**For Official Use** **ENV/JM/TG/EDTA(2001)1**

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development **22-Mar-2001**

**English - Or. English**

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

**ENV/JM/TG/EDTA(2001)1**
**For Official Use**

## Task Force on Endocrine Disrupters Testing and Assessment (EDTA) of the Test Guidelines Programme

# OECD REPORT OF THE INITIAL WORK TOWARDS THE VALIDATION OF THE RODENT UTEROTROPHIC ASSAY: PHASE ONE

**Third Meeting of the Validation Management Group for the Screening and Testing of Endocrine Disrupters (Mammalian Effects) to be held on 26-27 March 2001, beginning at 10:00 a.m. on 26 March**

**JT00104808**



**English - Or. English**

This document provides a consolidated report of the first phase work on the validation of the rodent uterotrophic assay. It contains the background on how the validation work was organised and carried out, and on the development and refinement of the standardised protocols used. The document also contains the details of the statistical analysis of the experimental work and the conclusions drawn from the initial phase one work. As this document provides a report of the first phase work only, it must be read in conjunction with reports of subsequent stages for a full appreciation of the OECD Validation work on the rodent uterotrophic assay.

Dr Joe Haseman of the US National Institute of Environmental Health Sciences was an independent consultant to the project, responsible for the overall statistical analysis of the phase one results. This overall analysis of the results was discussed at the second meeting of the Validation Management Group for the Screening and Testing of Endocrine Disrupters for Mammalian Effects (VMG-mammalian) on 20-21 January 2000. At the same meeting a report from a meeting of participating laboratories was discussed. The meeting of participating laboratories was held on 17 December 1999 to review the technical aspects of the protocol design and identify any practical difficulties encountered by the participating laboratories.

Invited experts, representatives of the lead laboratory and observers from participating laboratories and animal welfare organisations attended the second meeting of the VMG together with VMG members.

The VMG agreed that sufficient information had been obtained from this initial phase one work to confirm the reliability and transferability of the assay and to demonstrate that the assay was robust and reproducible for the detection of the strong oestrogen agonist – ethinyloestradiol (CAS No. 57-63-6). The VMG also agreed that the first phase work had demonstrated the feasibility of the protocol for detection of oestrogen antagonist as demonstrated by the results with the chemical ZM189, 154 (CAS No. 101908-22-9).

***ACTION REQUIRED:*** ***The Validation Management Group is invited to consider the consolidated report of the phase one work and confirm it as a comprehensive report of the work undertaken.***

# OECD REPORT OF THE INITIAL WORK TOWARDS THE VALIDATION OF THE RODENT UTEROTROPHIC ASSAY

## PHASE ONE

# FOREWORD

This document is the report of the first phase of the work on the OECD validation of the rodent uterotrophic assay. This phase was concluded in January 2000.

The document was drafted by the Secretariat. However, extensive parts were contributed by Dr Joe Haseman who provided the independent statistical analysis for the project. A scientific paper on the same work was also used *(Kanno J., Onyon L., Haseman J., Fenner-Crisp P., Ashby J. and Owens W. The OECD program to validate the rat uterotrophic bioassay to screen compounds for in-vivo estrogenic responses: Phase One. Environ. Hlth. Perspect. (2001) In press.)*.

This report provides a comprehensive summary of the phase one work carried out by the participating laboratories, including a detailed statistical evaluation of the results. The individual reports of the work carried out by participating laboratories and the raw data submitted are available directly from the Secretariat on request.

Contact for further details

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

Tel: 33-1-45-24-98-61
Fax: 33-1-45-24-16-75
www.oecd.org/ehs

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

**SUMMARY**

i) This report provides the results from an OECD inter-laboratory study conducted in 1999 to examine the reliability and transferability of a standardised OECD protocol for the rodent uterotrophic assay. This work is considered the first phase in a process to validate the rodent uterotrophic assay at the international level.

ii) The need to validate the rodent uterotrophic assay stems from the concerns that exist that ambient environmental levels of chemicals termed 'environmental oestrogens' may be causing adverse effects in both humans and wildlife due to the interaction of these chemicals with the endocrine system. Initial reviews of existing reports have found limited evidence for endocrine disruption in humans, but found several cases where local, high level exposures have resulted in effects in wildlife. In 1997, the OECD concluded that existing test methods were insufficient to identity such effects. As part of the OECD Test Guidelines Programme a *Special Activity on the Testing and Assessment of Endocrine Disrupters* was initiated to revise existing, and develop new OECD Test Guidelines for the testing of potential endocrine disrupters (http://www.oecd.org/ehs/endocrin.htm). A Task Force on Endocrine Disrupters Testing and Assessment (EDTA) was subsequently established to provide a focal point within OECD to consider and recommend priorities for the development of testing and assessment methods for endocrine disrupters.

iii) The rodent uterotrophic assay was one of three *in vivo* tests selected by the EDTA for international co-operative work. The principle of the rodent uterotrophic assay is that the uterus is under the control of oestrogens to stimulate and to maintain growth. If the endogenous source of this hormone is not available, either because of immaturity of the female animals, or because the animals have been ovariectomised, the animal requires an exogenous source to initiate or restore uterine growth. Chemicals that act as agonists may be identified as potential endocrine disrupters if they cause an increase in uterine weight or as antagonists if they cause a relative decrease when co-administered with a potent oestrogen. If successfully validated, the rodent uterotrophic assay may serve as a tool for the prioritisation of chemicals for further testing. Two animal models, the immature female rat and the adult ovariectomized rat, were used as the basis for the uterotrophic assay. Four standardised protocols (A, B, C and C') were compared.

iv) Dose-response data from nineteen participating laboratories using a high potency reference agonist, 17 alpha ethinylestradiol, otherwise referred to EE (CAS No. 57-63-6), and an antagonist, ZM 189,154 (CAS No. 101908-22-9) were used as the reference test substances for the phase one study.

