann, GR	Gavage/2% Cremophor EL aqueous solution	Altromin 1324 pellets, Altromin, GR	1
5	3	Methyl Testosterone	Wistar (AF) RJ: WI (IOPS AF)/ R. Janvier, FR	Gavage/ 0.5% methylcellulose aqueous solution	M20 contrôlé, Pietrement, FR	1
6	3	Tamoxifen	Wistar (AF) RJ: WI (IOPS AF)/ R. Janvier, FR	Gavage/ 3% ethanol and 0.5% methylcellulose aqueous solution	M20 contrôlé, Pietrement, FR	1
7	4	Genistein	Crj:CD(SD)IGS SPF/ Charles River, JP Atsugi Center	Gavage/ 1% CMC-Na aqueous suspension	CRF-1 pellet (irradiation sterilized), Oriental Yeast Co, JP	1
8	5	Ethinyl Oestradiol	Crj:CD(SD)IGS Rat, SPF/ Charles River, JP, Shiga Center	Gavage/Olive oil	MF pelleted diet, Oriental Yeast Co, JP	1
9	6	*p,p'*-DDE	Crl:CD®(SD)IGS BR/ Charles River, USA Kingston facility	Gavage/Corn oil	Certified Rodent Diet #5002, PMI Nutrition, USA	1
10	6	Nonylphenol	Crl:CD®(SD)IGS BR/ Charles River, USA Kingston facility	Gavage/Corn oil	Certified Rodent Diet #5002, PMI Nutrition, USA	1
11	7	*p,p'*-DDE	Crj:CD(SD)IGS SPF/ Charles River, JP Atsugi Center	Gavage/Corn oil	CE-2, Clea, JP	1
12	8	CGS 18302 B	Crj:CD(SD)IGS SPF/ Charles River, JP Atsugi Center	Gavage/Corn oil	MF Mash, Oriental Yeast Co, JP	1

Table 4 (continued). Laboratory parameters and conditions for Phase-2 updated TG 407 studies.

Study	Lab	Chemical	Rat strain/ Supplier	Administration/ Vehicle	Diet	Animals per cage
13	9	l-Thyroxine	Crj:CD(SD)IGS/ Charles River, JP	Gavage/0.5 % CMC-Na aqueous solution with 0.1 % Tween 80	CRF-1, Oriental Yeast Co, JP	1
14	10	Propylthiouracil	Sprague-Dawley (SD)/ Samtaco BIOKOREA	Gavage/Corn oil	Basal rat/mouse maintenance diet, Sam-Yang, Korea	
15	10	Tamoxifen	Sprague-Dawley (SD)/ Charles River, Korea	Gavage/Corn oil	Basal rat/mouse maintenance diet, Sam-Yang, Korea	1
16	11	Flutamide	Crj:CD IGS (SD)/Charles River, JP, Shiga Center	Gavage/2% Cremophor EL aqueous solution	CRF-1, Oriental Yeast Co, JP	1
17	12	Genistein	Alpk:AP$_f$SD (Wistar-derived)/ Alderley Park animal facility	Gavage/0.5 % CMC-Na aqueous solution	CT1, Special Diet Services, UK	5
18	12	Methyl Testosterone	Alpk:AP$_f$SD (Wistar-derived)/ Alderley Park animal facility	Gavage/0.5 % CMC-Na aqueous solution	CT1, Special Diet Services, UK	5
19	13	CGS 18302 B	HanIbm:WIST (SPF)/ RCC Ltd., SW	Gavage/0.5 % CMC-Na aqueous solution with 0.1 % Tween 80	Nafag No. 8900 FOR GLP, Eberle Nafag, SW	1
20	13	l-Thyroxine	HanIbm:WIST (SPF)/ RCC Ltd., SW	Gavage/0.5 % CMC-Na aqueous solution with 0.1 % Tween 80	Nafag No. 8900 FOR GLP, Eberle Nafag, SW	1

Table 5. Haematology and biochemical measurements made in Phase-2 of the updated TG407 studies

	Laboratory (Chemical)																		
Haematology and biochemical measurements	Lab 2 (Ethinyl Oestradiol)	Lab 5 (Ethinyl Oestradiol)	Lab 4 (Genistein)	Lab 12 (Genistein)	Lab 1 (Nonylphenol)	Lab 6 (Nonylphenol)	Lab 3 (Tamoxifen)	Lab 10 (Tamoxifen)	Lab 8 (CGS 18320B)	Lab 13 (CBS 18320B)	Lab 3 (Methyl T)	Lab 12 (Methyl Testosterone)	Lab 2 (Flutamide)	Lab 11 (Flutamide)	Lab 6 (DDE)	Lab 7 (DDE)	Lab 1 (Propylthiouracil)	Lab 10 (Propylthiouracil)	Lab 9 (l-thyroxine)	Lab 13 (l-thyroxine)
Haematology																				
haematocrit	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
haemoglobin conc.	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
erythrocyte count	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
mean erythrocyte volume	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
mean cell haemoglobin	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆
mean cell haemoglobin conc.	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
reticulocytes	◆	◆	◆		◆		◆		◆		◆		◆	◆		◆	◆		◆	
white blood cell count	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
platelet count	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
clotting time/potential	◆	◆	?	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
differential leucocyte	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
Clinical biochemistry																				
total cholesterol	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
triglycerides	◆	◆	◆	◆	◆	◆	◆			◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
total protein	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
albumin	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
globulin					◆	◆			◆	◆					◆		◆			◆
albumin/globulin ratio		◆	◆	◆		◆			◆	◆		◆		◆	◆	◆			◆	◆
alanine aminotransferase	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
aspartate aminotransferase	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
alkaline phosphatase	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
γ-glutamyl transpeptidase	◆	◆	◆	◆	◆	◆			◆	◆		◆	◆	◆	◆	◆	◆		◆	◆
bilirubin	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
creatinine	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
urea	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
glucose	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
sodium	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
potassium	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
calcium	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
chloride	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
inorganic phosphorus	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆

◆ - measurement was performed by the laboratory, but the methodology may not be standardized and the precise method may have been different from laboratory to laboratory.

Table 6. Organ weight measurements made in Phase-2 of the updated TG407 studies.

	Laboratory (Chemical)																			
Organ and Tissue Weights	Lab 2 (Ethinyl Oestradiol)	Lab 5 (Ethinyl Oestradiol)	Lab 4 (Genistein)	Lab 12 (Genistein)	Lab 1 (Nonylphenol)	Lab 6 (Nonylphenol)	Lab 3 (Tamoxifen)	Lab 10 (Tamoxifen)	Lab 8 (CGS 18320B)	Lab 13 (CBS 18320B)	Lab 3 (Methyl Testosterone)	Lab 12 (Methyl Testosterone)	Lab 2 (Flutamide)	Lab 11 (Flutamide)	Lab 6 (DDE)	Lab 7 (DDE)	Lab 1 (Propylthiouracil)	Lab 10 (Propylthiouracil)	Lab 9 (l-thyroxine)	Lab 13 (l-thyroxine)
adrenals	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
brain	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
epididymides [a]	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
heart	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
kidneys	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
liver	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
ovary	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
pituitary [b]	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
ventral prostate	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆			◆
dorsolateral prostate	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
whole prostate [c]		◆	◆	◆	◆	◆		◆	◆			◆		◆	◆	◆	◆	◆		
seminal vesicles	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
coagulating glands [d]	◆																			
spleen	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
testes	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
thymus	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
thyroid [b]	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
uterus	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆

◆ - measurement was performed by the laboratory, but the methodology may not be standardized and the precise method may have been different from laboratory to laboratory.

[a] sometimes weighed paired, sometimes single with the other side used directly for sperm parameters.

[b] Most often the pituitary and thyroid would be fixed, then dissection and trimming completed, then weighed – this was in order to avoid damaging the tissue and compromising the histopathology. This was sometimes done for other tissues, depending upon the practice of the particular laboratory.

[c] Most laboratories would, if it were weighed, weigh the whole prostate fresh, would fix it, then after dissection into the ventral and dorsolateral lobes, these tissues would be weighed.

[d] weighed separately from seminal vesicles.

Table 7. Histopathology measurements made in Phase-2 of the updated TG407 studies.

	Laboratory (Chemical)																			
Tissue Histopathology Analyses	Lab 2 (Ethinyl Oestradiol)	Lab 5 (Ethinyl Oestradiol)	Lab 4 (Genistein)	Lab 12 (Genistein)	Lab 1 (Nonylphenol)	Lab 6 (Nonylphenol)	Lab 3. (Tamoxifen)	Lab 10 (Tamoxifen)	Lab 8 (CGS 18320B)	Lab 13 (CBS 18320B)	Lab 3. (Methyl Testosterone)	Lab 12 (Methyl Testosterone)	Lab 2 (Flutamide)	Lab 11 (Flutamide)	Lab 6 (DDE)	Lab 7 (DDE)	Lab 1 (Propylthiouracil)	Lab 410 (Propylthiouracil)	Lab 9 (l-thyroxine)	Lab 13 (l-thyroxine)
adrenals	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
bone marrow	+	+	+	+	+	+	+	◆	+	+	+	+	+		+	+	+	◆		+
brain (including cerebrum, cerebellum and pons)	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
epididymides	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
heart	+	+	◆	+	+	+	◆	◆	◆	◆	◆	+	+		+	◆	+	◆	◆	◆
intestines (small)[a]	+	+	+	+	+	+	+	◆	+	+	+	+	+		+	+	+	◆		+
intestines (large)[a]	+	+	+	+	+	+	+		+	+	+	+	+		+	+	+			+
kidneys	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
larynx	+		+						+											
liver	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
lungs	+	+	+	+	+	◆	◆	◆	◆	◆	◆	+	+		◆	+	+	◆		◆
lymph nodes (route admin)	+	+	+	+	+	+	+		+	+	+	+	+		+	+	+			+
lymph nodes (systemic)	+	+	+	+	+	+	+		+	+	+	+	+		+	+	+			+
mammary gland (female)	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
mammary gland (male)	◆	◆	◆				◆		◆		◆		◆	◆		◆			◆	
ovaries	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
oviducts	◆		◆		◆	◆			+	+			◆		◆		◆		◆	+
pancreas	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆		◆
peripheral nerve	+	+	+	+	+		+	◆	+	+	+	+	+			+	+	◆		+
pituitary	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
prostate	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	DL	◆
seminal vesicles & coagulating glands.	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
spinal cord	+	+	+	+	+	+	+		+	+	+	+	+		+	+	+			+
spleen	◆	◆	◆	◆	+	+	◆	◆	◆	◆	◆	◆	◆		+	◆	+	◆	◆	◆
stomach	+	+	+	+	+	+	+	◆	+	+	+	+	+		+	+	+	◆		+
testes	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
thymus	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
thyroid	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
trachea	+	+	+		+	+	+	◆	+	+	+		+		+	+	+	◆		+
urinary bladder	+	+	+	+	+	+	+		+	+	+	+	+		+	+	+			+
uterus	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
vagina	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆
Other [b]	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓			✓		✓			✓
Obvious gross lesions	◆	◆	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆	◆	◆	◆		◆	◆

◆ - the tissue was fixed and the slides for this tissue were read by the laboratory, but the histopathological description and grading are not standardized and may have differed from laboratory to laboratory

\+ tissue was fixed, but not read, in the updated 407 study; DL – dorsolateral prostate lobe only.

[a] multiple sections to cover different regions (e.g., duodenum, ileum, and so on) were taken.

[b] ✓ Other tissues done by this laboratory besides those listed; may include additional measurements typically performed by the participating lab, but not formally required by the TG 407 protocol. Examples are the aorta, esophagus, Hardian glands, muscle - non-cardiac striated, salivary glands, and Zymbal's glands.

23. Final reports for TG 407 were not fully standardized, as to the terminology used in various sections, overall format, the detail of reporting for various methods, or the actual data reporting and format. These differences ranged from minor variations in the terms used for the same clinical chemistry assay to different descriptions and grading schemes for certain histopathological features. The latter in particular led to several inquiries by the Secretariat so that it could be ascertained that laboratories were indeed reporting the same effect in a target organ. In the Annexes, the descriptions of individual methods used for thyroid hormones and sperm counts and morphology have been included. In addition, in a following section, variations in statistical approaches among the laboratories have been reported.

Good Laboratory Practice compliance

24. The laboratories were encouraged to perform the updated TG 407 protocol in compliance with Good Laboratory Practices (GLP). From the reports received from the various laboratories, it appeared that 13 of the 20 studies were carried out in full compliance with GLP. Five of the studies were carried out under procedures that could be considered as generally complying with GLPs, but that would require a quality assurance audit for one or more aspects in order to bring the studies into full, confirmed compliance. Two studies were apparently not carried out in accordance with GLPs.

25. The GLP status of the studies is summarized in Table 8, and additional comments have been provided as to whether the test substance dosing solutions were analyzed to confirm that the nominal doses were indeed administered. This important study element was confirmed to have been done for 18 of the 20 updated TG 407 studies.

Selection of test substance doses

26. The Lead Laboratory, the participating laboratories, and other experts began a round of consultations beginning immediately after EDTA-4 to select the three doses to be used for each of the ten test substances. After compiling the available background data on the selected test substances, it became apparent that, for several substances, preliminary range-finding studies were necessary to approximate the no-observed-adverse effect dose and the maximum tolerated dose before undertaking the updated TG 407 studies. Five 7- or 14-day range finding studies were performed. Table 9 reports for each test substance whether proprietary data, literature values, and professional judgment were used to select the doses or whether range finding studies were conducted. In the cases of methyl testosterone and *p,p'*-DDE, different dose series were selected for males and females.

Table 8. Good Laboratory Practice (GLP) Compliance of the 407 Studies.

Chemical	Lab	GLP Compliance		Comments on GLP and substance and/or dosing sample analyses
		Yes[a]	No	
Ethinyl Oestradiol	2	✓		EE stock and dosing solution analyses were conducted and reported. For technical reasons, the dosing solutions were analysed at an outside, non-GLP compliant laboratory.
	5	✓		EE stability and stock and dosing sample analyses were stated to have been done, but specific results were not reported.
Genistein	4	✓		Genistein stock and dosing sample analyses were conducted to confirm levels, and the results were reported.
	12	✓		Genistein stock and dosing sample analyses were conducted to confirm levels, and the results were reported.
Nonylphenol	1	✓		Nonylphenol stock and dose sample analyses were conducted during the study, and results were reported.
	6	✓		Nonylphenol stock and dosing samples analyses were conducted during the study, and the analytical results were reported.
Tamoxifen	3		✓	According to the final report: "This study was not performed in compliance with Good Laboratory Practice in that it was not subjected to specific Quality Assurance inspections. It was performed according to standard operating procedures which were previously accepted and periodically inspected by Quality Assurance Unit."
	10		✓	The report did not state whether tamoxifen dosing sample analyses were done, and no analytical results were reported.
CGS 18320B	8	✓		CGS 18320B analyses for the test substance stability and dosing samples were conducted to confirm dosage levels, and the analytical results were reported.
	13	✓		CGS 18320B stock and dosing samples analyses were conducted to confirm dosage levels, and the analytical results were reported.
Methyl Testosterone	3		✓	According to the final report: "This study was not performed in compliance with Good Laboratory Practice in that it was not subjected to specific Quality Assurance inspections. It was performed according to standard operating procedures which were previously accepted and periodically inspected by Quality Assurance Unit."
	12	✓		Methyl testosterone stock and dosing sample analyses were conducted to confirm dosage levels, and the analytical results were reported.
Flutamide	2	✓		Flutamide stock and dosing solution analyses were conducted to confirm dosage levels and the analytical results were reported.
	11	✓		The report did not state whether Flutamide dosing sample analyses were done, and no analytical results were reported.
p,p'-DDE	6	✓		DDE test substance dosing samples were analysed for homogeneity and stability during the study, and the analytical results were reported.
	7		✓	DDE stability and test substance dosing samples were analysed, and the results were reported.
Propyl-thiouracil	1	✓		PTU test substance dosing samples were analysed during the study, and the results were reported.
	10		✓	The report did not state whether PTU dosing sample analyses were done, and no analytical results were reported.
l-Thyroxine	9		✓	Thyroxine test substance dosing samples were analysed, and the analytical results were reported.
	13	✓		Thyroxine analyses for the test substance dosing samples were conducted to confirm dosage levels, and the analytical results were reported.

[a] As stated and reported in the final report submitted to the OECD.

Table 9. Summary of the procedures used to select test substance doses for the updated TG 407 studies.

Test substances whose doses were selected based on literature values and professional judgments	Test substances whose doses were selected based on preliminary range finding studies
Ethinyl Oestradiol [a]	Genistein
Tamoxifen [a]	Nonylphenol
CGS 18320B [a]	Methyl Testosterone
Flutamide [a]	*p,p'*-DDE
Propylthiouracil	l-Thyroxine

[a] Proprietary data in pharmacological studies were the primary data source.

27. The doses for each test substance including any sex specific doses or modifications of those doses between studies or during a study are reported in Table 10. It should be noted that the animals were weighed periodically during the 28-day period as required by the TG 407 guidelines. These body weights were used to calculate the necessary gavage volume from the standard dosage solution for each animal, and the volumes were revised when new body weights were recorded. Thus, combined with the analyses of the dosing solutions, the dose were reasonably consistent and met the nominal targets across the 28-day period.

Table 10. The selected doses employed for each test substance in the updated TG 407 studies.

Chemical	Male Doses	Female Doses
Ethinyl Oestradiol	10, 50, 200 µg/kg/day	Same as for males
Genistein	120, 400, 1000 mg/kg/day	Same as for males
Nonylphenol	Study 1: 20, 80, 300/250[a] mg/kg/day	Same as for males
	Study 2: 20, 80, 200/150 [a] mg/kg/day	Same as for males
Tamoxifen	5, 30, 200 µg/kg/day	Same as for males
CGS 18320B	0.3, 3, and 30 mg/kg/day	Same as for males
Methyl Testosterone	10, 40, 200 mg/kg/day	10, 100, 600 mg/kg/day
Flutamide	1, 10, 100 mg/kg/day	Same as for males
p,p'-DDE	Study 1: 12.5, 50, 200/150 [a] mg/kg/day	Study 1: 6.5, 25, 100 mg/kg/day
	Study 2: 12.5, 50, 100 mg/kg/day	Study 2: 6.5, 25, 75 mg/kg/day
Propylthiouracil	0.1, 1, 10 mg/kg/day	Same as for males
l-Thyroxine	0.01, 0.1, 1 mg/kg/day	Same as for males

[a] The top dose was reduced during the study due to animal mortality, moribundity, or excessive weight loss.

Test Substance Supply

28. The European Chemical Industry Association (CEFIC) has previously supported the chemical supply for the uterotrophic and Hershberger validation programs. CEFIC also supported the extensive chemical supply needs of the updated TG 407 validation program by providing financial and managerial responsibility for a centralised chemical repository and the acquisition of the 10 test substances. TNO (The Netherlands) continued to serve as the centralised chemical repository as it had for other validation programs and phases. CEFIC arranged for the test substances to be purchased, donated, or acquired by synthesis and deposited at the central repository. Where test substances were included in more than one validation program, it was attempted to acquire sufficient quantities of test substances so that the same batch of test substance would be used for current and future use across validation programs.

29. After the participation of each laboratory was confirmed, the quantities of test substance that would be needed by each laboratory were calculated and some excess provided to account for possible errors and wastage. Test substances were shipped in compliance with regulatory and customs requirements of each nation where participating laboratories were located. Shipments were timed to arrive before the

study animals in order to avoid wastage, e.g., expiration of the time window for using immature animals. Other details of the substance supply and handling included:

- Participating laboratories were asked to ensure that personnel performing the necropsies would not know the identity or doses of the test substances administered to the animals.
- Generic Material Safety Data Sheets were prepared and provided for all test substances so that the health and safety of personnel at the laboratory would not be compromised. These Safety Data Sheets were provided in sealed envelopes to a specified individual at each laboratory who agreed to keep this envelope sealed unless in cases of emergency.

Group sizes

30. An important element of the updated TG 407 studies was the evaluation of the statistical power of the group size, particularly for those endpoints considered to be potentially endocrine sensitive. This was accomplished by running two parallel studies, either concurrently or in a block design. Each study was composed of 5 animals per sex per dose as is the practice under the current TG 407. Both studies were to be analyzed first as independent and individual studies.

31. The raw data from these two 'individual' studies were then combined so that it presented a study comprising 10 animals per sex per dose, and the data were statistically reanalyzed as a third study. The results were then compared for any differences in statistical significance, any inconsistencies, and in order to consider the overall value of possibly increasing the TG 407 group size to 10 animals per sex per dose in order to effectively identify endocrine-active substances.

32. The nomenclature used hereafter in this report refers to those studies using 5 animals per sex per dose as individual Subgroups A and B. When these two Subgroups are combined to form 10 animals per sex per dose, then the nomenclature used hereafter refers to the combined Subgroups.

33. Unfortunately, not all laboratories followed these directions for their studies or in all aspects of their studies. The actual status is as follows and is also summarized in Table 11.

- For 13 studies, the above instructions were followed and there are data for Subgroups which began with n=5 for each sex in each dose group (although some mortalities occurred or in some cases samples and tissues were insufficient or lost) and data for combined Subgroups with n=10 for each sex in each dose group.
- In 3 studies, the directions were followed in part, except for histopathology. In these studies, the original work was indeed performed in two individual Subgroups and later combined for almost all measurements. However, the fixed histopathological samples were reported and analysed as a single group of 10 per sex per dose (combined Subgroups). The primary cause appears to have been that the slides were sent to an outside pathologist without instructions to read and record the individual Subgroups separately. In these cases, the histopathology data for the individual Subgroups were not available for this report.
- Another 2 studies were performed, analysed, and reported as 10 animals per sex per dose, and no data or analyses for individual Subgroups were available.
- Another 2 studies were performed as a single study of 6 animals per sex per dose. Therefore, the available data were similar, but not identical to, a single Subgroup study for the corresponding duplicate study.

Statistical procedures used by the laboratories

34. The statistical analysis needs of the current and updated TG 407 data are relatively complex. TG 407 does not require that a particular statistical approach should or should not be used. Therefore, the following illustrative points are made to indicate why different laboratories have chosen or applied different statistical approaches.

- The data are of different types, indicating that different statistical approaches should be used, e.g., continuous data such as a tissue weight or serum clinical value and non-continuous parametric data such as the grades for a histological observations.
- There are different choices in approach, such as a strictly pairwise comparison approach between separate groups (e.g., t-test) or a multiple comparison approach across groups between a vehicle control and treatment groups (e.g., Dunnett's). The former will result in accepting a higher rate of false positives and a slightly greater chance of achieving statistical significance in marginal circumstances.
- The acceptable level for false positives and achieving significance, e.g., $p < 0.05$ or <0.01, and whether more than one level was analysed for statistical significance.
- A particular statistical approach may require certain assumptions be met, e.g., the distribution of the data to be normal or relatively homologous. This can require the use of transformation techniques, which can themselves vary depending upon the nature of the data distribution, e.g., a square root transformation is usually considered appropriate where the group variances are proportional to the group means, while a log transformation is usually considered appropriate where the group standard deviations are proportional to the group means. Thus, when normality or homogeneity of the data cannot be achieved, then an alternative statistical approach appropriate for non-normal data may be chosen.
- Certain procedures may be carried out in a given sequence, e.g., when there is a significant overall ANOVA F-test result, this may be followed by a multiple comparison using Tukey's technique. Should the conditions and assumptions of the first step not be met, there may then be an alternative approach to the second following step.
- There may be other elements such as the group size and the number of groups that suggest whether a particular statistical approach is appropriate or not, and
- There may still be several options and choices as to different details in these approaches.

Table 11. Status of the use of individual and combined Subgroups for comparison of the statistical power of the updated TG 407 Studies.

Chemical	Laboratory	Two Subgroups of 5		Comments
		Yes	No	
Ethinyl Oestradiol	2	✓		The laboratory updated the calculations for the combined Subgroups of 10 for the OECD Secretariat to include a decimal place for the means and to report the SDs.
	5	✓		
Genistein	4	✓		
	12	✓		
Nonylphenol	1	✓		
	6	✓		The histopathological results are reported for the combined Subgroups of n = 10 and not as individual Subgroups of n = 5.
Tamoxifen	3	✓		
	10		✓	Used a single group size for both males and females of n = 6.
CGS 18320B	8	✓		
	13		✓	This laboratory ran only a single group size of n = 10 and did not conduct the study with individual Subgroups of 5.
Methyl Testosterone	3	✓		
	12	✓		
Flutamide	2	✓		The laboratory updated the calculations for the combined Subgroups of 10 for the OECD Secretariat to include a decimal place for the means and to report the SDs.
	11	✓		
p,p'-DDE	6	✓		The histopathological results are reported for the combined Subgroups of n = 10 and not as individual Subgroups of n = 5.
	7	✓		
Propylthiouracil	1	✓		
	10		✓	Used a single group size for both males and females of n = 6.
l-Thyroxine	9	✓		
	13		✓	This laboratory ran only a single group size of n = 10 and did not conduct the study with individual Subgroups of 5.

35. The result is that a variety of statistical methods were used by the laboratories in the course of the updated TG 407 studies. An effort has been made to identify the procedures in each study, and these are summarised in Table 12.

Table 12. Statistical procedures used in Phase-2 of Updated TG407 studies.

	Laboratory (Chemical)																			
Statistical procedures	Lab 2 (Ethinyl Oestradiol)	Lab 5 (Ethinyl Oestradiol)	Lab 4 (Genistein)	Lab 12 (Genistein)	Lab 1 (Nonylphenol)	Lab 6 (Nonylphenol)	Lab 3(Tamoxifen)	Lab 10 (Tamoxifen)	Lab 8 (CGS 18320B)	Lab 13 (CGS 18320B)	Lab 3 (Methyl Testosterone)	Lab 12 (MT)	Lab 2 (Flutamide)	Lab 11 (Flutamide)	Lab 6 (DDE)	Lab 7 (DDE)	Lab 1 (Propylthiouracil)	Lab 10 (Propylthiouracil)	Lab 9 (l-thyroxine)	Lab 13 (l-thyroxine)
Assumption Checks:																				
Test for homogeneity of variance																				
Bartlett's (alpha=0.05)							**X**			**X**	**X**					**X**				**X**
Bartlett's (alpha=0.01)			**X**			**X**									**X**					
Bartlett's (alpha=?)		**X**												**X**					**X**	
None or not indicated	**X**			**X**	**X**				**X**			**X**	**X**				**X**			
Test for normality																				
Yes	**X**												**X**							
No or not indicated		**X**	**X**	**X**	**X**	**X**	**X**		**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**		**X**	**X**
Statistical Tests for Group Differences:																				
Test for significance – constant variance and normally-distributed data																				
Dunnett's	**X**		**X**		**X**		**X**				**X**		**X**	**X**			**X**			
ANOVA / Dunnett's		**X**				**X**			**X**	**X**					**X**	**X**			**X**	**X**
ANOVA or ANCOVA / t-tests				**X**								**X**								

ANOVA – analysis of variance; ANOVCA – analysis of covariance.

Table 12 continued. Statistical procedures used in Phase-2 of Updated TG407 studies.

Statistical procedures	Laboratory (Chemical)																			
	Lab 2 (Ethinyl Oestradiol)	Lab 5 (Ethinyl Oestradiol)	Lab 4 (Genistein)	Lab 12 (Genistein)	Lab 1 (Nonylphenol)	Lab 6 (Nonylphenol)	Lab 3(Tamoxifen)	Lab 10 (Tamoxifen)	Lab 8 (CGS 18320B)	Lab 13 (CGS 18320B)	Lab 3 (Methyl Testosterone)	Lab 12 (MT)	Lab 2 (Flutamide)	Lab 11 (Flutamide)	Lab 6 (DDE)	Lab 7 (DDE)	Lab 1 (Propylthiouracil)	Lab 10 (Propylthiouracil)	Lab 9 (l-thyroxine)	Lab 13 (l-thyroxine)
Test for significance – non-constant variance or non-normal data																				
Non-parametric ANOVA / Dunnett-type mean rank		X							X							X			X	
Dunnett-type mean rank			X																	
Welch test – p-value adjusted (Holm's)	X												X							
Steel's test														X						
Kruskal-Wallis / Dunn's test						X				X					X					X
Jonckheere's trend test						X									X					
Kruskal-Wallis / Mann-Whitney-Wilcoxon tests	X				X								X				X			
Mann-Whitney-Wilcoxon tests			X				X				X			X						
Additional Information on Statistical Procedures:																				
Data adjustment																				
Log transformation of certain variables	X						X				X		X							
Arcsine transformation of certain variables				X								X								
None or not indicated		X	X		X	X			X	X				X	X	X	X		X	X
Additional Tests																				
Regression – linear dose response						X									X					
Regression lack-of-fit						X									X					
Fisher's exact tests									X					X						
Test Type																				
One-sided																				
Two-sided	X		X	X	X		X			X	X	X	X			X	X			X
Unknown		X				X			X					X	X				X	

ANOVA – analysis of variance; ANOVCA – analysis of covariance.

36. The net result will be that there are likely subtle differences in the statistical sensitivity and in a few results among laboratories in this report due to the different statistical approaches. That is, if the different procedures were performed on the same data set, one laboratory might achieve statistical significance in a marginal circumstance while a second laboratory might fail to achieve statistical significance. As the raw data were not available and the task of statistical reanalysis was enormous, the Secretariat has used the statistical results of the individual laboratories for comparisons despite this potential shortcoming. However, it is judged that these instances should be limited and that other variables were likely to be much greater contributors to the variations observed.

RESULTS OF THE UPDATED TG 407 STUDIES

37. In this chapter, the results of the updated TG 407 studies with each test substance are summarized and the two studies are compared. The results for each test substance are summarized and compared for the body weights, haematology and clinical chemistry, absolute and relative organ weights, histopathology, thyroid hormones, sperm counts and morphology, and oestrous cycle as measured by vaginal smears. The overall results are then summarized and compared to available subchronic, chronic, and reproductive data on the test substances or similar chemicals in order to assess the ability of the updated TG407 studies to identify adverse effects on reproductive and endocrine tissues as well as other organs related to both their endocrine and non-endocrine mechanisms of action.

STUDIES OF ENDOCRINE ACTIVE COMPOUNDS

Ethinyl Oestradiol

Introduction

38. This section summarizes the results of the updated TG 407 studies with ethinyl oestradiol (EE) and compares these results with a 90-day study (26) and 1-generation reproductive study (27)(28)conducted with 17β-estradiol.

Background on EE and estrogens

39. EE is a potent pharmacological estrogen taken orally for birth control. It has a somewhat higher binding affinity for the estrogen receptor alpha than the endogenous 17β-estradiol with a relative binding affinity of log 2.28 (29). EE is slightly more potent when these compounds are compared in the uterotrophic bioassay. The EE first achieved a statistically significant increase in uterine weight at an oral gavage dose of 3 μg/kg body weight/day, and 17β-estradiol first achieved a statistically significant increase in uterine weight at an oral gavage dose of 100 μg/kg body weight/day (30).

Description of EE experiments

40. EE was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in both updated 407 studies. The EE was administered using edible oil vehicles (corn oil in study 1 and olive oil in study 2) in three doses to both sexes in each study: a low (L) dose of 10 μg/kg body weight/day, an intermediate (I) or mid-dose of 50 μg/kg body weight/day, and a high (H) dose of 200 μg/kg body weight/day. Both studies incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The individual data from the Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power.

Summary of the updated TG 407 results with EE

41. The two updated TG 407 produced concordant response profiles to EE, and the EE effect levels observed were similar. A clear pattern of estrogenic responses were observed in both female and male reproductive tracts as well as the male mammary gland. The first study observed adverse responses at the lowest EE dose (10 μg/kg body weight/day) so that no NOEL could be set; the second study observed similar responses at 50 μg/kg body weight/day, so that 10 μg/kg body weight/day in the second study would be judged a NOEL. The pattern of tissue weight changes and histopathological changes observed in

the updated TG 407 study were similar to those seen with 17β-estradiol taking into account the expected differences in potency. Therefore, it is concluded that the updated TG 407 successfully detected effects of EE consistent with the endocrine mechanism of action of an estrogen.

Mortality, clinical observations and body weights in EE studies

42. No mortalities were observed in the animals at any EE dose level, and there were no treatment-related clinical signs in either study. Food consumption was decreased in absolute terms at the high EE dose in both sexes in each study, but did not achieve statistical significance in any instance. No statistically significant differences were recorded between the control and any dose group in the functional observational battery and the motor activity assessment in either study. Male body weights were significantly decreased in both studies at the high EE dose in the combined Subgroups and in one of the individual Subgroups. Male body weights decreased more than 10% in both remaining individual Subgroups in both studies, but did not achieve statistical significance (Table 13). Female body weights did not change significantly. The mean body weight changes in the female individual Subgroups ranged from +3% to –9.7% of the controls (Table 13). Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with EE

43. Eight haematological and clinical chemistry parameters were significantly changed in the combined Subgroups of a sex in both studies (studies in agreement). These parameters were female erythrocyte counts, haemoglobin concentration, albumin, total protein, alkaline phosphatase, and aspartate aminotransferase, male total cholesterol levels, and female total cholesterol levels (Table 14). There were no parameters where the statistically significant changes in one study were directionally in conflict with the other study. For five other parameters, statistical changes were observed in one study with similar changes in the direction of the absolute values occurring in the second study. These parameters were male and female haematocrit values, female reticulocyte counts, and male triglycerides, aspartate aminotransferase, and haemoglobin concentration (Table 14). There were six parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 14). More detailed summaries of the haematological and clinical chemistry findings, including the combined Subgroups and individual Subgroups, are in Annex 4.

Table 13. Changes in body weights during 407 studies with EE.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
2	M	- - +2.9%	- - -5.1%	↓↓ **-12.5%**	- - -3.1%	- - -8.7%	↓↓ **-13.7%**	- - +9.3%	- - -1.7%	- - -11.9%
	F	- - +3.5%	- - +3.5%	- - +1.0%	- - +2.5%	- - +4.5%	- - +3.0%	- - +5.1%	- - +3.0%	- - -1.0%
5	M	- - -1.0%	- - -3.0%	↓↓ **-14.0%**	- - -1.7%	- - -2.6%	- - -12.4%	- - -0.4%	- - -3.4%	↓↓ **-15.6%**
	F	- - -0.6%	- - -0.8%	- - -6.9%	- - -1.2%	- - -4.2%	- - -9.7%	- - +0.1%	- - +2.9%	- - -3.8%

M: male; F: female; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

Table 14. Haematology and clinical chemistry results from 407 studies with EE.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 2	Statistically Significant in Laboratory 5
Significant Parameters in Agreement in Both Studies: ↓ F Haemoglobin concentration ↓ F Erythrocyte counts ↓ M Total cholesterol ↓ F Total cholesterol ↓ F Albumin ↑ F Total protein ↑ F Alkaline phosphatase ↓ F Aspartate aminotransferase	***Common measures, direction of percentage change similar, significant in first study only:*** ↑ F Reticulocyte counts **50.0** 42.9[a] ↑ M Triglycerides **42.6** 13.6 ↓ M Aspartate aminotransferase **-29.9** -15.9	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ M Haematocrit -5.7 **-8.6** ↓ F Haematocrit -1.3 **-5.1** ↓ M Haemoglobin conc. -3.6 **-5.7**
	Common measures, direction of percentage change differ, significant in first study only: ↑ M Reticulocyte counts **31.8** -5.7 ↑ M Prothrombin quick times **9.1** -3.4 ↑ F Prothrombin quick times **20.7** 2.5 ↓ M Total bilirubin **-36.4** 28.6 ↓ M Sodium **-1.4** 0.0	***Common measures, direction of percentage change differ, significant in second study only:*** ↑ M Alkaline phosphatase 0.0 **23.9**
	Significant measurement performed in first lab only: ↓ M Leucocyte values [b] ↓ M Atypical leucocyte values [b] ↓ M Lymphocyte values [b] ↓ M Basophil values [b]	***Significant measurement performed in second lab only:*** ↓ F Cholinesterase ↓ M ↓ F Albumin/globulin ratio ↑ M ↑ F γ-Glutamyl Transferase

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. Immediately below the specific parameter, the first study percentage change in the absolute value is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

[b] The decrease in leucocyte, atypical leucocyte, lymphocyte, and basophil values were judged to be spurious based on 1) the control means were driven by high values in one control animal in Subgroup A and 2) the treated animal values were within historical values.

Organ and tissue weights results with EE

44. There were concordant statistically significant changes in major organs in the combined Subgroups of both EE studies. In males, the absolute adrenal weights were increased and the absolute kidney weights decreased in both studies at the high EE dose (Table 15A). The relative weights of the liver and adrenals were increased in both studies at the high EE dose, and the relative brain weights were increased probably due to the significant decreases in body weights at the high EE dose (Table 15B). In females, the absolute liver weights were increased in both studies at the high EE dose (Table 15A), and the relative liver weights were increased at all three EE doses (Table 15B). In a number of cases, the

statistically significant changes in the combined Subgroups were not observed in the individual Subgroups (Tables 15A and B)

45. There were concordant statistically significant changes in the reproductive tracts of both males and females in the combined Subgroups of both studies. In males, there were statistically significant absolute and relative decreases in the seminal vesicles, the coagulating glands, and the ventral and dorsolateral prostate in both studies at the high EE dose. In most cases, statistical significance was also observed in both individual Subgroups (Tables 15A and B). The absolute weights of the epididymis was statistically decreased in one study and the absolute values were decreased by greater than 10% in the second study at the high EE dose, but the relative weights were not significantly different in either case, noting the decrease in male body weights (Tables 15A and B). In the females, the absolute and relative ovary weights were significantly decreased in the second study at the high EE dose, but not in the first study. The absolute and relative uterine weights were significantly increased in the second study. There was an absolute increase of greater than 10% in the uterine weights in the first study, and the relative uterine weight achieved a statistically significant increase at the high EE dose (Tables 15A and B). The means and standard deviations of the absolute weights and the individual relative weights for both studies including the combined Subgroups and individual Subgroups are in Annex 5.

Histopathology findings with EE treatment

46. Taking into account different subjective descriptions of the histopathological lesions in the two studies, the findings were concordant. The primary observation in the major organs was the changes in the male adrenal cortex, and this was observed in both studies (Table 16).

47. In the male reproductive tract, concordant with the tissue weights, atrophic changes in the seminal vesicles, coagulating glands, and the ventral and dorsolateral prostate were observed in both studies. The observations were first made at the intermediate EE dose, and the frequency and severity of the observations increased at the high EE dose (Table 16). Changes in the testes were described differently in the two studies. In the first study, germinal epithelial degeneration was noted at a low frequency at the high EE dose and Leydig cell atrophy was observed in one animal at the intermediate EE dose and 8 of 10 animals at the high EE dose. In the second study, the only observations recorded were spermatocyte degeneration in 1 and 2 of 10 animals at the intermediate and high doses, respectively (Table 16).

Table 15A. EE-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
2	M	Liver							▲	▲	
		Kidney			↓			↓			▽
		Adrenal		▲	↑↑		▲	↑↑	▲	▲	↑
		Brain									
		Spleen			▽			▽			
		Thymus		▽	▽		▽	▽			▽
		Testes							▲		
		Epididymis		▽	▽			▽	▲	▽	
		Seminal Vesicles		▽	↓↓		▽	↓↓		▽	↓↓
		Coagulating Glands			↓↓	▽	▽	↓↓	▲	▽	↓↓
		Ventral Prostate	▲	▽	↓↓	▲	▽	↓↓	▲	▽	↓↓
		Dorsolateral Prostate		▽	↓↓		▽	↓			↓↓
	F	Liver	▲	↑↑	↑↑	▲	↑	▲	▲	▲	↑
		Adrenal		▲	▲		▲	▲			▲
		Spleen		▲	▲		▲	↑		▲	
		Thymus				▽					
		Ovaries									
		Uterus	▲	▲	▲	▲	▲	▲			▲
5	M	Liver									
		Kidney			↓			↓↓			▽
		Heart			↓			↓			
		Adrenal		▲	↑		▲	↑↑		▲	▲
		Brain									
		Spleen									
		Thymus									
		Pituitary									
		Testes									
		Epididymes			↓			▽		▽	↓
		Seminal Vesicles	▽	↓	↓↓		▽	↓↓	▽	↓	↓↓
		Ventral Prostate		↓	↓↓		▽	↓↓	▽	▽	↓
		Dorsolateral Prostate		↓↓	↓↓		▽	↓↓		↓	↓↓
	F	Liver		↑	↑	▲	▲	▲		▲	▲
		Ovary			↓↓	▲	▲	▽	▽	▽	↓↓
		Uterus			↑↑			↑	↑		▲

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 15B. EE-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
2	M	Liver			↑	[b]	[b]	[b]		↑	
		Kidney									
		Adrenal		↑↑	↑↑		↑	↑	[b]	[b] ▲	[b] ▲
		Brain			↑			↑			▲
		Spleen						▲			
		Thymus					▽				
		Pituitary			↑	▲		▲		▲	▲
		Testes									▲
		Epididymes						▽		▽	
		Seminal Vesicles		▽	↓↓		▽	↓↓		[b] ▽	↓↓
		Coagulating Glands		▽	↓↓	▽	▽	↓		[b] ▽	↓↓
		Ventral Prostate		▽	↓↓	▲	▽	↓		[b] ▽	↓↓
		Dorsolateral Prostate		▽	↓↓			▽	▽	[b] ▽	↓↓
	F	Liver	↑	↑↑	↑↑		↑↑	↑↑	▲	▲	↑↑
		Adrenal		▲	▲		▲				▲
		Spleen			▲			↑↑			
		Thymus				▽					
		Ovaries					▽				▽
		Uterus	▲	▲	↑	▲	▲	▲			▲
5	M	Liver		↑↑	↑↑		↑↑	↑↑		▲	↑↑
		Kidney									
		Heart									
		Adrenal		▲	↑↑		▲	↑↑		▲	▲
		Brain			↑↑			↑			↑
		Spleen		▲	▲		▲			▲	▲
		Thymus			▲		▲	▲			▲
		Pituitary		↑↑	↑↑		↑↑	↑↑		↑	▲
		Testes			↑			▲			▲
		Epididymes									
		Seminal Vesicles		↓	↓↓		▽	↓↓	▽	▽	▽
		Ventral Prostate [a]		▽	↓↓		▽	↓	▽	▽	▽
		Dorsolateral Prostate		↓↓	↓↓		▽	↓↓		↓	↓
	F	Liver	↑	↑↑	↑↑		↑↑	↑↑		↑	↑↑
		Ovary			↓↓			▽	▽	▽	↓↓
		Uterus			↑↑			↑↑	▲		▲

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

[b] statistics not calculated, where a control or test substance group n = ≤ 4 (study B), then the power of the test was considered too low to provide meaningful information.

In females, the first study recorded finding an estrogenic pattern of changes in both the uterus (increased epithelial cell height) and vagina (keritinisation and mucification) at the low EE dose with the number of affected animals increasing in a dose related manner (Table 16). The second study recorded similar uterine and vaginal observations, but only at the high EE dose (Table 16). The observations in the ovaries appear to have been consistent taking into account differences in the descriptions: the follicles were described as atrophied in first study and, in the second study, the observations were described as an increased proportion of small follicles. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6. In all of the above cases, the frequency of findings was similar in the two individual Subgroups.

Thyroid hormone results with EE treatment

48. At one or more times, each of the three thyroid hormone values achieved a statistically significant difference in the two studies (Table 17). In several cases, these changes were not dose-related. Otherwise, none of the statistically significant changes that could be interpreted as dose-related in one study were replicated in the other study. In most cases, statistical significance was achieved only in the combined Subgroups. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with EE treatment

49. No significant changes were observed in either sperm numbers or morphology in either study, including both the combined Subgroups and the two individual Subgroups at any EE dose (Table 18). Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Estrous cyclicity

50. In one study, there was a clear dose-response in the loss of oestrous cyclicity as measured by vaginal cytology (smears) with 2/10 animals at the intermediate EE dose and 10/10 animals at high EE dose. In contrast, the vaginal cytology smears in the other study were judged originally to be representative of normal cycling animals in a diestrous stage. However, when the smear data were compared to the uterine and vaginal histopathology results, most animals were judged by these results to have been in oestrus or proestrus, not dioestrus. The cause of this apparent discrepancy was not resolved. The vaginal cytology results from both studies are in Annex 9.

Comparison of the updated TG 407 results from EE treatment with data from chronic, reproductive, and developmental studies

51. A one-generation study has been conducted on 17β-estradiol (E2) with a full complement of standard and estrogen sensitive endpoints using 30 animals per sex per dose, and where parallel groups of 10 animals per sex at the same dose were administered 17β-estradiol in a 90-day subchronic study (26)(27)(28). Administration was via the diet at four calculated doses of 3 μg E2/kg body weight/day in both sexes; 140 and 170 μg E2/kg body weight/day in males and females, respectively; 530 and 690 μg E2/kg body weight/day in males and females, respectively; and 3.16 and 4.12 mg E2/kg body weight/day in males and females, respectively. For comparative purposes, EE is more potent than E2 and the gavage versus dietary administration should also be taken into account. However, both are potent estrogens and the same basic toxicological profile would plausibly be expected.

Table 16. Significant histopathological findings after EE treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
2	M	**Adrenal**	Reduced zona fasculata vacuolisation	2/10	2/10	6/9	2/5	2/5	3/4			3/5
		Mammary glands	Increased ratio basophilic to acidophilic cells	2/9	5/10	8/9	2/4	2/5	5/5		3/5	3/4
		Testes	Germinal epith. degeneration			2/10			1/5			1/5
			Leydig cell atrophy		[a]	7/9[a]			3/4[a]		[a]	4/5
		Epididymis	Atrophy		[a]	[a]			[a]		[a]	
		Seminal Vesicles	Atrophy		4/10	8/10			3/5		4/5	5/5
		Coagulating Glands	Atrophy		1/9	10/10		1/4	5/5			5/5
		Ventral Prostate	Atrophy			6/9			3/4			3/5
		Dorsolateral Prostate	Atrophy			5/9			2/4			3/5
	F	**Ovaries**	Increased follicles			4/10			2/5			2/5
		Uterus	Increased epithelial cell height [b]	3/10	3/10	6/10	1/5	1/5	4/5	2/5	2/5	2/5
			Granulocytic infiltration increased [b]	7/10	6/10	5/10	5/5	4/5	4/5	2/5	2/5	1/5
			Dilation	1/10	1/10	3/10	1/5	1/5	2/5			1/5
		Vagina	Thickened epithelium [c]	1/10	4/10	10/10	1/5	3/5	5/5		1/5	5/5
			Keratinisation [c]	3/10	4/10	6/10	1/5	2/5	3/5	2/5	2/5	3/5
			Mucification [c]	1/10	1/10	4/10		1/5	2/5	1/5		2/5

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

[a] One male in these groups appeared to be sexually immature. Not only were the Leydig cells small, the germinal epithelium of the testes was obviously immature and the sperm counts of this individual were low. These individuals' observations were judged not to be treatment related, and they are not included.

[b] Control values were 0/10 for uterine epithelial cell height; 8/10 for > grade 1 granulocytic infiltration and treated numbers are animals > grade 1.

[c] Control values were 0/10 for thickened epithelium, keratinization, and Mucification.

Table 16 continued. Significant histopathological findings after EE treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
5	M	**Adrenal**	Cortical hypertrophy			3/10			2/5			1/5
		Testes	Spermatocyte degeneration		1/10	2/10			1/5		1/5	1/5
		Mammary	Atrophy		1/10	4/10			2/5		1/5	2/5
		Ventral Prostate	Atrophy		1/10	4/10			2/5		1/5	2/5
		Dorsolateral Prostate	Atrophy		1/10	3/10			2/5		1/5	1/5
		Seminal Vesicles	Atrophy		1/10	5/10			2/5		1/5	3/5
	F	**Ovary**	Atrophy			2/10			1/5			1/5
		Uterus	Squamous epithelial metaplasia			1/10			1/5			
			Increased glandular epithelial ht			1/10			1/5			
			Increased lumen epithelial height			6/10			4/5			2/5
		Vagina	Mucification			7/10			3/5			4/5

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Table 17. Thyroid hormone results after EE treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
2	M	T_3	↑								
		T_4									
		TSH		↑	↑	nc	nc	nc	nc	nc	nc
	F	T_3									
		T_4		↑	↑		↑	↑			
		TSH	↑	↑	↑		↑				
5	M	T_3									
		T_4			↑						
		TSH									
	F	T_3			↓						↓
		T_4									
		TSH	↑								

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in body weight ($p<0.05$); ↓↓ - statistically significant decrease in body weight ($p<0.01$); ↑ - statistically significant increase in body weight ($p<0.05$); ↑↑ - statistically significant increase in body weight($p<0.01$).

nc - statistics were not calculated as vehicle control for both Subgroups was reduced to ≤ 4 animals. However, a comparison of absolute values and SD's with those female combined Subgroups that did achieve statistical significance would indicate it is a reasonable assumption that statistical significance could have been otherwise achieved.

Table 18. Sperm parameters after EE treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
2	Sperm Count	Data were not analyzed for Combined Subgroups								
	Abnormalities	Data were not analyzed for Combined Subgroups								
5	Sperm Count									
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group; Not applicable – Analyses of individual Subgroups were not performed or the study was not conduced with individual Subgroups. ND – No analysis performed on these groups.

52. The selected findings in the two updated TG 407 studies with EE and the 90-day and 1-generation studies with E2 are compared in Table 19. The comparison should take into account the somewhat lower potency of E2, the extended time of the E2 study, and the dietary versus gavage administration differences.

Table 19. Comparison of updated TG 407 EE results with other estrogen studies.

Parameter	**EE – updated TG 407** (EE LOEL µg/kg/d)				**E2 – 90-day and reproduction studies** (E2 LOEL µg/kg/d)			
	Laboratory 2		**Laboratory 5**		**90-day (26)**		**Reproduction (27)(28)**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Body weight	↓ 200		↓ 200		↓ 140	↓ 140	↓ 140	↓ 140
Haematology and clinical chemistry								
Haematocrit			↓ 200	↓ 50	↓ 500	↓ 170		
Haemogloblin		↓ 200	↓ 200	↓ 50	↓ 500	↓ 700		
Erythrocyte counts		↓ 200		↓ 50	↓ 500	↓ 700		
Total cholesterol	↓ 10	↓ 200	↓ 10	↓ 200	↓ 3000	↓ 4000		
Organ and Tissue Weights	***(relative)***				***(relative P)***			
Liver	↑ 200	↑ 10	↑ 50	↑ 10				↑ 700
Adrenal	↑ 50		↑ 200				↑ 500	↑ 4000
Epididymis							↓ 3000	
Accessory Male Reproductive Tissues [a]	↓ 200 (S, C, V, D)		↓ 200 (V) ↓ 50 (S, D)				↓ 500	
Uterus		↑ 200		↑ 200				↑ 700
Ovaries				↓ 200				↓ 700
Histopathological Findings					***(for P animals)***			
Mammary gland	50		50				140	
Testes – interstitial/Leydig	200						500	
Testes – seminif./spermatocyte degeneration	200		Eq 50				500	
Epididymal –atrophy							500	
Prostate – atrophy	200		50				500	
Seminal vesicles – atrophy	50		50				50	
Uterus – hyperplasia/increased epithelial height		10		200				700
Ovaries – atrophic changes in follicles		200		200				140
Vagina		10		200				

[a] In the 407 studies these were individual tissues (S - seminal vesicles, C - coagulating glands, ventral (V) and dorsolateral prostate (D)). In the 90-day and reproductive studies, these male tissues were dissected and weighed as a single unit. Eq – equivocal, low rate of individuals affected.

Table 19 continued. Comparison of updated TG 407 EE results with other estrogen studies.