v) All laboratories and all protocols were successful in detecting increases in uterine weight using EE. No substantive performance differences between models were identified when levels of EE ranging from 0.01 – 10 microgram/kg per day were tested. Significant uterine weight increases were achieved under a variety of different experimental conditions (e.g., strain of rat, diet, housing conditions, bedding, vehicle, etc.). For each protocol, there was generally good agreement among laboratories with regard to the actual EE doses producing the first significant increase in uterine weights and also EE doses achieving the maximum uterine response doses within and among laboratories.

vi) An evaluation of the relative sensitivities of the four protocols, assessed by identifying the lowest dose at which statistical significance was first demonstrated found that Protocol A (using immature animals and exposure by oral gavage) was less sensitive than the other three in producing equivalent increases in uterine weight. This result was not unexpected given the fact that the other three protocols include exposure by sub-cutaneous injection. For the adult ovariectomised protocols, limited data suggest that using a seven day dosing regimen (Protocol C') may have a sensitivity advantage over a three day exposure period (Protocol C). However the consequence of other effects seen in the longer exposed animals such as reduced body weights have not yet been fully considered.

vii) The feasibility of the assay in identifying an oestrogen antagonist was successfully demonstrated using a potent reference compound (ZM 189.154) (with the exception of one laboratory) by a comparable reduction in uterine weights relative to groups receiving EE alone.

viii) To broaden the interpretation of results the information on dose-response was modelled using the Hill equation. The modelled data indicated general agreement with the overall statistical analysis and provided a basis for the calculation of doses of ethinyl oestradiol that are expected to produce specific levels of uterine weight increase.

ix) Very few protocol refinements were considered necessary to improve the conduct and performance of all 4 protocols being investigated.

x) Overall it was concluded from this first phase of the work that both models (immature rat and adult ovariectomised rat) appear robust, reliable and transferable across laboratories for high potency oestrogen agonists such as EE. Further work to confirm these findings with chemicals of weaker oestrogenic activity, and to further examine sensitivity, is needed to fully characterise the performance and necessary level of standardisation of the assay. This will be considered in later phases of the work.

## INTRODUCTION

1. The need to validate the rodent uterotrophic assay stems from the concerns that exist that ambient environmental levels of chemicals termed 'environmental oestrogens' may be causing adverse effects in both humans and wildlife due to the interaction of these chemicals with the endocrine system (1). Initial reviews of existing reports have noted limited evidence for endocrine disruption in humans, but noted several cases where local, high level exposures have resulted in effects in wildlife (2) (3) (4). In 1997, the OECD concluded that existing test methods were insufficient to identity such effects.

2. The OECD initiative to develop and validate *in vitro* and *in vivo* assays for the detection of chemicals that may interfere with the endocrine response was taken following the recommendations of a number of national, regional and international workshops (2) (5) and following a detailed OECD review of the status of existing test and research methods. This Detailed Review Paper focussed on test methods for sex hormone disrupting chemicals (6). As part of the OECD Test Guidelines Programme a *Special Activity on the Testing and Assessment of Endocrine Disrupters* was initiated to revise existing, and develop new OECD Test Guidelines for the testing of potential endocrine disrupters. Please see the following internet site for further details: (http://www.oecd.org/ehs/endocrin.htm).

3. A conceptual framework for the testing and assessment of chemicals is being developed to identify short and long-term assays of increasing complexity and detail to gather information on potential endocrine disrupters. The assays and techniques include: 1) structure activity relationships and *in vitro* assays that would identify a chemical based on certain intrinsic characteristics, e.g., oestrogen receptor binding affinity; 2) short-term *in vivo* assays to demonstrate relevant activity in the intact animal, e.g., the uterotrophic assay, and the Hershberger assay 3) long-term assays involving exposure to the test substance at different stages of the development of the animal, e.g., the two-generation reproductive assay.

4. The OECD framework aims to develop these assays as multipurpose tools rather than as a rigid scheme. The purpose and use of a bioassay may vary depending on the chemical substance and the already available toxicological data on that chemical. An early screen for one test substance could for another be a means to determine the test substance's mode of action (7). The uterotrophic assay, once validated, would fit within the framework as a short-term *in vivo* assay.

5. The rodent uterotrophic assay is based on the principle that the uterus is under the control of oestrogens to stimulate and to maintain growth. If the endogenous source of this hormone is not available the animal requires an exogenous source to initiate and or restore uterine growth.

6. The aim of the validation of the rodent uterotrophic assay is to develop a robust, reliable and relevant test method for the detection of chemicals that have the potential to act like and, consequently, interfere with endogenous female sex hormones. The rodent uterotrophic assay is therefore expected to identify chemicals which act like oestrogen agonists or antagonists (sometimes referred to as anti-oestrogens).

### Validation of test methods

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

8. Most experience with the validation of test methods has been in the development of alternative test methods leading to the reduction, refinement and/or replacement of methods using animals. It is in these areas that a considerable body of expertise of test method validation has been established. Some

countries and regions have formally incorporated this expertise in specialist organisations, to oversee and make recommendations on whether a particular test has been validated.

9. The Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Test Methods (8) provides the principles of validation which are followed by OECD. Work is underway to incorporate these principles into a revised OECD Guidance Document for the Preparation of Test Guidelines (Guidance Document no.76). The Solna principles are consistent with approaches used in Europe, particularly those of the European Centre for Validation of Alternative Methods (ECVAM) and in the US by the International Co-ordinating Committee on Validation of Alternative Methods (ICCVAM).

10. For any new or revised test method (animal or non-animal) to be valid for use for the testing and assessment of chemicals, it must meet the following minimum criteria:

- A rationale for the test method should be available. This should include a clear statement of the scientific need and regulatory purpose.
- The relationship of the endpoint(s) determined by the test to the *in vivo* biological effect and to the toxicity of interest must be addressed.
- The limitations of a method must be described, e.g., metabolic capability.
- A formal detailed protocol must be provided and should be readily available in the public domain. It should be sufficiently detailed to enable the user to adhere to it and it should include data analysis and decision criteria.
- Test methods and results should be available preferably in an independent peer reviewed publication. In addition, the result of the test should have been subjected to independent scientific review.
- Intra-test variability, repeatability and reproducibility of the test method within and amongst laboratories should have been demonstrated. Data should be provided describing the level of inter- and intra-laboratory variability and how these vary with time.
- The test method's performance must have been demonstrated using a series of reference chemicals preferably coded to exclude bias.
- The performance of test methods should have been evaluated in relation to existing relevant toxicity data as well as information from the relevant target species.
- All data supporting the assessment of the validity of the test methods including the full data set collected in the validation study must be available for review.
- Normally, these data should have been obtained in accordance with the OECD Principles of Good Laboratory Practice (GLP).