Parameter	EE – updated TG 407 (μg/kg/d)				E2 – 90-day and reproduction studies (μg/kg/d)			
	Laboratory 2		Laboratory 5		90-Day Study (26)		Reproduction Study (27)(28)	
	Male	Female	Male	Female	Male	Female	Male	Female
Sperm parameters					500			
Estrous cycle		10[a]		200		140		
Reproductive and developmental								
No litter produced by pair (reproductive failure)	Not applicable						700	
Implantation count/efficiency	Not applicable						↓ 170	
Pups per litter	Not applicable						↓ 170	
Vaginal opening (F1 female offspring)	Not applicable						Controls 31.1 d; 3 μg/kg/d ↓ 29.5 d; 170 μg/kg/d ↓ 22.3 d	

[a] 0/10 - 3/10 - 5/10 - 8/10 - cyclicity, number of non-diestric according to histopath for controls, L, I, and H doses, respectively.

Conclusions for the updated TG 407's performance with EE

53. The following conclusions are drawn from the updated TG 407 studies with EE and from the comparison of these results with 17β-estradiol studies:

- The updated TG 407 results indicated that the pattern of effects observed in the two EE studies were concordant and consistent across body weights, haematological and clinical chemistry, organ and tissue weights, histopathology and other parameters.
- The updated TG 407 detected clear effects in both the female and the male reproductive tracts.
- The updated TG 407 results indicated that the dose responses in the two EE studies were similar.
- The pattern of effects produced in the TG 407 studies with EE were also concordant with studies of the endogenous estrogen, 17β-estradiol, with animals of a similar age in 90-day studies.
- This pattern of effects is also consistent the reproductive and developmental effects of the endogenous estrogen, 17β-estradiol, observed in a 1-generation reproductive and development studies.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of EE consistent with the endocrine mechanism of action of an estrogen.

Genistein

Introduction

54. This section summarizes the results of the updated TG 407 studies with genistein (GN) and compares these results with a 1-generation reproductive range finding study, an unpublished 5-generation reproductive study, and *in utero* and developmental studies conducted with GN (31)(32)(33)(34).

Background on genistein

55. GN is a natural phytoestrogen contained in soy products. It has a lower binding affinity for the estrogen receptor alpha than the endogenous 17β-estradiol with a relative binding affinity of log -0.35 (35). In the uterotrophic bioassay, GN achieved a statistically significant increase in uterine weight at a dose of 20-60 mg/kg body weight/day (36). These data suggest that GN's potency that is about 20,000-fold less than EE and 17β-estradiol.

Description of genistein experiments

56. GN was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in both studies. The GN was administered using a corn oil vehicle in three doses to both sexes in each study: a low (L) dose of 120 mg/kg body weight/day, an intermediate (I) or mid-dose of 400 mg/kg body weight/day, and a high (H) dose of 1000 mg/kg body weight/day. Both studies incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. These individual data from these Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose at the end of both studies to assess the impact of increased group size on the statistical power.

Summary of the updated TG 407 results with genistein

57. It is tentatively concluded that the updated TG 407 successfully detected effects of GN consistent with the endocrine mechanism of action of an estrogen. One GN study observed an increase in the uterine weight of the high dose females, and both GN studies observed apparent changes in the synchronization of female reproductive tract tissues during the oestrous cycle. While 'normal' histology was observed for individual parameters, the overall synchronisation or orchestration of the oestrous cycle was altered. The subtle and atypical nature of the observations (interpretation of the synchrony of several tissues and cell types in the female reproductive tract to the oestrous cycle) necessary to arrive at this interpretation must be noted. The detection of weak estrogens is then likely to require increased vigilance and changes in the current practice of histopathological examinations of the female reproductive tract.

Mortality and body weights in genistein studies

58. No mortalities were observed in the animals at any GN dose level, and there were no treatment-related clinical signs in either study. Food consumption was not effected in either sex in the two studies. There were no statistically significant differences were recorded between the control and any dose group in the functional observational battery and the motor activity assessment.

59. Body weights were not affected in either sex in two studies at any GN dose including the combined Subgroups and all the individual Subgroups. All absolute changes were modest, random in direction, and within group-to-group variability (Table 20). Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Table 20. Changes in body weights during 407 studies with genistein.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
4	M	- - -0.2%	- - +1.5%	- - +0.2%	- - -1.7%	- - +1.2%	- - -2.9%	- - +1.2%	- - +1.7%	- - +3.7%
	F	- - -2.4%	- - -2.8%	- - 0.0%	- - +2.5%	- - -1.6%	- - -1.2%	- - -7.5%	- - -3.5%	- - +1.2%
12	M	- - +2.6%	- - -3.2%	- - -2.3%	- - +0.1%	- - -3.4%	- - -5.6%	- - +5.1%	- - -3.1%	- - +0.8%
	F	- - +0.6%	- - -2.1%	- - -1.7%	- - +0.3%	- - -2.8%	- - -2.0%	- - +0.9%	- - -1.4%	- - -1.4%

M: male; F: female; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

Haematology and clinical chemistry results with genistein

60. Two clinical chemistry parameters, female total protein concentration and total triglycerides, were significantly changed in the combined Subgroups in both studies (studies in agreement). There were no parameters where the statistically significant changes in one study were directionally in conflict with the other study. For other five parameters, statistical changes were observed in the second study with similar changes in the direction of the absolute values occurring in the first study. These parameters were male and female total cholesterol and female albumin, total bilirubin, and chloride (Table 21). There were three parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (Table 21). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Table 21. Haematology and clinical chemistry results from 407 studies with genistein.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	**Statistically Significant in Laboratory 4**	**Statistically Significant in Laboratory 12**
Significant Parameters in Agreement in Both Studies: ↑ F Total protein ↑ F Triglycerides	***Common measures, direction of percentage change similar, significant in first study only:***	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ M Total cholesterol -13.2 **-17.9**[a] ↓ F Total cholesterol -14.7 **-30.6** ↑ F Albumin 7.5 **6.1** ↑ F Total bilirubin 45.5 **49.0** ↓ F Chloride -0.3 **-2.1**
	Common measures, direction of percentage change differ, significant in first study only: ↑ M Total protein **3.3** -3.8 ↓ F Albumin/globulin ratio **-9.0** -0.8	***Common measures, direction of percentage change differ, significant in second study only:*** ↓ F Creatinine 3.3 **-15.1**
	Measurement performed in this lab only:	***Measurement performed in this lab only:*** ↑ M Prothrombin times

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. Immediately below the specific parameter, the first study percentage change in the absolute value is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Organ and tissue weights results with genistein

61. There were no significant changes in the absolute weights of any tissue in the first GN study; therefore, only the results of the second study are shown in Table 22A. The significant changes in relative weights in the two studies are shown in Table 22B. For the major organs, statistical changes in the liver, adrenal, kidney, and thymus weights were observed in one study, but not the other.

62. For the male reproductive tract, a significant decrease in the epididymis relative weight occurred in one study in one Subgroup, but those male reproductive tract tissues that were dramatically changed by EE in other 407 studies (e.g., seminal vesicles, dorsolateral prostate, and ventral prostate) were not affected by GN administration. For the female reproductive tract, a significant increase in the uterine weights was observed at the high dose in one study, when a clear outlier was removed from the mid-dose data. This is judged to be suggestive evidence of an estrogen in this study. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Histopathology findings with genistein treatment

63. The only histopathological observation from GN administration in major organs was a modest change in hepatic glycogen in the first study (Table 23). There were no histopathological observations in the male reproductive tract of either study that were judged to be treatment-related.

64. In females, the observed changes in the uterus and vagina in response to GN administration were characteristic of animals in estrous or metestrous in the second study (Table 23). These findings contrast with the vaginal smears which indicated diestrous. As with similar patterns in the EE studies, these findings are consistent with a mild exposure to estrogen. Again, the findings are individually part of natural pattern and not manifestly pathological; it is the difference in timing and synchronization that is the essential observation. In the first study, the pattern of vaginal vacuolization is suggestive of estrogens, but none of the uterine pattern of estrous or metestrous were noted in this study (Table 23).

65. In the GN studies cases, the frequency of findings was similar in the two individual Subgroups. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Thyroid hormone results with genistein treatment

66. At one or more times, two of the thyroid hormone values (T_3 and TSH) achieved a statistically significant difference in the two GN studies (Table 24). The TSH decrease in Subgroup A males the first study does not appear to be dose-related. The TSH increase in the females in the second study was not corroborated by a fall in T_4 levels; rather, absolute T_3 levels were unchanged and absolute T_4 levels were approximately 25% higher. There were no similar trends in the first study. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with genistein treatment

67. No significant changes were observed in either sperm numbers or morphology in either GN study in the combined Subgroups (Table 25). In the first study, there was statistically significant increase in sperm counts at the mid-dose of Subgroup A. This lacked corroboration and evidence for a dose relationship, and the finding was judged to be spurious. Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies are in Annex 8.

Table 22A. Genistein-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
12	M	Kidney									
		Adrenal			↑↑			▲			↑↑
		Thymus		↓	↓		▽	↓		↓	▽
		Epididymes									
		Seminal Vesicles	▲	▲		▲				▲	
		Ventral Prostate			▽			▽	▽	▽	▽
		Dorsolateral Prostate	▲	↑			▲		▲	↑	▲
	F	Liver			↑			↑			▲
		Adrenal	▲	▲	↑	▲	▲	↑	▲	▲	▲
		Thymus	▽	↓	▽	▽	▽				▽
		Thyroid	↑	↑↑	▲	▲	↑	▲	▲	▲	▲
		Ovaries				▲	▲				
		Uterus + cervix		▲	↑		↑	▲			▲

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 22B. Genistein-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
4	M	Liver		↑	↑						
		Epididymes									
		Ventral Prostate [a]			▽	▲			▽	▽	▽
		Dorsolateral Prostate	▲	▲	▲	▲	▲	▲		▲	
		Seminal Vesicles									
	F	Liver				↑	↑	↑			
		Ovaries									
		Uterus + cervix									
12[a]	M	Kidney	↑	↑↑	↑	↑	↑↑			↑	↑
		Adrenal			↑↑			↑			↑↑
		Thymus		↓	↓			▽		↓	▽
		Epididymes									
		Ventral prostate			▽			▽	▽		▽
		Dorsolateral Prostate		▲	▲		▲			▲	▲
		Seminal Vesicles		▲		▲				▲	
	F	Liver		↑↑	↑↑	↑↑	↑↑	↑↑			↑↑
		Adrenal	▲	↑	↑	▲	▲	↑	▲	▲	▲
		Thymus	↓	▽	▽	▽	▽				↓
		Thyroid	↑	↑↑	▲	▲	▲	▲	▲	▲	▲
		Ovaries				▲	▲				
		Uterus + cervix		▲	↑a		▲	▲		▲	▲

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

a The Syngenta CTL laboratory uses an ANOVA procedure adjusting organ weights with body weight as a covariate, a slightly different procedure than in most labs. Thus, the extraction below is not strictly comparable.

Table 23. Significant histopathological findings after genistein treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
4	F	**Liver**	Increased glycogen	3/10	5/10	8/10	2/5	2/5	4/5	1/5	3/5	4/5
		Uterus	Aplasia, unilateral		1/10			1/5				
		Vagina	Vacuolisation		2/10	2/10		1/5	1/5		1/5	1/5
12	M	**Kidney**	Intratubular microlithiasis		3/10	7/10		2/5	4/5		1/5	3/5
	F	**Uterus**	Dilatation		1/10	2/10		1/5				2/5
			Squamous epithelial metaplasia	3/10	3/10	6/10	3/5	1/5	4/5		2/5	2/5
			Endometrial epithelial hypertrophy	1/10	5/10	6/10	1/5	3/5	4/5		2/5	2/5
			Endometrial epithelial apoptosis/ vacuole.	2/10	7/10	7/10	2/5	4/5	5/5		3/5	2/5
			Increased stromal neutrophils	3/10	8/10	10/10	2/5	4/5	5/5	1/5	4/5	5/5
		Vagina	Keratinisation	1/10	4/10	3/10	1/5	3/5	3/5		1/5	

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Table 24. Thyroid hormone results after genistein treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
4	M	T_3									
		T_4									
		TSH								↑	
	F	T_3									
		T_4									
		TSH									
12	M	T_3									
		T_4									
		TSH									
	F	T_3									↓
		T_4									
		TSH			↑↑		↑	↑			a

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).

Shading indicates that no statistical significance was observed in the combined Subgroups or individual Subgroups of a given sex.

a Only 3 values were available at the high dose (versus expected n=5), and, although absolute values were double the control, they did not achieve significance.

Table 25. Statistically significant changes in sperm parameters after genistein treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
4	Sperm Count					↑				
	Abnormalities									
12	Sperm Count									
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group.

Estrous cyclicity

68. None of the females were judged to have exhibited an abnormal oestrous cycle in either study (i.e., absent or prolonged). However, in the second study, it was again noted, as with one of the EE study labs, that when compared to the uterine and vaginal histopathology results, most animals were judged by these results to have been in oestrus or proestrus, not dioestrus as judged by the vaginal smears. This lack of apparent synchronization among various tissues in the female reproductive tract has been further investigated in a later chapter. The vaginal cytology results from both studies are in Annex 9.

Comparison of the updated TG 407 results from genistein treatment with data from chronic, reproductive, and developmental studies

69. A 1-generation range finding reproductive study (31) and an unpublished 5-generation reproductive study (32) have been conducted with GN. Both studies included a full complement of standard and estrogen sensitive endpoints; the 1-gen used 10 dams per sex per dose and the 5-gen used 25 dams per sex per dose. Developmental studies based upon *in utero* and postnatal exposure have also been conducted (33)(34). Administered doses are noted in Table 26.

Table 26. Description of genistein reproductive and developmental studies.

Study	Route and timing	GN doses (mg/kg/d)
(31)	Dietary; F1 terminated pnd 50	Pregnant dams 0.3, 1.5, 6, 15, 33, and 80 mg/kg/d; females pups prior to puberty slightly more than 2X of the dams.
(32)	Dietary; offspring terminated pnd 140	0.4, 8, and 40 mg/kg/d
(33)	Dietary; F1 terminated females pnd 34 and males pnd 56	Approx 15 and 70 mg/kg/d
(34)	Dietary; sc and po during lactation and weaning; terminated ~ pnd 90	4 and 40 mg/kg/d

70. The selected findings in the two updated TG 407 studies with GN and the selected reproductive and developmental studies are compared in Table 27. The comparison should take into account the following important factors regarding dose and sensitivity:

- the limited GN doses in the reproductive and developmental studies, particularly the limited doses in the 5-generation study, where the maximum dose was approximately 40 mg/kg body weight/day in pregnant dams and, prior to puberty, would have reached approximately 80 mg/kg body weight/day in both sexes.
- the detection of an increase in uterine weights and other changes in the developmental studies occurred when exposure occurred during and prior to puberty, not from earlier *in utero* and lactational exposures at the GN doses administered.

Conclusions for updated TG 407 performance with genistein

71. The following conclusions are drawn from the updated TG 407 studies on GN and from the comparison of these results with reproductive and development studies on GN:

- The updated TG 407 results provided equivocal evidence of GN endocrine activity at 400 mg/kg body weight/day in the two studies. Changes in relative uterine weights were recorded in one study, and the synchronization of the female reproductive tract in the oestrous cycle appeared to be altered.

- The updated TG 407 detected no evidence in either study for effects in the male reproductive tract based on the results from organ and tissue weights, histopathology, and sperm parameters.
- The updated TG 407 did detect a statistically significant change in uterine weight in one study at the high dose and there was supporting evidence from the histopathology of changes in at least one tissue of the female reproductive tracts in both studies. As this evidence was modest, the physiological changes induced by GN administration may require a disciplined analysis of the data. The central problem is that the pattern of effects is not one of frank pathological changes and lesions in any tissue. Instead, the overall synchronization of a temporal sequence in several tissues during the oestrous cycle is disrupted so that 'normal' histology is observed for individual parameters, but the overall coordination of the oestrous cycle is impaired. Thus, one must carefully correlate the pattern in several cell types of the ovary, uterus, and vagina with that of the oestrous cycle and evaluate whether there are dissimilarities with the normal, expected sequence and its orchestrated pattern. This is not the current practice in most TG 407 studies.
- The available reproductive and development studies for GN also indicate little or no evidence for systemic toxicity at doses up to about 90 mg/kg body weight/day in adults, pups, or immature animals. These studies did observe kidney pathology, but previous studies indicated that the special phytoestrogen free diet was a partial contributor to these lesions.
- The described pattern of histopathological findings in F1 female reproductive tract at the higher dose GN reproductive study studies (~170 mg/kg body weight/day) appears similar to those described in TG 407 studies:

  "Both uterine and vaginal structures showed an inappropriate combination of changes that reflected influences of estrus, metestrus, and diestrus. Abnormal cellular maturation in the vaginas of 9 of 15 animals was observed and labeled as dysynchronous, a term that denotes marked departure from normal cyclic morphologic changes in the epithelium. This included hypertrophy of the mucinous layer that reflects the progesteronal influence during proestrus. The apparent increased progesteronal drive in the 1250 ppm dose group is consistent with the observation of persistent corpora lutea. In addition, 4 of 15 animals in the 625 ppm group were diagnosed with moderate vaginal dysynchrony in the absence of detectable ovarian degeneration." (31).

- Modest effects (acceleration of vaginal opening) were seen in the reproductive tracts of immature females in development studies when exposure to GN occurred during puberty at approximately 40 mg/kg body weight/day. However, the TG 407 females are mature and have established oestrous cycles, and these effects would not be measurable in the TG 407 animals.
- No male reproductive tract effects were seen in any of the reproductive or development studies in response to GN administration, concordant with the TG 407 findings.
- It is concluded that for GN, the updated TG 407 gave equivocal evidence for endocrine mechanism of action of an estrogen. However, the subtle and atypical nature of the observations (synchrony of several tissues and cell types in the female reproductive tract to the oestrous cycle) necessary to arrive at this conclusion must be noted. The detection of very weak estrogens such as GN is then likely to require increased vigilance and changes in the current practice of histopathological examinations of the female reproductive tract.

Table 27. Comparison of updated TG 407 GN results with other genistein studies.

Parameter	GN updated TG 407 studies (LOEL mg/kg/d)				GN 1- and 5-gen reproduction studies (LOEL mg/kg/d)		GN *in utero* exposure and developmental studies (LOEL mg/kg/d)	
	Laboratory 4		Laboratory 12		(31)	(32)	(33)	(34)
	Male	Female	Male	Female				
Haematology and clinical chemistry								
Total protein		↑ 1000		↑ 1000			Not done	Not done
Triglycerides		↑ 400		↑ 400				
Organ and Tissue Weights								
Liver					↑ M 80			
Kidney								
Adrenal								
Thymus								
Ventral Prostate					↓ 80			
Uterus				↑ 1000			↑ 70 at pnd 35	↑ 40 pnd 22
Histopathological Findings								
Mammary gland - hyperplasia					M and F 66	M 40-80	Not done	Not done
Dorsolateral prostate – hyperplasia					170			
Ventral prostate– decreased secretory activity					170			
Uterus – hyperplasia/ increased epithelial height				400	Dysnychronous pattern - 170			
Ovaries – atrophic changes in antral follicles					170			
Vagina - dysnychronous		400		400	66			
Sperm parameters					Not done		Not done	Not done
Estrous cycle							Not done	Not done
Reproductive and developmental								
Vaginal opening	Not applicable				80 mg/kg/d mean 3 d < control, but not stat. significant	↓ 40	↓ 70	↓ 40

[a] In the 407 studies these were individual tissues (seminal vesicles, coagulating glands, ventral and dorsolateral prostate. In the 90-day and reproductive studies, these male tissues were weighed as a single unit.

Nonylphenol

Introduction

72. This section summarizes the results of the updated TG 407 studies with nonylphenol (NP) and compares these results with a 90-day study (37) and 3-generation reproductive study conducted with NP (38)(39).

Background on nonylphenol

73. NP is a synthetic chemical intermediate used largely to produce surfactants. It has a lower binding affinity for the estrogen receptor alpha than the endogenous 17β-estradiol with a relative binding affinity of log -1.43 (29). In the uterotrophic bioassay, NP achieved a statistically significant increase in uterine weight at 75 mg/kg body weight/day (36), suggesting a potency of about 70,000-fold less than 17β-estradiol and EE.

Description of nonylphenol experiments

74. NP was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in both updated 407 studies. The NP was administered using a corn oil vehicle in three doses to both sexes in each study. Due to animal mortalities and clinical signs, the initial high dose was lowered in one study from 300 mg/kg/d to 250 mg/kg body weight/day. With further clinical signs and body weight loses in the first study; this led the second lab to again lower their initial high dose to 200 mg/kg body weight/day. Mortalities and clinical signs still occurred in the second study, and the high dose was lowered still further to 150 mg/kg body weight/day. Therefore, these studies are not fully comparable at the high dose, and more severe effects may plausibly be expected in the first study. The other doses were the same and were not modified: a low (L) dose of 20 mg/kg body weight/day and an intermediate (I) or mid-dose of 80 mg/kg body weight/day. Both studies incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The individual data from these Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power. In one study, the histopathology data were analyzed only for the combined Subgroups.

Summary of the updated TG 407 results with nonylphenol

75. The updated TG 407 did not detect estrogenic responses with NP at doses below a maximum tolerated dose. Equivocal evidence in uterine histopathology and decreases in the absolute weights of male accessory reproductive tissues were observed at doses which caused mortalities and other clinical signs in the animals. In other respects, the updated TG 407 was consistent with other NP studies for effects on major organs. The reproductive studies were able to indicate estrogenic effects only in immature females and changes in the time of vaginal opening and the length of the oestrous cycle in *in utero* exposed offspring; parameters in the P generation adults were not affected in the reproductive studies as with the updated TG 407 studies.

Mortality and body weights in nonylphenol studies

76. There was a single female mortality in both studies at the respective high NP dose level in each study, and this was assessed as treatment-related in both studies. Clinical signs were observed at the high dose in both studies. Food consumption was not statistically different in either sex in either of the two studies. In the study with the 200/150 mg/kg body weight/day high dose, female grip strength was significantly decreased, but the relationship to treatment was judged to be very unlikely in the absence of

other effects. Supporting this conclusion, there was no statistical change in this parameter in the other study where the NP dose was 300/250 mg/kg body weight/day. There were no other statistically significant differences recorded for any other functional observational battery and motor activity assessment parameters in either study.

77. Male body weights were significantly decreased at the high NP dose in the combined Subgroups and in one of the individual Subgroups in the study with the 300/250 mg/kg body weight/day dose (Table 28). Male body weights decreased more than 10% in the remaining individual Subgroup, but did not achieve statistical significance. At 200/150 mg/kg/day in the other study, male body weights were largely unchanged.

78. Female body weights did not change significantly or in absolute terms in this study or for either sex in the other study (Table 28). Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with nonylphenol

79. Five haematological and clinical chemistry parameters, female haemoglobin, haematocrit, erythrocyte counts, and globulin levels and male triglycerides, were significantly changed in the combined Subgroups in both studies (studies in agreement) (Table 29). There were no parameters where the statistically significant changes in one study were directionally in conflict with the other study. For five other parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values occurring in the other study. These parameters were female total cholesterol, triglycerides, albumin, and blood urea nitrogen and male sodium levels (Table 29). There were seven parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 29). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Organ and tissue weights results with nonylphenol

80. There were concordant statistically significant changes in major organs in the combined Subgroups of both NP studies. In both sexes, the absolute and relative kidney weights were increased at the high NP doses (Tables 30A and B). In both studies, the relative weights of the liver were increased in both sexes at the high NP doses. Due to the higher NP dose in the first study, the increase in the relative weights of the adrenals and decrease in the thymus of both sexes should be noted although these were not observed in the second study with the lower dose (Tables 30A and B).

81. In males, there were statistically significant absolute decreases in the seminal vesicles and the whole and dorsolateral prostate weights in the study with the high 300/250 mg/kg body weight/day NP dose. Although decreased, the relative weights of these tissues did not achieve statistical significance (Tables 30A and B). No significant changes were observed in these tissues at the 200/150 mg/kg body weight/day NP dose, although the absolute weights of the seminal vesicles decreased by over 10% (Table 30A).

82. There were no significant changes in the weights of the female reproductive tract tissues in either study. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Table 28. Changes in body weights during 407 studies with nonylphenol.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
1[a]	M	- - -1.4%	- - -2.5%	- - -1.9%	- - -4.6%	- - -4.2%	- - +0.4%	- - +2.0%	- - -0.8%	- - -4.3%
	F	- - +5.2%	- - -4.1%	- - -1.5%	- - +1.9%	- - -5.1%	- - -4.7%	- - +8.4%	- - -3.3%	- - +0.9%
6	M	- - -6.4%	- - -7.6%	↓↓ **-16.1%**	- - -9.5%	- - -4.5%	↓↓ **-17.9%**	- - +1.1%	- - -6.6%	- - -14.2%
	F	- - +2.9%	- - -0.5%	- - +1.3%	- - +3.1%	- - +1.3%	- - +1.1%	- - +1.7%	- - -4.1%	- - +1.0%

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

[a] The high dose in the study of laboratory 1 began at 200 and decreased to 150 mg/kg/d while the high dose in the study of laboratory 6 began at 300 and decreased to 250 mg/kg/d. Therefore, these are not fully comparable at the high dose.

Table 29. Haematology and clinical chemistry results from 407 studies with nonylphenol.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 1	Statistically Significant in Laboratory 6
Significant Parameters in Agreement in Both Studies: ↓ F Haemoglobin concentration ↓ F Haematocrit ↓ F Erythrocyte counts ↓ M Triglycerides ↑ F Globulin	***Common measures, direction of percentage change similar, significant in first study only:*** ↑ F Cholesterol **17.5** 5.1[a]	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ F Triglycerides -21.7 **-57.4** ↑ F Blood urea nitrogen 19.7 **46.2** ↓ F Albumin -5.3 **-11.9** ↑ M Sodium 0.1 **2.3**
	Common measures, direction of percentage change differ, significant in first study only: ↓ M Prothrombin quick times **-10.0** 4.3 ↑ F White blood cell counts **48.1** -27.0 ↑ M Alanine aminotransferase **20.3** -1.4	***Common measures, direction of percentage change differ, significant in second study only:*** ↑ M Blood urea nitrogen 7.4 **131.3** ↑ M Creatinine -2.6 **200.0** ↑ F Chloride -1.6 **2.8** ↓ M Alkaline phosphatase 8.7 **-24.2** ↓ F Sodium -1.2 **-2.1**
	Measurement performed in this lab only:	***Measurement performed in this lab only:*** ↓ F Albumin/globulin ratio

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. Immediately below the specific parameter, the first study percentage change in the absolute value is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Histopathology findings with nonylphenol treatment

83. Taking into account different descriptions of the histopathological lesions in the two studies, the findings in major organs were concordant. Corroborating the changes in organ weights, both studies observed histopathological lesions in the liver and the kidneys of both sexes at the high, and, in some instances, the mid- NP dose (Table 31).

84. In the male reproductive tract, the only findings were a single case of testicular atrophy and of seminal vesicle atrophy at the high 300/250 mg/kg body weight/day dose. In the female reproductive tract, one instance of uterine dilation was observed at the 200/150 mg/kg body weight/day NP dose and 3 cases of uterine luminal distention in the other study at the 300/250 mg/kg body weight/day NP dose. While suggestive, these findings were not judged to be conclusive for detecting a weak estrogen. More detailed

histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Thyroid hormone results with nonylphenol treatment

85. The T_3 and T_4 thyroid hormone values were significantly increased in the male combined Subgroups from the 200/150 mg/kg body weight/day NP dose, but no changes or absolute differences were observed these values in the 300/250 mg/kg body weight/day NP male dose group in the other study or in either of the female high dose groups (Table 32). Further, there was no corresponding increase in the TSH values. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with nonylphenol treatment

86. In the 200/150 mg NP/kg/d males of the first study, there were no statistical or absolute changes in either sperm numbers or morphology (Table 33). In the 300/250 mg NP/kg/d study, the sperm numbers from the mid- and high dose samples were significantly decreased, but not in a dose responsive manner and with group sperm number CVs in this study ranging from 50 to 120 (Table 33). The sperm morphology numbers in the second study showed a very high background of abnormalities, approaching 45%. However, there were no changes in sperm morphology observed in the second study. Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Estrous cyclicity

87. The vaginal smears from the 200/150 mg/kg body weight/day females when correlated with histopathological data were judged to reveal some lack of synchrony (see Table 95). In the 300/250 mg/kg study, however, there was no apparent lack of synchronisation. The vaginal cytology results are in Annex 9.

Table 30A. Nonylphenol-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
1	M	Liver			↑			▲			▲
		Kidney			↑↑			▲			↑
		Thymus			▽	▽	▽	▽			▽
		Epididymes									↓
		Seminal Vesicles			▽			▽			▽
		Whole Prostate	▲		▽				▲		▽
	F	Liver							↑		↑
		Kidney			↑↑			▲			↑↑
		Adrenal			↓						
		Heart			↓						
		Thymus	▲			▲			▲		▽
		Ovaries									
		Uterus and cervix		▽	▽	▲	▽			▽	▽
6	M	Liver									
		Kidney			↑↑			▲	↑		↑↑
		Heart			↓↓			↓↓			↓
		Adrenal			↑↑			↑↑			↑
		Spleen		▽	↓↓	▽	▽	↓↓			▽
		Thymus	↓	↓↓	↓↓	↓↓	▽	↓↓	▽	▽	↓
		Epididymes									
		Seminal Vesicles			↓↓		▽	↓↓		▲	▽
		Prostate			↓			▽			↓
		Ventral Prostate			▽		▲	▽		▽	▽
		Dorsolateral Prostate			↓			▽		▲	▽
	F	Liver			↑↑			↑↑			↑
		Kidney			↑↑			↑	▲		↑↑
		Adrenal			↑		▲	▲			↑
		Thymus		▽	↓			▽	▲	▽	▽
		Ovaries						▽		▽	
		Uterus		▽		▽	▽			▽	

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 30B. Nonylphenol-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
1	M	Liver		↑	↑↑		↑	↑↑			↑
		Kidney		↑↑	↑↑			↑		↑	↑↑
		Epididymes				▲	▲				
		Seminal Vesicles			▽			▽			▽
		Prostate	▲			▲	▲		▲		
	F	Liver		↑	↑↑					↑	↑↑
		Kidney		↑	↑			▲		↑	↑↑
		Adrenal									
		Heart									
		Brain		↑	↑↑			▲			
		Spleen	▽			↓		▽			
		Thyroid		▲	↑		▲	▲		▲	▲
		Thymus	▲							▲	
		Ovaries		▲						▲	
		Uterus and cervix		▽	▽	▲				▽	▽
6	M	Liver			↑			▲			
		Kidney			↑↑			↑			↑↑
		Heart									
		Adrenal	▲		↑↑	▲	▲	↑↑			↑
		Spleen			▽			▽			
		Thymus	▽	↓	↓↓	↓	▽	↓↓	▽	▽	↓
		Epididymes				▲		↑	▽		
		Seminal Vesicles			▽	▲	▽	▽		▲	▽
		Prostate			▽						▽
		Ventral Prostate		▲	▽		▲			▽	▽
		Dorsolateral Prostate			▽			▽		▲	▽
	F	Liver		↑	↑↑		↑	↑↑			↑↑
		Kidney			↑↑			↑↑			↑↑
		Adrenal			↑↑		▲	↑			↑
		Thymus		▽	↓		▽	↓	▲		▽
		Ovaries						▽			
		Uterus				▽	▽				

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 31. Significant histopathological findings after nonylphenol treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
1	M	Liver	Peliosis	2/10	4/10	2/10	2/5	2/5			2/5	2/5
			Vacuolisation, periportal		2/10	3/10		1/5	2/5		1/5	1/5
		Kidney	Tubular dilatation, multifocal			1/10						1/5
			Dilatation, pelvic		1/10	1/10		1/5	1/5			
	F	Liver	Peliosis	3/10	6/10	5/10	2/5	4/5	2/5	1/5	2/5	3/5
			Vacuolisation, periportal	1/10	2/10	2/10		1/5		1/5	1/5	2/5
		Kidney	Calcification	4/10		6/10	1/5		3/5	3/5		3/5
			Dilatation, pelvic			1/10			1/5			
			Chronic nephropathy			3/10			1/5			2/5
			Tubular dilatation, multifocal			6/10			2/5			4/5
		Uterus	Dilatation			1/10			1/5			
6	M	Liver	Hypertrophy, centrilobular			5/10	Not applicable			Not applicable		
			Inflammation	1/10	1/10	3/10						
		Kidney	Hyaline droplets		2/10	1/10						
			Dilatation, pelvic		1/10	1/10						
			Tubular nephrosis			9/10						
		Testes	Atrophy			1/10						
		Seminal vesicles	Atrophy			1/10						
	F	Liver	Inflammation	1/10		2/10						
		Kidney	Mineralization	1/10	1/10	1/10						
			Dilatation, pelvic		1/10	1/10						
			Tubular nephrosis	1/10		9/10						
		Uterus	Luminal distention			3/10						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.
Not applicable – Laboratory 6 did not report the histopathological findings by individual Subgroup.

Table 32. Thyroid hormone results after nonylphenol treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
1	M	T_3			↑↑			↑			
		T_4		↑↑	↑↑						↑↑
		TSH									
	F	T_3									
		T_4									
		TSH									
6	M	T_3									
		T_4									
		TSH									
	F	T_3									
		T_4						↑↑			
		TSH									

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).

Shading indicates that no statistical significance was observed in the combined Subgroups or individual Subgroups of a given sex.

Table 33. Statistically significant changes in sperm parameters after nonylphenol treatment.

Lab		L	I	H	L	I	H	L	I	H
1	**Sperm Count**	ND	ND		ND	ND		ND	ND	
	Abnormalities	ND	ND		ND	ND		ND	ND	
6	**Sperm Count**		↓↓	↓		↓	↓			

L – low dose group; I – intermediate or mid-dose group; H – high dose group; ↓ - statistically significant decrease sperm numbers ($p<0.05$); ↓↓ - statistically significant decrease sperm numbers ($p<0.01$); ND – No analysis performed on these dose groups with finding of no difference between controls and the high dose group animals.

Comparison of the updated TG 407 results from nonylphenol treatment with data from chronic, reproductive, and developmental studies

88. A 90-day NP study with of 10 animals per sex has also been published, using doses of 14 and 16, 45 and 50, and 130 and 150 mg NP/kg body weight/day, respectively, for males and females (37). A 3-generation study has been conducted with dietary administration of NP with a full complement of standard and estrogen sensitive endpoints using 30 animals per sex per dose, using doses of 8-19, 28-63, and 88-185 mg NP/kg body weight/day (38). Another 2-generation reproduction study with gavage administration of NP has also been published, using doses of 2, 10, and 50 mg NP/kg body weight/day (39).

89. The selected findings in the two updated TG 407 studies with NP reproductive and 90-day studies are compared in Table 34. The comparison should take into account the extended times of the 90-day and reproductive studies relative to the 407 and the dietary versus gavage administration differences.

Conclusions for 407 performance with nonylphenol

90. The following conclusions are drawn from the updated TG 407 studies with NP and from the comparison of these results with other NP studies:

- The updated TG 407 results indicated that maximum tolerated dose was exceeded at the top NP dose with several animal mortalities recorded.
- The updated TG 407 results were consistent with other studies in regards to NP effects on major organs, particularly, the kidneys.
- In the updated TG 407, increases in the absolute weights, but not relative weights, of several male accessory reproductive tissues were observed at a does judged to be above the maximum tolerated dose due to mortalities, clinical signs, and body weight differences. The reproductive studies with NP observed no changes in the tissue weights or the histopathology for any tissue in the male reproductive tract.
- The reproductive studies with NP observed an increase in uterine weights at pnd 21, but not in adult animals. This would be consistent with uterotrophic responses in immature animals at approximately the same dose. However, the TG 407 animals are sexually mature and are potentially not as sensitive, which is consistent with the absolute, but not statistically significant, decreases in uterine weight observed in these studies.
- Both reproductive studies with NP observed acceleration in the time of vaginal opening in the F1 and F2 generations. The young adult TG 407 animals are already sexually mature, and this measurement cannot be made with these animals.
- The effects on the synchronisation of the female reproductive tract tissues seen with EE and GN may have been seen in one NP study at a low frequency, but confirmation was absent at the higher NP dose in the second study. Therefore, this evidence is judged to be equivocal. This leads to the overall conclusion that the updated TG 407 did not observe effects consistent with estrogen at doses below the maximum tolerated dose.

Table 34. Comparison of updated TG 407 NP results with reproductive and developmental studies.

Parameter	NP – updated TG 407 (LOEL mg/kg/d)				NP – 90-day and reproduction studies (LOEL mg/kg/d)					
	Laboratory 1		Laboratory 6		90-day (37)		Reproduction (38)		Reproduction (39)	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Body weight			↓ 300/250		↓ 130			↓ 88-185		
Haematology and clinical chemistry										
Haematocrit		↓ 200/150		↓ 80						
Haemogloblin concentration		↓ 200/150		↓ 80						
Erythrocyte counts		↓ 200/150		↓ 300/250						
Triglycerides	↓ 80		↓ 80							
Organ and Tissue Weights										
Liver	↑ 80	↑ 80	↑ 300/250	↑ 80					↑ 50	
Kidney	↑ 80	↑ 80	↑ 300/250	↑ 300/250	↑ 150				↑ 50	
Adrenal										
Thymus			↓ 80	↓ 300/250						↓50
Thyroid									↑ 50	
Pituitary									↑ 50	
Uterus								↑ pnd 21 88-185		
Ovaries						↓ 150				↓50
Histopathological Findings										
Liver	80	80	80	80					50	
Kidney		200/150	300/250	300/250	150		8-19		50	
Uterus				Eq 300/250						
Sperm Numbers/Morphology							F2 only ↓ 88-185			
Estrous Cycle										
Reproductive Parameters										
Live pups per litter	NA	NA	NA	NA	NA	NA		↓ 88-185		
Vaginal opening	NA	NA	NA	NA	NA	NA		↑ 28-63		↑ 50
Estrous cycle: F1 and F2	NA	NA	NA	NA	NA	NA		↑ 88-185		

NA – Not applicable

Tamoxifen

Introduction

91. This section summarizes the results of the updated TG 407 studies with tamoxifen (TAM) and compares these results with reports of three oral dosing studies exposing male and female rats to tamoxifen for 3 months (2 studies) and 6 months (40)(41) and with a chronic carcinogenicity study (42).

Background on tamoxifen

92. Tamoxifen is a pharmaceutical antiestrogen sometimes administered to breast cancer patients. Tamoxifen has a lower binding affinity for the estrogen receptor alpha than the endogenous 17β-estradiol with a relative binding affinity of log 0.21, but when metabolized to 4-hydroxytamoxifen the binding affinity surpasses 17β-estradiol with a relative binding affinity of log 2.24 (29). In the uterotrophic bioassay, tamoxifen is a partial agonist/antagonist meaning that when administered alone there is a modest stimulation of the uterus. When Tamoxifen is coadministered with 17β-estradiol, the action of the 17β-estradiol is antagonized.

Description of tamoxifen experiments

93. TAM was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in the first study and to a single group of six animals per sex per dose in the second study. The TAM was administered using an aqueous solution with 3%ethanol and 0.5% methylcellulose in the first study and a corn oil vehicle in the second study. There were three doses in each study: a low (L) dose of 5 μg/kg body weight/day, an intermediate (I) or mid-dose of 30 μg/kg body weight/day, and a high (H) dose of 200 μg/kg body weight/day. The first study incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The functional observational battery and the motor activity assessment were not conducted in the second study. The individual data from the first study's Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power; this could not be done for the second study.

Summary of the TG 407 results with tamoxifen

94. The two updated TG 407 produced consistent response profiles to TAM in the female reproductive tract, and the TAM effect levels observed were similar. The pattern of tissue weight changes and histopathological changes observed in the updated TG 407 study were similar to those seen in subchronic studies with TAM, taking into account the higher doses used in these studies. Therefore, it is concluded that the updated TG 407 successfully detected effects of TAM consistent with the endocrine mechanism of action of an antiestrogen. In regards to the male reproductive tract, the first study detected both changes in tissue weights and histopathology, but the histopathological observations were not duplicated in the second study.

Mortality and body weights in tamoxifen studies

95. No mortalities were observed in the animals at any TAM dose level, and the only treatment-related clinical signs in either study were in a single female at the high dose in the first study. Food consumption was significantly reduced in the combined Subgroups at the high dose in both sexes in both studies, and at the mid-dose in females in the first study. The functional observational battery and the motor activity assessment were performed in only the first study (and none of these observations were

made in the second study), and there were no statistically significant differences recorded between the control and any dose group.

96. Body weights were significantly decreased in both sexes in both studies at the high TAM dose. The decreases were similar in magnitude in both studies and were slightly higher on a percentage basis in males than females in both studies. Female body weights were also significantly decreased at the mid-dose of TAM in the first study (Table 35). Detailed body weight means and standard deviations for both studies, including the combined Subgroups and individual Subgroups for the first study, are in Annex 3.

Table 35. Changes in body weights during 407 studies with tamoxifen.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
3	M	- - +5.1%	- - -1.6%	↓↓ **-15.4%**	- - +2.5%	- - -3.5%	↓↓ **-16.3%**	- - +7.8%	- - +0.3%	↓↓ **-14.2%**
	F	- - -3.6%	↓ **-7.1%**	↓↓ **-12.9%**	- - -6.2%	- - -5.4%	↓↓ **-13.8%**	- - -0.9%	↓ **-8.8%**	↓↓ **-12.0%**
10 [b]	M	- - -7.9%	- - -6.6%	↓ **-19.4%**	Not applicable			Not applicable		
	F	- - +0.3%	- - -6.2%	↓ **-14.3%**						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight($p<0.01$); - - : no statistically significant change.

[b] Laboratory 10 also observed statistically significant decreases in male and female body weights of a similar magnitude (i.e., with tamoxifen 19.4% in males and 14.3% in females). However, this lab used a group size of 6 in both the tamoxifen and PTU studies, and this lab did not conduct the subgroup experiments. The significance achieved in both studies was at $p<0.05$.

Haematology and clinical chemistry results with tamoxifen

97. Only one clinical chemistry parameter, male cholesterol, was significantly changed in the combined Subgroups in both studies (studies in agreement) (Table 36). There were no parameters where the statistically significant changes in one study were directionally in conflict with the other study. For seven other parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study. These parameters were male white blood cell counts, prothrombin times, and alkaline phosphatase levels, and female white blood cell counts, mean cell haemoglobin, and chloride and calcium levels (Table 36). There were three parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 36). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Table 36. Haematology and clinical chemistry results from 407 studies with tamoxifen.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 3	Statistically Significant in Laboratory 10
Significant Parameters in Agreement in Both Studies: ↓ M Total cholesterol	***Common measures, direction of percentage change similar, significant in first study only:*** ↑ M Prothrombin times **11.0** 13.5[a] ↓ F Mean cell haemoglobin concentration **-2.1** -2.8 ↑ M Alkaline phosphatase **28.5** 6.8 ↑ F Chloride **1.8** 0.8 ↓ F Calcium **-3.9** -17.0	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ M White blood cell counts -18.5 **-38.8** ↓ F White blood cell counts -4.3 **-23.5**
	Common measures, direction of percentage change differ, significant in first study only: ↓ F Total cholesterol **-45.7** 2.1 ↑ F Sodium **2.1** -0.5	***Common measures, direction of percentage change differ, significant in second study only:*** ↓ M Potassium -1.4 **-15.2**
	Measurement performed in this lab only: ↑ M Triglycerides ↓ F Phosphorus	***Measurement performed in this lab only:***

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. The first study percentage is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Organ and tissue weights results with tamoxifen

98. In the major organs, the absolute weights of a number of major organs decreased in the first study at the high dose of TAM, but none in the second study (Table 37A). Given the significant decreases in the body weights of both sexes, the relative organ weights need to be taken into account. The relative weight of the male heart was increased in both studies. The relative weights of the male brain, adrenals, and kidney were significantly increased in the first study, having the larger number of animals in the combined Subgroups (Table 37B). In females, the only change in the relative weights of the major organs was an increase in the brain and thymus weights in the first study (Table 37B).

99. In male reproductive tract, there were significant increases in the testes weights in both studies and significant decreases in the relative weights of the ventral prostate and the seminal vesicles in the first study and the whole prostate in the second study. The relative weights of the epididymis decreased in the second study, but were unchanged in the first study. The relative weights of the seminal vesicles and coagulating glands decreased in the first study, but were not measured in the second study (Table 37B).

100. In the female reproductive tract, absolute and relative uterine weights were significantly decreased at the high TAM dose in both studies. In the first study, potentially consistent with the partial agonist/antagonist action of TAM, the absolute and relative uterine weights were significantly increased at the low TAM dose. Absolute ovarian weights were decreased at the high TAM dose in both studies. The relative ovarian weights were significantly decreased in the second study. Although the absolute ovarian weights decreased over 10% in the first study, the decrease did not achieve statistical significance. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Histopathology findings with tamoxifen treatment

101. There were no histopathological changes reported in the major organs in the first study. In the male reproductive tract, the first study reported several treatment-related changes in the mammary gland, the prostate, the seminal vesicles, and possibly the testes at the mid- and high doses of TAM (Table 38). In the female, the first study reported treatment-related changes in the ovaries, uterus, and vagina at the mid- and the high doses of TAM (Table 38). Detailed histopathological summaries for this study, including the combined Subgroups and individual Subgroups, are in Annex 6.

102. Although not in the original report, the results of the histopathological observations of the major organs and the reproductive tracts in the second study were made available at the Secretariat's request. The second study described severe endometrial squamous metaplasia and atrophy of the endometrial glands and severe follicular cysts in the ovaries at the high TAM dose. Again, the number of individuals affected and grades of the observed severity were made available at the request of the Secretariat for this report (Table 38).

Thyroid hormone results with tamoxifen treatment

103. The T_4 thyroid hormone values were significantly increased in the female high TAM dose groups in both studies (Table 39). The T_3 thyroid hormone values were significantly increased in males in the second study, but not in the first study. There were no significant changes in the TSH values in either study (Table 39). Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with tamoxifen treatment

104. No significant changes were observed in either sperm numbers or morphology in either TAM study (Table 40). The means and standard deviations for sperm numbers and percentages for sperm morphology from the first study including the combined Subgroups and individual Subgroups are in Annex 8. The second study only reported these results graphically.

Table 37A. Tamoxifen-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
3	M	Liver							↑↑		
		Kidney									
		Heart							↑		
		Adrenals		▲	▲						▲
		Brain			↓						
		Pituitary			↓		▲	↓↓	▲	▲	▽
		Testes									▽
		Epididymis		↓↓	↓↓			↓		↓	↓↓
		Ventral Prostate			↓↓		▽	↓↓	▲	▲	▽
		Dorsolateral Prostate	▽	▽	↓↓	▽	▽	▽		▽	↓
		Seminal Vesicles		▽	↓↓	▲	▽	↓		▽	↓↓
	F	Liver	↓		↓↓			↓	▽	↓↓	↓↓
		Kidneys			↓↓						↓↓
		Heart			↓↓			↓			↓↓
		Adrenals		▽	↓↓	▽		▽		▽	↓
		Brain									
		Thymus			▲			▲			
		Pituitary		▽	↓		▽	▽		▽	▽
		Ovary	▽	▽	↓↓	↓	▽	↓↓			↓
		Uterus	↑↑	↓	↓↓	↑	▽	↓	▲	▽	↓
10	M [a]	Heart	▲			Not applicable			Not applicable		
		Testes									
		Epididymes	▲								
		Prostate	↓[b]	↓[b]	↓[b]						
	F	L Ovary	▽	▽	↓[b]						
		R Ovary	↓[b]	↓[b]	↓[b]						
		Uterus		▽	↓[b]						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

[a] Seminal vesicle and coagulating gland weights were not reported by laboratory 10.

[b] No indication in the final report that $p < 0.01$ was analyzed. Not applicable – only a single group of six animals was used in this study.

Table 37B. Tamoxifen-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
3	M	Kidney	↓		↑↑			↑	↓		
		Heart		↑	↑↑						▲
		Adrenals		▲	↑↑		▲	▲			↑↑
		Brain			↑↑			↑↑			↑
		Pituitary									
		Testes			↑			↑			
		Epididymis	↓↓						↓		
		Ventral Prostate			↓↓	▽	▽	↓↓	▲	▲	▽
		Dorsolateral Prostate	▽	▽	▽	▽	▽	▽	▽	▽	▽
		Seminal Vesicles		▽	↓↓	▲	▽	▽	▽	▽	↓↓
	F	Liver									
		Kidney									
		Heart							▽		
		Adrenals									▽
		Brain			↑↑			▲		↑	↑
		Thymus			↑			▲			▲
		Pituitary			▽			▽			▽
		Ovary	▽	▽	▽	▽	▽	▽			▽
		Uterus	↑		↓↓	↑↑	▽	▽	▲		↓
10	M [a]	Heart	▲		↑[b]	Not applicable			Not applicable		
		Testes	▲		↑[b]						
		Epididymes	▲	▲	↑[b]						
		Prostate	↓[b]	↓[b]	↓[b]						
	F	Ovaries	▽	▽	↓[b]						
		Uterus			↓[b]						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

[a] Seminal vesicle and coagulating gland weights were not reported by laboratory 10.

[b] No indication in the final report that $p < 0.01$ was analyzed. Not applicable – only a single group of six animals was used in this study.

Table 38. Significant histopathological findings after tamoxifen treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
3	M	**Testis**	Germinal epithelial vacuolation	1/10	2/10	2/10	1/5		1/5		2/5	1/5
			Germinal epithelial eosinophilia	1/10		3/10			2/5	1/5		1/5
		Mammary gland	Acinar atrophy		9/10	6/8		4/5	2/3		5/5	4/5
		Prostate gland	Decreased secretion		5/10	10/10		1/5	5/5		4/5	5/5
		Seminal vesicles	Decreased secretion		4/10	9/10		2/5	4/5		2/5	5/5
	F	**Ovary**	Interstitial gland hypertrophy		3/10	9/10		2/5	5/5		1/5	4/5
		Uterus	Epithelium and glandular hypertrophy		8/10	9/10		5/5	4/5		3/5	5/5
		Vagina	Mucoid metaplasia		6/10	10/10		4/5	5/5		2/5	5/5
			Epithelial hyperplasia			8/10			4/5			4/5
10	F	**Ovary**	Increased follicular cysts	0/6	1/6	6/6	Not applicable			Not applicable		
		Uterus	Endo. epithelium - metaplasia	1/6	0/6	6/6						
			Endo. Glandular atrophy	1/6	1/6	6/6						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Table 39. Thyroid hormone results after tamoxifen treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
3	M	T_3									
		T_4									
		TSH									
	F	T_3									
		T_4			↑		↑	↑			
		TSH									↓
10	M	T_3			↑	Not applicable			Not applicable		
		T_4									
		TSH									
	F	T_3									
		T_4		↑	↑						
		TSH									

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).

Shading indicates that no statistical significance was observed in the combined Subgroups or individual Subgroups of a given sex. Not applicable – only a single group of six animals was used in this study.

Table 40. Statistically significant changes in sperm parameters after tamoxifen treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
3	Sperm Count									
	Abnormalities									
10	Sperm Count				Not applicable, did not perform Subgroups			Not applicable, did not perform Subgroups		
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group; Not applicable – Analyses of individual Subgroups were not performed or the study was not conduced with individual Subgroups.

Estrous cyclicity results with tamoxifen treatment

105. The first study did not discuss or report the vaginal cytology observations; the second study indicated that no effects on the oestrous cycle were observed based on vaginal cytology, but did not report the data. These data were then specifically requested by the Secretariat, and both laboratories provided the data in a standard format. The vaginal cytology results from both studies are in Annex 9.

Comparison of the updated TG 407 results from tamoxifen treatment with data from other subchronic studies

106. The tamoxifen updated TG 407 studies are summarised in Table 41. Only summary reports of three subchronic tamoxifen studies were available based on a gavage study for three months at 2, 20, and 100 mg/kg/day using ten animals per sex per group; a gavage study for three months at 0.5 and 2 mg/kg/day using 15 females per group; and a gavage study for six months at 0.5, 0.8, 2.4, 4.8 and 9.6 mg/kg/day with 5 animals per sex per group sacrificed at 3 months and the remaining 20 animals per sex per group sacrificed at 6 months (40)(41). More recent, subchronic, chronic or reproductive studies were not found, but a chronic carcinogenicity study was discovered (42).