11. In 1998, the Joint Meeting of the OECD Chemicals Group and Committee and Working Party on Chemicals, Pesticides and Biotechnology (the Joint Meeting) decided that the criteria and approaches for the validation of alternative test methods should apply equally to the development of all toxicology tests, *in vitro* or *in vivo,* and that similar criteria would also apply to tests for ecotoxicological tests. The Joint Meeting agreed that flexibility should be shown when making decisions so that any validation work is appropriate for a specific purpose. The degree of flexibility should be identified on a case-by-case basis. Most importantly all decisions on the extent and design of the validation study should be fully transparent and documented.

12. Historically, most validation work has been conducted in formal stages such as test development, optimisation/prevalidation, peer review, assessment of readiness for validation and validation itself. The rodent uterotrophic assay has been in widespread use since the 1930's and is therefore different in certain aspects to the development of a new alternative test. Detailed consideration must be given to the experience gained over this period during its widespread use.

13. In considering how best to apply flexibility to the validation to the rodent uterotrophic assay, it was decided to tackle the OECD validation work in phases – to provide information that would address the Solna Principles but at the same time take into account the substantial experience in the use of the assay. Instead of rigidly defining the work - a number of phases were planned, each with the possibility of influencing the following one. In this way each phase of the work would use and build upon the results of earlier phases. This approach is represented in Figure 1 which shows how the assessment process of the relevance and reliability of new or significantly revised testing methods for hazard characterisation can be undertaken in a stepwise, yet flexible, manner while still providing the information necessary to address the Solna criteria and principles.

14. This report of the first phase of the OECD validation of the rodent uterotrophic assay forms an important part of understanding the rodent uterotrophic assay, the possibilities for its application, any limitations and its reliability and relevance. It forms the basis for the design and technical details of the follow-up work of the full validation of the assay, which is expected to continue with a second phase in 2000/2001.

**Phase one of the OECD work on the rodent uterotrophic assay**

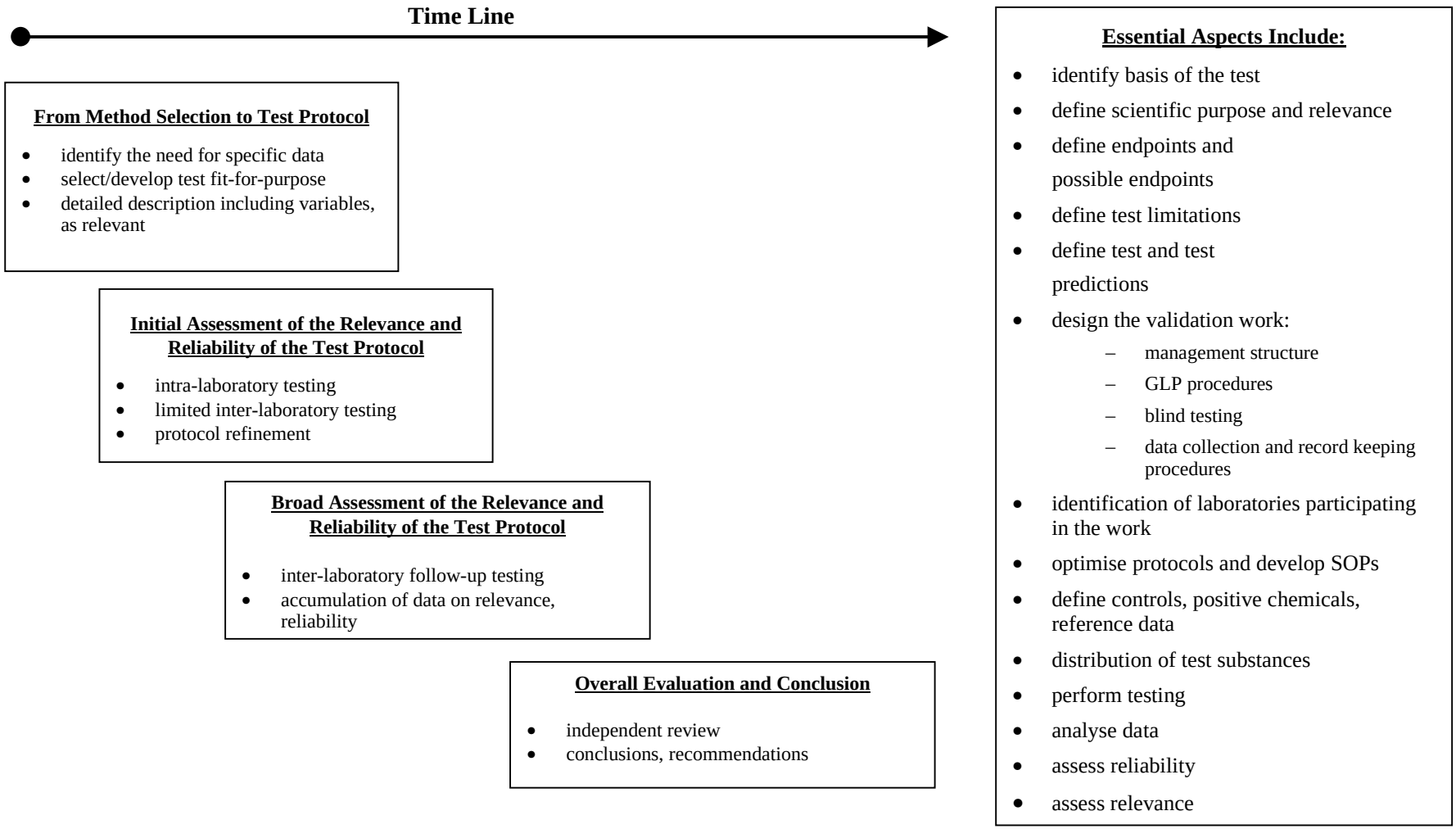
15. The aim of phase one of the OECD work on the rodent uterotrophic assay was to develop and refine a standardised test protocol for the assay, based largely on those already existing and the experience gained in their use. The aim was also to generate initial data on the inter- and intralaboratory variation in the use of the assay.