107. The summary descriptions of the subchronic studies are qualitative in nature. Neither quantitative values or statistical significance were described. In the first study, body weights were depressed in both sexes at all doses; the weights of ovaries, uterus, testes, seminal vesicles, and ventral prostate were depressed in all dose groups; and histopathological changes were confined to the reproductive tract tissues in both sexes. In particular, testicular sperm production was suppressed, the male accessory tissues were atrophic, the ovaries showed limited evidence for ovulation (fewer corpea lutea and follicular cysts), and the uterine epithelial and stoma tissues were regressed. The results of the second study were only said to confirm the results of the first study in the females. In the third study, body weights were decreased, including at the lower doses used in this study. The pattern of changes in male and female reproductive tissues occurred at the lower doses with less severe effects at 0.5 mg/kg/day.

Table 41. Summary of updated TG 407 TAM results.

Parameter	**TAM – updated TG 407** (LOEL μg/kg/d)			
	Laboratory 3		**Laboratory 10**	
	Male	**Female**	**Male**	**Female**
Body weight	↓ 200	↓ 30	↓ 200	↓ 200
Haematology and clinical chemistry				
Total cholesterol	↓ 30		↓ 30	
Organ and Tissue Weights	**(relative)**			
Kidney	↑ 200			
Heart	↑ 30		↑ 200	
Adrenal	↑ 200			
Thymus		↑ 200		
Prostate	↓ 200 (V)		↓ 5 (W)	
Epididymis			↑ 200	
Seminal Vesicles	↓ 200			
Uterus		↓ 200		↓ 200
Ovaries				↓ 200

[a] In the 407 studies these were individual tissues (ventral (V) and whole (W) prostate).
[b] Changes qualitatively reported only in the high dose group; number of affected individuals and grades were not reported.
NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.

Table 41 continued. Summary of updated TG 407 TAM results.

Parameter	TAM – updated TG 407 (LOEL μg/kg/d)			
	Laboratory 3		Laboratory 10	
	Male	Female	Male	Female
Histopathological Findings				
Mammary gland – acinar atrophy	30		NR	NR
Testes – epithelial changes	Eq 200		NR	
Prostate – decreased secretion	30		NR	
Seminal vesicles – decreased secretion	30		NR	
Uterus –atrophic changes in epithelium/ epithelial and glandular cell hypertrophy and hyperplasia		30		200[b]
Ovaries – atrophic changes in follicles/interstitial gland hypertrophy		30		200
Vagina – mucoid metaplasia		30		200
Sperm parameters				
Estrous cycle		NR		NR

108. The qualitative pattern of the TG 407 results is consistent with the summaries of preclinical studies of tamoxifen reported in the literature with the decreases in body weight and the limitation of histopathological findings to the reproductive tracts. The findings are consistent with descriptions of atrophy in the male reproductive tract, effects on the ovarian follicles, and atrophic changes in uterine tissues previously described. As these studies were of longer duration and conducted with higher doses, the 28-day TG 407 results are considered robust in their reproduction of the toxicological pattern in both males and females.

Conclusions for the updated TG 407 performance with tamoxifen

109. The following conclusions are drawn from the updated TG 407 studies with TAM and from the comparison of these results with subchronic studies:

- The updated TG 407 results indicated that the pattern of effects observed in the two TAM studies was consistent across body weights, organ and tissue weights, histopathology and other parameters.
- The updated TG 407 detected clear effects in both the uterus and the ovaries of the female reproductive tract. The first study detected clear histopathological effects in the male reproductive tract; these data were not observed in the second study. This is consistent with 200 μg/kg body weight/day being a minimal dose for the observation of effects in the male reproductive tract based on 3 and 6 month studies.
- The updated TG 407 results also indicated that the dose responses in the two TAM studies were similar.
- The pattern of effects produced in the TG 407 studies with TAM were consistent with summarized reports of subchronic studies in animals of a similar age.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of TAM consistent with the endocrine mechanism of action of an antiestrogen.

CGS 13820B

Introduction

110. This section summarizes the results of the updated TG 407 studies with CGS 18320B (CGS) and compares these results with exposure to this and other related non-steroidal aromatase inhibitors in both screening and subchronic studies (43)(44)(45)(46)(47)(48)(49)(50)(51)(52).

Background on CGS 18320B

111. CGS 18320B is an unmarketed pharmaceutical aromatase inhibitor; it was used in the updated TG 407 studies due to the unavailability of Fenarimol. Similar drugs to CGS in mechanism of action are: Fenarimol, Fadrozole, Letrozole, and Vorozole. These drugs are typically used as auxiliary therapeutic drugs in the treatment of mammary cancer to suppress the endogenous production of estradiol.

Description of CGS 18320B experiments

112. CGS was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in the first study and to a single group of 10 animals per sex per dose in the second study. The CGS was administered using a corn oil vehicle in three doses to both sexes in each study: a low (L) dose of 0.3 mg/kg body weight/day, an intermediate (I) or mid-dose of 3 mg/kg body weight/day, and a high (H) dose of 30 mg/kg body weight/day. Both studies incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The individual data from the individual Subgroups in the first study were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power; the second study used on a single group size of ten animals per sex per dose.

Summary of the updated TG 407 results with CGS 18320B

113. The two updated TG 407 studies produced consistent response profiles to CGS in the female reproductive tract, and the CGS dosage levels on tissue weight changes and histopathological observations were similar. The pattern of tissue weight changes and histopathological changes observed in the updated TG 407 studies were also similar to those seen in other screening and subchronic studies with CGS and other non-steroidal aromatase inhibitors. Therefore, it is concluded that the updated TG 407 successfully detected effects of CGS consistent with the endocrine mechanism of action of an aromatase inhibitor.

Mortality and body weights in CGS 18320B studies

114. No treatment-related mortalities occurred during either CGS study. Some clinical signs were observed in the high dose males in the first study only, and no other treatment-related clinical signs were noted for other dose groups in either study. Food consumption was significantly increased in the females at all CGS doses in both studies. The functional observational battery and the motor activity assessments in the first study recorded a statistically significant decrease in female grip strength at the high dose, and, in the second study, female mean horizontal (total distance) and mean vertical (number of rearings) activities were significantly decreased at the high dose. The mean body temperature in the high dose females was significantly decreased in both studies.

115. Female body weights were significantly increased (20-40% in absolute terms) at all CGS doses in both studies (Table 42). Male body weights did decrease by approximately 10% at the high dose in the first study, but decreased only modestly by 2.5% in the second study (Table 42). Detailed body weight

means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with CGS 18320B

116. Four haematological and clinical chemistry parameters, female while blood counts and sodium, potassium, and chloride levels, were significantly changed in the combined Subgroups in both studies (studies in agreement) (Table 43). In the case of alanine aminotransferase values, the statistically significant changes in one study were in directional conflict with the other study. For eleven other parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study (see Table 43). There were four parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 43). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Organ and tissue weights results with CGS 18320B treatment

117. In males, there were significant increases in the relative weights of the liver and the adrenals at the high CGS dose in both studies. In the first study, in parallel with the 10% decline in body weights, there were also significant increases in the brain, thyroid, and testes weights at the high CGS dose. In the second study, there was a significant increase in the relative weight of the heart at the high CGS dose (Table 44 B). In the accessory tissues of the male reproductive tract, there was a significant decrease in the relative whole prostate weights in the second study at the mid- and high CGS doses. Absolute declines in the relative seminal vesicles and the ventral prostate were recorded in the first study, but these values did not achieve statistical significance at either the mid- or high CGS doses (Table 44B).

118. In females, the large increase in body weights is paralleled by significant increases in the absolute weights of a number of major organs in both studies at the mid- and high doses of CGS (Table 44A). The relative weights of the liver were significantly increased at the mid- and high CGS doses in both studies; the relative kidney and heart weights were significantly increased at the high CGS dose in both studies, and the relative brain weights were significantly decreased at all CGS doses in both studies (Table 44B). In the female reproductive tract, the relative ovarian weights were significantly increased at all CGS doses in the first study, but no significant changes were observed in the second study (Table 44B). However, the uterine observations were concordant, as the relative weights decreased in both studies at the mid- and the high CGS dose.

119. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Table 42. Changes in body weights during 407 studies with CGS 18320B.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
8	M	- - -1.5%	- - -1.5%	- - -9.7%	- - -3.5%	- - -3.5%	- - -10.9%	- - +0.2%	- - +0.5%	- - -8.2%
	F	↑↑ **+23.6%**	↑↑ **+32.8%**	↑↑ **+26%**	- - +25.5%	↑↑ **+35.5%**	- - +28.7%	↑↑ **+20.8%**	↑↑ **+30.0%**	↑↑ **+22.8%**
13	M	- - -5.1%	- - +1.8%	- - -2.5%	Not applicable			Not applicable		
	F	↑↑ **+31.0%**	↑↑ **+41.2%**	↑↑ **+34.3%**						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

Table 43. Haematology and clinical chemistry results from 407 studies with CGS 18320B.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 8	Statistically Significant in Laboratory 13
Significant Parameters in Agreement in Both Studies: ↑ F White blood cell count ↓ F Potassium ↑ F Sodium ↓ F Chloride **Significant Parameters in Conflict Between Studies**: First Laboratory ↓ F Alanine aminotransferase Second Laboratory ↑ F Alanine aminotransferase	***Common measures, direction of percentage change similar, significant in first study only:*** ↓ M Haematocrit **-8.0** -2.1[a] ↓ M Erythrocyte counts **-6.1** -2.4 ↓ M Haemoglobin concentration **-5.7** -2.8 ↓ F Albumin **-16.6** -4.2 ↓ F Creatinine **-9.4** -14.6 ↑ F Phosphorus **30.0** 12.8 ↓ M Potassium **-11.5** -2.8	***Common measures, direction of percentage change similar, significant in second study only:*** ↑ F Alkaline phosphatase 21.6 **37.0** ↑ F Globulin 2.4 **14.4** ↑ M Sodium 1.2 **1.3** ↓ M Chloride -0.7 **-2.3**
	Common measures, direction of percentage change differ, significant in first study only: ↓ F Haematocrit **-6.1** 3.5 ↓ F Erythrocyte counts **-8.0** 2.4 ↓ F Haemoglobin concentration **-7.4** 1.6 ↓ F Total protein **-9.6** 1.5	***Common measures, direction of percentage change differ, significant in second study only:***
	Measurement performed in this lab only: ↑ F Reticulocyte value ↑ F Segmented neutrophil value ↑ F Prothrombin times	***Measurement performed in this lab only:*** ↑ M Triglycerides

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. Immediately below the specific parameter, the first study percentage change in the absolute value is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Histopathology findings with CGS 18320B treatment

120. In the major organs, histopathological findings were reported in the adrenals, liver, kidneys, and thyroid. Both studies described similar changes in the adrenal cortex, i.e., cortical vacuolisation in the first study in both males and females at the mid- and doses, and cortical lipid hyperplasia in the second studies in males at the mid- and high doses and in females at all doses (Table 45). This is consistent with the differences in specificity among aromatase inhibitors where high dosages may result in inhibition of other steroidogenic enzymes. In the second study, hepatic centrilobular hypertrophy was reported on both sexes at the high CGS dose, but no hepatic changes were reported in the first study (Table 45). Similarly, kidney findings were reported in males and females at the high CGS dose (Table 45). Thyroid findings were reported in both studies, but the descriptions do appear to be consistent as these were described as "increased small follicles" in the first study and as mild follicular hypertrophy in the second study.

121. In the male, there were findings in the first study of atrophy in the mammary glands, but not the second study. There was atrophy in one individual out of nine in several of the accessory sex glands in the first study, but there were no relevant findings in the second study.

122. In the female, both studies observed pituitary changes at all CGS doses. Hypertrophic changes in the female mammary gland were observed at all doses in both studies. In the female reproductive tract, both studies observed dramatic ovarian findings in all animals even at the low CGS dose, and atrophy of the uterine and vaginal tissues at all CGS doses. More details of the histopathology for both studies, including the combined Subgroups and individual Subgroups, are in Annex 6.

Thyroid hormone results with CGS 18320B treatment

123. The thyroid hormone data were not consistent between the two CGS studies. The T_3 and T_4 thyroid hormone values were significantly decreased in the male combined Subgroups at all CGS doses in the first study (Table 46). Although there was no corresponding increase in TSH levels or hepatic weights, the histopathological descriptions could be interpreted to indicate decreased thyroid activity. T_3 levels in females were also significantly decreased.

124. In the second study, T_3 levels were significantly increased in males at the high CGS dose, and T_3 and T_4 levels were significantly increased in females at the high dose and T_4 levels at the low and mid-CGS doses as well (Table 46). TSH values were significantly higher in females at the high dose, and the histopathological findings indicated low grade hypertrophy in a few individuals. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with CGS 18320B treatment

125. No significant changes were observed in either sperm numbers in either CGS study (Table 47). In the first study, there was statistically significant decrease in sperm morphological abnormalities at the low CGS dose. This observation lacked evidence for a dose relationship as no changes were seen at the mid- and high CGS doses, the direction of the change indicated it would not be considered adverse, and the finding was judged to be spurious random event. Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Estrous cyclicity results with CGS 18320B treatment

126. Almost all of the CGS-treated females in both studies were judged to have abnormal oestrous cycles at all doses as the vaginal smears indicated the females remained in diestrous for all five days prior to day 28 of the study. The vaginal cytology results from both studies are in Annex 9.

Comparison of the updated TG 407 results from CGS 18320B treatment with data from chronic, reproductive, and developmental studies

127. Non-steroidal aromatase inhibitors have been sought as treatments for estrogen-stimulated cancers, particularly mammary cancers. Published summaries of preclinical studies indicate that, in the female rat, there are three immediate and primary effects depending upon dose: atrophy of the uterine tissues observable in decreased weight and histopathological changes, diminished or even cessation of the oestrous cycle, and ovarian changes as pituitary stimulation occurs due to inhibited estrogen synthesis and ovulation is inhibited. In some studies, where reported, there were histopathological changes in the pituitary noted. Other effects depending upon the compound and the dose are an increase in the female body weight and, dependent the specificity of the compound, impairment of adrenal steroid synthesis resulting in feedback stimulation effects on that tissue (46)(47)(48)(49)(50)(51). The Secretariat was also provided a highly edited version of subchronic study on CGS 18320B that indicates a significant gain in female body weight with corresponding changes in most organ weights, a decrease in uterine weights and histopathological atrophy of that organ, and a possible increase in ovarian weights (but with no histological data on that tissue provided) (52).

128. There are also reports of decreases in the accessory male reproductive tract tissues. The pharmacological evidence and experimental data on the ventral prostate have been summarized (43). In addition, another screening assay in the adult, intact male has shown statistically significant weight decreases in several accessory male reproductive tract tissues with 15-day administration of two related aromatase inhibitors at high doses (44)(45). Therefore the results summarized in Table 48 for the two TG 407 CGS 18320B studies are consistent with the literature profile of expected effects.

Table 44A. CGS 18320B-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
8	M	Liver									
		Kidney					▽				
		Adrenals	▽		▲	▽		▲	▽		▲
		Brain									
		Thyroid			▲	▽	▽	▲	▲	▲	▲
		Testes							▲		
		Seminal Vesicles	▽	▽	↓	▽	▽	▽			▽
		Whole Prostate	▽	↓	↓↓	▽	▽	▽		▽	↓
		Ventral Prostate	↓	↓	↓	▽	▽	▽	▽	▽	▽
		Dorsolateral Prostate		▽	▽		▽	▽	▲		▽
	F	Liver	▲	↑↑	↑↑	▲	↑↑	↑↑	▲	↑↑	↑↑
		Kidney	▲	↑↑	↑↑	▲	↑↑	↑↑	▲	↑↑	↑↑
		Heart	▲	↑↑	↑↑	↑	↑↑	↑↑	▲	↑	↑↑
		Adrenals	▽	▲	▲	▽	▲	▲	▽		▲
		Brain		↑	↑						
		Pituitary	↓↓	▽	▽	▽		▽	▽	▽	▽
		Spleen	▲	↑↑	↑↑	▲	↑	↑	▲	↑↑	↑↑
		Thymus		↑↑	↑↑	▲	▲	↑		↑↑	↑
		Thyroid									
		Ovaries	↑↑	↑↑	↑↑	▲	↑↑	↑	↑↑	↑↑	↑↑
		Uterus/cervix	▽	↓↓	↓↓	▽	↓	↓↓	↓↓	↓↓	↓↓

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant. Not applicable – this study was not conducted with individual Subgroups, but only a combined group of ten animals per sex per dose.

Table 44A continued. CGS 18320B-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
13	M	Liver				Not applicable			Not applicable		
		Heart									
		Adrenals		↑	↑↑						
		Seminal Vesicles	▽		▽						
		Prostate [a]	↓	↓↓	↓↓						
	F	Liver									
		Kidney	↑	↑↑	↑↑						
		Heart	↑	↑↑	↑↑						
		Adrenals		▲	↑						
		Brain		↑↑	↑↑						
		Spleen		▲	↑						
		Thymus	▲	▲	▲						
		Ovaries	↑	↑	▲						
		Uterus/cervix	↓	↓↓	↓↓						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant. Not applicable – this study was not conducted with individual Subgroups, but only a combined group of ten animals per sex per dose.

[a] Laboratory 13 did not dissect and weigh the ventral and dorsolateral prostate separately.

Table 44B. CGS 18320B-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
8	M	Liver			↑↑			↑			
		Kidney									↑
		Adrenals			↑			↑	▽		▲
		Brain			↑↑			↑			▲
		Thyroid			↑			↑	▲	▲	▲
		Testes			↑↑			▲	▲	▲	▲
		Seminal Vesicles		▽	▽	▽	▽	▽			
		Whole Prostate		▽	▽	▽	▽	▽		▽	▽
		Ventral Prostate	↓	↓	▽	▽	▽	▽	▽	▽	▽
		Dorsolateral Prostate		▽	▽		▽	▽	▲		▽
	F	Liver		↑↑	↑↑			↑↑		↑↑	↑↑
		Kidney			↑			↑			↑↑
		Heart			↑↑			↑↑			↑↑
		Adrenals	↓↓	▽		↓↓	▽		↓	▽	
		Brain	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓
		Pituitary	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓
		Spleen		▲	↑		▲	▲		▲	▲
		Thymus		▲	↑		▲	▲		▲	▲
		Thyroid	↓	↓	▽	▽	▽	▽	▽	▽	▽
		Ovaries	↑	↑↑	↑↑	▲	▲	▲	▲	▲	↑
		Uterus/cervix	▽	↓↓	↓↓	▽	↓↓	↓↓	↓↓	↓↓	↓↓

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 44B continued. CGS 18320B-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
13	M	Liver			↑↑	Not applicable			Not applicable		
		Heart			↑						
		Adrenals	▲	▲	↑↑						
		Seminal Vesicles									
		Prostate [a]	▽	↓↓	↓↓						
	F	Liver		↑↑	↑↑						
		Kidney			↑						
		Heart			↑↑						
		Adrenals									
		Brain	↓↓	↓↓	↓↓						
		Spleen									
		Thymus									
		Ovaries									
		Uterus/cervix	↓↓	↓↓	↓↓						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant. Not applicable – this study was not conducted with individual Subgroups, but only a combined group of ten animals per sex per dose.

[a] Laboratory 13 did not dissect and weigh the ventral and dorsolateral prostate separately.

Table 45. Significant histopathological findings after CGS 18320B treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
8	M	**Adrenal**	Cortical cell vacuolation		1/10	7/9		1/5	4/5		0/5	3/4
		Thyroid	Increased small follicle		3/10	3/9		2/5	2/5		1/5	1/4
		Accessory sex glands	Low rates of atrophy at high dose (1/10 dorsal and ventral prostate, seminal vesicles, coagulating glands)									
		Mammary gland	Atrophy	5/10	7/10	8/9	2/5	5/5	4/5	3/5	2/5	4/4
	F	**Adrenal**	Cortical cell vacuolation		9/10	9/10		5/5	5/5		4/5	4/5
		Spleen	Increased extramed. haematopoiesis	1/10	10/10	10/10	1/5	5/5	5/5		5/5	5/5
		Pituitary	Hytropic degen. basophilic cells	8/10	9/10	10/10	3/5	4/5	5/5	5/5	5/5	5/5
		Thyroid	Increased small follicle[1]	2/10	1/10	6/10	1/5	1/5	4/5	1/5		2/5
		Ovary	Follicular cyst(s)	10/10	10/10	10/10	5/5	5/5	5/5	5/5	5/5	5/5
			Interstitial cell atrophy	10/10	10/10	10/10	5/5	5/5	5/5	5/5	5/5	5/5
		Uterus	Atrophy	7/10	10/10	10/10	2/5	5/5	5/5	5/5	5/5	5/5
		Vagina	Atrophy	7/10	10/10	10/10	2/5	5/5	5/5	5/5	5/5	5/5
		Mammary gland	Hypertrophy	9/10	9/10	10/10	5/5	5/5	5/5	4/5	4/5	5/5
13	M	**Liver**	Centrilobular hypertrophy			6/10	Not applicable			Not applicable		
		Kidney	Tubular atrophy			6/10						
			Tubular dilatation			2/10						
			Mineralization			1/10						
		Adrenal	Fatty cortical lipid hyperplasia[2]	5/10	9/10	10/10						
	F	**Liver**	Centrilobular hypertrophy	2/10		8/10	Not applicable			Not applicable		
		Kidney	Tubular dilatation			1/10						
			Mineralization			6/10						
		Pituitary	Hypertrophy in pars distalis	2/10	5/10	4/10						
		Thyroid	Follicular hypertrophy	2/10	2/10	3/10						
		Adrenal	Fatty cortical lipid hyperplasia[2]	9/10	10/10	10/10						
		Ovaries	Increased antral follicle size	10/10	10/10	10/10						
		Uterus	Atrophy	8/10	10/10	10/10						
		Mammary gland	Increased grade alveolar development	10/10	10/10	10/10						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.
Not applicable – this study was not conducted with individual Subgroups, but only a combined group of ten animals per sex per dose.

[1] Control thyroid – 2/10 in combined, 1/5 in Subgroup A, and 1/5 in Subgroup B.
[2] Control adrenal – 3/10 in combined in males; 5/10 in combined in females.

Table 46. Thyroid hormone results after CGS 18320B treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
8	M	T_3	↓↓	↓↓	↓↓	↓↓				↓	↓↓
		T_4	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓			↓
		TSH									
	F	T_3			↓↓	↓↓		↓			
		T_4									
		TSH	↑↑			↑↑					
13	M	T_3			↑↑	Not applicable			Not applicable		
		T_4									
		TSH									
	F	T_3			↑						
		T_4	↑↑	↑↑	↑↑						
		TSH			↑						

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).
Not applicable – this study was not conducted with individual Subgroups, but only a combined group of ten animals per sex per dose.

Table 47. Statistically significant changes in sperm parameters after CGS 18320B treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
8	Sperm Count									
	Abnormalities	↓								
13	Sperm Count				Not applicable, did not perform Subgroups			Not applicable, did not perform Subgroups		
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group; Not applicable – Analyses of individual Subgroups were not performed or the study was not conduced with individual Subgroups.

Table 48. Summary of the updated TG 407 CGS 18320B results

Parameter	CGS 18320B – updated TG 407 (LOEL mg/kg/d)			
	Laboratory 8		Laboratory 13	
	Male	Female	Male	Female
Body weight		↓ 0.3		↓ 0.3
Haematology and clinical chemistry				
White blood cell counts		↑ 3		↑ 0.3
Organ and Tissue Weights (relative)				
Liver	↑ 30	↑ 3	↑ 30	↑ 3
Heart		↑ 30	↑ 30	↑ 30
Adrenal	↑ 30		↑ 30	
Brain		↓ 0.3		↓ 0.3
Testes	↑ 30			
Prostate			↓ 3 (W)	
Uterus		↓ 3		↓ 0.3
Ovaries		↑ 0.3		
Histopathological Findings				
Liver – centrilobular hypertrophy			30	30
Kidney – epithelial changes			30	30
Adrenal – cortical changes	3	3	0.3	0.3
Mammary gland – atrophy (m)/ hypertrophy (f)	0.3	0.3	NR	0.3
Uterus –atrophic changes in epithelium and epithelial glands		0.3		0.3
Ovaries –cystic changes in follicles/interstitial gland atrophy		0.3		0.3
Vagina – atrophic changes in epithelium		0.3		NR

[a] In the 407 studies these were individual tissues (ventral (V) and whole (W) prostate).
NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.

Table 48 continued. Summary of the updated TG 407 CGS 18320B results.

Parameter	CGS 18320B – updated TG 407 (mg/kg/d)			
	Laboratory 8		Laboratory 13	
	Male	Female	Male	Female
Sperm parameters				
Estrous cycle		0.3		0.3

Conclusions for the updated TG 407 performance with CGS 18320B treatment

129. The following conclusions are drawn from the updated TG 407 studies with CGS and from the comparison of these results with several aromatase inhibitor studies:

- The updated TG 407 results indicated that the pattern of effects observed in the two CGS studies was largely consistent across body weights, organ and tissue weights, histopathology and other parameters. A possible exception is the data for the thyroid.
- The updated TG 407 detected clear effects due to CGS administration in both the uterus and the ovaries of the female reproductive tract in both studies, including changes in tissue weights and histopathology. The first study detected clear histopathological effects in the male mammary gland, but these data were not reported in the second study.

- The updated TG 407 results also indicated that the dose responses for different parameters in the two CGS studies were similar.
- The pattern of effects produced in the TG 407 studies with CGS was consistent with screening and reproductive studies in animals of a similar age.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of CGS consistent with the endocrine mechanism of action of an aromatase inhibitor.

Methyl Testosterone

Introduction

130. This section summarizes the results of the updated TG 407 studies with 17α-methyl testosterone (MT). No recent, major published studies were discovered for comparison with the TG 407 studies.

Background on Methyl Testosterone

131. Methyl testosterone is a potent pharmaceutical androgen that is orally active and whose mode of action is as an agonist for the androgen receptor.

Description of Methyl Testosterone experiments

132. MT was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in both studies. These individual data from these Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power. The MT was administered using 0.5% aqueous methylcellulose in the first study and a corn oil vehicle in the second study. For males in both studies, the three dose levels were: low (L) dose of 10 mg/kg body weight/day, an intermediate (I) or mid-dose of 40 mg/kg body weight/day, and a high (H) dose of 200 mg/kg body weight/day. For females, the three dose levels were: low (L) dose of 10 mg/kg body weight/day, an intermediate (I) or mid-dose of 100 mg/kg body weight/day, and a high (H) dose of 600 mg/kg body weight/day. Both studies incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment.

Summary of the updated TG 407 results with Methyl Testosterone

133. The two updated TG 407 produced consistent response profiles to MT in non-reproductive organs and in the male and female reproductive tracts. The dose responses and MT effect levels observed were similar in the two studies. These profiles are consistent with the effects of androgens, particularly on the feedback inhibition of the pituitary and the Leydig cells of the testes (the endogenous source of ~95% of the testosterone in the adult male rat). Therefore, it is concluded that the updated TG 407 successfully detected effects of MT consistent with the endocrine mechanism of action of an androgen agonist.

Mortality and body weights in Methyl Testosterone studies

134. No treatment-related mortalities occurred during either MT study. Some clinical signs were observed at the respective high dose in males and females in both studies; the primary clinical sign was increased salivation which one study suggested was due to the gavage administration. Food consumption was significantly increased in the first study in the males at the high MT dose and in females at all MT doses, but no significant overall changes were recorded in the second study. In the first study, there were no significant changes in the functional observational battery for both sexes and in motor activity for males, but the motor activity was statistically significantly decreased for all three female treatment groups

in the first study. In the second study, a number of statistically significant changes were recorded, although most were within historical ranges and were not attributed to treatment.

135. Female body weights were significantly increased at all MT doses in the first study, but only at the low and mid-doses in the second study; where the body weights were unchanged at the high dose (Table 49). Male body weights were significantly decreased at the respective MT high dose in the first study and at both the mid- and high doses in the second study (Table 49). Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with Methyl Testosterone

136. In eight instances, haematological and clinical chemistry parameters were significantly changed in the combined Subgroups in both studies (studies in agreement). These were male and female total cholesterol and triglycerides, male prothrombin times, and female alkaline phosphatase, glucose, and phosphorus levels (Table 50). For some twelve parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study (see Table 50). There were six parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 50). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Table 49. Changes in body weights during 407 studies with Methyl Testosterone.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
3	M	- - -0.6%	- - +1.3%	↓ **-7.8%**	- - -1.8%	- - -2.6%	- - -6.3%	- - +0.6%	- - +5.2%	↓ -9.4%
	F	↑↑ **+12.6%**	↑↑ **+16.1%**	↑↑ **+11.0%**	↑↑ **+14.6%**	↑↑ **+10.7%**	- - +7.5%	↑↑ **+10.4%**	↑↑ **+21.3%**	↑↑ **+14.4%**
12	M	- - -3.3%	↓↓ **-10.1%**	↓↓ **-16.6%**	↓ **-7.8%**	↓↓ **-11.6%**	↓↓ **-17.7%**	- - +1.5%	↓ **-8.5%**	↓↓ **-15.5%**
	F	↑↑ **+10.9%**	↑↑ **+13.8%**	- - +0.2%	↑↑ **+10.4%**	↑ **+12.0%**	- - -0.1%	↑ **+11.3%**	↑ **+15.6%**	- - +0.5%

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

Table 50. Haematology and clinical chemistry results from 407 studies with Methyl Testosterone.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 3	Statistically Significant in Laboratory 12
Significant Parameters in Agreement in Both Studies: ↑ M Prothrombin time ↓ M ↓ F Total cholesterol ↑ M ↑ F Triglycerides ↑ F Alkaline phosphatase ↓ F Glucose ↑ F Phosphorus	***Common measures, direction of percentage change similar, significant in first study only:*** ↓ F Mean cell haemoglobin concentration **-3.6** -2.1[a] ↓ M Albumin **-7.3** -3.0 ↓ F Albumin **-10.5** -2.4 ↓ F Aspartate aminotransferase **-11.3** -3.9 ↓ F Blood urea nitrogen **-28.3** -16.1	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ F White blood cell count -13.8 **-27.0** ↓ M Mean cell haemoglobin concentration -1.0 **-2.6** ↑ M Alkaline phosphatase 16.1 **13.6** ↓ M Creatinine -11.5 **-11.5** ↓ F Creatinine -19.9 **-19.1** ↓ M Glucose -5.4 **-21.0** ↑ M Phosphorus 8.2 **11.1**
	Common measures, direction of percentage change differ, significant in first study only: ↑ F Erythrocyte counts **6.8** -2.7 ↓ F Potassium **-12.7** 9.2	***Common measures, direction of percentage change differ, significant in second study only:*** ↓ M Haemoglobin concentration 2.9 **-3.2** ↓ F Haemoglobin concentration 2.7 **-4.0** ↓ M White blood cell count 10.1 **-20.9** ↓ M Aspartate aminotransferase 8.1 **-11.3**
	Measurement performed in this lab only: ↑ F % neutrophils ↓ F % lymphocytes	***Measurement performed in this lab only:*** ↓ M Monocyte value ↓ M Eosinophil value ↓ F Basophil value ↓ M ↓ F Large unstained cells ↑ F Prothrombin time ↑ F γ-Glutamyl transferase ↓ M ↓ F Creatinine kinase

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. The first study percentage is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Organ and tissue weights results with Methyl Testosterone treatment

137. There were concordant statistically significant changes in the absolute and relative weights of major organs in the combined Subgroups of both MT studies. In both sexes, the relative liver weights were significantly increased at the mid- and high MT doses, the relative kidney weights were significantly increased at various MT doses depending upon sex and study, and the relative thymus weights were significantly decreased at various MT doses depending upon sex and study (Tables 51A and B). In females, the adrenals were significantly decreased at the low and mid- MT doses in both studies. However, at the high MT dose, the relative adrenal weights were significantly increased in the second study and, although increased by over 10% in the first study, the increase did not achieve statistical significance (Table 51B). In females, the absolute weights of the heart were increased at the low and high MT dose in both studies, and also at the mid-MT dose in the first study (Table 51B). The relative brain weights in females were decreased at all MT doses in the first study, but only at the low and mid-dose in the second study (Table 51B). The relative thyroid weights were significantly increased in both sexes in the second study at the high MT dose, and although the absolute male relative thyroid values increased by 11% and the female values by 36% in the first study, these did not achieve statistical significance (Table 51B).

138. In males, there were statistically significant decreases in the absolute and relative weights of the testes in both studies at the mid- and high MT doses (Tables 51A and B). For relative weights of the accessory reproductive tissues, the ventral and dorsolateral prostate and the seminal vesicles were significantly increased in the first study. Although the absolute values of these tissues were increased by 10% or more in the second study, these did not achieve statistical significance (Table 51B). Also in the second study, the epididymes were significantly decreased, but the values for this tissue were basically unchanged in the first study; trends in the first study were similar but achieved statistical significance only at the low MT dose (Table 51B).

139. In females, the relative ovarian weights were significantly decreased at the mid- and high MT doses in both studies (Table 51B). In the second study, the relative uterine weights were significantly decreased at the low and mid-MT doses and significantly increased at the high MT dose (Table 51B).

140. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Histopathology findings with Methyl Testosterone treatment

141. Histopathological findings in major organs were concordant with most observed tissue weight changes in both studies. Both studies observed histopathological changes in the liver, kidneys, and adrenals, although the descriptions of these observations sometimes differed. Changes in thymic cortical tissue associated with androgens were observed in the first study. Both studies also observed modest hypertrophic changes in the thyroid follicles of both sexes at the high MT dose (Table 52).

142. In the male reproductive tract, both studies observed clear, but differently described, effects. In the first study, decreased Leydig cell numbers were described at the mid- and high MT doses, and degeneration of the germinal epithelium at the mid- and high MT dose. In the second study, complete arrest of spermatogenesis was recorded at the high MT dose. This is consistent with negative feedback on the pituitary with reduced LH and FSH stimulation which would be expected to result in atrophy of the Leydig cells and to inhibit spermatogenesis. In addition, the first study fixed and reviewed the male mammary glands, observing dose related increases in acinar secretion starting at the low MT dose.

143. In the female reproductive tract, both studies described changes in ovarian tissues indicating reduced activity or atrophy at the mid- and high MT doses. In the uterus, cervix, and vagina, the studies described changes at the mid- and high MT doses that suggest modest estrogenic-type activity in the tissues, which could possibly be due to the aromatization of the MT, paracrine signaling within the tissue by androgens, or both. In the female mammary glands, both studies described an apparent increase in secretory activity and glandular hyperplasia at the mid- and high MT doses.

144. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Thyroid hormone results with Methyl Testosterone treatment

145. The TSH values were significantly increased in both sexes and studies (Table 53). The dose-response was somewhat different between the studies. In the first study, male TSH values increased at all MT doses and female values at the mid- and high MT doses. In the second study, TSH values in both sexes were increased only at the high dose. The T_3 and T_4 thyroid hormone values varied (Table 53). T_3 values were significantly decreased in the second study females at all MT doses, while no other changes were recorded in the other groups (Table 53). T_4 values were increased in the first study in the high dose males and the mid- and high dose females. T_4 values were also increased in the second study in the females at all MT doses (Table 53). Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with Methyl Testosterone treatment

146. The sperm counts and morphology findings were not consistent between the studies. Significant changes were observed in both sperm numbers and morphology in the second MT study at the high MT dose (Table 54). In the first study, there was statistically significant decrease in sperm counts and a statistically significant increase in sperm morphology abnormalities at the mid-dose. However, this latter observation lacked further evidence for a dose-response relationship. Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Estrous cyclicity results with Methyl Testosterone treatment

147. The females in the first study were judged to have abnormal oestrous cycles in the mid- and high dose MT groups (i.e., absent or prolonged cycle). However, there were no details of the smears provided in the second study report, so no direct comparison can be made.

Summary of 407 results with Methyl Testosterone treatment

148. The results of the two updated TG 407 studies with methyl testosterone are summarized in Table 55. The pattern of changes in body weights, clinical chemistry parameters, organ and tissue weights, and histopathological findings are similar as are the overall dose responses of these endpoints. *In utero* screening studies have been performed with related compounds testosterone propionate (53) and trenbolone (54) indicating effects in offspring and in the Hershberger validation program with testosterone propionate (55) and MT (56). The latter studies indicated androgenic effects in the sensitive castrated animal occurred at 2-10 mg/kg/d MT, which compares favorably with results of the intact animals in Table 55. In addition, the possibility of oestrogenic action needs to be considered due to the possible conversion of the androgen by aromatase to relatively potent oestrogen analogues with the increase in uterine weights and the histopathological changes in the female reproductive tract and mammary gland.

Table 51A. Methyl Testosterone-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
3	M	Liver		▲	▲			▲		↑	▲
		Kidney		↑↑	↑↑		▲	↑↑		↑↑	↑↑
		Adrenals		▲	▲			▲	↓	↑	
		Brain			↓				↓↓		
		Spleen								▲	
		Thymus	▽		↓↓	▽		↓↓	▽		↓↓
		Testes		↓↓	↓↓		↓↓	↓↓		↓↓	↓↓
		Ventral Prostate			↑↑		▽	↑		▲	↑↑
		Dorsolateral Prostate			↑↑	▽	▽	↑↑	▲	▲	↑↑
		Seminal Vesicles	▽	↓	↑↑	▽	▽	↑↑	▽	▽	↑↑
	F	Liver	↑	↑↑	↑↑	▲	↑↑	↑↑		↑↑	↑↑
		Kidney	↑↑	↑↑	↑↑	↑↑	↑	↑↑	▲	↑↑	↑↑
		Heart	↑↑	↑	↑↑	↑			▲	▲	↑↑
		Adrenals	↓		↑			↑		▽	▲
		Brain									
		Pituitary		▽	▽	▲	▽	▽	▽	▽	▽
		Spleen	↑	▲		▲			▲	▲	▲
		Thymus		▽	↓↓		↓	↓↓			↓↓
		Thyroid		▲	▲	▲	↑	▲	▽		▲
		Ovary	▽	↓↓	↓↓	▽	↓↓	↓↓	▽	▽	▽
		Uterus + cervix	↓↓		▲	▽			▽	▲	↑

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 51A continued. Methyl Testosterone-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
12	M	Liver				▽					
		Kidney			↑↑			↑↑			↑↑
		Heart									↓
		Brain		↓↓	↓	↓	↓	↓		↓	↓
		Pituitary			↓			↓			↓
		Spleen		↓↓	↓↓	▽	▽	↓		↓↓	↓↓
		Thymus		↓	↓↓		↓	↓↓		▽	↓↓
		Thyroid			↑		▽	▲			▲
		Testes		↓↓	↓↓	↓	↓↓	↓↓		↓	↓↓
		Epididymis		↓↓	↓↓		↓↓	↓↓		↓↓	↓↓
		Right Cauda Epididymis	↓	↓↓	↓↓		↓	↓↓	▽	↓	↓↓
		Ventral Prostate		↓			▽		▽	▽	
		Dorsolateral Prostate		↓↓			↓			▽	
		Seminal Vesicles	▽	↓		▽	▽	▽		↓	
	F	Liver	↑	↑↑	↑↑	▲	↑↑	↑↑		↑↑	↑↑
		Kidney		↑↑	↑↑			↑↑	▲	↑↑	↑↑
		Heart	↑		↑	↑		↑			↑
		Adrenals	▽	↓	▲	▽	↓			▽	▲
		Brain			↓						
		Pituitary	▽	↓↓	↓↓	▽	↓↓	↓↓		▽	▽
		Thymus		↓↓	↓↓		↓	↓↓		↓	↓↓
		Thyroid			↑↑			↑	▽		↑↑
		Ovary		↓↓	↓↓		↓↓	↓↓		↓↓	↓↓
		Uterus/cervix	↓↓	↓	↑↑	▽	▽	↑↑	↓	▽	↑

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 51B. Methyl Testosterone-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
3	M	Liver		↑	↑↑			▲		↑↑	▲
		Kidney		↑↑	↑↑			↑↑		↑↑	↑↑
		Adrenals		▲	↑			▲	↓↓	▲	↑
		Brain									
		Spleen									
		Thymus	▽		↓↓	▽		↓↓	▽		↓↓
		Testes		↓↓	↓↓		↓	↓↓		↓↓	↓
		Ventral Prostate			↑↑		▽	↑↑		▲	↑↑
		Dorsolateral Prostate			↑↑	▽	▽	▲	▲		↑↑
		Seminal Vesicles	▽	↓	↑↑		▽	↑	▽	▽	↑↑
	F	Liver		↑↑	↑↑		↑↑	↑↑		▲	↑↑
		Kidney		↑	↑↑		▲	↑↑		▲	↑↑
		Heart									
		Adrenal	↓↓	↓↓	▲	↓	▽	▲	▽	↓↓	
		Brain	↓↓	↓↓	↓↓	↓↓	↓	↓	↓	↓↓	↓↓
		Pituitary		↓	↓			↓		↓	
		Thymus		↓↓	↓↓		↓↓	↓↓		↓	↓↓
		Thyroid	▽		▲		▲	▲	▽	▽	
		Ovary	↓↓	↓↓	↓↓	↓↓	↓↓	↓↓	▽	↓	↓
		Uterus + cervix	↓↓	▽	▲	▽	▽		↓		▲

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

Table 51B continued. Methyl Testosterone-treated body weight adjusted organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
12 [a]	M	Liver		↑↑	↑↑		↑	↑↑			↑↑
		Kidney	↑	↑↑	↑↑		↑↑	↑↑	↑		↑↑
		Heart			↑↑			↑↑			
		Adrenal			▲		▽	▲	▽	▽	
		Brain						▲			▲
		Pituitary		▽		▽	▽		▽	▽	▽
		Spleen								▽	
		Thymus			↓			↓			▽
		Thyroid			↑↑			↑↑			▲
		Testes		↓	↓↓			↓↓			↓↓
		Epididymis		↓↓	↓↓			↓↓		↓↓	↓↓
		Ventral Prostate		▽	▲		▽	▲	▽	▽	▲
		Dorsolateral Prostate		↓	▲		▽	▲	▽	▽	▲
		Seminal Vesicles		▽		▽	▽			↓	▲
	F	Liver		↑↑	↑↑		↑	↑↑		↑↑	↑↑
		Kidney			↑↑			↑↑		▲	↑↑
		Heart		↓	↑↑			↑↑		▽	
		Adrenal	↓↓	↓↓	↑	↓↓	↓↓		▽	↓	↑
		Brain	↓↓	↓↓		↓	↓↓		▽	▽	
		Pituitary	↓	↓↓	↓	↓↓	↓↓	↓		▽	▽
		Thymus		↓↓	↓↓		↓	↓↓		↓↓	↓↓
		Thyroid	▽		↑↑			↑	↓		↑↑
		Ovary		↓↓	↓↓	↓↓	↓↓	↓↓		↓	↓↓
		Uterus/cervix	↓	↓	↑↑	▽	▽	↑↑	↓	▽	↑

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant.

[a] The Syngenta CTL laboratory uses an ANOVA procedure adjusting organ weights with body weight as a covariate, a slightly different procedure than in most labs. Thus, the extraction below is not strictly comparable.

Table 52. Significant histopathological findings after Methyl Testosterone treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
3	M	**Liver**	Increased cytoplasmic eosinophilia		3/10	9/10			5/5		3/5	4/5
		Kidney	Tubular dilatation		1/10	4/10			3/5		1/5	1/5
			Cortical basophilic tubules	4/10[a]	9/10	9/10	2/5	5/5	4/5	2/5	4/5	5/5
			Microlithiasis/mineralization		3/10	5/10		1/5	2/5		2/5	3/5
			Increased hyaline droplets		2/10	5/10		1/5	2/5		1/5	3/5
		Adrenal	Increased zona fasicul. vacuolation		7/10	5/10		2/5	2/5		5/5	3/5
		Thymus	Reduced cortical tissue			5/10			1/5			4/5
		Thyroid	Follicular hypertrophy			3/10			1/5			2/5
		Testes	Leydig cell –low number (atrophy?)		8/10	10/10		4/5	5/5		4/5	5/5
			Degeneration germinal epithelium		1/10	2/10			1/5		1/5	1/5
			Germinal epithel, intracell. vacuoles		2/10	1/10		1/5	1/5		1/5	
		Epididymis	Oligospermia		2/10	3/10		1/5	2/5		1/5	1/5
			Intratubular spermatic debris		2/10	2/10		1/5	2/5		1/5	
		Mammary	Acinar secretion	2/10	2/9	10/10	1/5		5/5	1/5	2/5	5/5
	F	**Liver**	Increased cytoplasmic eosinophilia		6/10	6/10		2/5	3/5		4/5	3/5
		Kidney	Tubular dilatation			8/10			5/5			3/5
			Cortical basophilic tubules	4/10	4/10	8/10	2/5	2/5	5/5	2/5	2/5	3/5
		Adrenal	Increased zona fasicul. vacuolation		3/10	5/10		2/5	3/5		1/5	2/5
			Zona fasiculata eosinophilic droplets		7/10	9/10		4/5	4/5		3/5	5/5
		Spleen	Extramedullary haemopoiesis		1/10	9/10		1/5	5/5			4/5
		Thymus	Reduced cortical tissue	1/10		10/10			5/5	1/5		5/5
		Thyroid	Follicular hypertrophy		2/10	4/10		2/5	1/5			3/5
		Ovaries	Interstitial cell atrophy		6/10	10/10		4/5	5/5		2/5	5/5
			Follicular atresia	1/10	4/10	10/10	1/5	3/5	5/5		1/5	5/5
		Uterus	Epithelial hyperplasia	1/10	5/10	10/10		3/5	5/5	1/5	2/5	5/5
			Epithelial vacuolation			4/10			3/5			1/5
			Luminal dilatation		6/10	7/10		3/5	3/5		3/5	4/5
			Glandular dilatation		8/10	9/10		4/5	5/5		4/5	4/5
		Vagina	Epithelial hypertrophy/hyperplasia		5/10	10/10		3/5	5/5		2/5	5/5
			Eptih. mucoid metaplasia		10/10	10/10		5/5	5/5		5/5	5/5
		Mammary	Acinar secretion		8/8	9/9		3/3	5/5		5/5	4/4

[a] Control findings for basophilic tubules in the kidney were 4/10.

Table 52 continued. Significant histopathological findings after Methyl Testosterone treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
12	M	**Liver**	Inflammatory infiltration			2/10						2/5
		Kidney	Tubular basophilia	7/10	4/10	9/10	4/5	1/5	4/5	3/5	4/5	5/5
			Microlithiasis/mineralization		8/10	10/10		3/5	5/5		5/5	5/5
		Adrenal	Cortical vacuolation			9/10			4/5			5/5
		Thyroid	Follicular cell hypertrophy		3/5	9/10		2/5	5/5		1/5	4/5
		Testis	Arrested spermatogenesis			10/10			5/5			5/5
		Epididymis	Reduced spermatozoa			9/10			5/5			4/5
			Increased sperm precursor cells			10/10			5/5			5/5
	F	**Adrenal**	Cortical vacuolation			10/10			5/5			5/5
		Thyroid	Follicular cell hypertrophy			10/10			5/5			5/5
		Thymus	Depleted cortical lymphocytes			8/10			4/5			4/5
		Ovary	Reduced corpea lutea		10/10	10/10		5/5	5/5		5/5	5/5
			Increased med/large follicles		9/10	10/10		4/5	5/5		5/5	5/5
		Uterus	Endometrial epithel. hyperplasia		4/10	9/10		2/5	5/5		2/5	4/5
			Increased epithelial vacuolation		10/10	9/10		5/5	5/5		5/5	4/5
			Endometritis			5/10			2/5			3/5
		Vagina	Mucoid metaplasia		9/10	10/10		4/5	5/5		5/5	5/5
			Keratinisation		1/10						1/5	
		Cervix	Mononuclear cell infiltration	1/10	1/10	10/10	1/5	1/5	5/5			5/5
		Mammary gland	Secretory activity			10/10			5/5			5/5
			Hyperplasia		6/10	10/10		3/5	5/5		3/5	5/5

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Table 53. Thyroid hormone results after Methyl Testosterone treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
3	M	T_3									
		T_4			↑						
		TSH	↑	↑	↑↑	↑	↑	↑↑			
	F	T_3	↓	↓	↓↓					↓	↓↓
		T_4		↑↑	↑↑		↑	↑		↑	
		TSH		↑↑	↑↑		↑	↑↑		↑	↑↑
12	M	T_3							↓		
		T_4									
		TSH			↑						↑
	F	T_3									↑
		T_4	↑	↑↑	↑↑		↑↑	↑		↑	↑
		TSH			↑			↑			↑↑

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).

Table 54. Statistically significant changes in sperm parameters after Methyl Testosterone treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
3	Sperm Count		↓							
	Abnormalities		↑							
12	Sperm Count			↓↓			↓↓			↓↓
	Abnormalities			↑↑			↑↑			↑↑

L – low dose group; I – intermediate or mid-dose group; H – high dose group. ↓ - statistically significant decrease in sperm numbers or morphology ($p<0.05$); ↓↓ - statistically significant decrease in sperm numbers or morphology ($p<0.01$); ↑ - statistically significant increase in sperm numbers or morphology ($p<0.05$); ↑↑ - statistically significant increase in sperm numbers or morphology ($p<0.01$).

Table 55. Comparison of the updated TG 407 Methyl Testosterone results.

Parameter	**Methyl Testosterone – updated TG 407** (LOEL mg/kg/d)			
	Laboratory 3		**Laboratory 12**	
	Male	**Female**	**Male**	**Female**
Body weight	↓ 200	↑ 10	↓ 40	↑ 10, 100[a]
Haematology and clinical chem.				
Total cholesterol	↓ 10	↓ 100	↓ 10	↓ 600
Triglycerides	↑ 40	↑ 100	↑ 100	
Organ and Tissue Weights	(relative)			
Liver	↑ 40	↑ 100	↑ 40	↑ 100
Kidney	↑ 40	↑ 100	↑ 10	↑ 600
Heart				↑ 600
Adrenal	↑ 200	↓ 10, 100[b]		↓ 10, 100; ↑ 600[c]
Brain		↓ 10		↓ 10, 100[d]
Pituitary				↓ 10
Thymus	↓ 200	↓ 100	↓ 200	↓ 100
Thyroid			↑ 200	↑ 600
Testes	↓ 40		↓ 40	
Epididymis			↓ 40	
Prostate [c]	V ↑ 200 D ↑ 200			
Seminal vesicles	↓ 40; ↑ 200			
Ovaries		↓ 10		↓ 100
Uterus				↓ 10, 100; ↑ 600

[a] The female body weights decreased only at the low and the mid-doses and were not statistically significant at the high dose.

[b] The female relative adrenal weights in study 1 decreased only at the low and the mid-doses and were not statistically significant at the high dose.

[c] The female relative adrenal weights in study 2 significantly decreased at the low and the mid-doses and significantly increased at the high dose.

[d] The female relative brain weights in study 2 decreased only at the low and the mid-doses and were not statistically significant at the high dose.

NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.

Table 55 continued. Comparison of the updated TG 407 Methyl Testosterone results.

Parameter	Methyl Testosterone – updated TG 407 (mg/kg/d)			
	Laboratory 3		Laboratory 12	
	Male	Female	Male	Female
Histopathological Findings				
Liver – centrilobular hypertrophy	40	100	200	
Kidney – tubular changes	40	600	10	
Adrenal – cortical changes	40	100	200	600
Thymus – reduced cortical tissue	200	600		600
Thyroid – follicular hypertrophy	200	100	40	600
Testes – Leydig cells/germinal epithelium.	40		200	
Epididymis – decreased sperm	40			
Mammary gland – inc. secretion	10	100		
Uterus – multiple changes		100		100
Ovaries – multiple changes		100		100
Vagina – mucoid metaplasia		100		100
Sperm parameters			200	
Estrous cycle		100		

Conclusions for the updated TG 407 performance with Methyl Testosterone treatment

149. The following conclusions are drawn from the updated TG 407 studies with MT and from the comparison of these results with several aromatase inhibitor studies:

- The updated TG 407 results indicated that the pattern of effects observed in the two MT studies was largely consistent across body weights, organ and tissue weights, histopathology and other parameters. The pattern was generally consistent with androgen administration and possible aromatization of the androgen in the female to estrogenic substances.
- The updated TG 407 detected clear effects due to MT administration in both the male and female reproductive tracts in both studies, including changes in tissue weights and histopathology. There were histopathological effects in both the male and female mammary glands.
- The updated TG 407 results also indicated that the dose responses for different parameters in the two MT studies were similar.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of MT consistent with the endocrine mechanism of action of an androgen.