**Objectives of phase one**

16. Specifically, the first phase of this OECD validation work was designed to:

- demonstrate the dose-response relationship between uterine weight in immature female rodents and ovariectomized female rodents and oral or subcutaneous injection of the reference oestrogen 17alpha ethinyl oestradiol (EE) (CAS No. 57-63-6);
- enable variation between laboratories to be investigated;
- enable a comparison of the results from similar four protocols; i.e. the comparison of two animal models (immature female rats and mature ovariectomised rats) and comparison of two routes of exposure (oral gavage and sub-cutaneous injection);
- assist in selecting the appropriate reference dose of ethinyl oestradiol to use in a subsequent protocol(s) for investigating chemicals of unknown oestrogenic activity;
- confirm the anti-oestrogenic effects of the oestrogen antagonist known as ZM 189.154 (CAS No. 101908-22-9) and;
- enable necessary protocol refinements to be identified and discussed.

**Figure 1**
**Assessment Process of the Relevance and Reliability of New or Significantly Revised Testing Methods for Hazard Characterisation**



**Organisation of the validation project**

17. The National Co-ordinators of the Test Guidelines Programme established, in early 1998, a Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to provide a focal point within OECD to consider and recommend priorities for the development of testing and assessment methods for endocrine disrupters. Members of EDTA, having expertise in mammalian and ecotoxicological science, were nominated by Member countries, industry and environmental groups.

18. The EDTA subsequently set up two Validation Management Groups (VMG) one for mammalian test methods and one for ecotoxicology test methods. The role of both VMGs is to oversee and manage the conduct of the validation work at the practical level. A schematic diagram is provided in Figure 2, which describes the role and structure of these two OECD Validation Management Groups.

19. The VMG (mammalian) is made up of sixteen experts nominated by Member countries and non-government organisations. Membership of the VMG has a good balance of experts from disciplines including toxicology, endocrinology, test method development and validation work and is representative of the major OECD regions. Representatives of ICCVAM and ECVAM are formally part of the Group to provide additional independent objective review, given their extensive involvement in validation work of other tests and help regulatory acceptance, as necessary in future.

20. The first task of the VMG (mammalian) was to develop a standardised protocol(s) for the conduct of the rodent uterotrophic assay. After this, expressions of interest were sought from laboratories wishing to participate in the practical validation work. These laboratories were also invited to participate in meetings of the VMG, whenever appropriate. The final selection of participating laboratories was determined by the willingness of the laboratory to strictly follow the standardised OECD test protocol, provide their experimental data for independent statistical analysis, provide a report on actual test conditions at the time of the experiment and to provide a report of their experimental findings.

21. Nineteen laboratories from eight Member countries (Denmark, France, Germany, Japan, Korea, the Netherlands, the United Kingdom and the United States) participated in the phase one work. The number of laboratories from each country is shown in Table 1. Further details of the participating laboratories are given in Annex 1. Five laboratories were funded to participate through the European Chemical Industry Council (CEFIC) and one through the American Chemistry Council. Other laboratories participated on a voluntary and self-supporting basis. The laboratories included those who had many years of experience of using the assay and those who had not used the assay before.

22. The National Institute for Health Sciences, Japan took on the responsibility of Lead laboratory drafting the standard experimental protocol on behalf of the VMG; answering day to day technical questions from participating laboratories. The Secretariat provided the overall project co-ordination.

23. The test substances were donated by Schlering (EE) and AstraZeneca (ZM) from a single lot and were sent on request to the participating laboratory by Schelering and AstraZeneca, respectively.