Flutamide

Introduction

150. This section summarizes the results of the updated TG 407 studies with flutamide (FLU) and compares these results with studies based on in utero exposures with flutamide that resulted in the appearance of frank malformations in the male reproductive tracts of the offspring (57)(58)(59)(60).

Background on Flutamide

151. Flutamide is a pharmaceutical antiandrogen that has been used in the treatment of prostate cancer that acts as an antagonist for the androgen receptor.

Description of Flutamide experiments

152. FLU was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in both studies. The FLU was administered using a corn oil vehicle in three doses to both sexes in each study: a low (L) dose of 1 mg/kg body weight/day, an intermediate (I) or mid-dose of 10 mg/kg body weight/day, and a high (H) dose of 100 mg/kg body weight/day. Both studies incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The individual data from these Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power.

Summary of the updated TG 407 results with Flutamide

153. The two updated TG 407 produced consistent response profiles to FLU in non-reproductive organs and, particularly, in the male reproductive tract. The dose responses and FLU effect levels observed were similar in the two studies. These profiles are consistent with the effects of antiandrogens on the accessory tissue of the male reproductive tract both with *in utero* male development and in the pubertal and adult male. Histopathological changes in the testes were consistent with feedback inhibition of androgen signals to the hypothalamus and resulting pituitary stimulation of Leydig cells in the testes. Therefore, it is concluded that the updated TG 407 successfully detected effects of FLU consistent with the endocrine mechanism of action of an antiandrogen.

Mortality and body weights in Flutamide studies

154. No treatment-related mortalities occurred during either study. Some clinical signs were observed in the high dose males in the first study only, and no other treatment-related clinical signs were noted for other dose groups in either study. Food consumption was reduced in both sexes at the high dose in the first study, but was not statistically significant. Food consumption was unchanged in the second study. There were no statistically significant differences were recorded between the control and any dose group in the functional observational battery and the motor activity assessment in either study.

155. Male body weights were significantly decreased by approximately 15% at the high dose in the first study, but decreased only modestly by 5% in the second study (Table 56). Female body weights were not significantly changed in either study (Table 56). Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with Flutamide

156. In 13 cases, haematological and clinical chemistry parameters were significantly changed in the combined Subgroups in both studies (studies in agreement) (see Table 57). For ten other parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study (see Table 57). There were three parameters where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 57). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Table 56. Changes in body weights during 407 studies with Flutamide.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
2	M	- - -6.7%	- - -3.4%	↓↓ **-14.1%**	- - -9.4%	- - -3.2%	↓↓ **-16.5%**	- - -4.1%	- - -4.1%	- - -11.7%
	F	- - -2.4%	- - -1.0%	- - -2.9%	- - -2.4%	- - +1.5%	- - -3.9%	- - -2.9%	- - -3.4%	- - -2.0%
11	M	- - +4.2%	- - -0.4%	- - -5.3%	- - +2.6%	- - -3.2%	- - -11.0%	- - +6.0%	- - +2.6%	- - +0.8%
	F	- - +1.8%	- - +1.8%	- - -0.9%	- - +1.5%	- - +2.2%	- - +0.7%	- - +2.2%	- - +1.4%	- - -2.4%

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

Table 57. Haematology and clinical chemistry results from 407 studies with Flutamide.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 2	Statistically Significant in Laboratory 11
Significant Parameters in Agreement in Both Studies: ↓ M ↓ F Haemoglobin concentration ↓ M Haematocrit ↓ F Erythrocyte counts ↑ M ↑ F Reticulocyte values ↑ M ↑ F Total cholesterol ↑ M ↑ F Albumin ↓ M Alkaline phosphatase ↓ M Triglycerides ↑ F Blood urea nitrogen	***Common measures, direction of percentage change similar, significant in first study only:*** ↓ F Mean cell haemoglobin content **-2.6** -0.9[a] ↑ F Total bilirubin **80.0** cannot calculate values for second laboratory ↑ F Alanine aminotransferase **38.9** 15.4 ↓ M Potassium **-5.6** -6.4 ↓ M Phosphorus **-15.2** -1.2	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ F Haematocrit **-3.8** -5.5 ↓ M Erythrocyte counts **-3.7** -4.9 ↑ M Total protein 1.1 **9.3** ↓ M Aspartate aminotransferase **-14.8** -22.3 ↑ F Calcium 3.0 **3.9**
	Common measures, direction of percentage change differ, significant in first study only: ↓ F Triglycerides **-34.8** 9.7 ↓ M Creatinine **-25.4** 0.0	***Common measures, direction of percentage change differ, significant in second study only:*** ↓ F Glucose 4.8 **-15.5**
	Measurement performed in this lab only:	***Measurement performed in this lab only:*** ↑ M ↑ F Phospholipids ↑ M ↑ F Albumin/globulin ratio ↑ M ↑ F γ-Glutamyl transferase

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. The first study percentage is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Organ and tissue weights results with Flutamide treatment

157. There were concordant statistically significant changes in major organs in the combined Subgroups of both FLU studies. In both sexes, the absolute and relative liver weights were increased at the high FLU doses (Tables 58A and B). In both studies, the relative weights of the adrenals were increased in males at the high FLU dose. In the second study, the relative pituitary weight was significantly increased in males at the mid- and high doses of FLU, and the relative thymus weight was increased in males at the high FLU dose. In the first study, the relative brain weight was increased in females at the high FLU dose (Table 58B).

158. In male reproductive tract, the accessory reproductive tissues were significantly decreased in both absolute and relative weights. The epididymides were decreased in the first study at the mid- and high FLU dose, and also at the low FLU dose in the second study. Ventral prostate was significantly decreased

at the high dose in both studies, and also at the mid-dose of FLU in the second study. The dorsolateral prostate was measured only in the second study, and significantly decreased at the mid- and high FLU doses (Table 58B). The seminal vesicles were significantly decreased in both studies at both the mid- and high FLU doses (Table 58B). The first study also dissected the accessory tissues and weighed the entire male accessory reproductive organ complex or MARO, and the relative weights of the combined tissues were significantly decreased at the mid- and high FLU doses (Tables 58 A and B).

159. In the female reproductive tract tissues, there was a significant decrease in the relative weights of the ovaries in the second study, but not the first study. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Histopathology findings with Flutamide treatment

160. Taking into account differing descriptions of the histopathological lesions, the findings in major organs were concordant between the two studies. Both studies observed centrilobular hypertrophy in the liver of both sexes at the high FLU dose (Table 59). In males, both studies observed changes in the zona fasciculate of the adrenals (Table 59). Pituitary changes in males were described as hypertrophic basophilic cells and increased PAS-positive cells in the first study and clear cells in the anterior lobe of the pituitary in the second study (Table 59).

161. In the male reproductive tract, both studies observed Leydig/interstitial cell hypertrophy/hyperplasia at the high FLU dose. Both studies also observed atrophic changes in the epididymis, the ventral and dorsolateral prostate, and the seminal vesicles and coagulating glands at the high FLU dose. Together with the pituitary observations, these findings provide a consistent pattern of an antiandrogen capable of negative feedback on the pituitary and antagonism in target tissues.

162. In the female reproductive tract, both studies noted changes in the ovary described as increased interstitial glands in the first study and decreased numbers of corpea lutea n the second study.

163. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Thyroid hormone results with Flutamide treatment

164. The thyroid hormone findings between the two studies were in apparent conflict. In the first study, T_3 and T_4 thyroid hormone values were significantly decreased in males at the high FLU dose (Table 60). In the second study, T_3 and TSH values were significantly increased at the high FLU dose (Table 60). In the combined Subgroups, no statistically significant changes were found for any thyroid hormone value in the females in either study. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with Flutamide treatment

165. Significant decreases were observed in sperm numbers in both FLU studies (Table 61). The second study, also recorded a significant decrease in abnormal sperm morphology at the mid-FLU dose, but lacked a dose-response relationship and apparent biological plausibility. Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Table 58A. Flutamide-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
2	M	Liver			▲	▽					▲
		Kidney		↓	↓↓	↓	▽	↓↓			▽
		Heart			↓	▽	▽	▽			▽
		Adrenals		▲						▲	▲
		Brain									
		Spleen		▽	↓	▽	▽	▽	▲		▽
		Ventral Prostate	↓↓	↓↓	↓↓	↓↓	↓	↓↓	↓	↓	↓↓
		Seminal Vesicles	▽	↓↓	↓↓	▽	↓	↓↓		↓	↓↓
		Epididymis	▽	↓↓	↓↓		↓	↓↓	▽	↓	↓↓
		MARO	↓↓	↓↓	↓↓	▽	↓↓	↓↓	▽	↓↓	↓↓
	F	Liver			↑↑			↑			↑↑
		Brain			↓↓						
		Pituitary			↓	nc	nc	nc	nc	nc	nc
11	M	Liver	▲		↑↑			↑	▲	▲	↑↑
		Right Adrenal		▲	↑		▲				▲
		Left Adrenal			▲		▲				
		Pituitary			↑						▲
		Thymus	▲		▲		▽		▲	▲	▲
		Right Epididymis	↓	↓↓	↓↓	▽	↓	↓↓	↓↓	↓↓	↓↓
		Left Epididymis	↓	↓↓	↓↓	▽	▽	↓↓	↓↓	↓↓	↓↓
		Seminal Vesicles	▽	↓↓	↓↓	▽	▽	↓	▽	↓	↓
		Whole Prostate		↓↓	↓↓	▽	↓	↓↓		↓	↓↓
		Ventral Prostate		▽	↓↓	▽	▽	↓↓	▲	▽	↓↓
		Dorsolateral Prostate	▽	↓↓	↓↓	▽	↓	↓	▽	▽	↓
	F	Liver			↑↑			↑↑			↑↑
		Right Ovary			▽			▽			
		Left Ovary			↓			▽			▽

M: male; F: female; L: low dose, I: intermediate or mid-dose; MARO – male accessory reproductive organs; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant; ▽ : 10% or greater decrease, but not statistically significant. nc – statistics not calculated.

Table 58B. Flutamide-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
2	M	Liver			↑↑			↑↑			↑↑
		Kidney									
		Heart									
		Brain									
		Adrenals	▲	↑	↑↑	▲		▲	▲	▲	↑
		Spleen					▽				
		Ventral Prostate	▽	▽	↓↓	▽	▽	↓↓	▽	▽	↓↓
		Seminal Vesicles		↓↓	↓↓		↓	↓↓		▽	↓↓
		Epididymis		↓	↓↓		▽	↓		▽	↓↓
		MARO	▽	↓↓	↓↓	▽	↓↓	↓↓	▽	▽	↓↓
	F	Liver			↑↑			↑↑			↑↑
		Brain			↓						
		Pituitary				nc	nc	nc	nc	nc	nc
11	M	Liver		↑	↑↑			↑↑	↑	↑	↑↑
		Right Adrenal [a]		▲	↑↑		▲	↑		▲	▲
		Left Adrenal		▲	↑↑		▲	↑		▲	▲
		Pituitary		↑	↑↑		↑	↑↑			↑
		Thymus	▲		↑			▲	▲	▲	↑
		Epididymis	↓	↓↓	↓↓			↓↓	↓	↓↓	↓↓
		Seminal Vesicles	▽	↓↓	↓↓	▽	▽	↓	▽	▽	↓
		Ventral Prostate	▽	↓	↓↓	▽	▽	↓↓		▽	↓↓
		Dorsolateral Prostate	▽	↓↓	↓↓	▽	▽	↓	▽	▽	↓
	F	Liver			↑↑			↑↑			↑↑
		Right Ovary [a]			▽			▽			
		Left Ovary			↓			▽			▽

M: male; F: female; L: low dose, I: intermediate or mid-dose; MARO – male accessory reproductive organs; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲ : 10% or greater increase, but not statistically significant ▽ : 10% or greater decrease, but not statistically significant; nc- a control or test substance group n = ≤ 4 (study B), then the power of the test was considered too low to provide meaningful information.

[a] Individual organ weights and their statistics are given in the final report tables; but apparent statistics for the paired organs are reported in the summary.

Table 59. Significant histopathological findings after Flutamide treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
2	M	**Liver**	Hypertrophy			10/10			5/5			5/5
			Cytoplasmic change			10/10			5/5			5/5
			Single cell necrosis			1/10			1/5			
		Adrenal	Vacuolation microvesicular zona fasciculata			9/10			5/5			
		Pituitary	Hypertrophic basophilic cells			10/10			5/5			5/5
			Cellular inclusions		1/10	10/10		1/5	5/5			5/5
			Increased PAS-positive cells	1/10	10/10	10/10		5/5	5/5	1/5	5/5	5/5
		Testis	Leydig cell hypertrophy		9/10	10/10		4/5	5/5		5/5	5/5
		Epididymis	Decreased tubular diameter/size	1/10		9/10			4/5	1/5		5/5
			Increased interstitial tissue	1/10		9/10			4/5	1/5		5/5
			Spermatic debris	1/10	1/10	3/10			2/5	1/5	1/5	1/5
		Ventral prostate	Reduced secretory products	1/10	1/10	10/10	1/5		5/5		1/5	5/5
			Flattened epithelial cells	1/10		9/10	1/5		5/5			4/5
			Reduced acinar size			10/10			5/5			5/5
		Dorsolateral prostate	Reduced secretory products		1/10	8/10		1/5	5/5			3/5
			Flattened epithelial cells			8/10			3/5			5/5
			Reduced acinar size		1/10	10/10		1/5	5/5			5/5
		Seminal vesicles	Reduced secretory products			9/10			5/5			4/5
			Flattened epithelial cells			10/10			5/5			5/5
			Reduced luminal size			9/10			5/5			4/5
		Coagulating glands	Reduced secretory products			9/9			5/5			4/4
			Flattened epithelial cells			9/9			5/5			4/4
			Reduced acinar size			9/9			5/5			4/4
	F	**Liver**	Hypertrophy			10/10			5/5			5/5
			Cytoplasmic change			10/10			5/5			5/5
			Single cell necrosis			1/10			1/5			
		Ovaries	Increased interstitial glands	4/10	3/10	8/10	2/5	2/5	3/5	2/5	1/5	5/5
		Uterus	Squamous cell metaplasia	2/10	3/10	2/10	1/5	2/5	1/5	1/5	1/5	1/5

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Table 59 continued. Significant histopathological findings after Flutamide treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
11	M	**Liver**	Hypertrophy, centrilobular hepatocytes			10/10			5/5			5/5
		Adrenal	Hypertrophy zona fasciculata cortical cells			2/10			2/5			
		Pituitary	Clear cell, anterior lobe		1/10	10/10		1/5	5/5			5/5
		Testis	Interstitial cell hyperplasia			9/10			5/5			4/5
		Epididymis	Decreased sperm number		1/10	10/10		1/5	5/5			5/5
		Prostate	Atrophy		1/10	10/10		1/5	5/5			5/5
		Seminal vesicles	Atrophy			10/10			5/5			5/5
		Coagulating glands	Atrophy		1/10	10/10		1/5	5/5			5/5
		Mammary gland	Small alveolus/duct			10/10			5/5			5/5
	F	**Liver**	Hypertrophy, centrilobular hepatocytes			10/10			5/5			5/5
		Ovary	Decrease corpea lutea/cyst			1/10			1/5			

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Table 60. Thyroid hormone results after Flutamide treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
2	M	T_3			↓						
		T_4			↓↓			↓			
		TSH									
	F	T_3									
		T_4									
		TSH									
11	M	T_3			↑						↑
		T_4									
		TSH			↑						↑
	F	T_3									
		T_4									↓
		TSH									

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).

Shading indicates that no statistical significance was observed in the combined Subgroups or individual Subgroups of a given sex.

Table 61. Statistically significant changes in sperm parameters after Flutamide treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
2	Sperm Count			↓						
	Abnormalities									
11	Sperm Count			↓↓			↓↓			↓↓
	Abnormalities		↓			↓	↓			

L – low dose group; I – intermediate or mid-dose group; H – high dose group; ↓ - statistically significant decrease in sperm numbers or morphology ($p<0.05$); ↓↓ - statistically significant decrease in sperm numbers or morphology ($p<0.01$).

Estrous cyclicity results with Flutamide treatment

166. The females in both studies were judged to have exhibited normal oestrous cycles with no-treatment related effects.

Comparison of updated TG 407 results from Flutamide treatment with data from developmental and other studies

167. Flutamide has been extensively studied as a reference androgen receptor antagonist. *In utero* exposure causes the development of frank malformations in almost all tissues of the male reproductive tract that require androgen stimulation in a window of approximately gestational days 16-20 (57)(58)(59). These effects are observed at administered doses to the dams of 5 mg/kg body weight/day and higher. Similarly, flutamide will inhibit testosterone dependent sexual development and maturity of the male accessory reproductive tract tissues when administered to prepubertal males starting about postnatal day 25 and thereafter at similar to slightly higher doses (60)(61). In the Hershberger screening assay, flutamide is an effect antagonist of administered androgens across all accessory tissues tested, and this occurs reproducibly across laboratories in the Phase-1B validation studies of the Hershberger assay for five male accessory sex tissues (55).

168. The findings in both updated TG 407 studies are consistent with this body of evidence that the primary target of flutamide is the male accessory sex tissues whose growth is stimulated by testosterone and its metabolite dihydrotestosterone acting through the androgen receptor. The result of FLU administration is then decreased weight and histopathological atrophic changes in the target tissues (Table 62). Related findings in the pituitary and testes reflecting inhibition of the hypothalamic feedback controlling testosterone production were also observed, including effects on sperm production. In addition, the findings of both studies are consistent with changes in other organs (Table 62).

Table 62. Comparison of the updated TG 407 Flutamide results.

Parameter	**FLU – updated TG 407** (LOEL mg/kg/d)			
	Laboratory 2		**Laboratory 11**	
	Male	**Female**	**Male**	**Female**
Body weight	↓ 100			
Haematology and clinical chemistry				
Haematocrit	↓ 100		↓ 100	
Haemoglobin concentration	↓ 100	↓ 100	↓ 100	↓ 100
Erythrocyte counts		↓ 100	↓ 100	↓ 100
Total cholesterol	↑ 100	↑ 100	↑ 100	↑ 100
Triglycerides	↓ 100	↓ 100	↓ 100	
Albumin	↑ 100	↑ 100	↑ 100	↑ 100
Organ and Tissue Weights	(relative)			
Liver	↑ 100	↑ 100	↑ 100	↑ 100
Adrenals	↑ 10		↑ 100	
Pituitary			↑ 10	
Epididymis	↓ 10		↓ 10	↓ 10
Ventral Prostate	↓ 100		↓ 10	
Dorsolateral Prostate	NR		↓ 10	
Seminal Vesicles	↓ 10		↓ 10	
Ovaries		↓ 100		

Table 62 continued. Comparison of the updated TG 407 Flutamide results.

Parameter	FLU – updated TG 407 (LOEL mg/kg/d)			
	Laboratory 2		Laboratory 11	
	Male	Female	Male	Female
Histopathological Findings				
Liver – centrilobular hypertrophy	100	100	100	100
Adrenal – cortical changes	10		100	
Pituitary	10		100	
Testes	100		100	
Epididymis	100		100	
Ventral Prostate	100		100	
Dorsolateral Prostate	100		100	
Seminal vesicles	100		100	
Coagulating glands	100		100	
Ovaries – inc. interstitial glands and decreased corpea lutea		100		Eq 100
Sperm parameters	100		100	
Estrous cycle				

[a] In the 407 studies these were individual tissues (ventral (V) and whole (W) prostate).
NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.

Conclusions for the updated TG 407 performance with Flutamide treatment

169. The following conclusions are drawn from the updated TG 407 studies with FLU and from the comparison of these results with several aromatase inhibitor studies:

- The updated TG 407 results indicated that the pattern of effects observed in the two FLU studies were largely consistent across body weights, organ and tissue weights, histopathology and sperm number parameters. A possible exception is the data for the thyroid, which appeared to be contradictory between the studies.
- The updated TG 407 detected clear effects due to FLU administration in livers of both sexes, the male adrenal and pituitary, the ovary, and particularly the male reproductive tract in both studies. The overall pattern of effects indicates an antiandrogenic profile, including the ability to cause negative feedback on the FSH/LH secretion of the pituitary.
- The updated TG 407 results also indicated that the dose responses for different parameters in the two FLU studies were similar.
- The pattern of effects in the male reproductive tract accessory tissues was produced in the TG 407 studies with FLU and were consistent with a large body of data from in utero studies, pubertal studies, and other screening studies.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of FLU consistent with the endocrine mechanism of action of an anti-androgen.

p,p'-DDE

Introduction

170. This section summarizes the results of the updated TG 407 studies with *p,p'*-DDE or 1,1-dichoro-2,2-bis-(*p*-chlorophenyl)ethylene (DDE) and compares these results with impairment of the male reproductive tract development with *in utero* and pubertal exposures to *p,p'*-DDE (62)(63)(64), and in the Hershberger screening assay (65). The compound has also shown thyroid related effects in another endocrine screening assay with the intact male (66).

Background on p,p'-DDE

171. *p,p'*-DDE is an environmental persistent and bioaccumulative metabolic product of the banned pesticide DDT. *p,p'*-DDE is a ligand that competes with testosterone for binding to the androgen receptor (62), and has shown antiandrogenic properties on the male offspring, when administrated to pregnant dams (62)(63) and on pubertal development in males (64). *p,p'*-DDE is also positive in Hershberger bioassay in Sprague Dawley and Long Evans rats at 100-150 mg/kg/d (56)(65), and DDE administration has resulted in thyroid effects in an intact male screening assay (66).

Description of p,p'-DDE experiments

172. DDE was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in both studies. The DDE was administered using a corn oil vehicle in three doses to both sexes in each study. For males in the first study, the initial high DDE dose was lowered from 200 mg/kg/d to 150 mg/kg/d on day 7 during the study due to mortalities. The other two doses for males were not modified during the study: a low (L) dose of 12.5 mg/kg body weight/day and an intermediate (I) or mid-dose of 50 mg/kg body weight/day. For females in this first study, the three DDE doses throughout the study were: a low (L) dose of 6.5 mg/kg body weight/day, an intermediate (I) or mid-dose of 25 mg/kg body weight/day, and a high (H) dose of 100 mg/kg body weight/day. As there were continued clinical signs and body weight losses in the first study; the second study lowered the initial high DDE dose for males to 100 mg/kg/d and for females to 75 mg/kg/d. The low and intermediate doses for both sexes were the same as the first study. Therefore, these studies are not fully comparable at the high dose in either sex, and more severe effects may be plausibly expected in the first study. Both studies incorporated the complete updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The individual data from these Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power. In one study, the histopathology data were analyzed only for the combined Subgroups.

Summary of the updated TG 407 results with p,p'-DDE

173. Neither updated TG 407 assay provided clear evidence of the antiandrogenic activity of *p,p'*-DDE. Instead, both assays provided evidence that *p,p'*-DDE may impair the thyroid resulting in classical follicular hypertrophy based upon histopathological examination. The large degree of hepatic enlargement and changes in circulating thyroid hormone levels suggest these thyroid effects may be the result of increased elimination of circulating thyroid hormones resulting in increased pituitary stimulation of thyroid TSH secretion.

Mortality and body weights in p,p'-DDE studies

174. In the first DDE study with the greater respective male and female high dose, there were five male and three female mortalities; all were judged to be treatment related. There were no mortalities in the

lower dose groups or in the second study. Clinical signs were observed at the high dose in both studies, particularly, in males, but also in females. Food consumption was significantly decreased in both sexes at their respective high doses in the first study, but not in the second study. However, there were no statistically significant differences between the control and any dose group in any other functional observational battery and motor activity assessment parameters in either study.

175. There were no significant changes in the body weights of either sex in any dose group in the two DDE studies. (Table 63). Even in the high dose groups, the percentage changes versus controls were modest in both studies. In the first study with the higher respective doses, there was a +5.6% increase in the high dose females and only a -1.4% decrease in the high dose males. Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with p,p'-DDE

176. In ten instances with sexes considered separately, haematological and clinical chemistry parameters were significantly changed in the combined Subgroups in both studies (studies in agreement) (see Table 64). There were no parameters where the significant changes in one study were directionally in conflict with the other study. For seven other parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study (Table 64). There were some 13 instances where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 64). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Organ and tissue weights results with p,p'-DDE treatment

177. There were concordant statistically significant changes in major organs in the combined Subgroups of both DDE studies. The absolute liver weights were increased at all DDE doses in the males of both studies and at the mid- and high doses in the females in both studies (Table 65A). The relative liver weights were increased in the males in the second study at all DDE doses and the in the males in the first study and the females in both studies at the mid- and high doses of DDE (Table 65B). These increases were dramatic with absolute values more than doubling in the first study were the DDE doses were highest and nearly doubling at the somewhat reduced doses in the second study. The relative kidney weights were increased in males at the high DDE in both studies (Table 65B).

178. Statistically significant changes were observed in the thyroid. Absolute and relative thyroid weights were increased in females in laboratory 6 at the high PTU dose, and relative thyroid weights of both sexes were increased in laboratory 7 (Tables 65A and B).

Table 63. Changes in body weights during 407 studies with *p,p'*-DDE.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
6	M	- - +2.5%	- - +6.1%	- - -1.4%	- - +9.5%	↑ +11.5%	- - +2.6%	- - -3.8%	- - +1.2%	- - -5.0%
	F	- - +5.1%	- - +1.3%	- - +5.6%	- - +0.7%	- - -1.7%	- - +0.5%	- - +9.6%	- - +4.4%	- - +10.0%
7	M	- - -1.6%	- - -0.8%	- - -4.7%	- - -0.7%	- - -3.1%	- - -3.8%	- - -2.5%	- - +1.5%	- - -5.5%
	F	- - -0.6%	- - -1.2%	- - -3.6%	- - -0.7%	- - +1.3%	- - +2.3%	- - -0.7%	- - -3.6%	- - -9.4%

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

Table 64. Haematology and clinical chemistry results from 407 studies with *p,p'*-DDE.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 6	Statistically Significant in Laboratory 7
Significant Parameters in Agreement in Both Studies: ↑ M Activated prothrombin time/activated partial thromboplastin time ↑ M ↑ F Platelet count ↑ M ↑ F Cholesterol ↓ M Triglycerides ↑ M ↑ F Total protein ↓ M Albumin/globulin ratio ↑ M Calcium	***Common measures, direction of percentage change similar, significant in first study only:*** ↑ F Albumin **14.6** 7.9[a] ↑ M Blood urea nitrogen **21.1** 11.1	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ F Alkaline phosphatase -14.7 **-44.1** ↓ M Aspartate aminotransferase -30.2 **-24.1** ↓ M Glucose -10.8 **-22.1** ↓ F Glucose -7.3 **-18.6** ↑ M Sodium 1.1 **1.2**
	Common measures, direction of percentage change differ, significant in first study only: ↓ M Mean cell haemoglobin concentration **-2.7** 0.3 ↓ F Mean cell haemoglobin concentration **-1.8** 1.5 ↓ M Mean erythrocyte cell vol. **3.4** -0.4 ↓ F Haemoglobin concentration **-7.6** 1.4 ↓ F Haematocrit **-6.1** 0.2 ↓ F Erythrocyte counts **-7.7** 1.5 ↑ M Albumin **13.2** 0.0 ↑ F Phosphorus **10.6** -9.3 ↓ F Chloride **-3.3** 0.8	***Common measures, direction of percentage change differ, significant in second study only:*** ↓ F Triglycerides 0.0 **-49.2** ↑ M γ-Glutamyl transferase 12.0 **-9.3** ↓ F Alanine aminotransferase 2.9 **-19.4** ↓ F Blood urea nitrogen 17.2 **-33.3**
	Measurement performed in this lab only: ↑ M ↑ F Globulin	***Measurement performed in this lab only:***

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. The first study percentage is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

179. The only statistically significant changes in the reproductive tracts of either sex were an increase in the relative ovarian weights in the second study (Tables 65A and B). The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Histopathology findings with p,p'-DDE treatment

180. The histopathological findings in major organs corroborated the changes in organ weights with DDE administration. Both studies observed centrilobular hypertrophy in the liver of both sexes. All animals in the three male dose groups were affected in both studies, and all females of all dose groups in the second study and of the mid- and high DDE doses in the first study. In males, eosinophilic bodies in the kidneys were observed at the mid- and high DDE doses in the second study, and vacuolization and cortical hypertrophy in the adrenals at the mid- and high DDE doses in the first study. Other findings included hypertrophy of the clear cells in the pituitary in the first study at the mid- and high DDE doses. All of the above changes displayed a clear dose response (Table 66).

181. The major effect observed was hypertrophy of the thyroid follicular cells. This was observed in both sexes at the lowest DDE dose in the first study and the mid-DDE dose in the second study. The findings increased in a dose responsive manner. Together with the thyroid and liver weights, these data suggest an increased metabolism of circulating thyroid hormones resulting in pituitary stimulation leading to thyroid hypertrophy.

182. Several histopathological changes were recorded in the male reproductive tract in the high DDE dose group of the first study, including increased residual bodies and atrophy of several sex accessory tissues including the epididymis, seminal vesicles, and prostate. However, five of the male animals had died prematurely in the early phase of the study leading to a reduction of the high DDE dose and other animals had died subsequently in the later phase of the study. Upon examination of the individual animal data, it is clear that the changes in the male reproductive tract are associated with animal deaths including those with only short-term DDE exposure. This suggests possible autolysis or degeneration of the tissues due to the moribund state of the animals before the tissues could be fixed as the cause of the observations. In contrast, the thyroid changes are clearly associated with the surviving animals that had long-term exposure to DDE and underwent normal necropsy procedures (Table 67).

183. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Thyroid hormone results with p,p'-DDE treatment

184. The T_4 thyroid hormone values were significantly decreased in the male combined Subgroups in both studies at the high DDE dose. The TSH values were increased at all DDE doses in the males in the first study, but no statistically significant TSH changes occurred in the second study (Table 68). The T_3 thyroid hormone was significantly increased in males combined Subgroups in both studies at the high DDE, and also in the low and mid-doses in the second study (Table 68). In females, T_3 thyroid hormone was significantly increased at the high DDE dose in the second study. Absolute values of T_4 declined and TSH increased in the combined Subgroups of both studies, but did not achieve statistical significance in either study (Tables 15 and 16, Annex 4). The male thyroid hormones then provide support for the increases in the relative thyroid weights and the histopathology that would indicate thyroid hypertrophy in the DDE treated animals (Table 68), but the female data offer only tentative support. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with p,p'-DDE treatment

185. The only significant changes were observed in sperm numbers or morphology in either DDE study occurred at the low DDE dose in the first study that lacked evidence for a dose relationship as the absolute sperm numbers for the mid- and high DDE dose were similar or even higher (Table 69). Absolute values in the second study tended to be slightly higher with increased DDE doses. Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Estrous cyclicity results with p,p'-DDE treatment

186. Based on vaginal smears, the females in both studies were judged to have exhibited normal oestrous cycles.

Comparison of 407 results from p,p'-DDE treatment with data from chronic, reproductive, and developmental studies

187. The influence of *p,p'*-DDE administration on the male during periods of male reproductive tract development has been studied in short-term *in utero* and pubertal studies. These studies have indicated a reduced anogenital distance and retained nipples after *in utero* exposure during the window of male reproductive tract differentiation and development at 100 mg/kg/d DDE by gavage to the dam (62)(63)(64). *In utero* and lactationally exposed adult males had a low rate of hypospadias and reduced ventral prostate weights, but other tissues in the reproductive tract were not affected (64). DDE administration (100 mg/kg/d) after weaning (pnd 21-57) results in a delay in preputial separation (62). In the Hershberger assay, published data indicate effects on the ventral prostate and seminal vesicles at 100 mg/kg/d DDE (65), while in the OECD Phase-2 studies of the Hershberger validation program DDE has indicated statistically significant effects on one or more tissues in a range of 30-160 mg/kg/d DDE in nine laboratories (56). These latter studies have shown dramatic hepatic weight increases at DDE doses lower than those having effects on the male reproductive organs as an antiandrogen (5-16 mg/kg/d DDE). In short-term 14-day studies with DDE, hepatic weight increases began at 50 mg/kg/d and were parallel to statistically significant decreases in circulating T_4 beginning at the same dose level (66).

188. Collectively, these data suggest that DDE has a combination of toxicities, acting as both an antiandrogen and an antithyroid. The data suggest that the thyroid effects might precede the antiandrogenic effects, particularly in the intact male where the hypothalamic-pituitary-testes axis may be able to compensate for the antiandrogenic effects with increased testosterone production directed by the pituitary. This presents a somewhat difficult picture for the interpretation of the TG 407 data.

Table 65A. *p,p'*-DDE-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
6	M	Liver	↑	↑↑	↑↑	↑↑	↑↑	↑↑	▲	↑↑	↑↑
		Kidney			▲	▲	▲	▲			
		Thyroid		▲			↑				▲
		Epididymes	▽						▽		
		Ventral Prostate				▲			▽		▽
		Dorsolateral Prostate	▽			▽	▽				
		Seminal Vesicles				▲		▲	▽	▲	
	F	Liver	▲	↑↑	↑↑	▲	↑	↑↑	▲	↑	↑↑
		Thyroid			↑↑			↑	▲	▲	▲
7	M	Liver	↑↑	↑↑	↑↑	▲	↑↑	↑↑	↑	↑↑	↑↑
		Kidney								↑	▲
		Adrenal				▽			▲		▲
		Thyroid		▲	▲		▲	▲		▲	▲
		Epididymes						↓			
		Ventral Prostate		▲			▲		▽		
		Dorsolateral Prostate			▽		▽	▽	▽		▽
		Seminal Vesicles								▲	
	F	Liver	▲	↑↑	↑↑		↑	↑↑	▲	↑	↑↑
		Kidney									▲
		Thyroid	▽		▲			▲	▽	▽	

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant.

Table 65B. *p,p'*-DDE-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
6	M	Liver	▲	↑↑	↑↑	↑	↑↑	↑↑	▲	↑↑	↑↑
		Kidney			↑↑			↑↑			▲
		Thyroid		▲			↑				▲
		Epididymes	▽	▽		↓↓	↓↓	↓	▽		
		Ventral Prostate									
		Dorsolateral Prostate				▲	▽		▽		▽
		Seminal Vesicles	▽					▲			
	F	Liver	▲	↑↑	↑↑	▲	↑↑	↑↑	▲	↑	↑↑
		Thyroid			↑			▲			▲
7	M	Liver	↑↑	↑↑	↑↑	↑↑	↑↑	↑↑	↑↑	↑↑	↑↑
		Kidney			↑↑					↑	↑↑
		Adrenal			▲	▽			▲		↑↑
		Thyroid		▲	↑		▲	↑		▲	▲
		Epididymes									
		Ventral Prostate		▲			▲	▲	▽		
		Dorsolateral Prostate			▽		▽	▽	▽		▽
		Seminal Vesicles								▲	
	F	Liver	▲	↑↑	↑↑		↑	↑↑	▲	↑	↑↑
		Kidney			▲					↑	↑
		Thyroid	▽		▲			▲	▽	▽	▲
		Ovaries		▲	↑↑	▲		↑↑		▲	↑

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant.

Table 66. Significant histopathological findings after *p,p'*-DDE treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
6	M	**Liver**	Hypertrophy, centrilobular. hepatocytes	10/10	10/10	5/5	Not applicable			Not applicable		
			Hypertrophy diffuse			5/5						
		Thyroid	Follicular cell hypertrophy	6/10	9/10	5/5						
		Adrenal	Hypertrophy, cortical cells			3/5						
			Vacuolisation	2/10	5/10	5/5						
		Pituitary	Hypertrophy, clear cells		3/10	3/5						
		Brain	Congestion			4/5						
		Testes	Increased residual bodies			4/5						
	F	**Liver**	Hypertrophy, centrilobular. hepatocytes	7/10	10/10	7/7						
		Thyroid	Follicular cell hypertrophy	3/10	9/10	7/7						
7	M	**Liver**	Hypertrophy, centrilobular. hepatocytes	10/10	10/10	10/10	5/5	5/5	5/5	5/5	5/5	5/5
			Ground glass bodies	10/10	10/10	10/10	5/5	5/5	5/5	5/5	5/5	5/5
			Mitotic figures		5/10	6/10		2/5	2/5		3/5	4/5
		Thyroid	Follicular cell hypertrophy		10/10	10/10		5/5	5/5		5/5	5/5
			Follicular cell hyperplasia		7/10	10/10		3/5	5/5		4/5	5/5
		Kidney	Eosinophilic body	4/10	9/10	9/10	2/5	4/5	4/5	2/5	5/5	5/5
	F	**Liver**	Hypertrophy, centrilobular. hepatocytes	10/10	10/10	9/9	5/5	5/5	5/5	5/5	5/5	4/4
			Ground glass bodies	6/10	10/10	9/9	2/5	5/5	5/5	4/5	5/5	4/4
			Mitotic figures	8/10	10/10	9/9	5/5	5/5	5/5	3/5	5/5	4/4
		Thyroid	Follicular cell hypertrophy		3/10	6/9		1/5	3/5		2/5	3/4
			Follicular cell hyperplasia			2/9			1/5			1/4
		Adrenal	Hypertrophy zona fasciculata	ND	ND	2/9	ND	ND	0/5	ND	ND	2/4

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.
Not applicable – Laboratory 6 did not report the histopathological findings for the individual Subgroups, but only for the combined Subgroups.
ND – these tissues from the low and mid-dose groups were not read.

Table 67. Histopathological data in study 1 with death or termination of high dose DDE males in laboratory 6.

Animal Number	**IGI669M**	**IGI686M**	**IGI691M**	**IGI693M**	**IGI696M**	**IGI698M**	**IGI700M**	**IGI701M**	**IGI706M**	**IGI707M**
Death or Termination Status	SM	SD	SD	SD	SS	SS	SS	SS	SS	SM
Day of Death/Termination	18	6	7	28	Necropsy	Necropsy	Necropsy	Necropsy	Necropsy	7
Thyroid										
hemorrhage	-	-	-	3	-	-	-	-	-	-
hypertrophy/hyperplasia, follicular epithelium	-	-	-	-	3	3	4	2	3	-
Testes										
degeneration, unilateral, diffuse	<->	-	-	2	-	-	-	-	-	-
increased residual bodies	<->	2	2	1	-	-	-	-	-	1
Epididymides										
atrophy	2	3	3	-	-	-	-	-	-	2
Infiltration, mononuclear cell, focal	-	-	-	1	-	-	-	-	-	-
Necrotic germ cells/ residual bodies	1	3	2	1	-	-	-	-	-	1
Prostate										
atrophy	<3>	<2>	2	<2>	-	-	-	-	-	3
inflammation, chronic, focal/multifocal	-	<2>	-	-	-	-	-	-	-	-
Seminal vesicles										
atrophy	<2>	<2>	<3>	<2>	-	-	-	-	-	<3>
Coagulating glands										
atrophy	2	2	3	-	-	-	-	-	-	3

SS – Scheduled sacrifice
SD – Spontaneous death
SM – Sacrifice Moribund

< > Brackets indicates gross path observation; <-> indicates no histopathological observations consistent with gross pathology;
<3> indicates a pathological observation of grade 3 consistent with gross pathology
3 grade 3 histopathological observation, but no gross pathological observation

Table 68. Thyroid hormone results after *p,p'*-DDE treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
6	M	T_3			↑↑						
		T_4			↓↓			↓			
		TSH	↑	↑	↑	↑		↑↑			
	F	T_3							↑	↑	
		T_4									
		TSH		↑							
7	M	T_3	↑	↑↑	↑↑			↑↑	↑	↑	↑↑
		T_4			↓↓			↓↓			↓
		TSH									
	F	T_3			↑↑			↑			↑↑
		T_4									↓
		TSH				↑					

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase thyroid hormone levels ($p<0.01$).

Table 69. Statistically significant changes in sperm parameters after *p,p'*-DDE treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
6	Sperm Count	↓								
	Abnormalities									
7	Sperm Count									
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group; ↓ - statistically significant decrease in sperm numbers or sperm morphology ($p<0.05$).

Table 70. Comparison of updated TG 407 DDE results.

Parameter	DDE – updated TG 407 (LOEL mg/kg/d)			
	Laboratory 6		Laboratory 7	
	Male	Female	Male	Female
Body weight				
Haematology and clinical chemistry				
Haematocrit		↓ 100		
Erythrocyte counts		↓ 100		
Platelet counts	↑ 50	↑ 100	↑ 50	↑ 75
Total cholesterol	↑ 50	↑ 100	↑ 50	↑ 75
Triglycerides	↓ 200/ 150		↓ 50	↓ 75
Total protein	↑ 50	↑ 25	↑ 100	↑ 75
Organ and Tissue Weights	**(relative)**			
Liver	↑ 50	↑ 25	↑ 12.5	↑ 25
Kidney	↑ 200/ 150		↑ 100	
Thyroid		↑ 100	↑ 100	↑ 75
Ovaries				↑ 75
Histopathological Findings				
Liver – centrilobular hypertrophy	12.5	6.5	12.5	6.5
Kidney – various changes			12.5	75
Adrenal – cortical changes	12.5			Eq 75
Thyroid – follicular hypertrophy	12.5	Eq 6.5/ 25	50	Eq 25/ 75
Sperm parameters				
Estrous cycle				

[a] In the 407 studies these were individual tissues (ventral (V) and whole (W) prostate).
NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.

Conclusions for 407 performance with p,p'-DDE

189. The following conclusions are drawn from the updated TG 407 studies with DDE and from the comparison of these results with other DDE studies:

- The updated TG 407 results indicated that maximum tolerated dose was exceeded at the top DDE dose with the deaths of half of the males and several females in first study.
- The updated TG 407 results indicated that the pattern of effects observed in the two DDE studies were largely consistent across body weights, organ and tissue weights, histopathology and other parameters.
- The updated TG 407 detected clear effects due to DDE administration in the liver and in the thyroid, including changes in tissue weights, histopathology in the liver, and histopathology in the thyroid.
- Statistical changes in the thyroid hormone data in the males support the impact on the thyroid by DDE, but the data in females is tentative based on absolute trends, not statistically significant changes. Collectively, the organ weight, histopathology, and thyroid hormone results suggest

that DDE may accelerate the metabolism of circulating thyroid hormones, leading to increased stimulation of the thyroid gland and follicular hypertrophy.

- No changes in the tissue weights or the histopathology for any tissue in the male reproductive tract were observed with DDE in surviving male animals. The recorded changes in the high DDE dose males occurred in animals suffering premature mortality and may be attributed to the degeneration of the tissues before dissection and fixation.
- The pattern of effects produced in the TG 407 studies with DDE were consistent with studies showing DDE effects on the thyroid, but was not consistent with evidence in other studies indicating possible antiandrogenic effects on the male reproductive tract.
- It is concluded that the updated TG 407 successfully detected effects of DDE consistent with the endocrine mechanism of action on the thyroid, but did not detect effects consistent with antiandrogenic action.

Propylthiouracil (PTU)

Introduction

190. This section summarizes the results of the updated TG 407 studies with propylthiouracil (PTU) and compares these results with an unpublished 2-generation study conducted by the USEPA (67).

Background on PTU

191. PTU is a potent, classical thyroid toxicant that inhibits thyroid peroxidase. The action of thyroid hormone and its mechanistic pathway have been well characterized over the past 80 years (68)(69). PTU and related compounds have been used extensively in the investigation of thyroid toxicity and its different mechanisms (references in (69)(69) and see (70)(71)). The primary effects from PTU administration are a gross increase in thyroid weight and the appearance of a classical follicular hypertrophy due to stimulation by elevated TSH release from the pituitary in response to low levels of circulating T_3/T_4 hormones.

Description of PTU experiments

192. PTU was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in the first study and to a single group of six animals per sex per dose in the second study. The PTU was administered using distilled water as vehicle in laboratory 1 and corn oil as vehicle in laboratory 10 in three doses to both sexes in each study: a low (L) dose of 0.1 mg/kg body weight/day, an intermediate (I) or mid-dose of 1 mg/kg body weight/day, and a high (H) dose of 10 mg/kg body weight/day. The first study incorporated the full updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The functional observational battery and the motor activity assessment were not conducted in the second study. The individual data from the first study's Subgroups were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power.

Summary of the updated TG 407 results with PTU

193. The two updated TG 407 produced similar response profiles to PTU, and the PTU effect levels observed were similar. A clear pattern of thyroid responses was observed in both the female and the male. The NOEL could not be set as both studies reported thyroid histopathological changes at the lowest dose of 0.1 mg PTU/kg body weight/day used in the study. The pattern of tissue weight changes and histopathological changes observed in the updated TG 407 study were similar to those seen with PTU in two-generation reproductive toxicity study (67). The effect levels were within an order of magnitude

between the short-term TG 407 and the reproductive study for major organs and closer for the thyroid histopathological effects. Therefore, it is concluded that the updated TG 407 successfully detected effects of PTU consistent with the endocrine mechanism of action of a thyroid toxicant.

Mortality and body weights in PTU studies

194. No mortalities or abnormal clinical signs were observed in the animals at any PTU dose level in either study. Food consumption was significantly reduced in both sexes at the high dose in both studies. In the functional observational battery and the motor activity assessment from the first study, all of the few differences recorded in the functional battery were judged to be incidental in nature.

195. Body weights were significantly decreased in both sexes in both studies at the high PTU dose. The decreases were similar in magnitude in both studies and were slightly higher on a percentage basis in males than females in both studies (Table 71). Detailed body weight means and standard deviations for both studies, including the combined Subgroups and individual Subgroups for the first study, are in Annex 3.

Haematology and clinical chemistry results with PTU

196. In one case, the clinical chemistry values for female blood urea nitrogen were significantly changed in both studies (in agreement) (Table 72). There were no parameters where the statistically significant changes in one study were directionally in conflict with the other study. In some 15 other instances, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study (see Table 72). There were five instances where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 72). The differences in group sizes, n = 10 for the combined subgroups in the first study and n = 6 for the single group in the second study, should be recognized for its potential influence on these results. More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Table 71. Changes in body weights during 407 studies with PTU.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
1	M	- - +4.6%	- - -1.0%	↓↓ **-17.8%**	- - +6.3%	- - -0.1%	↓↓ **-18.1%**	- - +3.0%	- - -2.0%	↓↓ **-17.5%**
	F	- - -0.9%	- - -4.3%	↓↓ **-12.8%**	- - +1.6%	- - -1.0%	↓ **-10.1%**	- - -3.1%	- - -7.3%	↓↓ **-15.3%**
10	M	- - -0.9%	- - +0.7%	↓[a] **-16.8%**	Not Applicable			Not Applicable		
	F	- - +1.0%	- - +0.8%	↓[a] **-8.7%**						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight ($p<0.01$); - - : no statistically significant change.

[a] Laboratory 10 performed the statistical analysis for $p<0.05$ only, and not for $p<0.01$.

Table 72. Haematology and clinical chemistry results from 407 studies with PTU.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 1	Statistically Significant in Laboratory 10
Significant Parameters in Agreement in Both Studies: ↑ F Blood urea nitrogen	***Common measures, direction of percentage change similar, significant in first study only:*** ↓ F Haemoglobin concentration **-4.8** -4.7[a] ↓ F Haematocrit **-7.1** -4.1 ↓ F Erythrocyte counts **-6.4** -1.4 ↑ F Mean cell haemoglobin concentration **2.0** -0.5 ↑ M Albumin **4.2** 4.1 ↓ M Alkaline phosphatase **-19.5** -16.4 ↑ M Blood urea nitrogen **12.8** 16.2 ↑ M Creatinine **9.2** 1.8 ↑ F Sodium **0.8** 0.8 ↓ F Potassium **-12.5** -2.9	***Common measures, direction of percentage change similar, significant in second study only:*** ↓ F Platelet counts -4.4 **-17.3** ↑ F Albumin 2.0 **10.8** ↑ M Total protein 4.9 **10.6** ↑ M Cholesterol 18.2 **25.1** ↑ F Chloride 1.1 **2.3**
	Common measures, direction of percentage change differ, significant in first study only: ↓ M Platelet counts **-11.1** 5.5 ↓ M Alanine aminotransferase **-23.3** 12.2 ↓ F Alanine aminotransferase **-14.9** 49.5 ↑ F Creatinine **12.0** -3.8 ↓ M Potassium **-7.8** 9.2	***Common measures, direction of percentage change differ, significant in second study only:***
	Measurement performed in this lab only: ↑ F Reticulocyte count ↓ M ↓ F Triglycerides ↓ M Phosphorus	***Measurement performed in this lab only:***

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. The first study percentage is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Organ and tissue weights results with PTU treatment

197. There were numerous statistically significant changes in the absolute and relative weights of the major organs of both sexes in both PTU studies. These changes are attributed to the impact of the thyroid toxicity of PTU and the subsequent effects on overall growth of the young adult animals. In conjunction with the decreased body weights at the high PTU dose, there are decreases in the absolute weights of most major organs in both sexes including the liver, kidneys, heart, adrenals, spleen, and thymus. In many cases, the relative weights are also decreased. However, the relative weights of conserved organs in both sexes such as the brain and pituitary and the testes and accessory reproductive tissues in the male are significantly increased in one or both studies, reinforcing the interpretation that the underlying impact on the major organs is due to overall somatic growth (Tables 73A and B).

198. The absolute and relative thyroid weights were significantly increased at the mid- and high PTU doses in both sexes in the first study (Tables 73A and B). However, thyroid weights were not measured in the second study. The means and standard deviations of the absolute weights and the individual relative weights for both studies, including the combined Subgroups and individual Subgroups, are in Annex 5.

Histopathology findings with PTU treatment

199. Both studies reported follicular hypertrophy of the thyroid in both sexes at all doses of PTU (Table 74). The grade of severity increased dramatically with dose and involved all animals in the first study (Table 74). Although not in the original report, the results of the histopathological observations of the major organs and the thyroid were made available at the request of the Secretariat for this report. The second study described increasingly severe follicular hypertrophy in both sexes with increasing PTU dose (Table 74).

200. In the first study, significant findings in the pituitary of both sexes were reported at the mid- and the high PTU doses. Almost all animals were reported to have hyperplasia of the basophilic and chromophobe cells of the pituitary and reduced numbers of acidophilic cells.

201. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Thyroid hormone results with PTU treatment

202. The T_3, T_4, and TSH thyroid hormone values in the first study were concordant with the tissue weight and histopathological findings (Table 75). T_3 values in both sexes were significantly decreased at the high PTU dose. T_4 values in both sexes were significantly decreased at the mid- and high PTU doses, and also in females at the low PTU dose. TSH values in both sexes were significantly increased at the mid- and the high PTU doses. Thus, in the first study, the hormone values were generally equally sensitive as relative changes in the thyroid weight, but slightly less sensitive than histopathology in males.

203. In the second study, the hormone values appeared to be less sensitive than thyroid weights and histopathology, but due to variability (high coefficients of variation) apparently large differences in means did not reach statistical significance. T_3 and T_4 values were significantly decreased only at the high PTU dose in males. In females, absolute decreases were about 25% and 60%, respectively, which did not achieve statistical significance. The TSH values were significantly increased only at the high PTU dose in females, and, despite a 50% rise in absolute values, did not achieve statistical significance in males. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with PTU treatment

204. No significant changes were observed in either sperm numbers or morphology in either PTU study in the combined Subgroups (Table 76). Detailed means and standard deviations for sperm numbers and percentages for sperm morphology for the first study, including the combined Subgroups and individual Subgroups, are in Annex 8. However, the second study reported the sperm findings in graphical form, and absolute values for the means and standard deviations are not available.

Estrous cyclicity results with PTU treatment

205. All the female dose groups were judged to have exhibited normal oestrous cycle in the first study. In the second study, the number of days in metestrous was stated to be significantly decreased in females of the low dose group. However, due to the lack of a dose response relationship, this was assessed by the investigators as a spurious and incidental finding.

Comparison of 407 results from PTU treatment with data from chronic, reproductive, and developmental studies

206. A 2-generation reproductive toxicity study in rats has been conducted with PTU administered in the drinking water for the US Environmental Protection Agency using an updated battery of endpoints, including thyroid weights, histopathology, and thyroid hormones as well as subtle developmental endpoints such as anogenital distance, vaginal opening, and preputial separation (67). The F_0 generation was exposed for a total of 19-weeks. Estimated dosages for the F_0 generation were a low dose of 0.1 mg PTU/kg/d for both males and females; a mid- or intermediate dose of 0.2-0.5 mg PTU/kg body weight/day for males and 0.3-0.5 mg PTU/kg/d for females (indicated as a value of 0.4 in Table 77 for both sexes); and a high dose of 0.9-1.8 mg PTU/kg body weight/day for males (indicated as 1.1 in Table 77) and 1.1-1.7 mg PTU/kg body weight/day for females (indicated as 1.3 in Table 77). The selected findings in the two updated TG 407 studies with PTU and the selected reproductive study are compared in Table 77.

Table 73A. PTU-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
1	M	Liver			↓↓	▲		↓↓			↓
		Kidney			↓↓			↓↓			↓↓
		Heart			↓↓			↓↓			↓↓
		Adrenals			↓↓			↓			↓
		Spleen			↓↓			↓		▽	↓↓
		Thymus			↓↓			↓	▽	▽	↓
		Brain			↓↓						
		Thyroid		↑↑	↑↑	▽	↑↑	↑↑	▲	↑↑	↑↑
		Testes									
		Ventral Prostate									
		Seminal Vesicles		▽	↑	▽	▽		↑		↑
	F	Liver			↓↓			↓			↓↓
		Kidney			↓↓			↓↓			↓↓
		Heart			↓↓			↓			↓↓
		Brain									
		Adrenals			↓↓			↓			↓
		Spleen	↓		↓↓	↓	↓	↓↓			↓
		Thymus	▲	▽	▽	▲				▽	↓
		Pituitary									
		Thyroid		↑↑	↑↑	↑↑	↑↑	↑↑		↑↑	↑↑
		Uterus/cervix	↓↓	↓	↓↓	▽		▽	↓↓	▽	↓↓

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant.

[a] No indication in the final report that $p < 0.01$ was analyzed.

Table 73A continued. PTU-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
10	M	Liver			↓[a]	Not applicable			Not applicable		
		Kidney			↓[a]						
		Heart			↓[a]						
		Spleen			↓[a]						
		Lung			↓[a]						
		Thymus	↓[a]		↓[a]						
		Thyroid	Not measured								
	F	Liver			▽						
		Kidney			▽						
		Heart			↓[a]						
		Spleen			▽						
		Thymus			↓[a]						
		Thyroid	Not measured								

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose;↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant.

[a] No indication in the final report that $p < 0.01$ was analyzed.

Table 73B. PTU-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
1	M	Liver									
		Kidney			↓↓			↓			↓
		Heart		↓	↓↓		↓↓	↓↓			
		Adrenals			▽			▽			
		Spleen			▽			▽			▽
		Thymus			↓↓			▽	▽	▽	▽
		Brain			↑↑			↑↑			↑
		Thyroid		↑↑	↑↑	▽	↑↑	↑↑		↑↑	↑↑
		Testes			↑↑			↑↑			↑↑
		Ventral Prostate			↑			▲			▲
		Dorsolateral Prostate		▽	▲		▽				▲
		Seminal Vesicles			↑↑	▽	▽	▲	▲		↑↑
	F	Liver			↓↓						
		Kidney			↓↓						↓↓
		Heart			↓↓						
		Brain		↑	↑↑			↑			↑
		Adrenals			↓↓			▽			↓
		Spleen	▽		↓↓	▽	▽	↓↓		▲	↓
		Thymus	▽			▲				▽	↓
		Pituitary	▲	↑	↑↑	▲	▲	↑		↑	↑↑
		Thyroid		↑↑	↑↑	↑	↑↑	↑↑		↑↑	↑↑
		Uterus/cervix	↓↓			▽			↓↓		▽

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant.

Table 73B continued. PTU-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
10	M	Liver			↓[a]	Not applicable			Not applicable		
		Kidney			↓[a]						
		Heart									
		Spleen			↓[a]						
		Lung	↓[a]								
		Thymus	↓[a]	▽	↓[a]						
		Thyroid	Not measured								
	F	Liver		▽							
		Kidney	▽								
		Heart		▽	▽						
		Spleen	▲		▽						
		Thymus	▲	▽	↓[a]						
		Thyroid	Not measured								

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight ($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant.

[a] No indication in the final report that $p < 0.01$ was analyzed.

Table 74. Significant histopathological findings after PTU treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
1	M	Thyroid	Follicular hypertrophy Grade 1	4/10			Not applicable			Not applicable		
			Grade 2	2/10								
			Grade 3		7/10							
			Grade 4		3/10	10/10						
		Pituitary	hyperplasia chromophobe and basophilic cells - loss acidophilic		9/10	10/10						
	F	Thyroid	Follicular hypertrophy Grade 1	1/10								
			Grade 2	1/10	1/10							
			Grade 3		8/10	2/10						
			Grade 4		1/10	8/10						
		Pituitary	hyperplasia chromophobe and basophilic cells - loss acidophilic		10/10	10/10						
10	M	Thyroid	Follicular hypertrophy Grade 1	3/6	5/6		Not applicable			Not applicable		
			Grade 2		1/6	3/6						
			Grade 3			3/6						
	F	Thyroid	Follicular hypertrophy Grade 1	2/6	3/6							
			Grade 2		1/6	4/6						
			Grade 3			2/6						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.

Not applicable – laboratory 1 reported combined Subgroup and individual animal data, and histopathological data was not provided by individual Subgroup; laboratory 10 used a single group of 6 animals.

Table 75. Thyroid hormone results after PTU treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
1	M	T_3			↓↓			↓↓			
		T_4		↓↓	↓↓		↓↓	↓↓			
		TSH		↑↑	↑↑	↑	↑↑	↑↑		↑↑	↑↑
	F	T_3			↓↓			↓↓			↓
		T_4	↓	↓↓	↓↓		↓↓	↓↓		↓	↓↓
		TSH		↑↑	↑↑		↑↑	↑↑			↑↑
10	M	T_3			↓	Not applicable			Not applicable		
		T_4			↓						
		TSH									
	F	T_3									
		T_4									
		TSH			↑						

L – low dose group; I – intermediate or mid-dose group; H – high dose group. M – male; F – female; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).
Not applicable – the study did not perform a design with individual Subgroups, so those data are not available.

Table 76. Statistically significant changes in sperm parameters after PTU treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
1	**Sperm Count**	ND	ND		ND	ND		ND	ND	
	Abnormalities	ND	ND		ND	ND		ND	ND	
10	**Sperm Count**				Not applicable, did not perform Subgroups			Not applicable, did not perform Subgroups		
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group; Not applicable – Analyses of individual Subgroups were not performed or the study was not conduced with individual Subgroups. ND – No analysis performed on these individual Subgroups.

Table 77. Comparison of updated TG 407 PTU results with a 2-generation reproductive study.

Parameter	PTU – updated TG 407 (LOEL mg/kg/d)				PTU – Reproduction Study (67) (LOEL mg/kg/d)			
	Laboratory 1		Laboratory 10		F0		F1 and F2	
	Male	Female	Male	Female	Male	Female	Male	Female
Mortality							↑ 1.1	↑ 1.4
Body weight	↓ 10	↓ 10	↓ 10	↓ 10	↓ 1.1	↓ 1.4	↓ 1.1	↓ 1.4
Haematology and clinical chemistry								
Blood urea nitrogen		↑ 10		↑ 10				
Organ and Tissue Weights	(relative)				(relative)			
Liver		↓ 10	↓ 10	↓ 10				
Kidney	↓ 10	↓ 10	↓ 10	↓ 10				
Heart	↓ 1	↓ 10						
Adrenal		↓ 10						
Spleen		↓ 10	↓ 10				↓ 1.1	↓ 1.4
Thymus	↓ 10		↓ 10	↓ 10			↓ 1.1	↓ 1.4
Brain	↑ 10	↑ 1			↑ 1.1		↑ 1.1	↑ 1.4
Pituitary		↑ 1			↑ 0.4			
Thyroid	↑ 1	↑ 1	NR	NR	↑ 0.4	↑ 0.1	↑ 0.4	↑ 0.1
Testes	↑ 10				↑ 1.1			
Ventral Prostate/Seminal vesicles	↑ 10				↑ 1.1			
Histopathological Findings								
Thyroid – centilob. hypertrophy	0.1	0.1	0.1	0.1	0.4	0.4	0.4	0.4
Pituitary – hypertrophy basoph. And chromophobe cells, reduced number acidophilic cells	1	1						
Testes – germinal epithelial degeneration					0.1[a]		0.1[a]	
Sperm parameters								
Estrous cycle						↓ 1.4		
Thyroid Hormones								
T_3	↓ 10	↓ 10	↓ 10					
T_4	↓ 1	↓ 1	↓ 10	↓ 10	↓ 0.4	~ 0.4, 1.4[b]	↓ 0.4	↓ 0.4
TSH	↑ 1	↑ 1		↑ 10	↑ 0.4	↑ 0.4	↑ 0.4	↑ 0.4
Reproductive and developmental								
Pup Mortality	Not applicable						↑ 1.4	↑ 1.4
Pups per litter	Not applicable					↓ 1.4		↓ 1.4
Vaginal opening (female) and preputial separation (male)	Not applicable						↑ 0.4	↑ 0.4

[a] The 28-day 407 exposure may not be sufficient for the germinal epithelial lesion to manifest itself, as these lesions were found after 19-weeks (over 130 days) in the reproductive study.
NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.
[b] The F_0 females absolute values were decreased at these doses, but did not achieve statistical significance ($p<0.05$)

Conclusions for the updated TG 407 performance with PTU treatment

207. The following conclusions are drawn from the updated TG 407 studies with PTU and from the comparison of these results with several aromatase inhibitor studies:

- The updated TG 407 results indicated that the pattern of effects observed in the two PTU studies were largely consistent across body weights, organ and tissue weights, histopathology and other parameters. Although, the second study did not report the thyroid weights, the histopathological findings were consistent and in agreement with the first study.
- The updated TG 407 detected clear effects due to PTU administration on the thyroid, and this was reflected in retardation of somatic growth.
- The updated TG 407 results also indicated that the dose responses for PTU appeared to be similar for thyroid weights and histopathology.
- The thyroid hormone values were slightly less sensitive than tissue weights and histopathology. The statistical findings in the first study were concordant, and the absolute trends and statistical findings in the second study were concordant taking into account some variability in the measures in this laboratory as well as the smaller group size in achieving statistical significance.
- The available reproductive and development studies for PTU are highly concordant with the updated TG 407 studies. The pattern of effects on body weight and major organs were consistent and the LOELs are within an order of magnitude. The effects on thyroid weights and histopathology were identical and the correspondence of the histological LOELs between the updated TG 407 studies and the reproductive study was excellent.
- The thyroid hormones also were not more sensitive than traditional endpoints in the reproductive study. T_3 failed to respond statistically at any dose and T_4 decrease in the F_0 females did not achieve statistical significance at any dose.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of PTU consistent with the endocrine mechanism of action of a thyroid toxicant.

l-Thyroxine

Introduction

208. This section summarizes the results of the updated TG 407 studies with l-thyroxine (THY). No significant toxicity studies were found for comparison to the updated TG 407 studies.

Background on l-Thyroxine

209. l-Thyroxine is the natural T_4 product of the thyroid (tetraiodothyronine). While it is an agonist for the thyroid receptor isoforms, it is converted primarily in the peripheral tissues into T_3, which is a more potent agonist, and T_4 is thus typically considered the thyroid prohormone. Both hormones circulate in the bloodstream of the rat bound to low affinity hydrophobic sites such as on (pre)albumin. In humans, there is also a specific high affinity binding protein for T_4 and T_3, which reduces the availability for and therefore the rate of metabolism in humans versus the rat (68)(72)(73). Hence, the thyroid system and its feedback in rats is considered to be more susceptible to metabolic disturbance (72)(73)(74).

Description of l-Thyroxine experiments

210. THY was administered by gavage to two Subgroups of 5 animals per sex per dose for 28 days in the first study and to a single group of 10 animals per sex per dose in the second study. THY was administered using a vehicle of 0.5% CMC-Na aqueous solution with 0.1% Tween 80 in both studies. Three THY doses were administered to both sexes in each study: a low (L) dose of 0.01 mg/kg body weight/day, an intermediate (I) or mid-dose of 0.1 mg/kg body weight/day, and a high (H) dose of 1 mg/kg body weight/day. Both studies incorporated the complete updated TG 407 protocol, including the functional observational battery and the motor activity assessment. The individual data from the individual Subgroups in the first study were pooled into an overall combined Subgroup of ten animals per sex per dose to assess the impact of increased group size on the statistical power; the second study only used a single group size of 10 animals per sex.

Summary of the updated TG 407 results with l-Thyroxine

211. The two updated TG 407 produced similar response profiles to THY, and the THY effect levels observed were similar. The patterns of tissue and organ weight changes were consistent with an increased growth stimulus that would be expected from higher circulating THY levels. The findings of increased haematopoietic activity and cardiac myopathy are outcomes associated with hyperthyroidism (75). The histopathological changes in the thyroid are diagnostic of high circulating THY levels. The atrophy would be expected due to decreased pituitary stimulation. The thyroid hormone data were consistent with the profile and indicated significantly decreased TSH levels, but T_3 and T_4 were significantly elevated in only one study. Therefore, it is concluded that the updated TG 407 successfully detected effects of THY consistent with the endocrine mechanism of action of a thyroid toxicant.

Mortality and body weights in l-Thyroxine studies

212. No treatment-related mortalities or clinical signs were observed during either THY study. Food consumption was significantly increased in both sexes at the THY high dose in both studies. The functional observational battery and the motor activity assessments were negative in the first study at all doses. In the second study, a statistically significant decrease in mean vertical activity (total rearings) in the high dose males was judged to be spurious as there were no corroborating observations.

213. No significant changes in body weights were observed in males of either study at any THY dose. Female body weights were statistically increased in the second study (Table 78). Detailed body weight means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 3.

Haematology and clinical chemistry results with l-Thyroxine

214. In eight cases, haematological and clinical chemistry parameters were significantly changed in the combined Subgroups in both studies (studies in agreement) (see Table 79). In the case of male prothrombin times, the statistically significant changes in one study were directionally in conflict with the other study. For nine other parameters, statistical changes were observed in one study with similar directional changes occurring in the absolute values in the other study (see Table 79). There were eight instances where statistical significance was achieved in one study, but the absolute trends in the other study were in the opposite direction or unchanged (see Table 79). More detailed summaries of the haematological and clinical chemistry findings for both studies including the combined Subgroups and individual Subgroups are in Annex 4.

Organ and tissue weights results with l-Thyroxine treatment

215. Both THY studies reported similar profiles of absolute and relative weight changes. In both sexes of both studies, the absolute and relative kidney, heart, adrenal, and spleen weights were significantly increased at the high THY dose (Tables 80A and B), and in the females of the second study, the liver and heart relative weights were increased at the mid-THY dose. The absolute and relative weights of the liver were also significantly increased in both female high dose groups and the male high THY dose group of the second study, but not the first (Table 80B).

216. The relative thyroid weights where changed in the combined Subgroups only in the high THY dose females of the first study, where a significant decrease was observed (Table 80B). Detailed absolute organ and tissue weight means and standard deviations and relative organ and tissue weights for both studies including the combined Subgroups and individual Subgroups are in Annex 5.

Histopathology findings with l-Thyroxine treatment

217. The histopathological data corroborated the changes in organ weights. Both studies observed evidence of cardiac hypertrophy; the first study observed these changes in the high THY dose groups of both sexes and the second study observed these changes in the high THY dose male group (Table 81). Liver observations included diffuse hypertrophy and increased mitotic figures in both sexes in the first study at the high THY dose and decreased glycogen deposition at all THY doses and described hepatocellular dissociation at the high THY dose in both sexes in the second study (Table 81). Both studies described increased hematopoietic cells in the male spleen at the high THY dose, and the first study described increased hematopoietic activity in females in both the spleen and bone marrow at the high THY dose (Table 81). The second study described cortical cell hypertrophy in the female adrenals at the high dose. The first study also described hypertrophy of the pars distalis of the male pituitary at all THY doses (Table 81).

218. In the thyroid, both studies reported atrophic changes in the follicular epithelium that was described as a flattening of the cells (Table 81). This was reported at the mid- and high THY dose of both sexes in both studies. The dose response curve was slightly lower in the females of the second study, where two individuals out of 10 were reported to have follicular hypertrophy (Table 81). Collectively, the data suggest a hyperthyroid state in the animals with atrophy of the thyroid itself as a consequence of an increased negative feedback to higher centers and a corresponding reduced activity of the hypothalamic–pituitary-thyroid regulatory axis. More detailed histopathological summaries for both studies including the combined Subgroups and individual Subgroups are in Annex 6.

Table 78. Changes in body weights during 407 studies with l-Thyroxine.

Lab	Sex	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
		L	I	H	L	I	H	L	I	H
9	M	- - +0.8%	- - +3.6%	- - -2.4%	- - -1.2%	- - +4.7%	- - -3.1%	- - +2.8%	- - +2.6%	- - -1.7%
	F	- - -0.9%	- - +1.0%	- - +1.4%	- - -8.6%	- - -2.5%	- - -2.1%	- - +7.3%	- - +4.8%	- - +5.2%
13	M	- - -0.4%	- - +3.5%	- - +1.7%	Not applicable			Not applicable		
	F	- - +0.3%	- - +2.9%	↑ +6.8%						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in body weight ($p<0.05$); ↓↓ : statistically significant decrease in body weight ($p<0.01$); ↑ : statistically significant increase in body weight ($p<0.05$); ↑↑ : statistically significant increase in body weight($p<0.01$); - - : no statistically significant change.

[c] Although laboratory 13 did not run the study using two Subgroups, the increase in female body weights were consistent and similar in their high degree: 34.3 % at the low dose, 41.2 % at the intermediate or mid-dose, and 31.0 % at the high dose.

Table 79. Haematology and clinical chemistry results from 407 studies with l-Thyroxine.

Parameters Statistically Significant in both Studies Results in Agreement or Conflict	Statistically Significant in Laboratory 9	Statistically Significant in Laboratory 13
Significant Parameters in Agreement in Both Studies: ↓ M Mean cell haemoglobin ↓ M ↓ F Mean cell haemoglobin concentration ↓ F Total protein ↓ F Albumin ↓ F Blood urea nitrogen ↑ M ↑ F Phosphorus **Significant Parameters in Conflict Between Studies**: First laboratory ↓ M Prothrombin times Second laboratory ↑ M Prothrombin times	***Common measures, direction of percentage change similar, significant in first study only:*** ↓ M Cholesterol **-19.9** -12.4[a]	***Common measures, direction of percentage change similar, significant in second study only:*** ↑ F Alanine aminotransferase 23.4 **49.6** ↓ F Aspartate aminotransferase 10.7 **24.8** ↓ M Total protein -1.9 **-10.3** ↑ M Alkaline Phosphatase 7.6 **29.3** ↓ M Glucose -6.5 **-7.7** ↑ M ↑ F Bilirubin Cannot compare values ↓ M Calcium -0.4 **-1.8**
	Common measures, direction of percentage change differ, significant in first study only: ↑ F Haematocrit **5.6** -0.1 ↑ M Erythrocyte counts **4.9** -0.5 ↓ M Triglycerides **-58.6** 11.3 ↑ M Alanine aminotransferase **27.8** -30.3	***Common measures, direction of percentage change differ, significant in second study only:*** ↓ M Haemoglobin concentration 0.8 **-5.0** ↓ M Creatinine 0.0 **-16.7** ↓ F Creatinine -1.9 **-16.5** ↑ M Potassium -1.1 **8.0**
	Measurement performed in this lab only: ↑ M ↑ F Reticulocyte ratio	***Measurement performed in this lab only:*** ↑ F % Large unstained cells ↓ M Globulin

[a] When comparing laboratory results, the increase or decrease in a parameters is given as a percentage of the control in the high dose combined Subgroups. The first study percentage is given on the left and the second study is on the right. Where a parameter reached statistical significance, the percentage value is shown in bold; if not, the percentage value is in normal typeface.

Thyroid hormone results with l-Thyroxine treatment

219. The thyroid hormones provided evidence for effects from THY administration, but there was equivocal data in both studies. The TSH values were significantly decreased in both sexes in the first study and the males of the second study at the high THY dose (Table 82). The female TSH values in the second study were essentially unchanged. In the first study, in contrast to the TSH observations, T_3 and T_4 thyroid hormone values were erratic, e.g. changes occurred such as statistically significant decreases at the mid-dose. Absolute T_3 and T_4 values where higher in males, but again essentially unchanged in females. In the second study, however, both T_3 and T_4 thyroid hormone values were significantly increased in both sexes

at the mid- and high THY doses. Detailed T_3, T_4, and TSH means and standard deviations for both studies including the combined Subgroups and individual Subgroups are in Annex 7.

Sperm analyses results with l-Thyroxine treatment

220. No significant changes were observed in either sperm numbers or morphology in either THY study in the combined Subgroups (Table 83). Detailed means and standard deviations for sperm numbers and percentages for sperm morphology from both studies including the combined Subgroups and individual Subgroups are in Annex 8.

Estrous cyclicity results with l-Thyroxine treatment

221. The females were judged to have exhibited normal oestrous cycle in both studies according to the final reports submitted.

Comparison of 407 results from l-Thyroxine treatment

222. As no significant toxicity studies were found for comparison to the updated TG 407 studies, only the consistency of the two updated TG 407 studies is found in Table 84. The pattern of effects is relatively clear for hyperthyroid activity: the major organs are larger, there is evidence for cardiac myopathy, haematopoietic activity is increased, histopathology of the thyroid indicates atrophy, and there are changes in thyroid hormones. Collectively, these data support the overall conclusions that the studies are consistent and reproducible and indicate a pattern of hyperthyroid effects.

Table 80A. l-Thyroxine-treated absolute organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
9	M	Liver									
		Kidney			↑↑			↑↑			↑↑
		Heart			↑↑		▲	↑↑			↑↑
		Adrenal			↑↑				▲	▲	▲
		Spleen		▲	↑↑		▲	↑↑		▲	↑↑
		Thyroid	↑			▲		▽	▲		
	F	Liver			↑			▲			▲
		Kidney			↑↑	▽		↑			↑
		Heart			↑↑			↑↑			
		Adrenal		↑	↑↑		▲	↑↑			↑
		Spleen			↑↑			↑↑			↑↑
		Thyroid			▽			▽	▽	▽	▽
13	M	Liver				Not applicable			Not applicable		
		Kidney		↑	↑↑						
		Heart			↑↑						
		Adrenal			↑↑						
		Spleen		↑	↑↑						
		Thyroid	▽	▽							
	F	Liver		▲	↑↑						
		Kidney		▲	↑↑						
		Heart		↑	↑↑						
		Adrenal	▲	↑	↑↑						
		Spleen			↑↑						
		Thyroid	▲		▲						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant. Not applicable – this study was not conducted with individual Subgroups.

Table 80B. l-Thyroxine-treated relative organ and tissue weights - statistical significance and trends.

Lab	Sex	Organ or Tissue	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
			L	I	H	L	I	H	L	I	H
9	M	Liver									
		Kidney			↑↑			↑↑			↑↑
		Heart			↑↑			↑↑			↑↑
		Adrenal			↑↑			↑	▲		▲
		Spleen		▲	↑↑	▲	▲	↑↑			↑↑
		Thyroid	↑			▲	▽	▽	▲		
	F	Liver			↑↑			↑			▲
		Kidney			↑↑			↑↑	▽		↑↑
		Heart			↑↑			↑↑	▽		▲
		Adrenals		▲	↑↑		▲	↑↑	▽		▲
		Spleen			↑↑			↑↑			
		Thyroid			↓			▽	▽	▽	↓
13	M	Liver			↑↑	Not applicable			Not applicable		
		Kidney			↑↑						
		Heart			↑↑						
		Adrenal			↑↑						
		Spleen		▲	↑↑						
		Thyroid	▽	▽							
	F	Liver		↑	↑↑						
		Kidney			↑↑						
		Heart		↑	↑↑						
		Adrenal	▲	▲	↑↑						
		Spleen			↑↑						
		Thyroid	▲		▲						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ : statistically significant decrease in tissue weight ($p<0.05$); ↓↓ : statistically significant decrease in tissue weight ($p<0.01$); ↑ : statistically significant increase in tissue weight ($p<0.05$); ↑↑ : statistically significant increase in tissue weight($p<0.01$); ▲: 10% or greater increase, but not statistically significant; ▽: 10% or greater decrease, but not statistically significant. Not applicable – this study was not conducted with individual Subgroups.

Table 81. Significant histopathological findings after l-Thyroxine treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
9	M	**Thyroid**	Atrophy		3/10	10/10	Not applicable			Not applicable		
		Heart	Myocardial fiber hypertrophy			9/10						
			Myocardial degeneration/fibrosis			6/10						
		Spleen	Increased hematopoietic cells			4/10						
		Liver	Hypertrophy, hepatocyte, diffuse			10/10						
			Increase mitotic figure, hepatocyte			10/10						
			Hypertrophy, Kupffer cell			8/10						
			Extramedullary hematopoiesis			4/10						
		Kidney	Hyaline droplets			5/10						
		Mammary	Dilatation		1/10	5/10						
	F	**Thyroid**	Atrophy		8/10	10/10						
		Heart	Myocardial fiber hypertrophy			7/10						
		Spleen	Increased hematopoietic cells			3/10						
		Bone marrow	Increased hematopoietic cells			5/10						
		Liver	Hypertrophy, hepatocyte, diffuse			3/10						
			Increase mitotic figure, hepatocyte			5/10						
			Hypertrophy, Kupffer cell			2/10						
		Adrenal	Cortical cell hypertrophy			4/10						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose; Not applicable – the histopathology findings were reported only as combined Subgroups, not as individual Subgroups. Statistical analyses were not reported for the histopathological findings, although other data such as organ weights was analyzed by both individual Subgroups and combined Subgroups.

Table 81.continued Significant histopathological findings after l-Thyroxine treatment.

Lab	Sex	Organ or Tissue	Histopathological Finding	Combined Subgroups n = 10			Subgroup A n = 5			Subgroup B n = 5		
				L	I	H	L	I	H	L	I	H
13	M	**Thyroid**	Follicular cell atrophy		10/10	10/10	Not applicable			Not applicable		
		Heart	Inflammation with fibrosis			7/10						
		Liver	Glycogen deposition (decreasing trend)	9/10	7/10	0/10						
			Hepatocellular dissociation			10/10						
		Pituitary	Hypertrophy of pars distalis[1]	7/10	6/10	9/10						
		Spleen	Mean grade – extramedullary hematopoiesis	1.8	2.1	3.0						
	F	**Thyroid**	Follicular cell atrophy	2/10	10/10	10/10						
		Liver	Glycogen deposition (decreasing trend)	6/10	5/10	0/10						
			Hepatocellular dissociation			7/10						

M: male; F: female; L: low dose, I: intermediate or mid-dose; H: high dose.
[1] The control values for the pituitary were 5/10,

Table 82. Thyroid hormone results after l-Thyroxine treatment.

Lab	Sex	Thyroid Parameter	Combined Subgroups			Subgroup A			Subgroup B		
			L	I	H	L	I	H	L	I	H
9	M	T_3		↓↓			↓			↓↓	
		T_4									
		TSH			↓↓	↓↓	↓↓	↓↓			↓↓
	F	T_3		↓							
		T_4		↑							
		TSH			↓↓			↓			↓↓
13	M	T_3		↑	↑↑	Not applicable			Not applicable		
		T_4		↑↑	↑↑						
		TSH			↓						
	F	T_3		↑	↑↑						
		T_4		↑↑	↑↑						
		TSH									

M – male; F – female; L: low dose, I: intermediate or mid-dose; H: high dose; ↓ - statistically significant decrease in thyroid hormone levels ($p<0.05$); ↓↓ - statistically significant decrease in thyroid hormone levels ($p<0.01$); ↑ - statistically significant increase in thyroid hormone levels ($p<0.05$); ↑↑ - statistically significant increase in thyroid hormone levels ($p<0.01$).
Not applicable – the study did not perform a design with individual Subgroups, so those data are not available.

Table 83. Statistically significant changes in sperm parameters after l-Thyroxine treatment.

Lab	Parameter	Combined Subgroups			Subgroup A			Subgroup B		
		L	I	H	L	I	H	L	I	H
9	**Sperm Count**	ND	ND		ND	ND		ND	ND	
	Abnormalities	ND	ND		ND	ND		ND	ND	
13	**Sperm Count**				Not applicable, did not perform Subgroups			Not applicable, did not perform Subgroups		
	Abnormalities									

L – low dose group; I – intermediate or mid-dose group; H – high dose group; Not applicable – Analyses of individual Subgroups were not performed or the study was not conduced with individual Subgroups. ND – No analysis performed on these groups.

Table 84. Comparison of updated TG 407 l-Thyroxine results.

Parameter	THY – updated TG 407 (LOEL mg/kg/d)			
	Laboratory 9		Laboratory 13	
	Male	Female	Male	Female
Body weight				
Haematology and clinical chemistry				
Mean cell haemoglobin	↓ 1		↓ 1	
Mean cell haemoglobin conc.	↓ 1	↓ 1	↓ 1	↓ 1
Total protein		↓ 1		↓ 1
Blood urea nitrogen		↓ 1		↓ 1
Organ and Tissue Weights (relative)				
Liver		↑ 1	↑ 1	↑ 0.1
Kidney	↑ 1	↑ 1	↑ 1	↑ 1
Heart	↑ 1	↑ 1	↑ 1	↑ 0.1
Adrenal	↑ 1	↑ 1	↑ 1	↑ 1
Spleen	↑ 1	↑ 1	↑ 1	↑ 1
Thyroid		↓ 1		
Histopathological Findings				
Liver – centrilobular hypertrophy	1	1	1	1
Kidney -	1			
Heart – epithelial changes	1	1	1	
Adrenal – cortical changes		1		
Spleen – increased	1	1	1	
Bone marrow –atrophic changes in epith. and epithelial glands		1		
Thyroid – follicular atrophy	0.1	0.1	0.1	Eq 0.01/ 0.1
Thyroid Hormones				
T_3			↑ 0.1	↑ 0.1
T_4			↑ 0.1	↑ 0.1
TSH	↓ 1	↓ 1	↓ 1	

[a] In the 407 studies these were individual tissues (ventral (V) and whole (W) prostate).
NR – no histopathological results or observations for these tissues reported; Eq – equivocal, low rate of individuals affected.

Conclusions for the updated TG 407 performance with l-Thyroxine

223. The following conclusions are drawn from the updated TG 407 studies with THY:

- The updated TG 407 results indicated that the pattern of effects observed in the two THY studies were largely consistent across organ and tissue weights, histopathology and thyroid hormone parameters.
- The updated TG 407 detected clear effects due to THY administration in both studies. These included atrophic changes in the thyroids and characteristic hyperthyroid changes in the heart and spleen.
- The updated TG 407 results also indicated that the dose responses and severity of the effects were similar in the two THY studies.

- The pattern of expected changes in the thyroid hormone values was incomplete in both studies. The T_3 and T_4 increases were not observed in the first study, and the TSH declines were mild in the second study males and absent in the females.
- Therefore, it is concluded that the updated TG 407 successfully detected effects of THY consistent with the endocrine mechanism of action of an agent causing a pattern of hyperthyroidism in target tissues, causing negative feedback on the pituitary, and resulting in decreased activity or atrophy in the thyroid itself.

PERFORMANCE OF THE ADDED UPDATE ENDPOINTS FOR TG 407

224. This section reviews the performance of the added update endpoints in the TG 407 studies. The section evaluation and discussion is focused on (anti)oestrogen, (anti)androgen, and (anti)thyroid mechanisms, as detecting these mechanisms led to the consideration of the updates. The section concludes with recommendations for the inclusion of certain updates in a revised TG 407 guideline due to their apparent benefit in identifying potential endocrine modulators, the exclusion of other updates based on their lack of apparent benefit, and for possible continued work to refine and to improve certain updates to increase their benefits. In addition, several recommendations are made to potentially improve the conduct and consistency of histopathological examinations of endocrine target tissues.

225. The update endpoints are considered together with applicable current endpoints. As seen in Table 2, several tissues that are potential targets of endocrine modulation were weighed or histopathology examinations conducted in the current TG407. For example, histopathology of the uterus and gonads (testes and ovaries) are performed in the current TG 407. Therefore, both updated and current endpoints are considered together as part of an overall battery to enable the detection or 'flagging' of test substances that potentially modulate endocrine endpoints.

226. The review in this section addresses current TG 407 endpoints and proposed updates as follows:

- Absolute and relative tissue weights of the male reproductive tract, including the testes, the epididymides, the prostate whether whole, ventral, or dorsolateral, and the seminal vesicles and coagulating glands.
- Absolute and relative tissue weights of the female reproductive tract, including the ovaries and uterus.
- Absolute and relative pituitary weights and histopathology of the pituitary.
- Absolute and relative thyroid weights and histopathology of the thyroid.
- Absolute and relative adrenal weights and histopathology of the adrenals.
- Histopathology of the male and female reproductive tract tissues.
- Histopathology of the male and female mammary glands.
- Spermatology endpoints, including sperm count and morphology.
- Circulating primary thyroid hormones, including T_3, T_4, and TSH.
- Estrous cycle evaluation, using daily vaginal smears and an integrated histopathological evaluation of the female reproductive tract.

Absolute and relative tissue weights of the male reproductive tract

227. The absolute and relative tissues weights of the male reproductive tract were relevant and useful endpoints in the updated TG 407 studies. The absolute and/or relative weights of one or more tissues of the male reproductive tract were significantly changed in both studies with five test substances: EE, TAM,

CGS, MT, and FLU (Table 85). The seminal vesicles and coagulating glands as well as the separated ventral and dorsolateral prostate were often the most sensitive tissues for changes in absolute and relative weights in the male (see Tables 15A-B for EE, 37A-B for TAM, 44A-B for CGS, 51A-B for MT, and 58A-B for FLU. In the case of NP, decreases in the whole and ventral prostate were observed in the study that employed high NP doses. These findings are consistent with the plausible mechanism of action and previous data for the test substances:

- In EE-treated males, the absolute and relative decreases in several accessory tissues in the male reproductive tract are consistent with negative feedback inhibition on LH secretion and subsequent decreased testosterone production by Leydig cells.
- In TAM-treated males, a decrease in male tissue weights may be considered plausible either where TAM acts directly as a weak oestrogen at the tissue level, or to negative feedback on the pituitary with decreased secretion of LH/FSH and subsequent decreased testosterone production by Leydig cells.
- In CGS-treated males, previous data with aromatase inhibitors have shown decreases in the weights of male sex accessory tissues and parallel decreases in circulating levels of testosterone due to apparent disturbances in steroid synthesis beyond specific aromatase inhibition (44)(45).
- In MT-treated males, additional hormonal analyses in laboratory 3 indicated sharply reduced LH levels consistent with negative feedback that resulted in Leydig cell atrophy, reduced sperm production, decreased testicular and epididymal weights. Direct androgenic stimulation of the prostate weight is plausible as observed at the high doses. Histological correlates for oligospermia were present in both the testes and epididymides.
- In FLU-treated males, direct antiandrogen action is expected to lead to a decrease in both absolute and relative tissue weights of the target tissues by blocking the action of testosterone and dihydrotestosterone at the androgen receptor.

228. There was only one chemical where expected changes in the weights of male reproductive tract tissues did not occur. No statistically significant changes in male reproductive tract tissue weights were observed for the putative antiandrogen, *p,p'*-DDE (Tables 65 A-B) at sustained doses of 150 mg/kg/d in study 1 and 100 mg/kg/d in study 2. In contrast to these findings in the intact, sexually mature males of the TG 407, some modest antiandrogenic effects have been observed at these approximate doses with exposures *in utero* (62)(63)(64), before and during pubertal development (62), and in surgically castrated males absent homeostatic feedback in the Hershberger bioassay (56). Additionally, work in young adult males has indicated possible strain differences in sensitivity. DDE was positive as antiandrogen in young, sexually mature Long Evans male rats, but not in Sprague Dawley rats (65). As Sprague Dawley rats were used in both TG 407 studies (Table 4), this may contribute to the lack of observed activity.

229. In conclusion, the additional male reproductive tract tissue weights provided clear benefit for the detection of possible endocrine modulation. Therefore, these endpoints are recommended for inclusion in an updated TG 407 guideline. The incorporation of the dissection and separate weighing of both the ventral and dorsolateral lobes of the prostate is also recommended. Given the DDE results, the ability of the updated TG 407 to detect very weak antiandrogens may need continuing investigation.

Absolute and relative tissue weights of the female reproductive tract

230. The absolute and relative tissues weights of the female reproductive tract were relevant and useful endpoints in the updated TG 407 studies. The absolute and/or relative weights of the ovaries and uterus were significantly changed in both studies with four test substances: EE, TAM, CGS, and MT (Table 86). Significant changes suggestive of oestrogenic activity were observed in one study with GN, but were not observed in the second study. These findings are consistent with the plausible mechanism of and previous data for the test substances:

- In EE-treated females, the ovarian absolute and relative weights were significantly decreased in one study consistent with EE inhibition of LH/FSH secretion by feedback inhibition on the hypothalamic GnRH production. Uterine weights were increased consistent with direct EE action via the estrogen receptor on the uterus. The weights were consistent with histological correlates.
- In TAM-treated females, ovarian and uterine weights were decreased at high TAM doses. Histological correlates indicated stimulation of ovarian follicular interstitial cells and evidence of partial agonist activity of tamoxifen in the uterus with hyperplasia evident in the epithelial cells of that tissue.
- In CGS-treated females, ovarian weights were increased and uterine weights were decreased with histological correlates in both tissues as well as the vagina.
- In MT treated females, methyl testosterone is aromatisable to oestrogen so that the decrease in ovarian weights secondary to reduced pituitary LH secretion and increases observed in uterine weights along with the histological correlates support the plausibility of the observations.

231. In conclusion, the female reproductive tract tissue weights provided some benefits for the detection of possible endocrine modulation. Therefore, it is recommended that these endpoints be included in an updated TG 407 guideline. However, these TG 407 endpoints do not appear to be as sensitive as the uterine weights in the uterotrophic bioassay. The young adult females are intact and, with the hypothalamic-pituitary-ovarian feedback loop established, they can respond to test substances with some degree of adaptation and compensation. Further, they are cycling and, even though the attempt was made to necropsy during a particular phase of the oestrous cycle, the tissue weights of the female reproductive tract should then be more variable, increasing the CVs of the weights and decreasing their power. In the event that female animals would be sacrificed at a specific time point (e.g., day 29) rather than at a particular stage of the oestrus cycle, greater variability might be expected.

Absolute and relative pituitary tissue weights and pituitary histopathology

232. The absolute and/or relative pituitary weights were significantly changed in both studies only in EE-treated males (Tables 85 and 86). Significant changes in absolute and/or relative pituitary weights were observed in one, but not the other, laboratory with several test substances. In part, these findings may be attributed to the difficulty in dissecting the pituitary, trimming after fixation, and the small absolute weight of the tissue. This technical difficulty suggests a careful review of the plausibility of the results:

- In EE-treated males, the increase in relative weight is correlated with decreased body weight. This contrasts with a possible pituitary weight decrease due to negative feedback from EE and a subsequent decreased secretion of LH/FSH by the pituitary, although such a mechanism would impact only a subset of the cells. Histological correlates in the pituitary were absent.
- In TAM-treated males from laboratory 3, an absolute decrease in pituitary weight was not confirmed relative to body weight. A decrease may be considered to be plausible where TAM acts as an weak oestrogen in males rats due to negative feedback from TAM and a subsequent decreased secretion of LH/FSH by the pituitary. Histological correlates in the pituitary were again absent.
- In TAM-treated females from laboratory 3, an absolute decrease in pituitary weight was again not confirmed relative to body weight. A decrease may be considered to be plausible where TAM acts as a weak oestrogen in the hypothalamus due to negative feedback from TAM and a subsequent decreased secretion of LH/FSH by the pituitary. This would be at doses where TAM is clearly acting as an antagonist in other tissues. Histological correlates in the pituitary were again absent.

- In CGS-treated females from laboratory 8, the decrease relative to body weight appears to be due to sharply increased (~30%) body weights. The body weight decrease appears to be a result of oestrogen ablation, as oestrogen curtails body weight. The histological correlate would be hypertrophy of secretory cells, and this was indeed observed in both studies.
- In MT-treated females from laboratory 12, the absolute and relative decreases are plausible considering an increased negative feedback from MT and decreased secretion of LH/FSH. However, no histopathological correlate indicated decreased secretory activity.
- In FLU-treated females from laboratory 2, the decrease in absolute pituitary weight is inconsistent with an antiandrogen mode of action in the female, and no histological correlate was observed.

233. In conclusion, there is no consistent relationship between the observed statistically significant weight change in the pituitary and either the expected change based on the mode of action or the presence of supporting histopathological correlates. Therefore, this argues against the inclusion of pituitary weight as a relevant and reliable TG 407 update. Further, the additional handling from dissection and trimming of the pituitary could result in tissue damage, thereby compromising the potentially useful histological examination of the tissue.

234. In contrast, histopathology of the pituitary was often relevant in supporting findings of endocrine modulation in other tissues in the updated TG 407 studies. The histopathology of the pituitary corresponded with tissue weight changes and other findings with several compounds in both sexes (Tables 85 and 86). Pituitary changes were observed in both studies CGS in females and FLU in males. The plausibility of the former was previously noted, and the latter is consistent with increased demand for LH secretion by the pituitary. Thyroid toxicants also led to apparent stimulation of the pituitary in at least one study and sex with DDE (acting by liver enzyme induction and increased excretion of thyroid hormones) and PTU and potential negative feedback in the case of THY. Therefore, the inclusion of pituitary histopathology as a relevant and reliable TG 407 update is recommended with the interpretation carefully integrated with other observations.

Absolute and relative thyroid weights and thyroid histopathology

235. The absolute and relative tissues weights of the thyroid appeared to be relevant and useful endpoints in the updated TG 407 studies. The absolute and/or relative weights of the thyroid were dramatically increased in both sexes with PTU in the one laboratory that measured thyroid weight (Tables 85 and 86). Other laboratories observed statistical significance for thyroid weight with DDE treatment (laboratory 6 in females and laboratory 7 in males), MT treatment (laboratory 12 in both sexes), and THY treatment (laboratory 9 in females). The difficulty in dissecting and weighing the thyroid (e.g., protection from desiccation) as well as the small absolute weight of the tissue should be noted as with the pituitary.

236. Histopathology of the thyroid was clearly relevant and reliable for the detection of thyroid toxicants in the updated TG 407 studies. The histopathology of the thyroid corresponded with tissue weight changes and other findings with several compounds in both sexes (Tables 85 and 86). Consistent histopathological findings were observed in both studies with DDE, PTU and THY, and this included evidence of follicular hypertrophy with the first two substances and hypotrophy with the last one; CGS and MT also led to follicular hypertrophy similar to DDE and PTU. Thyroid histopathology was typically the most sensitive or among the most sensitive endpoints in detecting dose related changes from thyroid toxicants in the updated TG 407 studies. This is discussed further in the subsection on circulating thyroid hormones below.

237. In conclusion, the addition of thyroid weights and histopathology provided clear benefit for the detection of possible endocrine modulation. As the thyroid weight is somewhat less useful, care should be

exercised not to compromise the histopathology of the tissue. The trimming should then continue to be performed after fixation of the tissue and by carefully trained technicians. Therefore, both thyroid weights and histopathology are recommended for inclusion in an updated TG 407 guideline.

Absolute and relative adrenal weights and adrenal histopathology

238. The absolute and relative tissues weights of the adrenal glands appeared to be useful supplemental endpoints in the updated TG 407 studies. The adrenal weights were altered with several test substances in both studies, particularly in males with EE, CGS, and FLU. The organ weight changes were supported by corresponding histopathological observations, which often extended to females and to lower doses than significant changes in tissue weight. The histopathological observations were then the more sensitive technique with respect to determining a possible NOEL and LOEL in this tissue. The supplemental nature of the adrenal weights and histopathology should be recognized and changes in the adrenal parameters should by themselves not be considered diagnostic of potential endocrine modulation in the case of (anti)oestrogens, (anti)androgen, and thyroid toxicants. However, it is plausible that the adrenals would be a primary target of general steroidogenesis inhibitors or cholesterol esterase inhibitors, and, therefore, the adrenals might contribute to the diagnosis of those particular modes of action.

Histopathology of the male reproductive tract

239. The histopathology of the male reproductive tract tissues corresponded well with tissue weight changes, and histopathological observations were slightly more sensitive than tissue weight changes in the updated TG 407 studies with respect to determining a possible NOEL and LOEL in these tissue. Histopathological findings and their correlation with tissue weights were consistent between the two studies for each chemical in the case of EE, MT, and FLU, and evident in the first TAM study (Table 85). It should be recognized that the characteristic diagnostic changes observed in male reproductive tract in the updated TG 407 studies were often associated not with frank lesions, but more subtle changes in the activity of the tissues including the Leydig cells, the germinal epithelium of the testes, and the degree of secretory activity of the accessory tissues. In addition, when sperm production was altered as in the case of MT, this finding was confirmed in the testicular and epididymal histopathology.

240. While these findings support the utility and reliability of added histopathological updates in the male reproductive tract, these findings also indicate that, for endocrine-modulating chemicals, histopathologists must be aware of more than frank lesions in the tissues. They should be cognizant in their observations to carefully assess the sometimes subtle increases and decreases in the activity of the target tissues in male reproductive tract in order to detect possible endocrine modulation. Thus, the detection of weak androgens and antiandrogens is then likely to require increased vigilance and some possible modification of current practice for the histopathological examination of the male reproductive tract.

241. Neither of the two updated TG 407 assays provided clear evidence of the antiandrogenic activity of *p,p'*-DDE. No changes in tissues weights were recorded in the male reproductive tract in either study. Several histopathological changes were recorded in the male reproductive tract in the high DDE dose group of the first study, including increased residual bodies and atrophy of several sex accessory tissues including the epididymis, seminal vesicles, and prostate. However, when the individual animal data were examined, the changes were associated with premature deaths and possible autolysis or degeneration of the tissues due to the moribund state of the animals before the tissues could be fixed as the cause of the observations (Table 67). The same doses in the 100-150 mg/kg/d range are the approximate LOEL for effects on the male reproductive tract in the Hershberger studies in the OECD Phase-2 validation studies for the assay and in the published literature (56)(65). This suggests that the intact, young adult animals in

the updated TG 407 studies are somewhat less sensitive than the surgically castrated animals to antiandrogens as previously noted.

242. In conclusion, histopathology of the male reproductive tract tissue was beneficial for the detection of possible endocrine modulation, and it is recommended that these observations be included in an updated TG 407 guideline. A further suggestion is that appropriate experts should consult and possibly update publish recent review of the handling, processing, and interpretation of the male reproductive tract tissues (76)(77), focusing on the action of various endocrine mechanisms including oestrogens, androgens, antiandrogens, and possibly steroidogenesis inhibitors.

Histopathology of the female reproductive tract

243. The histopathology of the female reproductive tract tissues (ovaries, uterus, vagina) corresponded well with tissue weight changes. The histopathological observations were often more sensitive than tissue weight changes in the updated TG 407 studies. Histopathological findings were consistent between the two studies for each chemical in the case of EE, TAM, CGS, and MT (Table 86). Thus, potent compounds were easily and consistently detectable.

244. Two weak putative oestrogens were evaluated in the TG 407 studies, GN and NP. In the GN-treated females, histological correlates of oestrogenic activity were observed in laboratory 12 in the uterus and the vagina (Table 23 and Annex 6). However, only equivocal findings in the vagina were observed in laboratory 4. In the NP-treated females, the laboratories recorded observations only in a single female out of 10 (laboratory 1) and 3 out of 10 (laboratory 6) at the highest NP dose that would be considered consistent with oestrogenic activity (Table 31 and Annex 6). While these findings collectively support the utility of histopathological examination of the female reproductive tract, the GN and NP results indicate that routine histopathological examination of individual tissues may encounter difficulties in some cases with weak oestrogens.

245. The challenge for detecting weaker oestrogens in the female is analogous to the sometimes subtle changes in activity of the male reproductive tract. The findings were generally of 'normal' histological observations in a particular tissue for individual parameters. As further elaborated in the following subsection on the vaginal smears, the necessary interpretation involves the integration of data from several tissues and the recognition by the pathologists that the overall synchronisation or orchestration of the oestrous cycle has been altered. The target tissues that require an integrative approach include at a minimum the ovarian follicular thecal and granulosa cells which are that are involved in oestrogen production, the epithelial and epithelial gland cells of the uterus as well as the uterine stroma, and the vaginal epithelium. It is the subtle nature of the histopathological observations (interpretation of the synchrony of several tissues and cell types in the female reproductive tract to the oestrous cycle) that must be noted in order to interpret overall synchrony. As with the male tissues, the detection of weak estrogens is then likely to require increased vigilance and some possible changes in the current practice of histopathological examination of the female reproductive tract.

246. This leads to the suggestion for a consultation of appropriate experts to consult and to update a recent review of the histological examination of the female reproductive tract tissues (78). The focus of the new review should be the assessment of the overall synchronization of these tissues in the oestrous cycle and how to interpret potential alterations.

Table 85. Summary of current and updated TG 407 endpoints responsive to endocrine modulation in male rats

Chemical	Lab	Tissue Weights [a]								Sperm		Histopathology									Hormones		
		Pituitary	Thyroid	Testes	Epididymis	Prostate			Seminal vesicles coagulating glands	Count	Morphology	Pituitary	Thyroid	Testis	Epididymis	Prostate			Seminal vesicles coagulating glands	Mammary Gland	T_3	T_4	TSH
						W	VP	DL								W	VP	DL					
EE	2	X					X	X	X					X			X	X	X	X			X
	5	X			X		X	X	X					~			X	X	X	X		X	
Genistein	4																						
	12																			ND			
Nonylphenol	1																			ND	X	X	
	6					X		X		X										ND			
Tamoxifen	3	X		X	X		X	X	X					X		X			X	X			
	10		ND	X		X			ND											ND	X		
CGS 18320B	8					X	X		X				~							X	X	X	
	13					X			X											ND	X	X	
Methyl Testosterone	3			X	X		X		X	~ [b]	~ [b]		~	X	~					X		X	X
	12		X	X	X					X	X		X	X	X					ND			X
Flutamide	2				X		X		X	X		X		X			X	X	X		X	X	
	11				X	X	X		X	X		X		X	X	X			X	X	X		X
***p,p'*-DDE**	6											X	X	[c]	[c]	[c]			[c]	ND	X	X	X
	7		X										X								X	X	
Propylthiouracil	1		X									X	X							ND	X	X	X
	10		ND						ND				X							ND	X	X	
l-Thyroxine	9											X	X										X
	13												X							ND	X	X	X

Summary of findings: X – statistically significant finding in combined Subgroups; ~ – equivocal, low rate of histopathological findings in combined Subgroups, but findings consistent with endocrine mechanism; ND – Not done (only marked for essential target tissues for mechanism and the male mammary gland).
W – whole prostate; VP – ventral prostate; DL – dorsolateral prostate; TSH – thyroid stimulating hormone.

[a] Includes both absolute and relative changes.

[b] See Table 10, Annex 8, although the overall parameters did not achieve significance at the high dose, those individual animals with low sperm count had marked histopathological changes.

[c] See Table 67 in main report, tissues with observations were from moribund animals, some with limited time of exposure. Therefore, the changes are judged likely to be due to autolysis and not to be treatment related.

Table 86. Summary of current and updated TG 407 endpoints responsive to endocrine modulation in female rats.