**Figure 2**

**SCHEMATIC DIAGRAM TO DESCRIBE ROLE AND STRUCTURE OF OECD VALIDATION MANAGEMENT GROUP VMG (MAMMALIAN)**

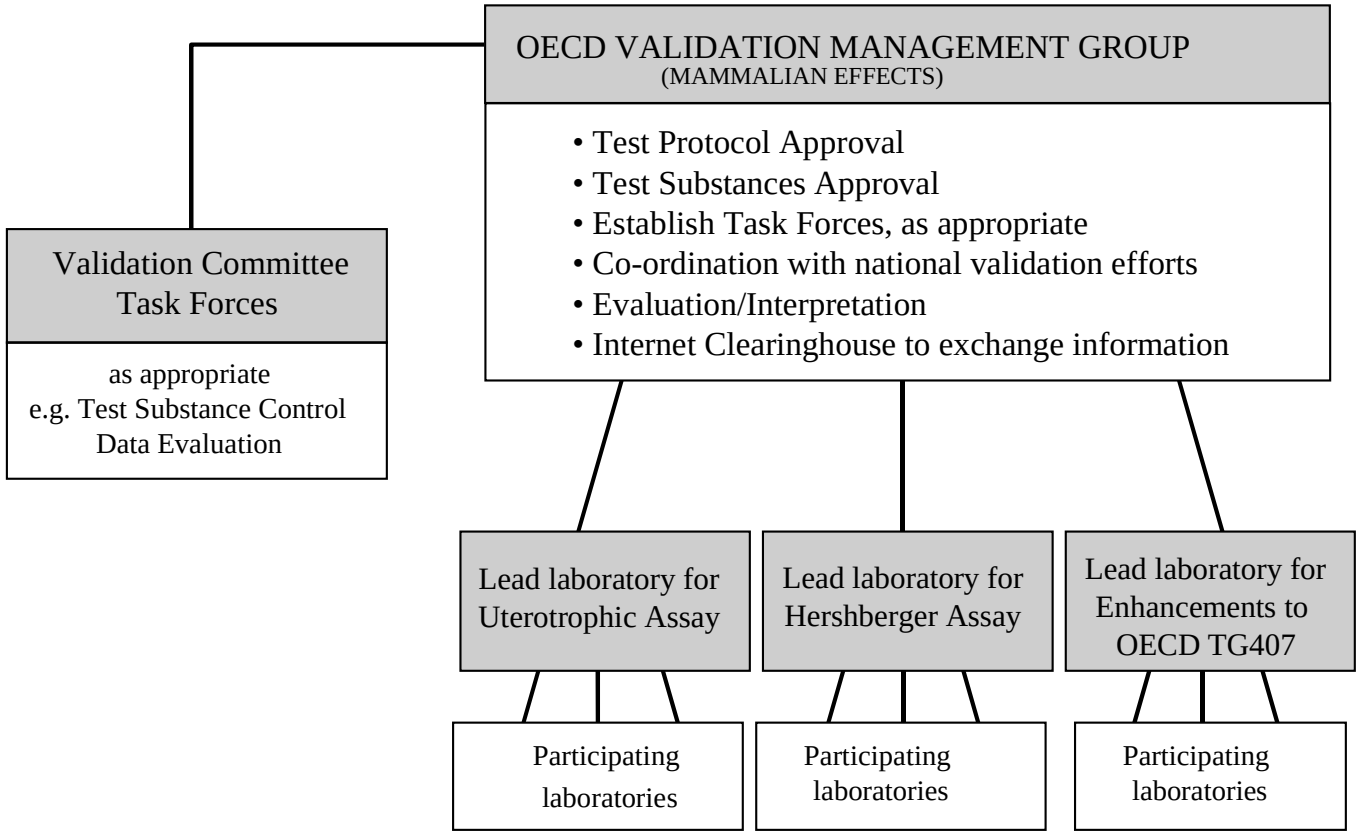


**Table 1: Number of laboratories participating in the phase one validation work**

| Country | Number of Laboratories |
|---|---|
| Denmark | 1 |
| Japan | 6 |
| Germany | 3 |
| France | 2 |
| Korea | 2 |
| Netherlands | 1 |
| United Kingdom | 2 |
| United States | 2 |
| Total | 19 |

24. Participating laboratories each wrote their own experimental protocol or standard operating procedure (SOP) based on the common agreed OECD standardised protocol(s). Consistency with the OECD standard protocol(s) was confirmed by the lead laboratory before the practical work commenced. Care was taken to ensure that full details of any aspects not already standardised were reported. Each laboratory tested ethinylestradiol (EE) at the agreed prescribed doses and tested the antagonist ZM189.54 at two prescribed doses of EE.

25. Experimental work took place over the period July 1999-September 1999. On completion of the experimental work, participating laboratories submitted their individual detailed data on a standardised Microsoft excel spreadsheet. This was then submitted to an independent statistician Dr Joe Haseman at the National Institute for Environment Health Sciences for an overall analysis of the laboratory results.

26. A meeting of participating laboratories was held in December 1999 following the completion of the experimental work. The aim of the meeting was to ensure that the VMG were advised of: a) any relevant practical experience from the participating laboratories that may help to refine the standardised protocol; and b) any practical issues which may influence the interpretation the data analysis.

27. Results from the participating laboratories, including the independent statistical evaluation were provided to the VMG for discussion at its second meeting on 20-21 January 2000.

## METHODS

### Test Method Development and Refinement

28. Historically, several candidate systems are available as screening tests for evaluation of oestrogenic activity: a vaginal cornification and keratinization response (9), a water imbibition response of the uterus after a single dose of the test compound (10), and a uterine tissue weight increase after several doses of the test compound (11-13), the so-called rodent uterotrophic assay.

29. The EDTA reached consensus to select the rodent uterotrophic assay following an expert workshop, held in Washington DC in 1998 which considered the recommendations made by the US EDSTAC, the OECD's Detailed Review Paper on the appraisal of test methods for sex-hormone disrupting chemicals (16) and the current application and use of the assay in OECD countries. Recommendations

made by the Workshop included that the uterotrophic assay should be considered a high priority for validation.

30. At the time this consensus was reached the rodent uterotrophic assay was already in widespread use, and had been since the 1930's. Its advantages included that its main endpoint (uterine weight) is a natural, biological target of endogenous oestrogen, the biological response is rapid and can be quantified and evaluated statistically. The conduct of the assay is within the means of most testing facilities without the use of specialised techniques.

31. A number of different protocols for the uterotrophic assay remain in use. A major barrier to the widespread and routine use of the assay is the number of major protocol designs encountered. These include:

- The use of immature animals which had not yet begun oestrus cycling;
- The use of surgically treated ovariectomised adult animals;
- The route of administration (oral, subcutaneous or intraperitoneal);
- The species used (i.e. rats and mice);
- The endpoints to be studied; and
- The length of exposure period.

32. In each case consensus was needed on whether there was an adequate scientific rationale preferring one design over another and whether changing the variable would lead to a reduction in the sensitivity of the assay. All things being equal, the basis for the consensus reached was that a standardised protocol should be developed based on issues of practicality – given the need for the assay to be reliable and transferable for use in different laboratories worldwide.

**Immature or adult animals**

33. The possibility of using either immature or mature animals creates two uterotrophic models, both based on a non-functional hypothalamic-pituitary-gonadal axis in order to ensure low levels of endogenous oestrogens. Low levels of endogenous oestrogens provides for a sensitive and consistent uterine response when testing chemicals. This applies both to agonists and antagonists of oestrogens. The immature model uses young female animals before significant ovarian oestrogen synthesis has begun and subsequent regulation by the hypothalamic-pituitary-gonadal axis. The adult model uses adult female rats following ovariectomy (OVX) and thus removing the primary source of oestrogen synthesis.