		Tissue Weights [a]					Histopathology						Hormones		
Chemical	**Lab**	**Pituitary**	**Thyroid**	**Ovaries**	**Uterus**	**Estrous**	**Pituitary**	**Thyroid**	**Ovaries**	**Uterus**	**Vagina**	**Mammary Gland**	**T_3**	**T_4**	**TSH**
EE	2				X	X			X	X	X			X	X
	5			X	X	X			X	X	X		X		
Genistein	4										~				
	12				X					X	X				X
Nonylphenol	1		X							~					
	6														
Tamoxifen	3	X		X	X				X	X	X			X	
	10		ND	X	X				X	X		ND		X	
CGS 18320B	8	X		X	X	X	X	X	X	X	X	X	X		
	13			X	X	X	X	~	X	X		X	X		
Methyl Testosterone	3			X		X		~	X	X	X	X	X	X	X
	12	X	X	X	X			X	X	X	X	X		X	X
Flutamide	2	X							X	~					
	11														
***p,p'*-DDE**	6		X					X							
	7							X					X		
Propylthiouracil	1		X				X	X					X	X	X
	10		ND					X				ND			X
l-Thyroxine	9		X					X							X
	13							X					X	X	

Summary of findings: X – statistically significant finding in combined Subgroups; ~ – equivocal, low rate of histopathological findings in combined Subgroups, but findings consistent with endocrine mechanism; ND – Not done (only marked for essential target tissues for mechanism); TSH – thyroid stimulating hormone.

[a] Includes both absolute and relative changes.

Histopathology of the male and female mammary glands

247. Histopathology of the female mammary gland observed consistent changes in both studies with CGS and MT (Table 86). Interestingly, no effects were seen in the one TAM study performing histopathological examination of the female mammary gland (Table 86). No other significant histopathological changes were observed in the female mammary gland in the other studies.

248. While the female mammary gland was removed and examined microscopically in 18 of the 20 studies, the same was not true of the male mammary gland. This tissue was examined in only 10 of the 20 studies (Table 85). This limited examination is unfortunate given the apparent responsiveness of the male mammary tissue to several of the test substances. In 6 of 10 studies where the male mammary gland was examined, histopathological changes were observed including both EE studies; the single TAM, CGS, and MT studies were the male mammary gland was examined; and one of the two FLU studies where this tissue was examined.

249. These findings support the utility and reliability of histopathological examinations of both the female and the male mammary glands in studies with suspected endocrine modulating compounds. The finding of changes in the male mammary gland with the modestly potent oestrogen, methoxychlor, in a published 407 study lends further support (19). The removal, preservation, and examination of mammary tissues from both sexes is then recommended for inclusion in an updated TG 407 guideline. A second point should be considered in this regard. Due to the limited experience of many pathologists with the male mammary gland, appropriate experts should publish a review of the handling, processing, and examination of the male mammary gland focusing on the reliable preparation of the tissue and the action of various endocrine mechanisms.

Sperm count and morphology

250. The inclusion of spermatology endpoints in the updated TG 407 provided little or no benefit in detecting the endocrine modulating activity of any test substance. The sperm endpoints were not sensitive and in some cases appeared to be highly variable and even unreliable. This lack of utility is plausible due to young age of the males soon after reaching sexual maturity, the short exposure time of 28-days relative to lengthier time of spermatogenesis in the male rat (approximately 60 days), and the technical needs of the assay in controlling the temperature and timing of the analysis at necropsy which may have contributed to significant CVs in some laboratories. In addition, the additional personnel time and resource usage for the sperm counts and morphology are particularly significant in light of the absence of any apparent value of the spermatology endpoints.

251. The spermatology findings of the updated TG 407 studies are summarized in Table 87. In several cases, apparent random, non-dose related incidents of statistical significance were found. With NP, the sperm counts in laboratory 6 were highly variable across dose groups and between subgroups (see Table 6, Annex 8), and findings of normal sperm morphology were consistently low in all control and dosage groups.

Table 87. Summary of updated TG 407 spermatology endpoints

Chemical	Lab	Sperm Counts	Sperm Morphology
Ethinyl Estradiol	2	No	No
	5	No	No
Genistein	4	No [a]	No
	12	No	No
Nonylphenol	1	No	No
	6	**Yes ↓** [a]	No
Tamoxifen	3	No	No
	10	No	No
CGS 18320 B	8	No	No [a]
	13	No	No
Methyl Testosterone	3	No [b]	No [b]
	12	**Yes ↓**	**Yes ↑**
Flutamide	2	**Yes ↓**	No
	11	**Yes ↓**	No [a]
***p,p'*-DDE**	6	No [a]	No
	7	No	No
Propylthiouracil	1	No	No
	10	No	No
l-Thyroxine	9	No	No
	13	No	No

[a] Statistical finding was not dose related.

[b] See Table 10, Annex 8, although overall parameters did not achieve significance in a dose related manner, particular individuals with low sperm counts in the mid- and high dose groups had marked, parallel histopathological observations.

252. With highly potent androgens (MT) and antiandrogens (FLU), spermatology findings offered limited consistency. However, significant findings were present only at the high doses with no impact on the determination of the NOEL or LOEL. With MT, one laboratory was able to correlate histopathological observations with large decreases in individual sperm counts and an increase in morphological abnormalities, although no overall statistically significant change was observed (Tables 10 and 11, Annex 8). The second study did observe statistically significant changes at the high MT dose (Table 12, Annex 8). With FLU, both laboratories observed dose related decrease in sperm counts at the high FLU dose, but only in the combined subgroups and neither individual subgroup in one study (Table 13, Annex 8). Morphological changes in sperm were not observed in one FLU study, and a statistically significant change at the mid-dose was not apparently dose-related in the second study.

253. The power of a method to detect change is related to its coefficient of variation (CV) in a given laboratory setting. The CVs for the sperm counts have been calculated and are reported in Table 88. A review of these data suggests that laboratory technique is a major variable. Laboratories 4, 5, and 8 had uniformly modest CV values. Other laboratories had more variable and sometimes much larger CV values. This indicates that any future studies dealing with sperm counts and, possibly, sperm morphology should begin with clearly defined standard operating procedures for such measurements that can achieve better and more consistent CVs than displayed in these studies.

254. In conclusion, the spermatology endpoints did not provide any substantive benefit for the detection of possible endocrine modulation, and the CVs indicate that substantial work might be needed to achieve reliable methodological practice among laboratories. Therefore, it is recommended that neither sperm counts or sperm morphology endpoints be included in an updated TG 407 guideline.

Table 88. Coefficients of variation for sperm count analyses.

Chemical	Lab	Control	Low	Mid	High
EE	2	25.0	25.6	31.5	31.7
	5	15.8	11.9	16.7	13.8
GN	4	11.5	19.2	17.0	16.2
	12	23.0	27.7	28.2	23.2
NP	1	29.7	ND	ND	19.6
	6	48.7	78.4	110.4	68.0
TAM	3	22.4	12.6	22.4	11.2
	10				
CGS	8	13.1	15.4	13.1	15.9
	13	28.4	23.3	22.3	13.8
MT	3	15.2	13.6	42.7	47.2
	12	38.6	56.7	24.5	72.5
FLU	2	29.4	43.0	25.0	42.3
	11	15.4	15.4	20.7	40.8
DDE	6	36.5	90.6	29.2	62.5
	7	25.5	19.2	23.3	18.0
PTU	1	20.9	ND	ND	21.4
	10				
THY	9	30.9	ND	ND	21.3
	13	22.8	31.8	27.0	19.3

Circulating thyroid hormones

255. The data for thyroid circulating hormones (T_3, T_4, TSH) indicate that the use of these endpoints to identify possible endocrine modulation needs detailed review and assessment. In a number of instances the hormones displayed statistical significance that was not dose-related, was absent any changes in thyroid tissue weight or histopathology, and there were sometimes inconsistent changes in one hormone with another (Table 89). In one case with CGS treatment, the hormonal data between laboratories were apparently contradictory with statistically significant increases in one study and statistically significant decreases in the second study for all three circulating thyroid hormones (see Table 89).

256. One potential difficulty with the performance of the thyroid hormones is that the literature indicates considerable variability in circulating thyroid hormone levels (72)(73). The sources of variability include:

- Temporal variations in diurnal cycle (where differences in the time of necropsy among laboratories and even within a laboratory might yield different circulating thyroid levels) as well as changes in the values during the female oestrus cycle;
- Handling stress on the animals, such as the movement of animals to the necropsy room immediately prior to necropsy; and
- Differences in anesthesia and blood sampling among laboratories.

257. Another difficulty in evaluating the thyroid hormone results was the use of different assay kits and the reporting of the results in different units by the various laboratories. The methods used by each laboratory have been included in Annex 7, when the methods were clearly included in the final report. The original units for the various hormones reported by each laboratory are noted in Table 90, and the conversion factors used are reported in the footnotes of Table 90.

Table 89. Summary of current and updated TG 407 thyroid endpoint results.

Chemical	Lab	T_3	T_4	TSH	Dose Related Thyroid Hormone Findings in Other Studies	
					Tissue Weight	Histopathology
Ethinyl Estradiol	2	No	**↑ F**	**↑ M - ↑ F**	**↑ M rl**	No changes
	5	**↓ F**	**↑ M**	No [a]	No changes	No changes
Genistein	4	No	No	No [a]	No changes	No changes
	12	No	No	**↑↑ F**	Non-dose related ↑ F	No changes
Nonylphenol	1	**↑ M**	**↑ M**	No	No changes	No changes
	6	No	No	No	No changes	No changes
Tamoxifen	3	No	**↑ F**	No	No changes	No changes
	10	**↑ M**	**↑ F**	No	Not measured	No changes
CGS 18320 B	8	**↓↓ M - ↓↓ F**	**↓↓ M**	No [a]	**↑ M rl**	No changes
	13	**↑↑ M - ↑ F**	**↑↑ F**	No	No changes	**Slight Hypertrophy in F**
Methyl Testosterone	3	**↓↓ F**	**↑ M - ↑↑ F**	**↑↑ M - ↑↑ F**	No changes	**Slight Hypertrophy in M & F**
	12	No	**↑↑ F**	**↑↑ M - ↑ F**	**↑ M ab/rl, ↑ F ab/rl**	**Slight Hypertrophy in M Hypertrophy in F**
Flutamide	2	**↓ M**	**↓↓ M**	No	No changes	No changes
	11	**↑ M**	No	**↑ M**	No changes	No changes
***p,p'*-DDE**	6	**↑↑ M**	**↓↓ M**	**↑ M**	**↑ F ab/rl**	**Hypertrophy in M & F**
	7	**↑↑ M - ↑↑ F**	**↓↓ M**	No	**↑ M rl**	**Hypertrophy in M & F**
Propylthiouracil	1	**↓↓ M - ↓↓ F**	**↓↓ M - ↓↓ F**	**↑↑ M - ↑↑ F**	**↑↑ M ab/rl, ↑↑ F ab/rl**	**Severe Hypertrophy M & F**
	10	**↓ M**	**↓ M**	**↑ F**	Not measured	**Severe Hypertrophy M & F**
l-Thyroxine	9	No [a]	No [a]	**↓↓ M -↓↓ F**	**↓ F rl**	**Atrophy M & F**
	13	**↑↑ M - ↑↑ F**	**↑↑ M - ↑↑ F**	**↓ M**	No changes	**Atrophy M & F**

↑ ↓ significant increase or decrease, respectively at $p < 0.05$, ↑↑ ↓↓ significant increase or decrease, respectively at $p < 0.01$;
M – male, F – female; for weights ab – absolute, rl – relative.
[a] Statistical finding was not dose related.

Table 90. Measurement units for T3, T4, and TSH among the 13 laboratories performing updated 407 studies.

Lab	T_3/T_4				TSH		
	ng/dl[a]	µg/dl[b]	ng/ml	nmol/l[c]	ng/ml	µg/l[d]	µIU/ml
1				T_3, T_4 ✓		✓	
2				T_3, T_4 ✓		✓	
3			T_3, T_4 ✓		✓		
4	T_3 ✓	T_4 ✓			✓		
5	T_3 ✓	T_4 ✓			✓		
6	T_3, T_4 ✓				✓		
7			T_3, T_4 ✓		✓		
8							
9	T_3 ✓	T_4 ✓			✓		
10	T_3 ✓	T_4 ✓[e]					✓
11	T_3 ✓	T_4 ✓			✓		
12				T_3, T_4 ✓	✓		
13				T_3, T_4 ✓	✓		

[a] To convert ng/dl to ng/ml, divide by one hundred.

[b] To convert µg/dl to ng/ml, multiply by ten.

[c] To convert nmol/l of T_3 to ng/ml; MW of T_3: 651.01 so 1 nmol/L = 651.01 ng/L = 0.651 ng/mL so multiply by 0.651
To convert nmol/l of T_4 to ng/ml MW of T_4 776.93 so 1 nmol/L = 776.93 ng/L = 0.777 ng/mL = so multiply by 0.777

[d] To convert µg/l of TSH to ng/ml, both the numerator and denominator are 3 orders of magnitude apart, the number should then be the same and only the units change.

[e] Units corrected after inquiry by the Secretariat

258. The individual means have then be converted were possible into common units and these values are reported in Table 91. The following observations are made based upon these values:

- With one exception, the T_3 values are as expected approximately 1-4% of the T_4 values. This is consistent with published ratios, and the primary production of T_4 by the thyroid and the subsequent conversion in target tissues via deiodinase to T_3.
- The one exception is laboratory 10. Here, the reported units appear to be in error as the T_4 values are approximately one thousand-fold less than in the other laboratories. However, this potential error could not be confirmed with the laboratory by the Secretariat.
- The overall agreement among the laboratories for the T_3 and T_4 values in Table 91 is relatively good, being within the same order of magnitude other than for a) the noted exception and b) dose groups receiving sufficient quantities of potent thyroid toxicants (e.g., PTU and THY).
- Conversion of TSH values for laboratory 10 was not accomplished, as conversion factors for international units were not available to complete those calculations.
- For TSH, there was overall agreement in the values with the possible exception of laboratory 12. An inquiry by the Secretariat to this laboratory revealed that they had conducted TSH analysis for a number of years and that these reported values were consistent with their historical baselines. This latter laboratory also employs a somewhat unique strain of rat that is bred in their own facility (see Table 4).
- Sexual differences and trends in circulating thyroid hormone concentrations were observed as has often been reported in the literature. In general, T_3 levels were higher in females; T_4 levels were almost always higher in males; and TSH levels were generally higher in males.

259. This leads to the conclusion that the data, although variable, do have underlying consistency, and that a more detailed assessment of their reliability and relevance is appropriate. This leads to several questions.

260. Were the hormonal data consistent with changes in the thyroid weights and observations of the thyroid histopathology (see Table 89)?

- In the case of PTU treatment, laboratory 1 observed statistically significant decreases in T_3 and T_4 in both sexes and a statistically significant increase in TSH. Absolute changes in the hormonal values were very large (Table 17, Annex 7). Laboratory 10 observed statistically significant decreases in T_3 and T_4 in males, but not females. A statistically significant increase in TSH in females only at the high dose. Absolute changes were large, including a 75% drop in TSH in males that did not achieve statistical significance (Table 18, Annex 7). This overall profile is consistent with the action of PTU.
- In the case of THY treatment, laboratory 9 observed statistically significant decreases in TSH in both sexes. However, there were no obvious absolute changes in either T_3 or T_4 even though the latter was the administered test substance (Table 19, Annex 7). Laboratory 13 observed a significant decrease in TSH in males, while increases in T_3 and T_4 were significant, and absolute changes in hormonal values were large in both sexes (Table 20, Annex 7). The TSH decrease is consistent with the action of THY; the unchanged values for T_3 or T_4 in one study appear inconsistent.
- In the case of DDE treatment, laboratory 6 observed a significant increase in T_3, a significant decrease in T_4, and a significant increase in TSH in males. Absolute trends in females were similar, and TSH was significantly decreased at the mid-dose, but not the high dose (Table 15, Annex 7). Laboratory 7 observed a significant increase in T_3 in both sexes and a significant decrease in T_4 in males. Absolute values of TSH were increased in males and absolute T_4 values were decreased in females, but these changes did not achieve statistical significance (Table 16, Annex 7). Combined with the large increases in liver weights, the T_4 and TSH changes are consistent with the induction of hepatic enzymes by DDE and leading to an accelerated excretion of thyroid hormones. A potential rationale for the increase in T_3 in both studies might be that, under conditions of strong liver enzyme induction when the animal could no longer maintain T_4 levels, increased T_3 synthesis has been observed in some cases (79).
- In the case of MT treatment, both laboratories observed significant increases in TSH in both sexes and T_4 in females. T_4 was significantly increased in one study, and the absolute values of T_4 were increased by approximately the same degree in the second study without achieving statistical significance (Tables 11 and 12, Annex 7). These changes are consistent with a stimulation of the thyroid and anabolic metabolism by high levels of androgens.

Table 91. Comparison of TG 407 circulating thyroid hormone values when converted to common units.

Chem	Lab	Sex	T_3 (ng per dL)				T_4 (ng per dL)				TSH (ng per mL)			
			C	L	I	H	C	L	I	H	C	L	I	H
EE	2	M	150.4	205.7	176.4	166.7	7226.1	7847.7	7459.2	8158.5	4.31	6.27	9.23	7.3
		F	99.6	95.0	99.0	106.8	4817.4	5749.8	6293.7	6837.6	2.5	3.05	4.36	4.34
	5	M	39	43	33	37	5230	5290	5510	6310	19.2	20	16.6	17.9
		F	55	48	47	44	4740	4960	5280	5030	20.1	26.5	23.6	17.7
GN	4	M	116.1	119.7	118.2	119.8	7200	7700	6700	7400	6.48	5.22	7.29	5.83
		F	129.1	108.8	114.8	121.1	5000	4400	4300	4600	2.67	4.4	4.02	3.22
	12	M	99.0	93.7	88.5	98.3	4996.1	4957.3	5097.1	5042.7	0.11	0.13	0.07	0.14
		F	84.6		90.5	89.8	4226.9		4483.3	5237.0	0.06		0.11	0.16
NP	1	M	91.1	83.3	98.3	112.6	3677.5	3888.9	4352.0	4714.8	7.44	6.73	7.1	6.94
		F	113.9	112.6	119.8	113.3	3264.2	3319.3	3328.7	3419.6	5.58	6.25	6.28	6.66
	6	M	70.6	73.2	60.8	72.3	5200	5300	4600	4700	8.3	8.7	8.6	9.3
		F	75.6	75.9	69.9	66.3	2600	2900	2500	3400	6.4	5.9	6.2	6.8
CGS	8	M	75	63	64	62	4400	3150	3260	2810	11.18	9.35	8.76	7.49
		F	65.6	55.6	58.9	50.1	2040	2590	2630	2560	5.77	7.8	5.89	6.55
	13	M	127.6	126.3	162.8	184.9	7233.9	6488.0	6674.4	6643.4	22.31	22.68	20.57	22.77
		F	181.0	220.7	227.9	250.6	5376.8	8111.9	8236.2	8515.9	29.02	33.22	29.83	40.27
Tam	3	M	35	28	25	36	2559	2213	2280	2840	3.53	4.5	5.3	2.96
		F	44	40	46	52	1693	1687	2281	3201	4.51	3.17	4.31	2.47
	10	M	10.4	9.77	14.43	119.54	4010	3940	4390	4170				
		F	124.7	111.8	117.09	113.76	2550	2450	3830	5260				
MT	3	M	31	32	36	34	2474	2460	2826	3052	4.92	8.58	8.81	11.97
		F	49	35	34	28	2159	2356	2858	2754	4.51	4.13	12.19	17.99
	12	M	107.4	95.0	108.7	103.5	5268.1	5485.6	5742.0	5594.4	0.22	0.13	0.24	0.42
		F	98.3	102.9	95.0	104.8	4001.6	4949.5	5920.7	5563.3	0.08	0.12	0.12	0.51
Flu	2	M	121.7	121.1	122.4	108.1	3807.3	4149.18	4164.72	2975.91	2.93	3.23	3	3.64
		F	126.3	125.6	127.6	124.3	2789.43	2657.34	2672.88	2486.4	1.77	1.6	1.23	2.32
	11	M	110.8	120.3	121.8	128.3	4570	4660	4620	4220	7.98	10.61	9.23	10.34
		F	110.9	106.9	111.4	122	3140	2660	2670	2240	7.12	6.96	6.44	8.62
DDE	6	M	90.3	85	96.9	115	5500	5100	4900	3600	8.6	13.9	13.5	18.2
		F	84.6	94.8	91.2	97.9	3500	3800	3800	3100	5.3	6.1	7.5	7.3
	7	M	70	110	130	170	10520	11070	8910	7480	25.7	29.2	27	31.7
		F	90	90	100	140	9960	10840	10110	7540	31.3	36.3	34.8	36
PTU	1	M	63.8	83.3	74.2	17.6	3822.84	4226.88	1585.08	621.6	6.47	7.77	30.7	58.4
		F	89.2	81.4	82.0	26.7	3457.65	2812.74	1437.45	761.46	5.8	6.7	20.6	64.2
	10	M	112.21	107.35	123.85	75.16	4090	3750	3370	90	0.018	0.029	0.057	0.027
		F	111.7	106.28	133.68	87.94	4190	3650	3240	1410	0.02	0.011	0.01	0.038
Thy	11	M	76.21	68.3	51.03	92.1	5140	5490	5390	5080	13.41	10.09	8.66	2.84
		F	67.13	58.69	47.84	61.6	3210	3190	4410	3400	5.11	5.04	4.73	2.97
	13	M	117.8	134.1	160.1	337.2	6565.7	7863.2	10046.6	13426.6	28.42	27.13	26.78	20.09
		F	164.1	158.2	192.0	302.1	4071.5	4763.0	6930.8	8850.0	35.53	35.91	33.55	34.33

Table 92. Coefficients of variation (CV) for circulating thyroid hormones.

Hormones			T_3	T_4	TSH
Overall CV Averages			**18.7%**	**20.6%**	**39.2%**
Chem	**Lab**	**Sex**			
EE	**2**	**M**	19.2%	12.4%	42.1%
		F	25.2%	17.9%	35.3%
	5	**M**	37.0%	16.3%	23.9%
		F	16.2%	17.3%	23.0%
GN	**4**	**M**	13.6%	19.3%	68.0%
		F	18.8%	24.9%	64.8%
	12	**M**	15.5%	15.7%	56.1%
		F	10.9%	18.1%	44.2%
NP	**1**	**M**	16.0%	12.4%	21.8%
		F	18.1%	20.2%	14.2%
	6	**M**	15.6%	21.4%	29.2%
		F	14.2%	20.4%	25.5%
Tam	**3**	**M**	34.0%	17.7%	53.5%
		F	25.0%	26.5%	48.7%
	10	**M**	6.9%	12.1%	66.7%
		F	8.8%	18.6%	37.5%
CGS	**8**	**M**	12.0%	17.1%	42.9%
		F	17.9%	27.9%	16.8%
	13	**M**	23.6%	15.5%	23.0%
		F	24.5%	13.7%	29.6%
MT	**3**	**M**	33.0%	16.1%	38.4%
		F	27.5%	14.2%	36.1%
	12	**M**	16.0%	14.9%	58.8%
		F	14.2%	17.0%	56.9%
Flu	**2**	**M**	6.8%	12.0%	78.4%
		F	9.8%	20.1%	53.8%
	11	**M**	11.5%	15.6%	27.6%
		F	17.0%	32.9%	24.0%
DDE	**6**	**M**	13.3%	15.4%	34.0%
		F	15.2%	24.6%	27.1%
	7	**M**	20.1%	13.7%	24.9%
		F	19.8%	28.1%	32.2%
PTU	**1**	**M**	37.8%	17.9%	42.3%
		F	31.6%	20.5%	30.4%
	10	**M**	18.3%	63.3%	104.4%
		F	17.7%	53.7%	37.3%
Thy	**9**	**M**	15.4%	21.0%	35.5%
		F	20.2%	24.9%	24.6%
	13	**M**	17.8%	9.6%	19.3%
		F	12.0%	23.5%	15.9%

261. Were the hormonal data in the above instances more or less sensitive than the thyroid weights and observations of the thyroid histopathology?

- For PTU, histopathology was clearly more sensitive than any circulating hormone in laboratory 10. In laboratory 1, histopathology was clearly changed in both sexes at the lowest PTU dose, where only T_4 in females was significantly decreased. Thus, histopathological observations and T_4 hormonal changes were equally sensitive in this laboratory.

- For THY, histopathology was clearly more sensitive than any circulating hormone in laboratory 13. In laboratory 9, histopathology was clearly changed in both sexes at mid- or intermediate THY dose. T_4 changes in the female were observed at the same dose, but TSH changes only at the high dose. Thus, histopathological observations and T_4 hormonal changes were equally sensitive in this laboratory.
- For DDE, histopathology was most sensitive with females in both studies. T_4 was as sensitive as histopathology in males in laboratory 6, and TSH was as sensitive as histopathology in males in laboratory 7.
- For MT, T_4 in laboratory 12 (females) and TSH in laboratory 3 (males) were potentially more sensitive than histopathology.
- In no case did T_3 give an indication of greater or equivalent sensitivity.

262. Were there incidents when the hormonal data were clearly contradictory? Yes, in two cases:

- Both T_3 and T_4 were statistically significant in opposite directions in the case of CGS.
- T_3 was significantly significant in opposite directions in the case of FLU.

As shown in Table 92, the CVs for the circulating thyroid hormones can vary among laboratories. As the discussion above indicated that T_4 and TSH might be the most useful thyroid hormones, so these CVs are noteworthy. First, it is clear that T_4 CVs are typically about one half or less of TSH CVs. This indicates that, with all else being equal, changes in the T_4 values might be less variable and provide more power. This is consistent with observations comparing the sensitivity of the hormonal values with the histopathology.

263. The following conclusions are offered:

- In this study T_3 thyroid hormone levels do not appear to provide information of sufficient reliability and sensitivity when compared with thyroid histopathology and the other two thyroid hormones. But there is evidence from other studies that rodents sometimes tend to overcompensate in response to certain anti-thyroid compounds (e.g. liver enzyme inducers), resulting in increased T_3 but quite normal T_4 and TSH levels. Therefore, retained serum samples might be submitted for T_3 analyses.
- The T_4 thyroid hormone levels may provide relevant information, and the results suggest that T_4 thyroid hormone levels could be equally sensitive with histopathology in determining a NOEL and LOEL. Given the moderate CVs displayed in the updated TG 407 studies, the T_4 thyroid hormone levels could provide modest power. Therefore, an expert review of factors such as the daily timing of necropsy, the necropsy animal handling and blood sampling procedures, and the analytical kits and procedures is recommended with the aim of improving the performance and utility of the T_4 thyroid hormone endpoint. With refinement, T_4 thyroid hormone levels would then be recommended for potential inclusion in an updated TG 407 guideline as supplementary endpoint. That is, where existing information would indicate the test substance is a thyroid toxicant or the histopathology of the thyroid indicated possible effects, then retained serum samples might be submitted for T_4 thyroid hormone analyses.
- The TSH hormone levels may also provide relevant information. However, the higher CVs displayed in the updated TG 407 studies, question the power and, thus, the overall sensitivity of this endpoint. Consideration of the TSH endpoint should await a similar expert review of factors impacting TSH variability and the need for further evaluation and additional mechanistic support from TSH, if significant changes in T_4 hormone levels are observed. Again, in the conditions noted for T_4 immediately above, retained serum samples might be submitted for TSH analyses.

Estrous cycle evaluation

264. Vaginal smears were incorporated in the updated TG 407 protocol in Phase-2 to facilitate the necropsy of female animals in dioestrus. This was intended to reduce wide variations in the weight and histopathology of the female reproductive tract tissues during the normal cycle. The updated TG 407 protocol specified that animals would be smeared daily from study day 24 and necropsied on, or after, study day 28, with the day determined by vaginal smears indicating that an individual animal was in dioestrus. This provided for a minimum of 5 days of smearing, but complicated logistics and assessment of treatment-effects with the staggered necropsy. If animals had not entered dioestrus by day 32, then they were to be necropsied regardless of their stage.

265. There were procedural variations in the way the smears were performed (Table 93). The largest variation existed with the day of smearing start with 4 laboratories extending the duration of smearing by starting earlier (days 18-21). Only one laboratory started smearing late (day 26), allowing only 3 smears prior to earliest termination. One laboratory removed animals on day 27 (only 4 smears), if they were found to be in dioestrus. There was a single laboratory that failed to smear animals for a two day period, which may reflect a weekend on study. Finally, there were two laboratories where a small proportion of animals (max 2/20) were not terminated on the first reported day of dioestrus after day 28, but on the second. Both laboratories removed the remainder of animals on day 1 dioestrus. This suggests the laboratories were likely to be making individual judgements relating to the smears and the cycle of the animal.

266. The final reports varied in the description of the vaginal smears and the inclusion of the data and their interpretation in the final report (Table 94). Five laboratories did not discuss the vaginal smear results, and 9 laboratories did not include the individual animal smear data in the final reports. No standardised reporting procedure for the smear data was provided, resulting in reporting variations when these data were included in the final reports. Due to the short duration of vaginal smears (5 days barely covers one full cycle in the female rat), no oestrus cyclicity analysis was included in the reports (although by default computerised analysis was included with the individual animal data for one laboratory). Therefore, the Secretariat requested these data in a standardised form from several laboratories (see Annex 8). In this section, the interpretation of the smears led to categorision into 3 groups: 1) normal or suggestive of normal cycle; 2) definitely abnormal or 3) equivocal and unable to categorise.

Table 93. Individual laboratory procedures for vaginal smears.

Laboratory	Starting day for smears	Comments
Protocol	Day 24	Kill in dioestrus day 28 onwards (day 5 smearing)
1	Day 24	Per protocol
2	Day 24	1/20 animals not removed on first day of dioestrus day 28 onwards
	Day 18	
3	Day 24	No smearing performed days 27, 28 (1 study)
4	Day 24	Per protocol
5	Day 21	Per protocol
6	Day 24	Per protocol
7	No information available	
8	Day 26	Last animals removed on day 31
9	Day 24	Per protocol
10	Day 21	N=6, Smears conducted per protocol
11	Day 20	2/20 animals not removed on first day of dioestrus day 28 onwards
12	Day 24	Per protocol
13	Day 24	Proportion of animals killed on day 4 smearing (day 27)
NB: The day of study start was harmonised across the laboratories where differences existed in study day numbering		

Table 94. Individual laboratory protocols for vaginal smears.

Lab	Individual smear data reported?	Text conclusions in final report on smear data
1	No	No substance-related effects were observed
2	Yes	All females were sacrificed in dioestrus. There were no statistical differences between treatment groups
		All females were sacrificed in dioestrus. There were no statistical differences between treatment groups
3	No	The information was not presented in the report but kept in the study file
4	No	None of the females exhibited apparent abnormal oestrous cycle
5	No	Two rats given 50 µg/kg EE showed continuous diestrous or other abnormalities, and these abnormalities were observed in all rats at 200 µg/kg EE
6	Yes	No specific text in results
7	No	No specific text in results
8	No	No specific text in results
9	No	No specific text in results
10	No	No significant change was detected in oestrous cycle compared with controls.
11	Yes	No clear changes related to FLU administration of flutamide were observed
12	No	No specific text in results
13	Yes	In all treatment groups, no abnormalities were observed
		In all treated female groups no oestrous cycle was observed during 5 consecutive days before sacrifice

267. Several laboratories noted that, upon histopathological examination, one or more tissues of the female reproductive tract where not concordant with the vaginal smears, i.e., the tissue observations in one or more individuals were not consistent with dioestrus. This has led to an attempt to carefully correlate the synchronisation of the female reproductive tract tissues with the oestrus cycle based upon the histopathology observations. However, this effort encountered varied approaches to reporting the histological sections of the female reproductive tract:

1. In 10 cases, the histological sections for female reproductive tract tissues were read from the vehicle control and high dose animals only, the intermediate doses were not reported where no abnormal findings had been observed in the high dose animals.
2. Five laboratories did use the sections of vagina and or uterus/cervix to assess the histological stage of the cycle (**data highlighted in bold in Table 95**). For control animals, there was only a mismatch of 1/ 50 animals between the terminal smear and the histological assessment of the stage of cycle. At the high dose, 39/50 animals had histological stages that were mismatched with the smear. The inter-dose values were 11/49 and 22/50 at the low and mid doses respectively. This data are strongly suggestive that that interpretative staging of the female reproductive tract tissues would be useful in determining treatment related effects on the synchronisation of the tissues in the oestrous cycle.
3. Many laboratories, however, did not indicate in the reports that staging of the tissues was attempted to assess concordance with the smear. These laboratories simply reported sections as normal or no abnormality (*data highlighted in italics in Table 95*). It would appear that the pathologists judged a tissue with normal features for ANY stage of cycle as no abnormalities detected (NAD). Thus, for a majority of studies, it is not possible to determine how frequently smears in dioestrus did or did not correlate with histology in the female reproductive tract was consistent with the animal being in dioestrus.
4. There were several studies, where it has not been possible to obtain individual smear data in time to allow comparisons to be made.

268. Where the individual smear data was available, the distribution of control animal smears was comparable with the duration of the cycle in that stage. On any given day, approximately one quarter of animals were in proestrus or oestrus, and a smaller percentage of animals were in metoestrus (Table 96). The latter is typically only of approximately 6 hours duration. However, there was a large range in the frequencies reported among the laboratories (Table 96). There are two possible reasons:

1. Individual labs varied in their readings and judgements of the stage cycles, probably as a result of differing interpretation. For example, metoestrus was recorded in 2% of readings at one laboratory and 26.6% in another.
2. In some experiments, 2-3 animals in a Subgroup of 5 had identical smear patterns, i.e., the animals' cycles appear to have been synchronised. Although the concept for the synchronisation of cage mates is widely discussed, these data were insufficient to make any determination.

269. In order to estimate whether the animals' cycles were normal, it must first be assumed that the animals on study would display regular cycles. All animals were approximately 7 weeks old when placed on study, and with dosing and acclimatisation would have been 11-12 weeks old at study completion. First oestrus is usually coincident with vaginal opening, and this occurs at approximately 34 days of age (5 weeks). The animals would then have had approximately 2-3 weeks to establish a regular cycle before the initiation of dosing and with 4 weeks of dosing to follow. Thus, the animals in the absence of a treatment related effect should have established regular cycles before smearing was begun, and the assumption that they animals would have had regular cycles appears to be valid.

Table 95. Correlation of oestrus staging by vaginal smear with histological findings in the female reproductive tract

		Vehicle Control			Low-Dose			Mid-Dose			High-Dose		
Chemical	**Lab**	Number sacrificed in dioestrus by smear	Number abnormal histology	Number stage matched histology	Number sacrificed in dioestrus by smear	Number abnormal histology	Number stage matched histology	Number sacrificed in dioestrus by smear	Number abnormal histology	Number stage matched histology	Number sacrificed in dioestrus by smear	Number abnormal histology	Number stage matched histology
Ethinyl oestradiol	**2**	**10**[a]	**0**	**10**	**10**	**1**	**7**	**8**	**3**	**7**	**10**	**8**	**2**
	5								4			5	
Genistein	4	10			10	ND		10	ND		10	2	
	12	**9**	**0**	**9**	**8**	**0**	**8**	**5**	**4**	**10**	**3**	**3**	**2**
Nonylphenol	1	*10*[b]	*0*		*10*	*ND*		*10*	*ND*		*6/9*	*0*	
	6	*10*	*0*		*10*	*ND*		*10*	*ND*		*10*	*0*	
Tamoxifen	**3**	**10**	**0**	**10**	**10**	**0**	**9**	**10**	**8**	**0**	**10**	**10**	**0**
	10	6		0	6		1	6		1	6	6	
CGS18320B	8		0			7			10			10	
	13	10	2		10	10		10	10		10	10	
Methyl Testosterone	**3**	**6**	**0**	**5**	**9**	**0**	**7**	**9**	**9**	**0**	**9**	**10**	**0**
	12		0			0			10			10	
Flutamide	**2**	**10**	**0**	**10**	**10**	**0**	**9**	**10**	**10**	**10**	**10**	**10**	**9**
	11	*10*	*0*		*10*	*ND*		*10*	*ND*		*10*	*0*	
DDE	6	10			10	ND		10	ND		7		
	7	10				ND			ND			10	
Propylthio-uracil	1	*10*	*0*		*10*	*ND*		*10*	*ND*		*10*	*0*	
	10					ND			ND				
L-thyroxine	9	*10*	*0*		*10*	*ND*		*10*	*ND*		*10*	*0*	
	13	*10*	*0*		*10*	*ND*		*10*	*ND*		*10*	*0*	

[a] Data in bold typeface indicates the labs evaluated the concordance between the vaginal smears and the histology of the female reproductive tract.

[b] Data in italic typeface indicates the labs did not report staging the histology of the female reproductive tract according the oestrus cycle.

Table 96. Assessment of the oestrous cycle in vehicle control animals

	Proestrus	Oestrus	Metoestrus	Dioestrus
Range	0-50%	0-50%	0-40%	10-70%
Mean	24.2%	26.4%	13.1%	36.3%

270. The major shortcoming with the data set is that, with a minimum of 5 days of vaginal smears (day 24 to termination), animals will have completed only one full cycle. This is sufficient to reveal only dramatic alterations in or even suspension of the cycle, and modest perturbations in the length of a stage or even the entire cycle are impossible to infer. Three laboratories smeared for up to 14 days prior to termination. With the observation of approximately three full cycles, it becomes possible to assess the normality the cycle or lack thereof as patterns have the opportunity to be repeat. Data from 14 days of observation often resulted in a finding of abnormal cycles, when compared to a second study that conducted vaginal smears for a more limited time (Table 97).

271. Overall, the observation of cycle abnormalities typically had strong histological correlates in the female reproductive tract with compounds such as EE, GN, TAM, MT, and FLU within an individual laboratory. Although the duration of smearing was sub-optimal for assessment of the oestrus cycle the data suggest that there were abnormalities of the cycle with NP where there was an absence of histological changes in the female reproductive tract. It is noteworthy that major disturbances of the thyroid hormonal axis with PTU and THY and modest disturbances with DDE did not noticeably alter the oestrus cycle.

272. The central issue is the possible value of the vaginal smears and the staging assessment of the female reproductive tract tissues with weak oestrogens, i.e., GN and NP. The cycle changes and histopathological data in these cases suggest that some utility in conducting vaginal smears. However, the data suggest greatest utility lies in the pathologists staging of the tissues and comparing the histological findings with the smears. The data further suggest that proper evaluation of the cycle requires a more detailed and thorough assessment of the histology of the female reproductive tract. The premise is that only a detailed assessment of those histological features directly related to the tissues' (ovaries, uterus, cervix, and vagina) responses to oestrogens through the oestrous cycle will be able to determine whether the synchronization of the reproductive tract has or has not been disturbed. Importantly, the various tissue components in the reproductive tract can only respond to an oestrogenic stimulus with a limited repertoire. Furthermore, the individual elements of this repertoire are natural features, not abnormal. All of these normal features will be found at some point in the natural cycle. It is only by careful assessment and staging of the detailed features diagnostic of oestrogen that any perturbation can be distinguished. This additional care in examination and data interpretation may require a change in procedures and some specialised understanding. This suggests additional training in interpreting of smear data and histopathology of the female reproductive tract outside the current TG 407 and similar to reproductive studies would be desirable as the successful detection of weak oestrogens would appear to depend heavily on the expertise of the reporting personnel.

Table 97. Number of animals with abnormal oestrous cycle as determined by smear on day 28 of study,at sacrifice (day 28-32variable) or at day 14 of smearing

		1st day terminations				Last date terminations				Up to 14 days smearing (if available)			
Chemical		Control	Low	Mid	High	Control	Low	Mid	High	Control	Low	Mid	High
EE	2	1	5	7	4	4	4	6	7				
	5												
GN	4	1	0	1	0	1	0	1	0				
	12	1	3	6	9	3	6	8	10				
NP	1	3	4	3	7	5	4	3	7				
	6	1	3	7	6	3	4	8	7				
TAM	3	0	0	5	10	1	0	9	10				
	10	2	6	5	6	2	6	6	6				
CGS	8												
	13	6	10	10	10	6	10	10	10				
MT	3	7	9	10	8	7	9	10	9				
	12	0	4	6	10	1	4	6	10	1	6	9	10
FLU	2	0	1	0	2	0	1	0	1	0	1	0	1
	11	2	2	2	7	3	4	3	10	5	5	6	8
DDE	6	1	1	2	1	2	2	4	2				
	7												
PTU	1	0	5	2	1	0	4	0	2				
	10												
THY	9	3	6	5	7	4	6	6	7				
	13	8	8	9	7	8	8	9	7				

273. Are vaginal smears and the complex, staggered necropsy necessary? Greater clarification of the TG 407 objectives and the technical feasibility of the procedures may be necessary. First, an answer would depend on the imperative for the updated TG 407 to detect very weak oestrogens. The more complex and intensive procedure would be applicable only if the answer were affirmative. Second, this also might depend on whether structural or *in vitro* information would be used to prescreen compounds, so that the complex and intensive procedure would only be applied to positive candidates. Finally, this would also depend on what would be the most successful course for the detection of oestrous cycle disturbance:

1. an entirely random presentation where necropsy is performed on a single day with the females in various cycle stages or
2. the use of smears to present the control animals in a more standardised presentation that would, in theory, make the detailed comparisons less challenging.

274. In conclusion, the value of vaginal smears and a more detailed and extensive histopathological staging of the female reproductive tract in the detection of weak oestrogens remains unresolved. This reemphasises the previous recommendation for a consultation of appropriate experts and the publication a review of the histological examination of the female reproductive tract tissues. The needed outcome is the likely feasibility and recommended best approach for an assessment of the overall synchronization of these tissues in the oestrous cycle and how to interpret potential alterations in the context of an updated TG 407 guideline.

Discussion and recommendations

275. Prior to recommending the inclusion or exclusion of particular proposed endpoints, a discussion of current histopathological practice in TG 407 studies is in order. In almost every instance where a proposed update appears to provide beneficial information to assist in the identification of possible endocrine modulators, the histopathological evaluation of tissues appears to play an essential role. In several areas, the male reproductive tract, the female reproductive tract, the male and female mammary glands, and possibly the pituitary, there appears to some lack of guidance and inexperience on how to read and interpret tissues, differences in the use of terminological that makes comparisons of studies difficult, and some differences in grading that may inhibit consistency and sensitivity. Therefore, high quality and thorough histopathological examination appears to be essential to using the TG 407 guideline to identify possible endocrine modulators. In addition, full reporting of the target tissues with consistent terminology, criteria, and grading is needed.

276. This leads to two major recommendations. First, consultations of pathological experts are needed to provide guidance on the detection of endocrine modulators, particularly those with weak potencies, in three areas: the male reproductive tract, the female reproductive tract, and the male and female mammary glands. Similar consultations may be useful to consider in a fourth area: the pituitary. The objective would be guidance on the range of observations that should be made (e.g., specific observations of the Leydig cells and secretory activity of ventral prostate), the detection of subtle changes in these target tissues, and particularly the apparent need for a mandatory oestrus cycle staging of the female reproductive tract. In the case of the adrenals and thyroid, current practice appears to be adequate based on the performance in these studies, and no consultation is recommended in those areas.

277. Second, the degree of blinding of the pathologist in TG 407 studies needs to be thoroughly considered. As noted, the histopathological changes are often subtle, making the recognition of changes and changes in normal patterns more technically challenging. Together with improved guidance on practice, identification of substances is likely to be improved where the pathologist is informed of 1) structural, *in vitro*, or other *in vivo* alerts as the possible endocrine activity of a test substance and 2) changes in both the significance and trends in the absolute and relative weights of particular tissues.

278. In conclusion, the following recommendations are made as to the addition of updates to the current TG 407 guideline.

279. The following endpoints should be included in an updated TG 407 guideline for the detection of toxicants with possible (anti)oestrogenic, (anti)androgenic, and antithyroid modes of action:

- Absolute and relative tissue weights of the male reproductive tract, including testes, epididymides, ventral and dorsolateral prostate and seminal vesicles with coagulating glands.
- Absolute and relative tissue weights of the female reproductive tract, including the paired ovaries and uterus.
- Histopathology of the pituitary.
- Absolute and relative thyroid weights and histopathology of the thyroid. The recommendation for the addition of the thyroid weights is qualified in that the dissection and trimming procedures must not compromise the more valuable thyroid histopathology.
- Histopathology of the male and female reproductive tract tissues (adding cervix and vagina and mandatory staging assessment).
- Histopathology of the male and female mammary glands.

280. T_3 , T_4 and TSH determinations need not be carried out on a routine basis, but they may be useful as optional endpoints in an updated TG 407 guideline for the confirmation of toxicants with antithyroid modes of action. Analysis of retained serum or blood samples might be useful in some cases where thyroid histopathology was positive to more closely define the mode of action. However, widespread and routine analysis appears to be redundant to the histopathology.

281. The following endpoints were not included in the protocol for these validation studies and therefore not investigated systematically. It is proposed that they should be included in an updated TG 407 guideline for the detection of toxicants with weak oestrogenic modes of action after histopathological practice to improve the observation and interpretation of these endpoints. After more experience has been gained a final decision should be made whether or not to retain these endpoints in the test guideline. Basis for this decision should be whether these investigations help to identify weak oestrogens, i.e. those not being clearly positive by the above mentioned obligatory endpoints (cf. para 278) but which may be identified by higher tier test systems. The aim of this review is to define best current practices and approaches for:

- Estrous cycle evaluation, using daily vaginal smears and an integrated histopathological evaluation of the female reproductive tract.

282. The following endpoints may be included in an updated TG 407 guideline as a supplemental endpoint to assist the detection of potential endocrine modulators. Positive findings should not be used in isolation to identify possible endocrine modulators, since they may result either from direct effects on the pituitary-adrenal axis or rather stem from unspecific stress caused by the chemical.

- Absolute and relative adrenal weights and histopathology of the adrenals.

283. The following endpoints are not recommended for inclusion in an updated TG 407 guideline. These studies have not shown these endpoints to be reliable or beneficial endpoints for the identification of potential endocrine modulators:

- Absolute and relative pituitary weights
- Spermatology endpoints, including sperm count and morphology.

ASSESSMENT OF GROUP SIZE AND POWER FOR THE UPDATED TG 407

284. This section assesses the utility of increasing the group size or the number of animals per group for detecting endocrine mediated effects in the updated TG 407 studies. The assessment is conducted in three steps: 1) basic statistical power calculations, 2) a review of the coefficients of variation found for current and proposed updated endpoints, and 3) a review of the results achieved with Subgroups of 5 animals per sex and with combined Subgroups of 10 animals per sex. The assessment concludes that, for potent test substances, 5 animals per sex are sufficient to detect endocrine modulation with (anti)oestrogenic, (anti)androgenic, and thyroid toxicants. However, in order to consistently detect modestly and weakly potent compounds, 10 animals per sex may be necessary in conjunction with improvements in practice such as histopathological examination discussed in the previous section of this report.

285. As in the previous section, the proposed update endpoints are considered together with applicable current TG 407 endpoints. The assessment does not consider here the role of statistical methods (e.g., the choice of a paired t-test method or Dunnett's method). In that regard, the power considerations in this section focus largely on continuous data (e.g., tissue weights, hormonal levels, and sperm numbers) and not discontinuous data (e.g., histological grades). The assessment is based on the assumption that there are three major inherent factors that impact the ability of endpoint to achieve statistical significance in response to endocrine modulation. These three factors are:

1. The magnitude of the change from starting value (where the starting value represents the vehicle control parameter) induced in the given endpoint by the test substance. It is the magnitude of this change that is analyzed for statistical difference between the vehicle control group and the treatment groups. The percent change of an endpoint is used in the power calculations.
2. The coefficient of variation (CV) for the endpoint; a lower CV improves the ability to statistically detect significance and a higher CV diminishes the ability. It is assumed that there are two basic subfactors in the CV: a) the inherent biological variation in the endpoint itself and b) the impact of laboratory technique such as the dissection, tissue handling, or the analytical method.
3. The sample size or the number of animals per group. There is generally an increase in power, particularly, with small group sizes as animals are added. However, this increase diminishes and is not linear, so that the benefit of adding further animals to large group sizes may be comparatively small in some circumstances.

Power calculations

286. The objective of the power calculations is to produce a technical basis to initially assess the possible differences between individual Subgroups of 5 animals per sex and combined Subgroups of 10 animals per sex in achieving statistical significance with a large number of different chemical and biological measurements.

287. The power calculations were performed under the following conditions: 1) the statistical method was Dunnett's (one-sided); 2) a total of four groups were used with three treatment and one vehicle control groups; the significance level was alpha or $p < 0.05$; and 3) the highest dose was presumed to have changed versus the control with the low and intermediate doses similar to the control (i.e., the condition considered to have the lowest overall parameter response that would likely be to achieve significance in real studies). Five values for percentage or magnitude of the parameter increase were used in the power calculations: 10%, 20%, 30%, 40%, and 50%. Four values for the CV were used in the power calculations: 10, 15, 20, and 25. The latter was the highest value used, although it was recognized that some values

could be higher (e.g., the mean CV for the TSH levels was 39.2, see Table 92). The power calculations were run for the two group sizes of interest: n = 5 and n = 10.

Table 98. Power calculations for group size in the updated TG 407 studies.

Approximate power (%) for detecting top dose effect				
% parameter change in the top dose group	**CV**	**n = 5**	**n = 10**	**% Increase in Power**
10	10	28.0	51.6	84.3%
	15	15.7	26.5	68.8%
	20	10.4	17.3	66.3%
	25	8.1	11.6	43.2%
20	10	73.2	96.6	32.0%
	15	42.0	71.4	70.0%
	20	26.6	46.7	75.6%
	25	18.9	33.1	75.1%
30	10	95.1	100.0	5.2%
	15	70.2	95.0	35.3%
	20	47.2	77.6	64.4%
	25	33.0	57.5	74.2%
40	10	99.4	100.0	0.6%
	15	86.8	99.5	14.6%
	20	65.6	93.0	41.8%
	25	47.9	78.3	63.5%
50	10	99.9	100.0	0.1%
	15	95.9	100.0	4.3%
	20	79.3	98.2	23.8%
	25	61.8	90.6	46.6%

288. The results of the calculations are shown in Table 98 and graphically in Figure 1. The results in Table 1 have also been used to calculate the increase in power provided by using 10 animals versus five (see the far right column in Table 1).

289. Table 1 and Figure 1 immediately illustrate two conclusions: 1) doubling the number of animals in the TG 407 never doubles the power in any circumstance presented; the power is always increased by less than 90%, and 2) groups sizes of 5 and 10 yield basically equivalent power (<15% increase in power) under some circumstances such as increases in the parameter of 30, 40, and 50% with a CV value of 10 and increases of 40 and 50% with a CV value of 15. Thus, increasing the group size for a parameter that responds with a large dynamic range (a change of 30-40%) and that can be measured accurately (a CV of 10-15) provides limited benefit. However, in other circumstances, an increase in group size does provide a varying degree of benefit by increasing power.

290. A third conclusion should also be explicitly stated. An alternative strategy to increasing the number of animals is to increase the statistical power by improving the measurement of the parameter. In other words, reducing the CV for the measured could increase the power, such as improving the laboratory dissection of a tissue or the performance of an analytical technique.

Coefficients of variation observed in the updated TG 407 studies

291. The calculated CVs for several current and updated absolute tissue weight endpoints are reported in Tables 99A-99J. The CV means, medians, minimums and maximums for the vehicle control groups and for all four groups (combining control and treatment groups) are reported in Table 100. The tissues include current, standard weights such as liver and adrenals, small and difficult to dissect tissues such as

the pituitary and thyroid, the various tissues in the male reproductive tract (testes, epididymides, ventral and dorsolateral prostate, and seminal vesicles), and female reproductive tract (ovaries and uterus).