34. While recognising animal welfare concerns both for and against the mature animal model and also issues of practicality, it was agreed by the VMG that both models should be included in the first phase of the validation work, because there was insufficient information to make a judgement on the most reliable and/or relevant model for the future.

**Route of administration**

35. Route of administration is a key issue affecting the sensitivity of the assay particularly for chemicals that need metabolism to active forms. The most sensitive route chosen might vary depending on the nature of the chemical being tested. Initially it was decided to focus efforts on the sub-cutaneous route of exposure only – but it was subsequently agreed that the oral route should be included for comparison and to provide a linkage to subsequent work on the assay using test chemicals metabolised by different routes.

**Preferred species**

36. The uterotrophic response has been used to evaluate oestrogenic activity using a number of mammalian and avian species. As the rat has become the preferred species for reproductive and developmental toxicity testing, the rat was chosen as the test species for the phase one and subsequent phases of the validation work within OECD. This decision was made with the acknowledgement that there is no information indicating that the mouse ought to be ruled out from future use.

**Endpoint to be studied**

37. Weight of the uterus was considered the major endpoint for inclusion in the assay but it was decided that further consideration was needed on whether this should be measured wet, or after the luminal fluid of the uterus had been removed (blotted). It was not decided whether other additional measures of cell growth such as BrdU staining, would increase sensitivity. Inclusion of such measures was left to the individual laboratories to carry out as supplementary work.

**Other variables**

38. In addition to the major protocol variables discussed above, test vehicle and animal diet were other aspects of the development of the standardised protocol that were very carefully considered. Several authors have reported that contaminated diets, presumably due the presence of phytoestrogens, could influence the baseline uterine weight (14-18).

39. The phytoestrogen content of commonly used laboratory diets remain largely unknown. As an alternative to specifying that a particular synthetic diet should be used, each laboratory was instructed to record full details of the diet used and retain a sample of the diet for further study and analysis if necessary.

40. To evaluate whether the test vehicle might also influence the results, both an untreated control and a vehicle control were included in the standardised protocol to allow for the detection of any significant contamination of the vehicle with phytoestrogen(s). If significant variations in the uterine weights from the vehicle control group were observed, the contributions of the vehicle to the increased uterine weight could then be investigated further. Accordingly, all laboratories were required to record full details of the vehicle used and retain a sample for subsequent investigation.

41. Apart from these points other characteristics of each of the four protocols were standardised as far as possible, recognising that it would not be practical to standardise every aspect. In many respects the VMG considered it was ultimately undesirable to standardise every aspect because if the rodent uterotrophic assay was to be adopted as an OECD Test Guideline, as much flexibility as possible should be maintained to ensure wide use while still ensuring that the assay will measure the effect of interest. The degree of standardisation of the four protocols achieved is summarised in Table 2.

42. Depending on the outcomes of the first phase, further refinement of the protocol would be considered.

**The Phase One Experimental Design**

43. Nineteen laboratories participated in the validation study following undertaking one or more of the four protocols. Protocols A and B use the immature animal model and exposure by oral gavage and subcutaneous injection respectively and Protocols C and C' use the ovariectomised adult animal model (OVX) with exposure by sub-cutaneous injection and exposure for 3 days and 7 days respectively.

**Table 2: Degree of standardisation for the protocols used in the phase one of the validation work**

| | |
|---|---|
| **Standardised aspects** | • Same lot of test substances<br>• Doses of test substance(s)<br>• Number of test groups (including vehicle and untreated controls<br>• Total volume of test substance and vehicle administered<br>• Body weight range of animals<br>• Age of supplied animals<br>• Conditions of supply of animals ( with or without foster dam)<br>• Acclimatisation of animals before exposure to test substance<br>• Age of animal at start of treatment<br>• Number of animals per cage (max. 6 animals per cage)<br>• Endpoints to be measured – blotted and uterine weight (to nearest 0.1 mg)<br>• Excision of uterus and procedure for blotting uterus<br>• Animal housing conditions , lighting, interval, room clean-up, temperature etc<br>• Reporting of experimental results using a standardised spreadsheet |
| **Factors that varied from laboratory to laboratory** | • Strain of rat<br>• Vehicle for delivery of test substance<br>• Bedding Material<br>• Rodent diet<br>• Method of humane killing<br>• Identification system for individual animals and groups |
| **Optional additional parts** | • Food consumption (grams per rat per day)<br>• Histopathology of vagina, uterus and ovaries, including optional use of BrdU staining as a measure of cell division |

44. Copies of the essential core standardised OECD protocols for protocols A, B and C are provided in Annex 2. Protocol C' differed from protocol C only with respect to the duration of response, and is therefore included as a variation in the description of Protocol C. The key aspects of the four protocols followed in the phase one validation work are given in Table 3.

**Table 3: Key aspects of the four protocols followed in the phase one validation work**

| | |
|---|---|
| **Protocol A** | Immature female rodents between 19 and 20 days old , before the onset of puberty with exposure by oral gavage for three days |
| **Protocol B** | Immature female rodents between 19 and 20 days old, before the onset on set of puberty with exposure by s.c. injection for three days |
| **Protocol C** | Mature female rodents ovariectomised at 6 weeks old or greater with exposure by s.c. injection for three days |
| **Protocol C'** | Mature female rodents ovariectomised at 6 weeks old or greater with exposure by s.c. injection for seven days |

**Description of the work carried out in phase one by the participating laboratories**

45. The lead laboratory was the National Institute of Health Sciences of Japan. A total of 19 laboratories from Denmark, France, Germany, Japan, Korea, the Netherlands, the United Kingdom, and the United States participated in the phase one experimental work. Laboratories were free to choose which of the four protocols they would follow for the phase one work. Some laboratories conducted a single experiment while others conducted several experiments according to more than one of the standardised protocols, enabling a comparison of the performance of the different protocols to be investigated. In total, 41 separate experiments were undertaken.

46. 16 laboratories from 7 countries carried out Protocol A; 12 laboratories from 6 countries carried out Protocol B; 9 laboratories from 3 countries carried out Protocol C; and 4 laboratories from one country carried out the extended Protocol C'. This participation provided information on which to compare different laboratories' performance of the same protocols as well as comparisons of the different protocols by the same laboratory.

47. Table 4 shows which protocols were followed by which laboratories.

**Table 4: The protocols followed by the 19 participating laboratories in the phase one validation work (see Annex 2 for full description of laboratories)**

| | **Laboratory** | **Protocol A** | **Protocol B** | **Protocol C** | **Protocol C'** |
|---|---|---|---|---|---|
| 1 | Japan Bioassay Research Centre | √ | √ | √ | √ |
| 2 | Sumitomo | √ | √ | √ | |
| 3 | Mitsubishi | √ | √ | √ | √ |
| 4 | BASF | √ | | | |
| 5 | Exxon Biosciences | √ | | | |
| 6 | CIT | √ | | | |
| 7 | FDSC | √ | √ | √ | √ |
| 8 | AstraZeneca[1] | √ | √ | √ | |
| 9 | CITI[2] | √ | √ | √ | |
| 10 | Korea (Chung)[3] | √ | √ | | |
| 11 | IET | √ | √ | √ | √ |
| 12 | WIL Research | √ | √ | | |
| 13 | Bayer | √ | | | |
| 14 | Rhone Poulenc[4] | √ | | | |
| 15 | TNO | √ | √ | | |
| 16 | Free University of Berlin | √ | | | |
| 17 | Denmark | | √ | | |
| 18 | Korea (Park) | | √ | √ | |
| 19 | Huntingdon | | | √ | |
| | **Total** | 16 | 12 | 9 | 4 |

[1] AstraZeneca is now Syngenta
[2] CITI is now Chemicals Evaluation and Research Institute (CERI)
[3] Research Institute of Chemical Technology, Korea
[4] Rhone Poulence is now Aventis

**Test System**

48. The test system chosen for the validation work was the laboratory rat. Literature reports have not shown any consistent different uterotrophic responses among different strains. Therefore laboratories were encouraged to use the strain of rat that they commonly used and therefore had the most historical control data.

49. The immature protocol specifies the age of the animals before exposure to the test substance can commence. This is specified as 19 or 20 days of age (day of birth counted as day 1). Such precision was necessary to unify the groups of animals as the same age and consequently being likely to be within the same pre-pubertal developmental period. Having good historical control data on the developmental milestones of the strain of animals being used was also considered essential, as was having good animal suppliers who could supply animals of precise ages. The danger of not ensuring synchronisation of the developmental period is that animals which had began puberty could be producing endogenous oestrogens which could lead to an increase in control uterine weights and possibly an overall decreased sensitivity of the assay.

50. For the adult OVX animals, ovariectomy was to occur at 6 weeks of age, or later, with a minimum period of 1 week after surgery before administration of the reference compounds.

51. Limiting the weight variability of the rats was another precaution taken to limit the chances of older, possibly pubertal, animals being inadvertently included in the immature study. It was stated that the body weight variation should not exceed log or be within +/- 5g of the mean weight. Allocation of animals into experimental groups was also done in principle by randomisation with the overall criterion that all groups of animals had the same mean weight population, within the 5% probability level.

**Chemicals**

52. The reference oestrogen agonist was 17alpha-ethinyl estradiol (EE) (CAS No. 57-63-6), and the reference oestrogen antagonist was the pure antagonist ZM 189,154 (ZM) (CAS No. 101908-22-9). The same lot of each chemical was distributed directly from the chemical suppliers. As only limited amounts of ZM 189.154 were available, only two experimental groups were included in each protocol using this test substance in combination with the positive reference chemical EE. This would provide an indication of the feasibility of the assay to detect oestrogen antagonists.

53. The doses of EE and ZM to be administered were specified to ensure that results could be statistically compared. For EE, a series of seven doses in half-log steps from 0.01 to 10 µg/kg/day was specified for both administration by oral gavage (p.o.) and sub-cutaneous injection (s.c.). This range of doses was taken from literature reports (16) and confirmed by the experience of other researchers in the field. The objective was to achieve a complete dose-response curve for EE if this was possible.

54. For the ZM, the reference EE dose specified was, in Protocol A, 3.0 µg/kg/day and, in Protocols B, C and C', 0.3 µg/kg/day with two ZM doses, 0.1 and 1.0 µg/kg/day to be co-administered. The two doses of ZM for each protocol were selected from the literature (19) after agreeing that the maximum antagonist effects would probably best be demonstrated near to the top of the ascending dose-response curve but before a stable maximum had been achieved. A lower dose of EE was also included near to the expected limit of detection of the assay for EE.

55. The experimental design for Protocol A (p.o.) exposure and Protocols B, C and C' (s.c. injection) are shown in Table 5.

56. EE and ZM 189154 were dissolved in a minimal amount of 95% ethanol and diluted to the final working concentration in the test vehicle (e.g. corn, arachis, sesame or olive oil). For the ZM189154, gentle heating up to 60 Cº was needed for dissolution.

**Experimental Conditions**

57. If adopted as an OECD Test Guideline the uterotrophic assay will be widely practised and therefore participating laboratories will use their traditional diet, vehicle, and housing procedures. The approach adopted in the phase one validation work followed this practice. Participating laboratories were encouraged to obtain their experimental animals from their historical animal suppliers. General instructions on acclimation and housing were provided. However, specific instructions were needed on arrangements for animal supply in accordance with the age of animals to be used. Immature animals were to be transported together in litters accompanied by the dam or a foster dam. Alternatively, the rodents were scheduled to arrive as a litter when they were 17 days old. A room temperature of 22 ± 3 Cº was specified together with a relative humidity 30-70%; artificial lighting with a 12 hour light and 12 hour dark cycle; and feed and drinking water (tap or filtered) provided *ad libitum*.

58. Each laboratory recorded the specific experimental and housing conditions and strain of rat used and retained samples of vehicle and diet for subsequent investigation, if necessary. The same test vehicle was used for EE and ZM. Individual animals were uniquely identified, e.g. by ear tags or tail tattoos, and each group coded, e.g. by a letter and a colour on housing cages.