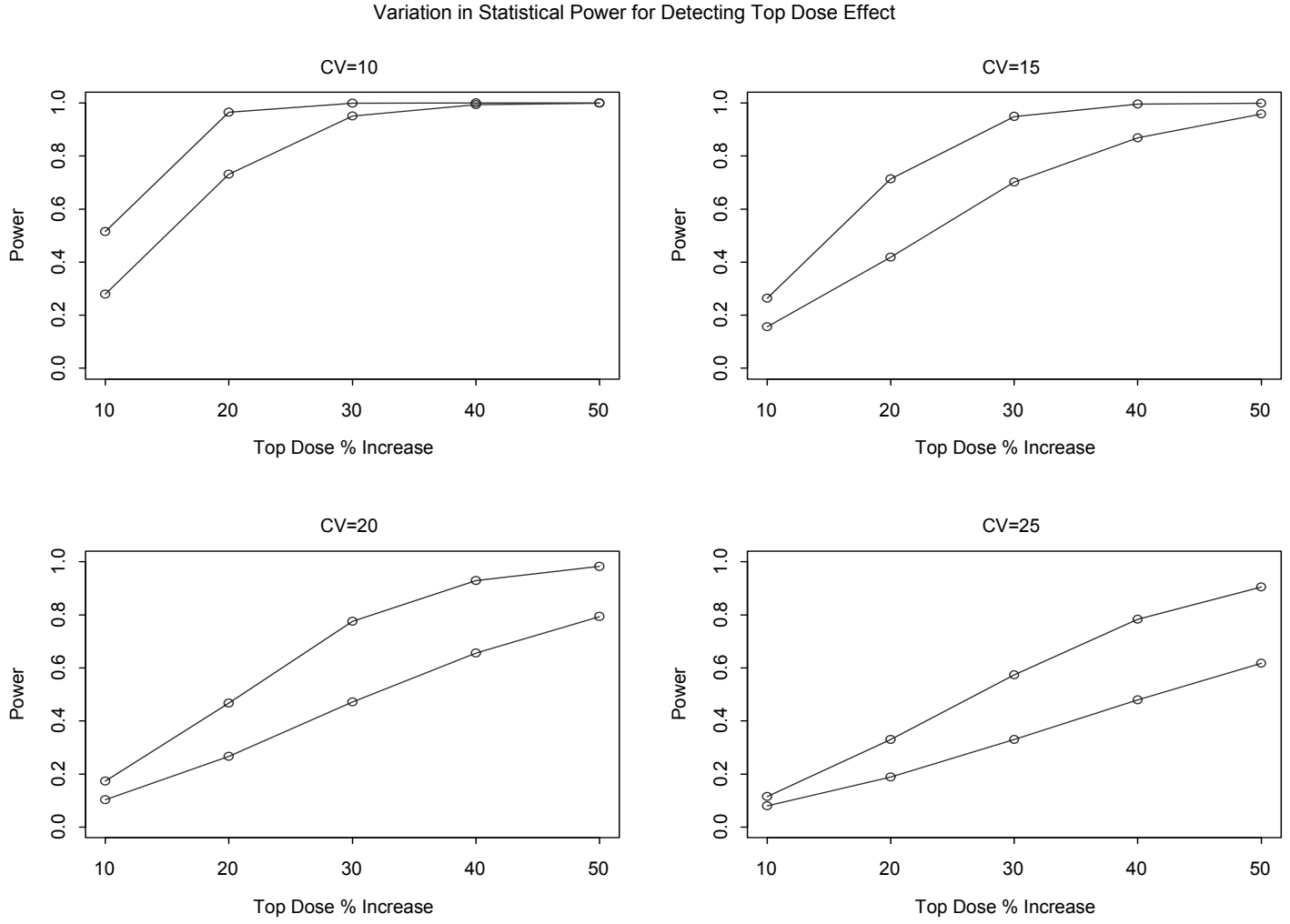

Figure 1. Comparison of power for groups sizes of n=5 and n=10 to detect a significant difference in the updated TG 407 studies.

Conditions are a one-sided Dunnett's analysis with the third of three treatment doses significant versus the control. Calculations are with different coefficients of variation and percentage increase in a parameter. Red line and circles are n=10 groups size and blue n=5.

Table 99A. Coefficients of variation for liver weights in the updated TG 407 studies.

Chemical	Lab	Liver - Male				Liver - Female			
		Control	Low	Mid	High	Control	Low	Mid	High
EE	2	10.7%	9.8%	15.1%	14.0%	9.4%	12.3%	13.2%	13.0%
	5	11.7%	14.2%	9.0%	13.8%	5.6%	13.9%	11.5%	12.1%
GN	4	10.0%	8.4%	14.7%	10.4%	9.7%	12.0%	11.2%	10.5%
	12	14.5%	8.1%	14.0%	6.7%	16.5%	8.2%	7.2%	13.1%
NP	1	10.0%	9.0%	6.3%	11.8%	9.1%	10.2%	11.8%	12.6%
	6	8.3%	9.4%	12.8%	14.0%	7.6%	5.0%	8.6%	14.3%
TAM	3	11.2%	13.8%	9.3%	4.9%	9.3%	6.4%	8.6%	6.2%
	10	15.2%	10.7%	11.0%	6.5%	12.2%	9.3%	9.3%	9.1%
CGS	8	12.3%	13.0%	10.4%	13.5%	5.7%	9.3%	10.0%	14.1%
	13	11.9%	8.5%	7.1%	11.1%	14.9%	7.9%	7.6%	7.2%
MT	3	12.2%	16.8%	14.9%	14.6%	10.1%	8.6%	17.3%	15.0%
	12	12.1%	11.3%	8.7%	8.7%	10.3%	5.6%	11.8%	8.0%
FLU	2	9.5%	13.3%	14.1%	9.7%	10.1%	11.3%	10.3%	5.6%
	11	11.6%	11.3%	7.9%	11.1%	8.1%	12.2%	10.2%	9.4%
DDE	6	13.7%	10.8%	9.6%	6.7%	8.0%	11.6%	11.6%	16.1%
	7	11.1%	8.1%	8.2%	7.3%	10.4%	10.6%	12.9%	19.4%
PTU	1	11.4%	13.2%	12.7%	8.7%	8.3%	7.6%	9.7%	10.5%
	10	11.1%	18.7%	8.7%	38.0%	11.5%	15.8%	9.6%	14.9%
THY	9	11.9%	8.4%	18.1%	7.9%	11.3%	13.7%	12.3%	10.4%
	13	10.4%	7.6%	7.4%	9.3%	10.3%	7.4%	9.0%	10.4%

Table 99B. Coefficients of variation for adrenal weights in the updated TG 407 studies.

Chemical	Lab	Adrenals - Male				Adrenals - Female			
		Control	Low	Mid	High	Control	Low	Mid	High
EE	2	12.6%	15.1%	12.5%	16.1%	15.2%	11.2%	17.8%	26.3%
	5	11.2%	14.9%	22.6%	18.5%	16.0%	14.8%	13.3%	15.6%
GN	4	14.3%	16.7%	16.1%	18.3%	11.9%	10.9%	13.4%	16.7%
	12	11.3%	12.5%	13.7%	8.2%	11.4%	15.4%	13.9%	19.3%
NP	1	11.5%	14.2%	14.5%	16.6%	6.9%	10.8%	10.2%	9.5%
	6	5.5%	12.1%	12.3%	13.8%	20.3%	16.4%	10.1%	18.5%
TAM	3	13.5%	20.4%	15.5%	11.5%	11.6%	9.5%	11.3%	12.5%
	10 [a]								
CGS	8	15.2%	16.5%	12.9%	27.5%	11.0%	13.6%	19.8%	22.6%
	13	14.2%	12.6%	11.8%	18.2%	12.7%	21.9%	19.2%	15.9%
MT	3	11.8%	21.7%	17.5%	15.8%	9.8%	7.1%	14.3%	21.5%
	12	14.8%	16.3%	20.0%	10.5%	9.7%	18.2%	20.0%	14.5%
FLU	2	14.2%	11.9%	17.8%	16.2%	18.4%	15.8%	17.1%	16.3%
	11	11.4%	9.5%	13.3%	14.5%	12.2%	11.5%	11.1%	9.8%
DDE	6	15.9%	16.4%	14.3%	18.8%	14.5%	17.4%	14.5%	8.6%
	7	17.2%	10.6%	21.9%	11.0%	12.5%	13.2%	8.1%	25.0%
PTU	1	17.2%	12.1%	17.5%	14.3%	15.3%	11.7%	10.8%	15.1%
	10 [a]								
THY	9	13.3%	15.7%	19.7%	8.3%	10.8%	11.9%	11.0%	10.9%
	13	10.4%	20.0%	11.4%	11.6%	16.7%	12.6%	12.1%	12.8%

[a] Individual, not paired, means and SDs, were reported (and individual data were not provided to calculate these).

Table 99C. Coefficients of variation for pituitary weights in the updated TG 407 studies.

Chemical	Lab	Pituitary - Male				Pituitary - Female			
		Control	Low	Mid	High	Control	Low	Mid	High
EE	2	19.0%	25.5%	26.0%	25.5%	16.2%	24.6%	15.4%	17.7%
	5	8.2%	8.0%	11.7%	18.4%	18.3%	10.1%	14.7%	14.5%
GN	4	14.2%	8.6%	12.1%	11.7%	14.5%	14.2%	14.5%	8.5%
	12	40.0%	20.0%	25.0%	20.0%	25.0%	16.7%	18.2%	16.7%
NP	1	12.6%	8.1%	18.9%	21.1%	12.1%	13.1%	12.5%	17.1%
	6	22.2%	25.0%	11.1%	22.2%	33.3%	30.0%	33.3%	40.0%
TAM	3	22.2%	22.2%	22.2%	16.7%	11.1%	22.2%	12.5%	28.6%
	10 [a]								
CGS	8	13.6%	11.4%	15.9%	9.9%	13.8%	13.6%	11.2%	7.8%
	13	23.5%	20.2%	24.1%	21.6%	29.4%	7.7%	16.4%	22.8%
MT	3	28.6%	28.6%	25.0%	28.6%	22.2%	22.2%	28.6%	42.9%
	12	22.2%	25.0%	28.6%	14.3%	22.2%	12.5%	14.3%	14.3%
FLU	2	7.3%	22.0%	18.2%	12.0%	17.7%	15.0%	11.8%	16.0%
	11	12.3%	9.4%	13.6%	9.2%	10.9%	10.3%	14.0%	20.7%
DDE	6	33.3%	11.1%	30.0%	11.1%	33.3%	16.7%	16.7%	10.0%
	7	16.0%	20.0%	13.3%	16.7%	7.2%	12.7%	12.2%	13.3%
PTU	1	14.8%	13.7%	15.1%	13.2%	25.4%	15.7%	14.6%	13.1%
	10 [a]								
THY	9	10.3%	9.1%	9.3%	7.3%	13.6%	15.1%	9.2%	13.2%
	13	26.3%	24.5%	17.3%	22.6%	10.6%	22.1%	17.1%	39.2%

[a] The resolution of the means and SDs were too low (0.01 grams) to calculate realistic CV values.

Table 99D. Coefficients of variation for thyroid weights in the updated TG 407 studies.

Chemical	Lab	Thyroid - Male				Thyroid - Female			
		Control	Low	Mid	High	Control	Low	Mid	High
EE	2	18.8%	13.1%	15.6%	8.8%	11.1%	12.4%	8.3%	12.9%
	5	15.1%	21.2%	14.8%	21.8%	14.1%	14.5%	18.1%	15.3%
GN	4	29.3%	27.9%	20.4%	16.1%	19.1%	14.1%	26.3%	19.4%
	12	21.4%	20.0%	7.1%	13.3%	25.0%	13.3%	18.8%	13.3%
NP	1	17.3%	15.4%	17.9%	14.3%	16.0%	13.1%	12.6%	16.0%
	6	19.1%	22.6%	25.7%	26.9%	17.1%	30.3%	39.4%	36.3%
TAM	3	16.7%	35.0%	20.0%	16.7%	21.4%	20.0%	23.1%	20.0%
	10 [a]								
CGS	8	19.7%	21.6%	19.2%	21.2%	13.0%	16.1%	24.2%	16.7%
	13	28.1%	20.6%	18.6%	19.3%	26.9%	19.7%	16.2%	11.8%
MT	3	16.7%	35.3%	22.2%	35.0%	35.7%	28.6%	22.2%	31.6%
	12	13.3%	20.0%	14.3%	11.8%	15.4%	33.3%	7.1%	16.7%
FLU	2	45.0%	15.6%	35.6%	41.3%	40.0%	47.5%	18.8%	45.0%
	11	12.1%	21.9%	14.9%	17.6%	11.5%	17.1%	18.4%	13.5%
DDE	6	21.0%	19.2%	14.9%	16.4%	17.9%	11.1%	15.4%	16.4%
	7	24.2%	24.7%	35.5%	22.2%	34.8%	22.0%	19.2%	28.8%
PTU	1	21.5%	19.2%	19.7%	19.3%	16.6%	33.4%	23.2%	23.8%
	10 [a]								
THY	9	17.6%	20.3%	17.7%	13.1%	12.4%	19.7%	11.0%	13.0%
	13	17.6%	19.5%	23.9%	20.2%	18.6%	16.8%	12.8%	14.2%

[a] Thyroid weights were not recorded.

Table 99E. Coefficients of variation for testicular weights in the updated TG 407 studies.

Chemical	Lab	Testes - Male			
		Control	Low	Mid	High
EE	2	11.7%	16.5%	24.7%	14.5%
	5	5.7%	7.4%	8.3%	5.7%
GN	4	10.9%	5.9%	7.5%	4.5%
	12	5.6%	5.5%	9.6%	6.6%
NP	1	5.0%	6.9%	9.9%	5.8%
	6	7.0%	26.7%	5.7%	8.6%
TAM	3	8.2%	9.9%	7.3%	7.4%
	10 [a]				
CGS	8	8.9%	11.4%	6.8%	6.1%
	13	7.2%	9.0%	3.3%	7.8%
MT	3	5.3%	7.5%	18.4%	27.3%
	12	5.9%	6.1%	5.9%	17.0%
FLU	2	6.0%	6.8%	8.7%	17.7%
	11 [a]				
DDE	6	6.7%	6.0%	6.8%	11.5%
	7	8.6%	10.0%	8.8%	9.3%
PTU	1	5.6%	8.3%	6.9%	7.2%
	10 [a]				
THY	9	12.9%	5.7%	7.1%	10.1%
	13	9.2%	10.7%	6.4%	9.5%

[a] Individual, not paired, means and SDs, were reported.

Table 99F. Coefficients of variation for epididymidal weights in the updated TG 407 studies.

Chemical	Lab	Epididymides - Male			
		Control	Low	Mid	High
EE	2	9.1%	9.8%	24.1%	24.3%
	5	5.5%	6.5%	10.9%	8.5%
GN	4	6.5%	7.3%	8.4%	6.1%
	12	6.7%	5.5%	5.3%	6.2%
NP	1	8.6%	7.8%	7.6%	7.2%
	6	7.9%	15.6%	9.5%	11.0%
TAM	3	9.1%	7.1%	4.7%	6.9%
	10 [a]				
CGS	8	6.6%	9.9%	5.4%	6.4%
	13	7.7%	7.8%	6.5%	8.0%
MT	3	6.2%	10.9%	20.7%	20.7%
	12	6.3%	9.1%	5.6%	10.7%
FLU	2	9.6%	15.0%	12.1%	24.2%
	11 [a]				
DDE	6	6.2%	31.3%	5.9%	4.1%
	7	6.8%	7.4%	9.1%	7.5%
PTU	1	8.9%	6.6%	9.0%	7.8%
	10 [a]				
THY	9	13.2%	7.1%	5.9%	11.6%
	13	12.1%	9.1%	10.0%	8.8%

[a] Individual, not paired, means and SDs, were reported.

Table 99G. Coefficients of variation for prostate weights in the updated TG 407 studies.

Chemical	Lab	Ventral or Whole [a] Prostate - Male				Dorsolateral Prostate - Male			
		Control	Low	Mid	High	Control	Low	Mid	High
EE	2	18.0%	19.6%	18.7%	36.8%	15.5%	15.7%	28.9%	35.2%
	5	16.1%	29.0%	31.0%	40.9%	14.8%	18.9%	19.4%	56.3%
GN	4	24.5%	24.0%	20.4%	17.4%	31.3%	11.1%	19.5%	21.1%
	12	19.1%	19.2%	22.0%	13.9%	15.0%	15.8%	15.7%	16.1%
NP	1 [a]	16.2%	13.3%	19.2%	20.2%				
	6	39.6%	52.7%	50.4%	38.6%	28.4%	29.2%	25.6%	44.7%
TAM	3	27.4%	12.3%	22.4%	22.0%	43.2%	28.3%	32.4%	32.1%
	10 [a]	11.3%	12.3%	34.4%	8.5%				
CGS	8	19.4%	21.3%	24.3%	14.1%	18.4%	27.1%	20.6%	19.7%
	13 [a]	20.0%	17.5%	13.2%	14.5%				
MT	3	21.2%	14.7%	30.8%	15.6%	38.7%	22.7%	29.7%	18.1%
	12	30.3%	31.7%	27.2%	34.9%	11.0%	19.2%	17.3%	22.2%
FLU	2	18.1%	13.3%	22.1%	48.4%				
	11	21.2%	25.0%	24.0%	29.4%	23.1%	20.6%	22.2%	12.5%
DDE	6	19.9%	18.5%	20.1%	27.9%	34.3%	43.1%	41.2%	26.7%
	7	15.5%	17.9%	16.8%	16.4%	13.5%	12.6%	23.3%	18.0%
PTU	1	18.8%	20.5%	18.5%	23.6%	23.7%	11.4%	17.7%	15.8%
	10 [a, b]								
THY	9	10.0%	14.3%	20.0%	20.8%	19.2%	16.2%	14.3%	17.0%
	13 [a]	22.2%	19.5%	20.0%	14.6%				

[a] Whole prostate weights, highlighted in light gray to distinguish

[b] SDs recorded were extremely small so that calculated CVs appear unrealistically low and are not reported here.

Table 99H. Coefficients of variation for seminal vesicle weights in the updated TG 407 studies.

Chemical	Lab	Seminal Vesicles - Male			
		Control	Low	Mid	High
EE	2	19.7%	16.5%	18.2%	32.6%
	5	8.4%	16.7%	35.8%	51.5%
GN	4	13.4%	20.6%	17.6%	12.3%
	12	20.2%	16.1%	22.9%	23.2%
NP	1	23.0%	6.0%	18.6%	18.2%
	6	11.2%	22.1%	20.3%	36.9%
TAM	3	20.3%	9.4%	24.7%	14.4%
	10 [a]				
CGS	8	13.3%	20.5%	16.7%	12.1%
	13	19.6%	17.8%	16.7%	16.3%
MT	3	21.0%	20.9%	51.8%	19.1%
	12	24.2%	20.2%	15.3%	29.1%
FLU	2	16.3%	14.5%	20.7%	14.4%
	11	13.2%	15.0%	16.1%	11.8%
DDE	6	18.4%	10.9%	14.6%	15.2%
	7	13.5%	13.8%	14.5%	15.6%
PTU	1	21.2%	24.9%	23.0%	19.6%
	10 [a]				
THY	9	11.9%	6.4%	22.6%	12.6%
	13	17.0%	21.6%	12.2%	10.5%

[a] Seminal vesicle weights were not recorded.

Table 99I. Coefficients of variation for ovarian weights in the updated TG 407 studies.

Chemical	Lab	Ovaries - Female			
		Control	Low	Mid	High
EE	2	13.9%	6.4%	12.3%	18.1%
	5	25.7%	20.1%	14.2%	23.2%
GN	4	15.0%	12.0%	16.3%	17.8%
	12	13.7%	13.6%	12.9%	20.0%
NP	1	12.7%	11.8%	17.5%	16.9%
	6	15.1%	30.5%	17.6%	34.8%
TAM	3	33.3%	16.9%	10.4%	18.6%
	10 [a]				
CGS	8	11.2%	15.6%	7.4%	17.1%
	13	15.0%	18.9%	11.3%	10.7%
MT	3	17.9%	12.3%	29.3%	20.6%
	12	11.8%	8.7%	19.6%	26.0%
FLU	2	15.1%	17.7%	18.6%	11.4%
	11	14.4%	15.3%	17.3%	17.9%
DDE	6	16.2%	16.5%	22.1%	21.8%
	7	9.6%	13.5%	13.0%	18.4%
PTU	1	20.4%	16.1%	35.7%	23.0%
	10 [a]				
THY	9	17.4%	8.5%	10.5%	15.6%
	13	21.6%	8.4%	17.9%	21.3%

[a] Individual, not paired, means and SDs, were reported.

Table 99J. Coefficients of variation for uterine weights in the updated TG 407 studies.

Chemical	Lab	Uterus - Female			
		Control	Low	Mid	High
EE	2	15.0%	63.7%	22.3%	49.0%
	5	14.8%	11.4%	7.5%	12.7%
GN	4	20.5%	15.0%	18.0%	20.5%
	12	19.3%	13.4%	16.5%	12.8%
NP	1	41.9%	54.0%	22.5%	17.5%
	6	13.6%	17.8%	15.9%	21.6%
TAM	3	14.0%	38.7%	14.6%	13.6%
	10	23.7%	30.8%	23.5%	10.0%
CGS	8	9.3%	19.2%	12.0%	11.0%
	13	7.7%	25.5%	19.6%	18.2%
MT	3	39.2%	15.0%	25.7%	46.1%
	12	31.9%	30.1%	15.3%	19.7%
FLU	2	13.2%	30.8%	7.7%	49.1%
	11	14.6%	11.6%	17.5%	14.3%
DDE	6	11.6%	21.5%	19.1%	21.1%
	7	13.4%	15.2%	12.7%	17.0%
PTU	1	17.3%	15.0%	10.1%	10.3%
	10	17.1%	13.9%	17.1%	20.0%
THY	9	15.4%	18.4%	45.9%	40.8%
	13	22.2%	20.5%	16.7%	12.3%

Table 100. Means, medians, minimums, and maximums for coefficients of variation.

Tissue	Vehicle control group mean CVs				Overall mean CVs Combining vehicle and treatment groups			
	mean	median	min	max	mean	median	min	max
Liver - M	11.54	11.50	8.3	15.2	11.30	11.05	4.9	38.0
Liver - F	9.92	9.90	5.6	16.5	10.54	10.30	5.0	19.4
Adrenal - M	13.08	13.40	5.5	17.2	14.72	14.30	5.5	27.5
Adrenal - F	13.16	12.35	6.9	20.3	14.17	13.35	6.9	26.3
Pituitary - M	19.26	17.50	7.3	40.0	18.03	17.75	7.3	40.0
Pituitary - F	18.71	16.95	7.2	33.3	17.71	15.05	7.2	42.9
Thyroid - M	20.81	18.95	12.1	45.0	20.57	19.40	7.1	45.0
Thyroid - F	20.37	17.50	11.1	40.0	20.13	17.50	7.1	47.5
Testes	7.67	7.00	5.0	12.9	9.11	7.45	3.3	27.3
Epididymides	8.06	7.70	5.5	13.2	9.58	7.85	4.1	31.3
Ventral Prostate	21.50	19.65	10.0	39.6	23.58	21.00	10.0	52.7
Dorsolateral Prostate	23.58	21.15	11.0	43.2	23.31	20.15	11.0	56.3
Whole Prostate	17.56	18.10	11.3	22.2	18.94	17.80	8.5	48.4
Seminal Vesicles	16.99	17.70	8.4	24.2	18.71	17.30	6.0	51.8
Ovaries	16.67	15.05	9.6	33.3	16.94	16.40	6.4	35.7
Uterus	18.79	15.20	7.7	41.9	20.69	17.20	7.5	63.7

M – male; F – female.

292. The following conclusions are drawn from these and other data:

- The overall CV results are reasonable based upon expectations. Larger and easier to dissect tissues tend to have lower CVs; examples are the liver, testes, and epididymides with mean CVs ranging from about 8 to 11. As tissues become smaller and more difficult to dissect, the CVs increase. The adrenal mean CVs are 12-14, the pituitary 18-19, and the thyroid 20-21. Fluid-filled male accessory tissues that are difficult to dissect, such as the prostate lobes, have higher CVs with mean CVs approaching 24.
- The CVs for the uterus and for the male accessory tissues are consistent with those observed in the uterotrophic and Hershberger validation programs (36)(56).
- As with the uterotrophic and Hershberger validation programs, a review of Tables 99A-J indicates apparent laboratory-related differences in CVs in several instances. This is again consistent with previous conclusions (36)(56) that laboratory technique in the dissection and tissue handling was an important variable for the CV values. This emphasizes the need for technical training and monitoring the performance of the laboratories that conduct the TG 407 and other assays.
- The updates tend to have somewhat higher CV values as they are smaller tissues, several of these tissues are more difficult to dissect, and often they may contain fluid. Therefore, based on the calculations in Table 98 for a CV of 20, the increase in group size to 10 would increase the power by 20-70% for these endpoints depending upon the observed magnitude of change.
- The CV values indicate that the previous power calculations are representative. The range in the CV values there was 10 to 25. The lowest median values are about 8 for the testes and the epididymides. The highest median values are for the ventral and dorsolateral prostate of 20-21.
- The power calculations are also applicable to T_3 and T_4 analyses, but do not cover the larger CVs for TSH (see Table 92).

293. The key conclusion is that increased group sizes of 10 animals per sex per group can in theory moderately improve the power of both the current and the updated TG 407 endpoints. However, the magnitude of this increase in power will depend on the percent change that a given chemical induces at a give dose, the CV of a given tissue measurement, and the how the capability of a given laboratory influences this CV. The magnitude of this increase in power or an improved ability to achieve statistical

significance will then vary not only by endpoint, but by the test substance's mechanism, potency, and dose and to some degree, by laboratory. The sensitivity was investigated here in terms of statistically significant differences from the control group. An alternative approach, e.g. the benchmark calculation, which uses the complete dose response assessment to calculate the critical dose level for each endpoint, may be applied to evaluate the data more efficiently. This may also help to overcome the need for doubling the animal numbers in the protocol, which may only marginally increase power.

Comparison of results from individual Subgroups and combined Subgroups

294. The number of parameters and the corresponding number of opportunities for statistical significance needs to first be appreciated. The updated TG 407 contains approximately 80 measured or graded parameters, including the functional observational battery and so on. This number of parameters could potentially expand with additional histopathology of certain tissues or the measurement of specialized hepatic enzymes (e.g., certain P450s). This base number is then doubled to approximately 160 when both sexes are considered, and then expands another three-fold to 480-500 when three dose groups are considered. This means the opportunity exists for 480-500 occurrences of statistical significance in each study.

295. Besides true results of statistical significance, there are false results of two types: false positives and false negatives. The false positive rate is basically established by the alpha or p value, i.e., for $p < 0.05$ about 1 in 20 cases of findings of statistical significance is accepted to be an erroneous false positive. The false negative rate, where a true positive is missed and not found to be significant, is heavily influenced by the power of a study. Thus, there should be more findings of significance with group sizes of 10, with fewer instances in which the group of 10 failed to find statistical significance (a false negative).

296. The use of 10 animals per sex per group could then change the assessment of the results in two principle ways. First, in the instance where statistical significance was achieved for an endpoint in one individual Subgroup but not the other, the combined Subgroups would either achieve significance and substantiate the statistical significance of the one individual Subgroup or fail to achieve significance and thereby fail to substantiate the finding. Second, the combined Subgroups could achieve significance where neither of the two individual Subgroups achieved significance. If this occurred in the high dose group, then an endpoint would be significant with n=10 where it was undetected in both groups of n=5. If this occurred in the intermediate or low dose group, the NOEL and LOEL for that endpoint would be lowered.

297. Tables 101-110 present a comparison of the individual Subgroup results and those of the combined Subgroups for all 10 test substances. Where individual Subgroups are available in both laboratories, there is a Table A and B for both laboratories testing a substance. However, laboratory 10 only used a group size of 6 animals per sex and laboratory 13 did not divide their study into two individual Subgroups (see Table 11). Therefore, results from only 16 of the 20 studies are presented. Tables exclude the functional observational battery results, histopathology, and some other data in favor of continuous data such as body weights, blood and chemical chemistry, absolute and relative organ weights, hormonal values, and sperm counts. Further, although the laboratories did use the same test substance and doses, the laboratories did not always use the same animal strain or laboratory diet (see Table 4) or statistical methods (see Table 12). So there are uncontrolled variables to acknowledge in making this comparison.

298. Each Table provides for three doses and both males and females for each dose. Each Table is then divided into two sections. The upper section considers the situation, where one of the two individual Subgroups (n=5) achieved statistical significance and the other did not. If the combined Subgroups (n=10) achieved significance, then it is considered to have substantiated the individual Subgroup achieving significance and the appropriate table cell is marked with a '+'. If the combined Subgroup, however, did not achieve significance, the table cell is marked '-'. The lower section considers the situation where

neither individual Subgroup achieved significance, but the combined Subgroup did. This table cell is marked with an 'O' indicated that only the combined Subgroup achieved significance.

299. The Tables are arranged in the same order as the test substance sections (Table 101 is for EE, Table 102 is for GN, and so on).

Table 101A. Comparison of statistical significance between combined Subgroups and individual Subgroups with EE in Laboratory 2.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body Weight ↓					+	
T_4, total thyroxine ↑				+		+
TSH ↑				+		
Liver, absolute weight ↑				+		+
Liver, relative weight ↑			-	+		
Kidneys, absolute weight ↓					+	
Adrenals, relative weight ↑			+			
Brain, relative weight ↑					+	
Thyroid, relative weight ↑					+	
Spleen, absolute weight ↑						+
Dorsolateral prostate, relative wt. ↓		NA		NA	+	NA
Effects statistically significant in neither individual Subgroup at given dose						
TSH ↑						O
Liver, relative weight ↑		O				
Pituitary, relative weight ↑					O	
Uterus, relative weight ↑	NA		NA		NA	O

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

Table 101B. Comparison of statistical significance between combined Subgroups and individual Subgroups with EE in Laboratory 5.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight ↓					+	
T_4, total thyroxine ↑						+
Liver, relative weight ↑			+			
Kidney, absolute weight ↓					+	
Heart, absolute weight ↓					+	
Adrenal, absolute weight ↑					+	
Adrenal, relative weight ↑					+	
Pituitary, relative weight ↑					+	
Epididymes, abs. weight ↓					+	
Seminal vesicles, abs. weight ↓		NA	+	NA		NA
Seminal vesicles, relative weight ↓		NA		NA	+	NA
Ventral prostate, relative weight ↓		NA		NA	+	NA
Dorsolateral prostate, abs. wt. ↓		NA	+	NA		NA
Dorsolateral prostate, relative wt. ↓		NA	+	NA		NA
Ovary, absolute weight ↓	NA		NA		NA	+
Ovary, relative weight ↓	NA		NA		NA	+
Uterus, absolute weight ↑	NA	-	NA		NA	+
Uterus, relative weight ↑	NA		NA		NA	+
Effects statistically significant in neither individual Subgroup at given dose						
T_3, total T_3 ↓					O	
Liver, absolute weight ↑				O		O
Testes, relative weight ↑		NA		NA	O	NA
Seminal vesicles, relative weight ↓		NA	O	NA		NA
Ventral prostate, abs. weight ↓		NA	O	NA		NA

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

Table 102A. Comparison of statistical significance between combined Subgroups and individual Subgroups with Genistein in Laboratory 4.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Liver, relative weight ↑		-		-		-
Sperm counts ↑		NA	-	NA		NA
Effects statistically significant in neither individual Subgroup at given dose						
Liver, relative weight ↑			O		O	

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

Table 102B. Comparison of statistical significance between combined Subgroups and individual Subgroups with Genistein in Laboratory 12.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
T_3 ↓						-
TSH ↑				-		
Liver, absolute weight ↑						+
Liver, relative weight ↑		-		+		
Kidney, relative weight ↑	+				+	
Adrenal, absolute weight ↑					+	+
Adrenal, relative weight ↑						+
Thymus, absolute weight ↓			+	+ *	+	
Thymus, relative weight ↓					+	-
Thyroid, absolute weight ↑		+				
Effects statistically significant in neither individual Subgroup at given dose						
Adrenal, relative weight ↑				O		
Thyroid, relative weight ↑		O		O *		
Uterus, absolute weight ↑	NA		NA	-	NA	O
Uterus, relative weight ↑	NA		NA		NA	O

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 103. Comparison of statistical significance between combined Subgroups and individual Subgroups with Nonylphenol in Laboratory 1.

		Low dose		Mid Dose		High Dose	
		male	female	male	female	male	female
Effects statistically significant in one of two individual Subgroups (n = 5)							
T_3	↑					+	
T_4	↑			+			
Liver, absolute weight	↑		-				-
Liver, relative weight	↑			+	+		+
Kidney, absolute weight	↑		-			+	+
Kidney, relative weight	↑			+	+		+
Spleen, relative weight	↓		-				
Epididymes, absolute weight	↓		NA		NA	-	NA
Effects statistically significant in neither individual Subgroup at given dose							
Liver, absolute weight	↑		-			O	
Adrenals, absolute weight	↓						O
Heart, absolute weight	↓						O
Brain, relative weight	↑				O		O
Thyroid, relative weight	↑						O

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

Table 103B. Comparison of statistical significance between combined Subgroups and individual Subgroups with Nonylphenol in Laboratory 6.

		Low dose		Mid Dose		High Dose	
		male	female	male	female	male	female
Effects statistically significant in one of two individual Subgroups (n = 5)							
Body weight	↓					+	
T_4	↓						-
Liver, relative weight	↑				+		
Kidney, absolute weight	↑					+	
Adrenals, absolute weight	↑						+
Spleen, absolute weight	↓					+	
Thymus, absolute weight	↓	+					
Thymus, relative weight	↓	-					+
Epididymes, relative weight	↑		NA		NA	-	NA
Seminal vesicles, abs. weight	↓		NA		NA	+	NA
Prostate, absolute weight	↓		NA		NA	+	NA
Sperm counts	↓		NA	+	NA	+	NA
Effects statistically significant in neither individual Subgroup at given dose							
Liver, relative weight	↑					O	
Thymus, absolute weight	↓			O			O
Thymus, relative weight	↓			O			
Dorsolateral prostate, abs. wt.	↓		NA		NA	O	NA

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

Table 104. Comparison of statistical significance between combined Subgroups and individual Subgroups with Tamoxifen in Laboratory 3.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight, ↓				+		
T_4 ↑				-		+
TSH ↓						-
Liver, absolute weight ↓				-		
Kidney, absolute weight ↓						+
Kidney, relative weight ↓	+					
Kidney, relative weight ↑					+	
Adrenals, absolute weight ↓						+
Adrenals, relative weight ↑					+	
Brain, relative weight ↑				-		+
Pituitary absolute weight ↓					+	
Testes, relative weight ↑		NA		NA	+	NA
Epididymes, absolute weight ↓		NA	+	NA		NA
Epididymes, relative weight ↓	+	NA		NA		NA
Ventral prostate, absolute ↓		NA		NA	+	NA
Ventral prostate, relative ↓		NA		NA	+	NA
Dorsolateral prostate, absolute ↓		NA		NA	+	NA
Seminal vesicles, relative wt. ↓		NA		NA	+	NA
Ovary, absolute weight ↓	NA	-	NA		NA	
Uterus, absolute weight ↑	NA	+	NA		NA	
Uterus, relative weight ↑	NA	+	NA		NA	
Uterus, relative weight ↓	NA		NA		NA	+
Effects statistically significant in neither individual Subgroup at given dose						
Liver, absolute weight ↓		O *				
Heart, relative weight ↑			O		O	
Brain, absolute weight ↓					O	
Pituitary absolute weight ↓						O
Thymus, relative weight ↑						O
Uterus, absolute weight ↓	NA		NA	O	NA	

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 105. Comparison of statistical significance between combined Subgroups and individual Subgroups with CGS 18320B in Laboratory 8.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight, ↑		+				+
T_3 ↓	+	-	+		+	
T_4 ↓	+		+			
Liver, relative weight ↑				+	+	
Kidney, relative weight ↑					-	
Heart, absolute weight ↑		-				
Adrenals, relative weight ↑					+	
Brain, relative weight ↑					+	
Thymus, absolute weight ↑				+		
Whole prostate, absolute weight ↓		NA		NA	+	NA
Ovaries, absolute weight ↑	NA	+	NA		NA	
Ovaries, relative weight ↑	NA		NA		NA	+
Uterus, absolute weight ↓	NA	-	NA		NA	
Uterus, relative weight ↓	NA	-	NA		NA	
Effects statistically significant in neither individual Subgroup at given dose						
Brain, absolute weight ↑				O		O
Pituitary, absolute weight ↓		O *				
Spleen, relative weight ↑						O
Thyroid, relative weight ↓		O		O *		
Thymus, relative weight ↑						O
Testes, relative weight ↑		NA		NA	O	NA
Whole prostate, absolute weight ↓		NA	O	NA		NA
Ventral prostate, absolute weight ↓	O	NA	O	NA	O	NA
Ventral prostate, relative weight ↓	O	NA	O	NA		NA
Seminal vesicles, absolute wt ↓		NA		NA	O	NA
Ovaries, relative weight ↑	NA	O	NA	O	NA	
Sperm abnormalities, ↓	O *	NA		NA		NA

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 106A. Comparison of statistical significance between combined Subgroups and individual Subgroups with Methyl Testosterone in Laboratory 3.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight ↑						+
T_3 ↓				+		+
T_4 ↑						+
TSH ↑	+		+		+	
Liver, absolute weight ↑			-			
Liver, relative weight ↑				+	+	
Kidney, absolute weight ↑		+	+			
Kidney, relative weight ↑			+			
Adrenals, absolute weight ↑			-			+
Adrenals, relative weight ↓	-	+		+ *		
Adrenals, relative weight ↑					+	
Brain, absolute weight ↓	-					
Pituitary, relative weight ↓				+		+
Thymus, absolute weight ↓				-		
Dorsolateral prostate, relative wt ↑		NA		NA	+	NA
Ovaries, absolute weight ↓	NA		NA	+	NA	+
Ovaries, relative weight ↓	NA	+	NA		NA	
Uterus, absolute weight ↓	NA		NA		NA	-
Effects statistically significant in neither individual Subgroup at given dose						
Body weight ↓					O	
T_3 ↓		O				
T_4 ↑					O	
Liver, absolute weight ↑		O				
Kidney, relative weight ↑				O		
Adrenals, absolute weight ↓		O *				
Brain, absolute weight ↓					O	
Spleen, absolute weight ↑		O *				
Seminal vesicles, absolute wt ↓		NA	O	NA		NA
Seminal vesicles, relative wt ↓		NA	O	NA		NA
Uterus, absolute weight ↓	NA	O *	NA		NA	
Sperm counts ↓		NA	O	NA		NA
Sperm abnormalities ↑		NA	O	NA		NA

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 106B. Comparison of statistical significance between combined Subgroups and individual Subgroups with Methyl Testosterone in Laboratory 12.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight ↓	-					
T_3 ↓	-					
T_4 ↑						-
Liver, relative weight ↑				+		
Kidney, absolute weight ↑				+		
Kidney, relative weight ↑	+		+			
Heart, absolute weight ↑	+ *					
Heart, relative weight ↑		+			+	+
Adrenals, absolute weight ↓				+ *		
Adrenals, relative weight ↓		+				
Adrenals, relative weight ↑						+
Brain, absolute weight ↓	-					
Brain, relative weight ↓		+		+		
Pituitary, absolute weight ↓				+		+
Pituitary, relative weight ↓		+		+		+
Spleen, absolute weight ↓			+			
Thymus, absolute weight ↓			+			
Thymus, relative weight ↓					+	
Thyroid, relative weight ↑					+	
Testes, absolute weight ↓	-	NA		NA		NA
Epididymis, relative weight ↓		NA	+	NA		NA
Dorsolateral prostate, abs. wt. ↓		NA	+ *	NA		NA
Ovary, relative weight ↓	NA	-	NA		NA	
Uterus, absolute weight ↓	NA	+	NA		NA	
Uterus, relative weight ↓	NA	+	NA		NA	
Effects statistically significant in neither individual Subgroup at given dose						
T_4 ↑		O				
Liver, absolute weight ↑		O				
Heart, relative weight ↓				O		
Brain, absolute weight ↓						O
Thyroid, absolute weight ↑					O	
Testes, relative weight ↓		NA	O	NA		NA
Ventral prostate, absolute wt ↓		NA	O *	NA		NA
Dorsolateral prostate, rel. wt. ↓		NA	O *	NA		NA
Uterus, absolute weight ↓	NA		NA	O	NA	
Uterus, relative weight ↓	NA		NA	O	NA	

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 107A. Comparison of statistical significance between combined Subgroups and individual Subgroups with Flutamide in Laboratory 2.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight, ↓					+	
Adrenals, relative weight ↑					+	
Seminal vesicles, relative wt. ↓		NA	+	NA		NA
MARO, relative weight ↓		NA	+	NA		NA
Effects statistically significant in neither individual Subgroup at given dose						
T_3 ↓					O	
Kidney, absolute weight ↓	-		O			
Heart, absolute weight ↓					O	
Adrenals, relative weight ↑			O			
Brain, absolute weight ↓					O	
Brain, relative weight ↓						O
Pituitary, absolute weight ↓					O	
Spleen, absolute weight ↓					O	
Epididymes, relative weight ↓		NA	O	NA		NA
MARO, absolute weight ↓	O	NA		NA		NA
Sperm count ↓		NA		NA	O	NA

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at dose.

Table 107B. Comparison of statistical significance between combined Subgroups and individual Subgroups with Flutamide in Laboratory 11.

		Low dose		Mid Dose		High Dose	
		male	female	male	female	male	female
Effects statistically significant in one of two individual Subgroups (n = 5)							
T_4	↓						-
Liver, relative weight	↑	-		+			
Adrenal (l), relative weight	↑					+	
Adrenal (l), relative weight	↑					+	
Pituitary, relative weight	↓			+			
Thymus, relative weight	↑					+	
Epididymes (r), absolute weight	↓	+	NA		NA		NA
Epididymes (l), absolute weight	↓	+	NA	+	NA		NA
Epididymes, relative weight	↓	+	NA	+	NA		NA
Seminal vesicles, absolute wt.	↓		NA	+	NA		NA
Dorsolateral prostate, abs. wt.	↓		NA	+	NA		NA
Sperm abnormalities	↓		NA		NA	-	NA
Effects statistically significant in neither individual Subgroup at given dose							
Adrenal (r), absolute weight	↑					O	
Pituitary, absolute weight	↓					O	
Seminal vesicles, relative wt.	↓		NA	O	NA		NA
Ventral prostate, relative wt.	↓		NA	O	NA		NA
Dorsolateral prostate, rel. wt.	↓		NA	O	NA		NA
Ovary (l), absolute weight	↓	NA		NA		NA	O
Ovary (l), relative weight	↓	NA		NA		NA	O

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at dose.

Table 108A. Comparison of statistical significance between combined Subgroups and individual Subgroups with *p,p'*-DDE in Laboratory 6.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
Body weight ↓			-			
T_3 ↑				-		-
T_4 ↓					+	
TSH ↑	+				+	
Liver, absolute weight ↑	+					
Liver, relative weight ↑	-					
Kidney, relative weight ↑					+	
Thyroid, absolute weight ↑			-			+
Thyroid, relative weight ↑			-			
Epididymis, relative weight ↓	-	NA	-	NA	-	NA
Effects statistically significant in neither individual Subgroup at given dose						
T_3 ↑					O	
TSH ↑			O			
Thyroid, relative weight ↑						O
Sperm counts ↓	O *	NA		NA		NA

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 108B. Comparison of statistical significance between combined Subgroups and individual Subgroups with *p,p'*-DDE in Laboratory 7.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
T_3 ↑	+		+			
T_4 ↓						-
TSH ↑		-				
Liver, absolute weight ↑	+					
Kidney, absolute weight ↑			-			
Kidney, relative weight ↑			-			-
Adrenal, relative weight ↑					-	
Thyroid, relative weight ↑					+	

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

Table 109. Comparison of statistical significance between combined Subgroups and individual Subgroups with Propylthiouracil in Laboratory 1.

	Low dose		Mid Dose		High Dose	
	male	**female**	**male**	**female**	**male**	**female**
Effects statistically significant in one of two individual Subgroups (n = 5)						
T_3 ↓					+	
T_4 ↓			+		+	
TSH ↑	-			+		
Kidney, relative weight ↓						+
Heart, relative weight ↓			+		+	
Adrenal, relative weight ↓						+
Pituitary, relative weight ↑				+		
Spleen, absolute weight ↓		+ *		-		
Thymus, relative weight ↓						-
Thyroid, absolute weight ↑		-				
Thyroid, relative weight ↑		-				
Seminal vesicles, abs. wt. ↑	-	NA		NA		NA
Seminal vesicles, rel. wt. ↑		NA		NA	+	NA
Uterus, absolute weight ↓	NA	+	NA		NA	+
Uterus, relative weight ↓	NA	+ *	NA		NA	
Effects statistically significant in neither individual Subgroup at given dose						
T_4 ↓		O				
Liver, relative weight ↓						O
Heart, relative weight ↓						O
Brain, absolute weight ↓					O	
Brain, relative weight ↑				O		
Thymus, relative weight ↓					O	
Ventral prostate, rel. wt. ↑		NA		NA	O	NA
Uterus, absolute weight ↓	NA		NA	O	NA	

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.

* No dose response relationship was observed at higher doses.

Table 110. Comparison of statistical significance between combined Subgroups and individual Subgroups with l-Thyroxine in Laboratory 9.

	Low dose		Mid Dose		High Dose	
	male	female	male	female	male	female
Effects statistically significant in one of two individual Subgroups (n = 5)						
TSH ↓	-		-			
Liver, relative weight ↑						+
Heart, absolute weight ↑						+
Heart, relative weight ↑						+
Adrenal, absolute weight ↑				+		
Adrenal, relative weight ↑					+	+
Thyroid, relative weight ↓						+
Spleen, relative weight ↑						+
Effects statistically significant in neither individual Subgroup at given dose						
T_3 ↓				O *		
T_4 ↑				O *		
Liver, absolute weight ↑						O
Adrenal, absolute weight ↑					O	
Thyroid, absolute weight ↑	O *					

↑ Statistically increased; ↓ Statistically decreased; + Finding in one individual subgroup (n = 5) substantiated by the combined Subgroups (n = 10); - Finding in one individual subgroup (n = 5) not substantiated by the combined Subgroups (n = 10); O – Finding only in combined Subgroups at the dose.
* No dose response relationship was observed at higher doses.

300. A review of Tables 101-110 indicates that basic expectations about the effect of increasing the group size on power are met. In all of the 16 studies with individual Subgroups, there were several parameters where one individual Subgroup achieved statistical significance and the other did not. In the majority of these cases, the combined Subgroups substantiated the finding, i.e., achieved statistical significance. However, there are consistently instances where the combined Subgroups did not substantiate findings in an individual Subgroup. In the case of GN in laboratory 4 (Table 102A), all four instances of statistical significance in individual Subgroups were not substantiated in the combined Subgroups. In addition, in all 16 studies with individual Subgroups, the combined Subgroup detected a change that was not observed in an individual Subgroup, detected a change at a lower dose, or both. Thus, the results are consistent with the power calculations that the larger group size will improve the chance of achieving statistical significance.

301. This greater statistical power does not change the opportunity for a false positive result which is determined by alpha which is typically < 0.05. There are consistent instances where statistical significance was observed, but was not dose-related. This is noted by the asterisks in Tables 101-110, and these are presumed to be false positives. In addition, statistical significance is not an automatic determinant that an observation does or does not have biological and toxicological relevance. Regardless of the group size and statistical significance, there is the continuing need to subject both the presence or absence of statistical significance to professional judgment and review, e.g., was the finding significant but within historical baselines or were the results of toxicological relevance despite the statistical findings, negative or positive?

Conclusions of the participating laboratories

302. This section reviews the conclusions of the participating laboratories that were contained in the final reports submitted to the Secretariat from the 16 laboratories having individual Subgroups. Fifteen laboratories reached conclusion or made comments about the possible benefit of having 5 or 10 animals per group; one laboratory did not.

303. With potent test substances (EE, TAM, CGS, MT, FLU, and PTU), there were nine studies with individual Subgroups. In eight of these studies, the laboratories drew conclusions or made comments about the group size. The conclusions and comments were uniform: 5 animals per sex per group were adequate to detect the possible endocrine modulation of these potent test substances. The increase to 10 animals per sex per group tended to increase the number of endpoints achieving statistical significance (for both the current general toxicological and the current-updated endocrine endpoints), often substantiated the findings when one Subgroup achieved significance, and sometimes reduced the NOEL for an endpoint. In the ninth study not specifically drawing conclusions or making comments on group size in its final report (laboratory 8 with CGS), a review of Tables 42, 44, 45, 46, and 47 indicate that 5 animals per groups were easily sufficient to detect possible endocrine modulation in females and in the male mammary gland from the aromatase inhibitor, but that other male parameters required 10 animals per group. Therefore, this study is considered consistent with those of other potent compounds in respect to group size.

304. With less potent test substances (GN, NP, DDE, and THY), there were seven studies with individual Subgroups, where the laboratories drew conclusions or made comments about the group size. The conclusions and comments of each laboratory in regards to group size are noted.

- GN – laboratory 4. No significant changes except increases in the male relative liver weights were detected, and this required 10 animals per group per sex.
- GN – laboratory 12. Significant changes in the uterine absolute and relative weights were detected, and this required 10 animals per group per sex.
- NP – laboratory 1. At a maximum dose of 200/150 mg/kg/d, the updated TG 407 detected effects in the kidneys and the liver, but did not detect effects suggesting possible endocrine modulation. The detection of the effects in the kidneys and the liver required 10 animals per group per sex.
- NP – laboratory 6. At a maximum dose of 300/250 mg/kg/d, the updated TG 407 detected a significant decrease in absolute tissue weights of male accessory reproductive tissues and histopathological changes in the uterus. The detection of these effects required 10 animals per group per sex. Histopathological data were reported only for combined Subgroups.
- DDE – laboratory 6. Changes in liver weights were detected in individual Subgroups (5 animals per sex per group). Changes in relative thyroid weight and in male circulating thyroid hormones required 10 animals per sex per group. Histopathological data were reported only for combined Subgroups.
- DDE – laboratory 7. Changes in liver weights and circulating thyroid hormones were detected in individual Subgroups (5 animals per sex per group). Changes in relative thyroid weight required 10 animals per sex per group. Histopathological data clearly indicated thyroid hypertrophy in individual Subgroups (5 animals per sex per group).
- THY – laboratory 9. Changes in several major organ absolute and relative weights were detected in individual Subgroups in both sexes. Additional changes were detected or confirmed or NOELs lowered with combined Subgroups. Circulating thyroid hormone changes were detected in individual Subgroups in both sexes. Histopathological data were reported only for combined Subgroups. Therefore, individual Subgroup sizes of 5 would have been sufficient to detect general and possible endocrine modulating effects, but the conclusions would have been reinforced with larger group sizes.

305. In conclusion, for a number of different endocrine mechanisms, high potency compounds are consistently detected with 5 animals per sex per group.

Discussion and conclusions

306. The outcome of the theoretical power calculations, that increased group size will to some degree improve detection, are supported by the data. The data also appear consistent with the theoretical calculations in other respects:

- doubling the group size does not provide a directly proportional improvement in power;
- a group size of 5 is sufficient with potent compounds that result in major changes in target organs or parameters; and
- a group size of 10 improves the detection of less potent compounds where the changes are of lower magnitude in target organs or parameters in the young adult animals used here, but for which effects may show up in higher tier test systems.

307. The observed CVs for the updated endpoints in these studies are not small. As with suggestions to seek improvements in histopathological practice in the previous chapter, improvements in SOPs and laboratory technique that would reduce the CVs are also likely improve the power of the current and updated TG 407 to some degree.

308. Animal moralities were encountered with two of the less potent compounds, NP and DDE. This indicates that increasing the dosage administered as a means to assist detection was not viable in these cases. Furthermore, the current TG 407 is typically conducted with the aim of arriving at the maximum tolerated dose for the high dose group.

309. Ultimately, the decision to increase or not increase the group size is more complicated. First, it is a value judgment that weighs the benefit of greater power achieved with increased group size against the animal welfare concerns of doubling the animal numbers. Second, the benefit of detection would depend upon the overall chemical testing strategy employed by regulatory jurisdictions. Under current proposals in some jurisdictions, reproductive and developmental studies would be default requirements for high production volume chemicals. A group size of 5 would be adequate to prioritise potent compounds, and there would be little or no apparent benefit in these cases to using groups of 10 animals. In these same proposals, the TG 407 might be one of the few default requirements for chemicals with low production volumes. In these cases, other tools for prioritization such as exposure assessments, structure-activity relationships and *in vitro* assays need to be considered in assessing any benefit of using groups of 10 animals. The sensitivity was investigated here in terms of statistically significant differences from the control group. An alternative approach, e.g. the benchmark calculation, which uses the complete dose response assessment to calculate the critical dose level for each endpoint, may be applied to evaluate the data more efficiently. This may also help to overcome the need for doubling the animal numbers in the protocol, which may only marginally increase power.

UPDATED TG 407 REPORT DISCUSSION

310. The OECD inter-laboratory studies were conducted in 2000-2001 to assess whether potential updates added to the current TG 407 can reliably detect strong and weak endocrine-modulating substances without interference with or compromise of the current TG 407. These studies comprised Phase-2 of the updated TG 407 program, which was proceeded by feasibility and exploratory studies in Phase-1. Overall responsibility for these studies, as well as other mammalian and ecotoxicological assays for endocrine issue, lay with the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA), which was established to develop and validate new and improved methods to identify and to assess substances acting through endocrine mechanisms. Specific design and management responsibilities for the TG 407 studies lay with the Validation Management Group-mammalian (VMG-mammalian), which reported to the EDTA.

311. The current 28-day repeat dose study (TG 407) provides information on the possible health hazards likely to arise from repeated exposure over limited period of time and includes observations on general, clinical, sensory, and neurological signs; haematology and clinical biochemistry; and gross pathology, weights, and histopathology for major organs and tissues (1). The TG 407 is intended to provide information on the major toxic effects of a test substance, indicate target organs; provide an estimate of a no-observed-adverse-effect level of exposure for certain toxic effects, and potentially provide a maximum tolerated dose. As a result of its broad spectrum of endpoints and observations, TG 407 is frequently part of regulatory data requirements as a pivotal study for the hazard assessment of new chemicals in many member countries.

312. The present TG 407 (without updates) is not primarily considered a definitive test, but is designed to identify possible chemical hazards from the toxic effects observed in the study. That means, the results from the updated TG 407 may be similarly used for initial hazard identification and "to raise a red flag" about potential toxicity concerns, which amongst others may relate to endocrine modulation. The resulting data, if not self-evident for clear adverse effects, would need to be evaluated in the context of other hazard and exposure data using a weight of the evidence approach. Any concerns raised by the weight of the evidence may then need to be specifically addressed in subsequent studies by further characterising the observed effects for hazard identification or for risk assessment purposes. Therefore, any indications of toxicity mediated through endocrine mechanisms seen in the updated TG407 protocol may need to be followed up and investigated in greater depth in tests designed specifically to examine adverse reproductive or developmental or carcinogenic effects and to more fully characterize their hazard.

Phase-1 overview

313. The principle for consideration of the TG 407 updates was that the protocol should incorporate the necessary array of endpoints sensitive to (anti)oestrogenic, (anti)androgenic, and thyroid toxicant mechanisms. This array should first include the primary endocrine organs that a) produce oestrogen, androgen, and thyroid hormones and b) respond to their homeostatic feedback systems. This would include the pituitary, thyroid, and gonads and the male and female reproductive tracts such as the prostate, epididymides, uterus, and mammary glands as the primary target organs. These evaluations would utilize traditional toxicological techniques such as tissue weights and full microscopic histopathology. In addition, other potential endpoints in involved in endocrine regulation and responses could be entertained and assessed, such as circulating hormone levels, sperm parameters, and oestrous cyclicity. After a series of consultations, it was originally proposed to add the following updates to the current TG 407 for Phase-1 feasibility investigations:

- Organ and tissue weights: Add to the current TG 407 endpoints of adrenal, testes, and epididymidal weights, the following tissues: seminal vesicles with coagulating glands, whole

prostate with the possibility to measure the ventral and dorsolateral prostate separately, paired ovaries, uterus, and thyroid.

- Histopathology: Perform additional histopathology on the pituitary, vagina, one epididymis, seminal vesicles with coagulation glands, and mammary gland in both sexes.
- Serum hormone levels: Add analysis of pituitary hormone and target organ hormones in the serum, including: leutinising hormone (LH), follicle stimulating hormone (FSH), prolactin, testosterone, 17β-estradiol, corticosterone, T_3, T_4, and thyroid stimulating hormone (TSH).
- Sperm production and quality parameters: Add analyses of sperm number, morphology, and motility endpoints.
- Estrous cyclicity: Add daily vaginal cytological smears to cover at least two full cycles in order to evaluate the stage of the oestrous cycle each day in order o assess the number of days per cycle and its regularity.

314. The Phase-1 studies were conducted in the summer and fall of 1999. A lead laboratory from Germany was designated to conduct the full updated TG 407 protocol including all current endpoints with the reference substance, FLU. In addition, five laboratories in Japan and one in Korea volunteered to study the TG 407 updates, but not necessarily performing the complete protocol, using FLU, EE, PTU, methoxychlor, TAM, and MT. A meeting of representatives from all participating laboratories and other expert toxicologists was held at Wuppertal, Germany, in December, 1999 to present and to discuss the Phase-1 results (18). Based on their recommendations, the VMG and EDTA agreed to proceed with Phase-2 with some modifications to the protocol, e.g., except for the circulating thyroid hormones, other hormonal analyses were dropped due to their variability, limited utility, and costs.

Phase-2 overview

315. The proposed updates to the TG 407 are intended to provide the endpoints that can detect presumed (anti)oestrogenic, (anti)androgenic, and thyroid toxicant mechanisms of action. The current endpoints and updates added in Phase-2 include additional tissue weights and histopathology of target organs in the male and female reproductive tracts; tissue weight and histopathology of the pituitary; tissue weight and histopathology of the thyroid; circulating levels of T_3, T_4, and thyroid stimulating hormone (TSH); sperm counts and morphology; and staging of the oestrous cycle using vaginal smear cytology (Table 2). Chemicals that potentially act as endocrine modulators may be identified by the TG 407 where they induce a statistically significant increase in the weights, pathological change of the target tissues, or affect other endpoints. These updates had been chosen based upon the initial feasibility and exploratory studies in Phase-1. In addition, the power of the current and update endpoints was to be studied by conducting the study as two individual Subgroups of 5 animals per sex per dose as is done in the current TG 407. These individual Subgroups would then be combined in a statistical reanalysis to represent a combined Subgroup size of 10 animals per sex per dose.

316. The TG 407 animals are sexually mature, young adults with the hypothalamic-pituitary-ovarian feedback loop established and intact. These mature and intact animals can respond to test substances with some degree of adaptation and compensation. This contrasts with the uterotrophic bioassay for (anti)oestrogens and the Hershberger bioassay for (anti)androgens where the animals are immature or ovariectomized/castrated. Thus, these compromised animals may be more sensitive to active test substances. This was suggested by the results of Phase-1, where several laboratories in Japan and Korea employed doses of the test substances that were active in the uterotrophic or Hershberger assays and little or no activity was observed with the TG 407 updates. Further, the mature females are cycling in the TG 407, and the tissue weights of the female reproductive tract should then be more variable, increasing the CVs of the ovarian and uterine weights. Thus, the degree of sensitivity of the TG 407 model needs to established relative to these other bioassays for (anti)oestrogens and (anti)androgens. For thyroid toxicants, the young adult rat appears to be very sensitive relative to humans, an intact animal is necessary

to observe pituitary stimulation of the thyroid, and 28-days appears to be sufficient time for thyroid toxicants to induce effects at sufficient doses. Thus, the TG 407 is plausible as a sensitive model for thyroid toxicants, but this sensitivity needs to be demonstrated.

317. Phase-2 of the TG 407 studies involved 13 laboratories from France, Germany, Japan, Korea, Switzerland, the U.K., and the U.S. These laboratories were from both the public and private sectors (Table 3). The lead laboratory was Bayer HealthCare AG, Wuppertal, Germany. Ten substances were tested in duplicate studies carried out in separate laboratories for a total of 20 studies in Phase-2. These test substances were distributed from a central repository. The laboratories were allowed to utilize the rat strain, diet, bedding, vehicle, and other procedures normally employed in the conduct of current TG 407 studies (Table 4). The laboratories were requested to conduct the full updated TG 407 so as to assess whether any interference or compromise would be encountered with the functional observation battery or any other current protocol requirements (Tables 5, 6, and 7). Almost, all laboratories conducted the studies under GLP conditions (Table 8). Data and final reports were requested from each laboratory to be provided to the Secretariat for archiving and for this report.

318. The 10 test substances included an array of potent and weak endocrine-active substances. Ethinyl oestradiol (EE), genistein (GN), and nonylphenol (NP) were selected as oestrogens. Tamoxifen (TAM) was selected as an antiestrogen. CGS 13820B (CGS) was selected as an aromatase inhibitor. Methyl testosterone (MT) was selected as an androgen. Flutamide (FLU) and *p,p'*-DDE (DDE) were selected as antiandrogens. Propylthiouracil (PTU) and l-thyroxine (THY) were selected as thyroid toxicants. In several cases, dose selection required preliminary range finding studies (Table 9). For each of these test substances, three doses were selected with the intent to provide a no-effect dose, a dose causing moderate effects, and a maximum tolerated dose (Table 10).

319. Specific goals of Phase-2 were then to evaluate current and updated endpoints and procedures in order to increase the sensitivity of the TG 407 to toxic effects presumed to occur via endocrine modulation:

- Further evaluate both current and added hormone-sensitive organ weights and histopathology in the male and female reproductive tracts as well as the pituitary and thyroid;
- Further evaluate cauda epididymidal sperm counts and morphology;
- Further evaluate circulating thyroid hormone levels including T_3, T_4 and TSH;
- Further evaluate staging of the oestrous cycle using vaginal smear cytology to standardize female necropsy timing and thereby minimize oestrous cycle-induced variations in uterine and ovarian weights and histopathology; and
- Further evaluate the utility and benefit of any greater power (statistical probability to detect a positive finding) provided by increasing the number of animals from the number in the current TG 407 from five per sex per dose to ten per sex per dose.

320. Overall, the TG 407 data and findings generally were in good agreement between the two laboratories conducting studies on the same substance at the same doses. A direct comparison has been conducted in the case of each test substance in the chapters of the main body of the report and in the Annexes. The statistical changes in body weights, haematology and clinical chemistry, and tissue weights; the changes in tissue histopathology; and other observations were generally consistent. The magnitude and shape of the dose response curve as well as the doses at which significant responses were observed were again similar between the two laboratories taking into account modest biological variability, strains, laboratory technique and other variables.

Ability of the updated TG 407 to identify (anti)oestrogens

321. The validation programme successfully achieved the goal of demonstrating the ability of the updated TG 407 protocol to detect potent (anti)oestrogenic substances. Both of the laboratories employing a test substance observed effects in target organs with EE, TAM, and CGS. Moreover, both laboratories were able to detect the activity of the potent test substances with group sizes of only five animals per sex per dose. The findings were detected using the tissue weights and histopathology of the current and updated tissues of the male and female reproductive tracts.

322. For EE, a clear pattern of estrogenic responses was observed in both female and male reproductive tracts as well as the male mammary gland. In males, there were statistically significant absolute and relative decreases in the seminal vesicles, the coagulating glands, and the ventral and dorsolateral prostate in both studies at the high EE dose with statistical significance observed for one or more tissues in both individual Subgroups of both studies. In the females, ovarian weights were significantly decreased in one study and the uterine weights significantly increased in the other study. Histopathological correlates were seen in the respective tissues in both sexes in both studies, and also in the male mammary gland. However, statistical significance in pituitary tissue weights was often inconsistent with the expected mechanism of action of a test substance, and sperm counts and morphology were insensitive to EE, TAM, and CGS administration. Oestrus was also affected in both studies.

323. For TAM, a pattern of responses were observed in both female and male reproductive tracts as well as the male mammary gland in one study. In males, there were significant increases in the testes weights in both studies and significant decreases in the relative weights of the ventral prostate and the seminal vesicles in the first study and the whole prostate in the second study. The relative weights of the seminal vesicles and coagulating glands also significantly decreased in the first study, but these tissues were not measured in the second study. Histopathological correlates were observed in the first study, but were not reported in the second study. Histopathological changes were also observed in the male mammary gland in the first study. In females, absolute and relative uterine weights were significantly decreased at the high TAM dose in both studies. Histopathological correlates were reported in females in both studies.

324. For CGS, a pattern of antioestrogenic effects were observed in the female animals. Body weights, typically restricted by oestrogen in adult females, increased dramatically. Relative ovarian weights were significantly increased in one study, but did not achieve significance in the second study. Uterine relative weights significantly decreased in both studies. Histopathological correlates were several and consistent in both studies in the pituitary, hypertrophy of the mammary gland, increased ovarian follicular size and pathology, and atrophy of the uterine and vaginal tissues at all CGS doses. Oestrus was also affected in both studies.

325. In contrast to the potent (anti)oestrogenic substances, findings with the updated and current endpoints were limited in the cases of the weak oestrogens GN and NP. For GN, the evidence of estrogen activity was equivocal. One study detected a statistically significant increase in uterine weight. Both studies did observe a possible desynchronisation of the female reproductive cycle among the tissue, but only when the data were reinterpreted by an expert pathologist. These changes were observed at 400 mg/kg/d. The subtle and atypical nature of the observations (asynchrony of several tissues and cell types in the female reproductive tract to the oestrous cycle) necessary to arrive at this conclusion must be noted. The detection of very weak estrogens such as GN may then require increased vigilance and changes in the current practice of histopathological examination of the female reproductive tract.

326. For NP, the updated TG 407 did not detect estrogenic responses with NP at doses below a maximum tolerated dose. Equivocal evidence in uterine histopathology and decreases in the absolute

weights of male accessory reproductive tissues were observed at doses which caused mortalities and other clinical signs in the animals. The effects on the synchronisation of the female reproductive tract tissues seen with EE and GN may have been seen in one NP study at a low frequency, but confirmation was absent at the higher NP dose employed in the second study. Therefore, this evidence was judged to be equivocal.

Ability of the updated TG 407 to identify (anti)androgens

327. The validation programme successfully achieved the goal of demonstrating the ability of the updated TG 407 protocol to detect potent (anti)androgenic substances. Both of the laboratories employing a test substance observed effects in target organs with MT and FLU. Moreover, both laboratories were able to detect the activity of the potent test substances with group sizes of only five animals per sex per dose. The findings were detected using the tissue weights and histopathology of the current and updated tissues of the male and female reproductive tracts.

328. For MT, effects were observed in both sexes. The absolute and relative weights of the testes were significantly decreased in both studies. Relative weights of the accessory reproductive tissues, the ventral and dorsolateral prostate and the seminal vesicles were significantly increased in the first study, but although the absolute values of these tissues were increased by 10% or more in the second study, they did not achieve statistical significance. Histological correlates were seen in both studies, and in the one study that reviewed the male mammary gland, clear histological changes were observed. In females, the relative ovarian weights were significantly decreased in both studies, and the uterine weights were significantly increased in one study. Both studies observed ovarian atrophy and apparent stimulation of the uterus, cervix, and vagina in the histopathological examinations.

329. For FLU, antiandrogenic effects were observed in the male reproductive tract in both studies. Absolute and relative weights of the epididymides, ventral prostate seminal vesicles were significantly decreased in both studies. The dorsolateral prostate was significantly decreased in the second study, but was not measured in the first study. Atrophy of these tissues was observed in the histopathological examination of both studies. Both studies also observed Leydig/interstitial cell hypertrophy/hyperplasia at the high FLU dose.

330. Sperm counts and morphology were found to be insensitive and variable with MT as with many potent test substances. When findings were present in sperm counts and morphology, these findings occurred at higher doses than in tissue weight and histopathology observations.

331. In contrast to the potent (anti)androgen FLU, no antiandrogenic findings were present in either laboratory in the case of DDE. For DDE, neither study provided evidence of antiandrogenic activity in the male reproductive tract, despite mortalities in both sexes indicating that the maximum tolerated dose was exceeded in one study. Several histopathological changes had been recorded in the male reproductive tract in the high DDE dose group of the first study. However, five of the male animals had died prematurely in the early phase of the study. Upon examination of the individual animal data, changes in the male reproductive tract were clearly associated with animal deaths including those with only short-term DDE exposure, suggesting autolysis as the cause rather than a treatment related effect (Table 67). Both DDE studies used Sprague-Dawley rats, which have been shown to be less sensitive to DDE at the administered doses than Long Evans rats, and this may have contributed to the lack of observed antiandrogenic effects.

Ability of the updated TG 407 to identify thyroid toxicants

332. The validation programme successfully achieved the goal of demonstrating the ability of the updated TG 407 protocol to detect thyroid toxicants. Both of the laboratories employing a test substance

observed effects in the thyroid with PTU and THY. Both laboratories were able to detect the activity of these test substances with group sizes of only five animals per sex per dose. For PTU, a clear pattern of thyroid hypertrophic responses was observed in both the female and the male histopathology. The NOEL could not be set as both studies reported thyroid histopathological changes at the lowest dose of 0.1 mg PTU/kg bw/d. The effect levels were within the same order of magnitude between the short-term TG 407studies and a 2-generation reproductive and developmental study for both increases in thyroid weight and for the histopathological changes. For THY, the findings in both studies of increased haematopoietic activity and cardiac myopathy are associated with hyperthyroidism. In the thyroid itself, atrophic changes were observed in both sexes in both studies at the intermediate and high THY doses.

333. In addition, thyroid modulation by MT and DDE were detected in both respective studies. The effects of MT were to induce follicular hypertrophy in both sexes in both studies and with increases in circulating T_4 and TSH. The thyroid effects of DDE including follicular hypertrophy in both sexes in both studies and hormonal changes that were preceded in dose by significant hepatic enlargement and histopathological changes. This suggests that the thyroid effects may be the result of increased elimination of circulating thyroid hormones resulting in increased pituitary stimulation of thyroid TSH secretion. Thus, the ability of the TG 407 to detect thyroid modulation appears to be robust, to extend beyond classical potent compounds, and to include indirect mechanisms.

334. Thyroid histopathology was consistently the most reliable and most sensitive endpoint for the detection of thyroid modulation. Thyroid weight was reliable, but was somewhat less sensitive when compared to thyroid histopathology. Circulating thyroid hormone levels (T_3, T_4, and TSH) were not always reliable and sensitive, but the standard operating procedures for blood sampling and for thyroid hormone analyses were not standardised to reduce stress induced variability and to reduce analytical variability, respectively. The Phase-2 results suggest then that further work may be useful with T_3, T_4 and TSH hormone levels as optional and supplementary endpoints. Where a compound is suspected of thyroid activity or changes are noted in thyroid histopathology, T_3, T_4 and TSH analyses of the retained serum samples may be warranted. The CVs of the T_3 and T_4 measurements were modest, averaging 20-21 across laboratories, but the TSH CV was nearly double that value. The primary concerns at this time are better understanding and possible standardisation of necropsy sampling conditions and analytical methods. The T_3 hormone levels did not appear sufficiently sensitive and yielded several false positives, but there is evidence from other studies that rodents sometimes tend to overcompensate in response to certain anti-thyroid compounds (e.g. liver enzyme inducers), resulting in increased T_3, but quite normal T_4 and TSH levels

Lack of interference with the current TG 407

335. The ability of laboratories to perform the entire TG 407 protocol was not negatively impacted by the updates. Laboratories were able to conduct the functional observation and motor activity batteries without interference. Study time for some females to reach diestrous was extended beyond 28-days in the studies, sometimes extending the work of the necropsy staff into weekends.

Reproducibility and reliability of TG 407 for systemic and organ toxicity

336. The findings in these studies also support the reproducibility and reliability of the traditional use of the TG 407 in flagging possible systemic and target organ toxicity. For each of the ten test substances, changes in body weights, haematology and clinical chemistry parameters, and current absolute and relative tissue weight endpoints, and the histopathology endpoints have been compiled into comparative tables in each of the test substance chapters. In addition, where other toxicological data were available on the test substance or a close mechanistic relative, these same endpoints have been compared with TG 407 data at the end of each test substance chapter.

337. These comparisons show that, within the TG 407 studies and with other toxicological data, these endpoints are fundamentally reproducible and that the TG 407 is basically a reliable predictor for other studies. The changes in body weight were a reproducible effect between the TG 407 studies, noting that, based on the power calculations, the percent change in body weight often had to be > 10% to achieve significance even with the low CV for this measurement. Similar body weight changes were observed in longer and more detailed toxicological assays. For haematological and clinical chemistry parameters, four basic classifications were entertained: 1) the parameters was statistically significant in both studies, 2) the parameter was statistically significant in one study and the absolute trend in the other study was in agreement, 3) the parameters was statistically significant in one study and the absolute trend in the other study was unchanged or moved in the opposite direction, and 4) the direction of statistical significance was in the opposite direction between the two studies. Within the TG 407 studies, the bulk of the changes fell into the first two categories. These data were often not published for other toxicological assays, so no comparison here was consistently made and no conclusion is offered. For tissue weights, such as liver, kidneys, and adrenals, the TG 407 studies were fundamentally reproducible and matched findings in longer and more detailed toxicological assays. Histopathological findings were also similar both among the TG 407 studies for a test substance and with longer and more detailed histopathological studies. Due to the longer time of administration, LOEL doses in these non-TG 407 studies were typically lower, as would be expected.

Variability of the proposed updates

338. An important factor affecting sensitivity of the endpoint is its variability as represented by the coefficient of variation (CV). A lower CV improves the ability to statistically detect significance, and a higher CV diminishes that ability. The CV values of the proposed enhancement tissue weights were slightly higher than a number of current tissues. For example, current tissues, that are larger and easier to dissect, such as liver and testes, had lower mean CVs ranging from about 8 to 11. As tissues became smaller and more difficult to dissect, CVs increased. Mean CVs were 12-14 for paired adrenals, 18-19 for pituitary, and 20-21 for the thyroid. Fluid-filled male accessory tissues that are difficult to dissect, such as the prostate lobes, had higher mean CVs approaching 24. The CVs for the uterus and for the male accessory tissues were consistent with those observed in the uterotrophic and Hershberger validation programs. In addition, as also with the uterotrophic and Hershberger studies, CV values varied and appeared to be related with the individual performing laboratories. This suggests that laboratory technique is a possible variable and could be connected to the ability to detect weakly acting substances.

339. The comparison of the histopathological observations reveals a lack of standardization in both terminology and grading. Individual pathologists often described tissue changes in different ways and used differing terminologies. For example, with FLU, pituitary changes in males were described as hypertrophic basophilic cells and increased PAS-positive cells in the first study and as increased clear cells in the anterior lobe of the pituitary in the second study (Table 59). Only a few of the final reports contained representative photographs of the key histopathological findings that could be used as a basis to assess the consistency of the observations. Additional remarks on the histopathology are found in the discussion on the male and female reproductive tracts below.

340. As noted, sperm numbers and morphology were relatively insensitive even in the case of potent compounds administered at high doses (Table 87). In addition, some laboratories experienced CV values greater than 30 (Table 88). When combined with the power calculations (Table 98 and Figure 1), this indicates that these endpoints would not be generally useful. Given the labour intensity of these endpoints and the need for care in sample timing and handling, sperm count and morphology appear to have little benefit for the TG 407 assay.

Power, group size, and number of animals

341. The utility of increasing the group size or the number of animals per group for detecting endocrine mediated effects were assessed between individual Subgroups of 5 animals per sex and combined Subgroups of 10 animals per sex. Power is a statistical description of the chance that an assay will detect a positive effect. Power is determined by the magnitude of an effect, the variability or CV of an effect, the number of observations, and other parameters, such as the statistical method used. The assessment of the TG 407 studies included theoretical power calculations, calculations of CVs, and a review of actual statistical significance comparisons between individual Subgroups of 5 animals per sex per dose and combined Subgroups of 10 animals per sex per dose. The assessment concludes that, for potent test substances, 5 animals per sex are sufficient to detect endocrine modulation with (anti)oestrogenic, (anti)androgenic, and thyroid toxicants. It is important to recognize that the power calculations imply that the same degree of benefit cannot be generalized across different endpoints. The increased power will depend upon the variability of a particular endpoint, or its inherent CV, as well as the magnitude of change induced by a test substance. Moreover, while the animal number may be doubled (increased by 100%), the increased power provided is at most about 80% and is sometimes far less (Table 98). Second, given that CV is a major contributor to power, improvements in laboratory skill and better standardization of histopathological examination, description, and terminology should also be entertained and may have higher priority than increasing animal numbers.

342. Ultimately, the decision to increase or not increase the group size may be even more complicated. First, it is a value judgment that weighs the benefit of greater power achieved with increased group size against the animal welfare concerns of doubling the animal numbers. Second, the benefit of detection would depend upon the overall chemical testing strategy employed by regulatory jurisdictions. Under current proposals in some jurisdictions, reproductive and developmental studies would be default requirements for high production volume chemicals. A group size of 5 would be adequate to prioritise potent compounds, and there would be little or no apparent benefit in these cases to using groups of 10 animals in light of other required tests. In these same proposals, the TG 407 might be one of the few default requirements for very low production volume chemicals. In these cases, the role of exposure assessments and other techniques such as structure-activity relationships and *in vitro* assays need to be considered in assessing any benefit of using groups of 10 animals. An alternative approach, e.g. the benchmark calculation, which uses complete dose response assessments to calculate the critical dose level for each endpoint, may use the data obtained with 5 animals per sex per dose more efficiently. This is likely a better way forward than doubling the animal numbers in the protocol, which may only marginally increase power.

Toxicological reproducibility of the TG 407 updates

343. A comparison has been made between the results of the updated TG 407 and other toxicological assays with the same or similar endocrine active test substances in most cases. At least some data were found for all substances except methyl testosterone and l-thyroxine. These comparisons support the ability of the TG 407 to detect systemic, target organ and endocrine-related toxicities for potent chemicals. Where chronic studies or reproductive and development studies easily detected effects, concordant results were typically found in the current TG 407 studies for systemic, target organ, and endocrine toxicities. As expected, the chronic circumstances in the reproductive and developmental studies usually resulted in the findings at lower doses than in the TG 407 studies, but not in all cases such as with PTU. These comparisons are summarised at the end of each test substance chapter in the body of this report.

Successful and reliable endpoints recommended for inclusion in an updated TG 407

344. The Phase-2 results support the inclusion of several current endpoints and related updates in a revised guideline as useful for the detection of endocrine modulating substances in young adult animals. Tissue weights and histopathology of the male and female reproductive tracts as well as histopathology of the male and female mammary glands clearly contributed to the detection of (anti)oestrogenic and (anti)androgenic substances with potent compounds in a reliable manner and with less potent compound in substance-specific cases. While there was no consistent relationship between pituitary weight change and an expected mode of action, histopathology of the pituitary was often relevant in providing support for findings in other tissues. Likewise, adrenal weights and histopathology also provided supplemental and supporting information for findings in other tissues. However, pituitary histopathology, adrenal weights, or adrenal histopathology should not by themselves be considered diagnostic of potential endocrine modulation in the case of (anti)oestrogens, (anti)androgen, and thyroid toxicants.

345. The Phase-2 results clearly support the inclusion of thyroid weights and histopathology in a revised guideline as able to detect thyroid toxicants of various potencies. As the thyroid weight is somewhat less sensitive, care should be exercised not to compromise the histopathology of the tissue. Thyroid trimming should then continue to be performed after fixation of the tissue and by carefully trained technicians.

Proposed endpoints not recommended for inclusion in an updated TG 407

346. The Phase-2 results do not support the inclusion of the spermatology endpoints. The spermatology endpoints provided no substantive benefit for detecting possible endocrine modulation. Sperm effects were found only with a few potent compounds at doses above those where other endpoints had already identified clear effects, and the CVs indicate that substantial work might be needed to achieve reliable methodological practice among laboratories.

Other recommendations

347. The Phase-2 results indicated that further work may be useful with T_3, T_4 and TSH hormone levels as optional and supplementary endpoints. The CVs of the T_3, and T_4 measurements were modest, averaging 20-21 across laboratories, but the TSH CV was nearly double that value.

348. The Phase-2 results suggest that a review of the histopathology procedures for the male and female reproductive tracts and for the mammary glands of both sexes has merit. The central problem confronting the histopathologists with (anti)oestrogens and (anti)androgens is that the pattern of effects is typically not one of frank pathological changes and lesions in any tissue. In the male, the pattern may be one of modest atrophy or hypertrophy or changes in particular target cells, such as the Leydig cells. In the female, the overall synchronization of a temporal sequence in several tissues during the oestrous cycle may be altered so that 'normal' histology is observed for individual parameters, but the overall coordination of the oestrous cycle is impaired. Thus, the effects are subtle and not necessarily atypical, and one must carefully correlate the pattern in several cell types of the ovary, uterus, cervix and vagina with that of the oestrous cycle in order to evaluate whether there are dissimilarities with the normal, expected sequence and its orchestrated pattern.

349. It is unclear at this time whether vaginal smears and the attempt to necropsy the females at a specific time in the oestrous cycle are essential to the outcome of such a staging and its assessment. This may extend the range of the time of test substance administration by several days, amounting to an increase in approximately 15% in some animals. This adds considerable labour as well, not only with the smears and their reading, but in conducting the necropsies over an extended period. Further, the impact of the

necropsy on the ability to discern significant changes in uterine and ovarian weights has not been conclusively resolved. The changes in weights with potent compounds appear to be of sufficient magnitude that no real benefit is provided by staging in these cases. In contrast, for the marginal findings in the case of genistein to be detectable, staging appears to have been essential.

350. Such a careful staging of the female reproductive tract tissues is not the current practice in TG 407 studies and was not conducted as part of these update studies. As this appears to be the most promising possibility at this time to reliably identify weak oestrogens with the TG 407 protocol, it may be worthwhile to conduct pilot studies with specific compounds. Due to the costs and animal use, major studies should await positive results from such feasibility studies.

351. This discussion also raises a question about the degree of blinding of the pathologist in TG 407 studies. As noted, the histopathological changes in the male and female reproductive tracts are often subtle, making the recognition of changes in the context of normal patterns technically challenging. Together with improved guidance on histopathological practice for these tissues, identification of substances is likely to be improved where the pathologist is informed of 1) structural, *in vitro*, or other *in vivo* alerts as the possible endocrine activity of a test substance and 2) changes in both the significance and trends in the absolute and relative weights of particular tissues.

CONCLUSIONS

352. In conclusion, the following endpoints are recommended for inclusion in an updated TG 407 guideline for the detection of toxicants with possible (anti)oestrogenic, (anti)androgenic, and antithyroid modes of action for young adult animals:

- Absolute and relative tissue weights of the male reproductive tract, including testes, epididymides, ventral and dorsolateral prostate and seminal vesicles with coagulating glands.
- Absolute and relative tissue weights of the female reproductive tract, including the paired ovaries and uterus.
- Histopathology of the pituitary.
- Absolute and relative thyroid weights and histopathology of the thyroid. The recommendation for the addition of the thyroid weights is qualified in that the dissection and trimming procedures must not compromise the more valuable thyroid histopathology.
- Histopathology of the male and female reproductive tract tissues (adding cervix and vagina so that a mandatory staging assessment of the oestrous cycle can be realised).
- Histopathology of the male and female mammary glands.
- If there are indications that the test substance is a thyroid toxicant or thyroid histopathology indicates effects, it could be appropriate to consider performing T3, T_4 and TSH analyses on retained serum samples, if the necropsy sampling and analytical questions noted in this report are resolved.

353. The decision whether or not to increase the group size from 5 animals per sex per dose to 10 animals per sex per dose requires wider input and consideration. In particular, the need for sensitivity and increased power should be evaluated in the context of an overall testing hierarchy. Such a decision should also take into account alternatives for improving the power of current endpoints by improving technique and practice as well as possibilities noted above such as the staging of the female reproductive tract cycle.

REFERENCES

1. OECD. (1995). OECD Guideline for the Testing of Chemical. TG 407. Repeated Dose 28-day Oral Toxicity Study in Rodents. Adopted on 27th July 1995.
2. OECD. (1992). Chairman's Report of the Meeting of the ad hoc Working Group of Experts on Systemic Short-term and (Delayed) Neurotoxicity.
3. EC/EEA/OECD/WHO. (1997). European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife. Environment and Climate Research Programme of DG XII, European Commission report EUR 17549. Workshop held 2-4 December 1996 at Weybridge, UK, Brussels: European Union,127 pp.
4. CSTEE. (1999). CSTEE Opinion on Human and Wildlife Health Effects of Endocrine Disrupting Chemicals, with Emphasis on Wildlife and on Ecotoxicology Test Methods. Report of the Working Group on Endocrine Disrupters of the Scientific Committee on Toxicity, Ecotoxicity and the Environment (CSTEE) of DG XXIV, Consumer Policy and Consumer Health Protection. Brussels, Belgium.
5. Cooper RL, RJ Kavlock. (1997). Endocrine disruptors and reproductive development: A weight of the evidence overview. J. Endocrinol. 152:159-166.
6. Crisp TM, Clegg ED, Cooper RL, Wood WP, Anderson DG, Baetcke KP, Hoffmann JL, Morrow MS, Rodier DJ, Schaeffer JE, Touart LW, Zeeman MG, Patel YM. (1998). Environmental endocrine disruption: An effects assessment and analysis. Environ. Health Perspect. 106(Suppl. 1):11-56.
7. NRC. (1999). Hormonally Active Agents in the Environment. Washington, DC: National Academy Press.
8. IPCS. (2002). Global Assessment of the State-of-the-Science of Endocrine Disruptors. Damstra T, Barlow S, Bergman A, Kavlock R, Van der Kraack G, ed. International Programme on Chemical Safety. WHO/PCS/ECD/02.2 Geneva, Switzerland.
9. OECD. (1998). Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10th-11th March 1998, ENV/MC/CHEM/RA(98)5.
10. USEPA. (1998). Endocrine Disrupter Screening and Testing Advisory Committee (EDSTAC) Final Report EPA/743/R-98/003. [http://www.epa.gov/scipoly/oscpendo/history/finalrpt.htm].
11. Gray Jr, LE, Kelce WR, Wiese T, Tyl R, Gaido K, Cook J, Klinefelter G, Desaulniers D, Wilson E, Zacharewski T, Waller C, Foster P, Laskey J, Reel J, Giesy J, Laws S, McLachlan J, Breslin W, Cooper R, DiGiulio R, Johnson R, Purdy R, Mihaich E, Safe S, Sonnenschein C, Welshons W, Miller R, Mcmaster S, Colborn T. (1997). Endocrine screening methods workshop report: Detection of estrogenic and androgenic hormonal and antihormonal activity for chemicals that act via receptor or steroidogenic enzyme mechanisms. Reprod. Toxicol. 11:719-750.
12. Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Gray LE, Kaatari S, Lucier G, Luster M, Mac MJ, Maczka C, Miller R, Moore J, Rolland R, Scott G, Sheehan DM, Sinks T, Tilson H. (1996). Research needs for the risk assessment of health and environmental effects of endocrine disruptors: A report of the U.S. EPA-sponsored workshop. Environ. Health Perspect. 104(Suppl. 4): 715-740.
13. Ankley G, Mihaich E, Stahl R, Tillitt D, Colborn T, McMaster S, Miller R, Bantle J, Campbell P, Denslow N, Dickerson R, Folma L, Fry M, Giesy J, Gray LE, Guiney P, Hutchinson T, Kennedy S, Kramer V, LeBlanc G, Mayes M, Nimrod A, Patino R, Peterson R, Purdy R, Ringer R, Thomas P, Touart L, Van der Kraak G, Zacharewski T. (1998). Overview of a workshop on screening methods for detecting potential (anti-) estrogenic/androgenic chemicals in wildlife. Environ. Toxicol. Chem. 17:68-87.

14. DeVito M, Biegel L, Brouwer A, Brown S, Brucker-Davis F, Cheek AO, Christensen R, Colborn T, Cooke P, Crissman J, Crofton K, Doerge D, Gray E, Hauser P, Hurley P, Kohn M, Lazar J, McMaster S, McClain M, McConnell E, Meier C, Miller R, Tietge J, Tyl R. 1999. Screening methods for thyroid hormone disruptors. Environ Health Perspect 107:407-415.

15. OECD. (1998). Detailed Review Paper on the Appraisal of Test Methods for Sex Hormone Disrupting Chemicals. OECD Monograph No 21.

16. OECD. (1999). First Meeting of the OECD Validation Management Committee for the Screening and Testing of Endocrine Disrupting Substances - Mammalian Effects (February, 1999). OECD document ENV/JM/TG/EDTA(99)1/FINAL.

17. OECD. (1999). Summary record of the third meeting of the OECD Task Force on Endocrine Disrupters (EDTA). OECD document ENV/JM/TG/EDTA/M(99)1.

18. OECD. (2000). Summary report of the meeting to discuss the Phase-1 results from the enhanced TG407 procedure; Wuppertal, Germany, December 15, 1999. Meeting Document 5. 2nd Meeting of the Validation Management Group for Task Force on Endocrine Disrupters Testing and Assessment (EDTA), 20-21 January 2000, Paris, France.

19. Andrews P, Freyberger A, Hartmann E, Eiben R, Loof I, Schmidt U, Temerowski M, Becka M. (2001). Feasibility and potential gains of enhancing the subacute rat study protocol (OECD test guideline no. 407) by additional parameters selected to determine endocrine modulation. A pre-validation study to determine endocrine-mediated effects of the antiandrogenic drug flutamide. Arch. Toxicol. 75:65-73.

20. Okazaki K, Okazaki S, Nishimura S, Nakamura H, Kitamura Y, Hatayama K, Nakamura A, Tsuda T, Katsumata T, Nishikawa A, Hirose M. (2001). A repeated dose (28 days) oral toxicity study of methoxychlor in rats, based on the 'Enhanced OECD test guideline 407'for screening for endocrine-disrupting chemicals. Arch. Toxicol. 75:513-521.

21. Okazaki K, Imazawa T, Nakamura H, Fukukawa F, Nishikawa A, Hirose M. (2002). A repeated 28-day oral dose toxicity study of 17α-methyltestosterone in rats, based on the 'Enhanced OECD test guideline 407' for screening for endocrine-disrupting chemicals. Arch. Toxicol. 75:635-642.

22. Toyoda K, Shibutani M, Tamura T, Koujitani T, Uneyama C, Hirose M. (2000). Repeated dose (28 days) oral toxicity study of flutamide in rats, based on the draft protocol for the 'Enhanced OECD test guideline 407'for screening for endocrine-disrupting chemicals. Arch. Toxicol. 74:127-132.

23. OECD. (2000). Summary record of the second meeting of the Validation Management Group on Screening and Testing for Endocrine Disrupters for Mammalian Effects. OECD document ENV/JM/TG/EDTA/M(2000)1/REV1.

24. OECD. (2000). Status of the work on the enhancement of Test Guideline 407. OECD document ENV/JM/TG/EDTA/(2000)6.

25. OECD. (2000). Summary record of the fourth meeting of the Task Force on Endocrine Disrupters (EDTA). OECD document ENV/JM/TG/EDTA/M(2000)2.

26. Biegel LB, Flaws JA, Hirshfield AN, O'Connor JC, Elliott GS, Ladics GS, Silbergeld EK, Van Pelt CS, Hurtt ME, Cook JC, Frame SR. (1998). 90-Day feeding and one-generation reproduction study in Crl:CD BR rats with 17ß-estradiol. Toxicol. Sci. 44:116-142.

27. Biegel LB, Cook JC, Hurtt ME, O'Connor JC. (1998). Effects of 17ß-estradiol on serum hormone and estrus cycle in female Crl:CD BR rats: Effects on parental and first generation rats. Toxicol. Sci. 44:143-154.

28. Cook JC, Johnson L, O'Connor JC, Biegel LB, Krams CH, Frame SR, Hurtt ME. (1998). Effects of dietary 17ß-estradiol exposure on serum hormone and testicular parameters in male Crl:CD BR rats. Toxicol. Sci. 44:155-168.

29. Blair R, Fang H, Branham WS, Hass B, Dial SL, Moland CL, Tong W, Shi L, Perkins R, Sheehan DM. (2000). Estrogen receptor relative binding affinities of 188 natural and xenochemicals: Structural diversity of ligands. Toxicol. Sci. 54:138-153.

30. Odum J, Lefevre PA, Tittensor S, Paton D, Routledge EJ, Beresford NA, Sumpter JP, Ashby J. (1997). The rodent uterotrophic bioassay: Critical protocol features, studies with nonylphenols, and comparison with a yeast estrogenicity assay. Reg. Toxicol. Pharmacol. 25:176-188.

31. Delclos KB, Bucci TJ, Lomax LG, Latendresse JR, Warbritton A, Weis CC, Newbold RR. (2001). Effects of dietary genistein exposure during development on male and female CD (Sprague-Dawley) rats. Repro. Toxicol. 15:647-663.

32. Delclos KB, Weis CC, Olson G, Bucci TJ, Newbold RR. (2003). A five generation reproductive toxicity assessment of the soy isoflavone genistein in CD Sprague-Dawley rats. Toxicologist 72(S-1):76.

33. Casanova M, You L, Gaido KW, Archibeque-Engle S, Janszen DB, d' Heck HA. (1999). Developmental effects of dietary phytoestrogens in Sprague-Dawley rats and interactions of genistein and daidzein with rat estrogen receptors α and β *in vitro*. Toxicol. Sci. 51:236-244.

34. Lewis RW, Brooks N, Milburn GM, Soames A, Stone S, Hall M, Ashby J. (2003). The effects of the phytoestrogen genistein on the postnatal development of the rat. Toxicol. Sci. 71:74-83.

35. Branham WS, Dial SL, Moland CL, Hass BS, Blair RM, Fang H, Shi L, Tong W, Perkins RG, Sheehan DM. (2002). Phytoestrogen and mycoestrogen binding to the rat uterine estrogen receptor. J. Nutr. 132:658-664.

36. OECD. (2003). OECD report of the validation of the rat uterotrophic bioassay: Phase-2. Testing of potent and weak oestrogen agonists by multiple laboratories. OECD/ENV/JM/TG/EDTA(2003)1.

37. Cunny HC, Mayes BA, Rosica KA, Trutter JA, Van Miller JP. (1997). Subchronic toxicity (90-day) study with *para*-nonylphenol in rats. Reg. Toxicol. Pharmacol. 26:172-178.

38. Chapin RE, Dulaney J, Wang Y, Lanning L, Davis B, Collins B, Mintz N, G Wolfe G. (1999). The effects of 4-nonylphenol in rats: A multigeneration reproduction study. Toxicol. Sci. 52:80-91.

39. Nagao T, Wada K, Marumo H, Ysohimura S, Ono H. (2001). Reproductive effects of nonylphenol in ras after gavage administration: A two-generation study. Repro. Toxicol. 15:293-315.

40. Tucker MJ, Adam HK, Patterson JS. (1984). Tamoxifen. In: Lawrence DR, McLean AEM, Weatherall M. (eds). Safety testing of new drugs. Academic, Orlando, FL. Pp 125-161.

41. Wogan GN. (1997). Review of the toxicology of tamoxifen. Sem. Oncology 24 (Suppl 1):S1-87-S1-97.

42. NTP. (1997). Final report on the reproductive toxicity of tamoxifen citrate (CAS #54965-24-1) administered by gavage to Sprague-Dawley rats. NTP-RACB-95003/PB97-199723. National Toxicology Program, Research Triangle Park, NC.

43. Turner KJ, Morley M, Atanassova N, Swanston ID, Sharpe RM. (2000). Effect of chronic administration of an aromatase inhibitor to adult male rats on pituitary and testicular function and fertility. J. Endocrinol. 164:225-238.

44. O'Connor JC, Frame SR, Ladics GS. (2002). Evaluation of a 15-day screening assay using intact male rats for identifying steroid biosynthesis inhibitors and thyroid modulators. Toxicol. Sci. 69:79-91.

45. O'Connor JC, Cook JC, Stone TW, Makovec GT, Frame SR, Davis LG. (1998). An ongoing validation of a Tier I screening battery for detecting endocrine-active compounds (EACs). Toxicol. Sci. 46:45-60.

46. Nunez SB, Blye RP, Thomas PM, Reel JR, Barnes KM, Maelly JD, Cutler GB Jr. (1996). Recovery of reproductive function in rats treated with the aromatase inhibitor Fadrozole. Reprod. Toxicol. 10:373-377.

47. Bhatnagar AS, Häusler A, Schieweck K, Lang M, Bowman R. (1990). Highly selective inhibition of estrogen biosynthesis by CGS 20267, a new non-steroidal aromatase inhibitor. J. Steroid Biochem. Molec. Biol. 37:1021-1027.

48. Bhatnagar AS, Häusler A, Schieweck K, Browne LJ, Bowman R, Steele RE. (1990). Novel aromatase inhibitors. J. Steroid Biochem. Molec. Biol. 37:363-367.

49. Bhatnagar AS, Schieweck K, Häusler A, Browne LJ, Steele RE. (1989). Inhibitors of estrogen biosynthesis: Preclinical studies with CGS 16949A, a new nonsteroidal aromatase inhibitor. Proc. R. Soc. Edinb. 95B:293-303.

50. Steele RE, Mellor LB, Sawyer WK, Wasvary JIM, Browne LJ. (1987). *In vitro* and *in vivo* studies demonstrating potent and selective estrogen inhibition with the non steroidal aromatase inhibitor CGS 16949A. Steroids 50:147-161.

51. Dukes M, Edwards PN, Large M, Smith IK, Boyle T. (1996). The preclinical pharmacology of 'Arimidex' (Anastrozole; ZD1033) – a potent, selective aromatase inhibitor. J. Steroid Biochem. Molec. Biol. 58:439-445.

52. Richig JW, et al. (1987). 13-Week oral toxicity study in rats. CGS 18320B (MIN 861022). Research Department, Pharmaceuticals Division Ciba-Geigy Corporation, Summit, New Jersey (unpublished data).

53. Wolf CJ, Hotchkiss A, Ostby JS, LeBlanc GA, Gray Jr LE. (2002). Effects of prenatal testosterone propionate on the sexual development of male and female rats: A dose-response study. Toxicol. Sci. 65:71-86.

54. Wilson VS, Lambright C, Otsby J, Gray Jr LE. (2002). In vitro and in vivo effects of 17β-trenbolone: A feedlot effluent contaminant. Toxicol. Sci. 70:202-211.

55. OECD. (2002). Final OECD report of the initial work towards the validation of the rat Hershberger assay. Phase 1. Androgenic response to testosterone propionate and anti-androgenic effects of flutamide. ENV/JM/TG/EDTA(2002)1/REV2/ADD1.

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

57. Imperato-McGinley J, Sanchez RS, Spencer JR, Yee B**,** Vaughan ED. (1992). Comparison of the effects of 5α-reductase inhibitor Finasteride and the antiandrogen Flutamide on prostate and genital differentiation: Dose-response studies. Endocrinology 131:1149-1156.

58. McIntyre BS, Barlow NJ, Foster PMD. (2001). Androgen-mediated development in male rat offspring exposed to flutamide in utero: permanence and correlation of early postnatal changes in anogenital distance and nipple retention with malformations in androgen-dependent tissues. Toxicol. Sci. 62:236-249.

59. Neri R, Florance K, Koziol P, van Cleave, S. (1972). A biological profile of a nonsteroidal antiandrogen, SCH 13521 (4'-nitro-3'-trifluoromethylisobutyranilide). Endocrinology 91:427-437.

60. Peets EA, Henson MF, Neri R. (1974). On the mechanism of the anti-androgenic action of flutamide (α-α-α-trifluoro-2-methyl-4'-nitro-m-propionotoluidide) in the rat. Endocrinology 94:532-540.

61. Viguier-Martinez MC, de Reviers MTH, Barenton B, Perreau C. (1983). Effect of a non-steroidal antiandrogen, Flutamide, on the hypothalamo-pituitary axis, genital tract and testis in growing male rats: Endocrinological and histological data. Acta Endocrinol. 102:299-306.

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

63. You L, Casanova M, Archibeque-Engle S, Sar M, Fan L-Q, d'A Heck H. (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.

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

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

66. O'Connor JC, Frame SR, Ladics GS. (2002). Evaluation of a 15-day screening assay using intact male rats for identifying antiandrogens. Toxicol. Sci. 69:92-108.

67. Wolfe GW, Pepperl S, Greenlaw J. (2002). Two-generation reproduction toxicity study of propylthiouracil when administered to Sprague-Dawley rats in the drinking water. Study performed for the US Environmental Protection Agency by TherImmune Research Corporation, TRC Study No. 7244-601. Gaithersburg, MD, USA.

68. Capen CC. (1997). Mechanistic data and risk assessment of selected toxic endpoints of the thyroid gland. Toxicl. Pathol. 25:39-48.

69. Capen CC, Martin SL. (1989). The effects of xenobiotics on the structure and function of thyroid follicles and C-cells. Toxicol. Pathol. 17:266-293.

70. MacKenzie CG. (1947). Differentiation of the antithyroid action of thiouracil, thiourea, and PABA from sulfonamides by iodide administration. Endocrinology 40:137-151.

71. Richards JB, Ingbar SH. (1959). The effects of propylthiouracil and perchlorate on the biogenesis of thyroid hormone. Endocrinology 65:198-210.

72. Döhler KD, Wong CC, von zur Mühlen A. (1979). The rat as model for the study of drug effects on thyroid function: Consideration of methodological problems. Pharmac. Ther. 5:305-318.

73. Davies DT. (1993). Assessment of rodent thyroid endocrinology: Advantages and pitfalls. Comp. Haematology Int. 3:142-152.

74. Hill RN, Erdreich LS, Paynter OE, Roberts PA, Rosenthal SL, Wilkinson CF. (1989). Thyroid follicular cell carcinogenesis. Fund. Appl. Toxicol. 12:629-697.

75. Greaves P. (1990). Cardiovascular system, heart and pericardium, cardiac weight changes, and cardiac hyptertrophy. In: *Histopathology of preclinical studies*. Greaves, P, ed. pp. 231-233. Elsevier, Amsterdam.

76. Foley GL. (2001). Overview of male reproductive pathology. Toxicol. Path. 29:49-63.

77. Creasy DM. (2001). Pathogenesis of male reproductive toxicity. Toxicol. Path. 29:64-76.

78. Davis BL, Travlos G, McShane T. (2001). Reproductive endocrinology and toxicological pathology over the life span of the female rodent. Toxicol. Path. 29:77-83.

79. McClain RM, Levin AA, Posch R, Downing JC. (1989). The effect of phenobarbital on the metabolism and excretion of thyroxine in rats. Toxicol. Appl. Pharm. 99:216-228.

ACKNOWLEDGEMENTS

A program of the magnitude of the updated TG 407 studies could only be undertaken with the willing participation and collective contributions of many parties. Special acknowledgement is due to the 13 participating laboratories in Phase-2 who contributed freely of their time and resources in the organisation and conduct, analyses of these validation studies. The contribution of CEFIC is acknowledged for the contribution of funds for and management of the chemical repository, particularly Dr. Elard Jacob (BASF, Germany) for the oversight of the chemical repository. The contributions of the governments of Japan and Korea and again of CEFIC are acknowledged in financially supporting laboratories to conduct this work. Many scientists and laboratories contributed their data and comments to assist the VMG-mammalian in the design and conduct of these studies. Special thanks are acknowledged to the Lead laboratory (Dr. A. Freyberger, Bayer HealthCare AG, Germany), consulting toxicologists (Dr. R. Bars of Bayer CropScience AG, Germany; Dr. P. Gelbke of BASF, Germany; Dr. N. Wallis of Syngenta, UK), and members of the Mammalian Validation Management Group. The financial support of CEFIC, ACC, and JCIA is acknowledged to support the secondment of Dr. W. Owens of Procter & Gamble, USA, to the OECD Secretariat for the purpose of the compilation and writing of this report.