**Test Procedure**

59. Animals were allowed to acclimatise to laboratory conditions for 2-3 days before beginning the test procedure. Mature animals had an additional period of post-operative recovery of one week.

60. The test substance (EE) was administered to the animals in the same sequence daily for 3 consecutive days by either oral gavage or sub-cutaneous injection. Body weights were measured daily and the volume of test substance adjusted as necessary to maintain the same daily dose. On the fourth experimental day the animals were humanely killed in the same sequence and the uteri carefully dissected.

61. Historically, most published uterotrophic results describe measuring absolute uterine weights following careful blotting of the uterus after the wall has been nicked or split to allow the luminal contents to drain out. The rationale given for measuring blotted uterine weight data is usually that the wet weights were more variable, and that this variability was increased by the possible loss of luminal fluid during dissection and tissue handling.

62. For purposes of the initial validation, it was decided to include both wet and blotted uterine weight endpoints and to therefore establish their variability in the four “standardised” protocols. The uterine nicking and blotting technique was adopted in all protocols. Both wet and blotted weights were recorded to the nearest 0.1 mg in all protocols. As several laboratories were performing the assay for the first time and in order to standardize procedures, a videotape of the ovariectomy procedure and the uterine dissection and preparation procedures was prepared and distributed to the participating laboratories.

63. Additional measurements were made by some laboratories, such as food consumption and histopathology of the uterus, the vagina and ovaries. To assist comparability of any histopathology undertaken, guidance was given in the protocol on appropriate tissue cross sections to make.

**Table 5: Experimental Groups and Design for Protocols A, B, C and C'**

| Groups | N = | Dose | | Route | Maximum total p.o. volume/day/rat | Dose | | Route | Maximum total s.c. volume/day/rat |
|---|---|---|---|---|---|---|---|---|---|
| | | EE (microgram/kg) | ZM189154 (milligram/kg) | | | EE (microgram/kg) | ZM189154 (milligram/kg) | | |
| 1 (untreated control) | 6 | 0 | 0 | Not-applicable | Not-applicable | 0 | 0 | Not-applicable | Not-applicable |
| 2 (vehicle control) | 6 | 0 | 0 | p.o. | 5ml/kg/day | 0 | 0 | s.c. | 4ml/kg/day |
| 3 | 6 | 0.01 | 0 | p.o. | 5ml/kg/day | 0.01 | 0 | s.c. | 4ml/kg/day |
| 4 | 6 | 0.03 | 0 | p.o. | 5ml/kg/day | 0.03 | 0 | s.c. | 4ml/kg/day |
| 5 | 6 | 0.10 | 0 | p.o. | 5ml/kg/day | 0.10 | 0 | s.c. | 4ml/kg/day |
| 6 | 6 | 0.30 | 0 | p.o. | 5ml/kg/day | 0.30 | 0 | s.c. | 4ml/kg/day |
| 7 | 6 | 1.00 | 0 | p.o. | 5ml/kg/day | 1.00 | 0 | s.c. | 4ml/kg/day |
| 8 | 6 | 3.00 | 0 | p.o. | 5ml/kg/day | 3.00 | 0 | s.c. | 4ml/kg/day |
| 9 | 6 | 10.00 | 0 | p.o. | 5ml/kg/day | 10.00 | 0 | s.c. | 4ml/kg/day |
| 10 | 6 | 3.0 | 0.1 | p.o. | 5ml/kg/day | 0.3 | 0.1 | s.c. | 4ml/kg/day |
| 11 | 6 | 3.0 | 1.0 | p.o. | 5ml/kg/day | 0.3 | 1.0 | s.c. | 4ml/kg/day |

N = number of rats per experimental group

**Table 6: Rat strains, vehicle, animal diets, and bedding used by participating laboratories**

| | Laboratory | Strain | Vehicle | Diet | Bedding | Animals/Cage |
|---|---|---|---|---|---|---|
| 1 | Japan Bioassay | SD Crj CD (SD) IGS | Corn Oil | CRF-1 PD | None | 3 |
| 2 | Sumitomo | SD Crj CD (SD) IGS | Corn Oil | CRF-1 PD | None | 2 |
| 3 | Mitsubishi | SD Crj; CD (SD) IGS | Sesame Oil | MF | Autoclaved Hardwood | 3 |
| 4 | BASF | Wistar Rj WI (SpF) Han | Olive Oil | GKD | SNIFF ( type ¾) | 3 |
| 5 | Exxon Biosciences | SD CD (SD) IG SBR | Corn Oil | PMI CRD | None | 2 |
| 6 | CIT France | SD CD (SD) GS BR | Corn Oil | AO4C PD | Autoclaved sawdust | 2 |
| 7 | FDSC-Japan | SD Crj; CD (SD) IGS | Sesame Oil | CE2 | None | 3 |
| 8 | AstraZeneca (Syngenta) | SD Alpk ApfSD | Peanut Oil | SDSRM1 | Shredded paper | 3 |
| 9 | CITI – Japan | SD CD(SD) IGS | Olive Oil | MF | None | 1 |
| 10 | Chung – Korea | SD CRJ CD | Corn Oil | Cheil CRC | None | 2 |
| 11 | IET-Japan | Wistar JCL | Corn Oil | MF | None | 3 |
| 12 | WIL Research | SD Crl CD (SD) IGS BR | Corn Oil | PMI CRD 5002 | Ground corn cobs | 3 |
| 13 | Bayer | Wistar HSB/CpbWU | Corn Oil | Altromin 1324 | Low dust wood granules | 3 |
| 14 | Rhone Poulenc (Aventis) | SD ICO-OFA SD (OIPS) caw | Corn Oil | AOC4 CRC | None | 2 |
| 15 | TNO – NL | Wistar Crl (Wi) WU BR | Corn Oil | SDS RM3 | None | 6 |
| 16 | Berlin | Wistar HSD CPD Wa | Peanut Oil | Altromin | Wood Chip | 3 |
| 17 | Denmark | Wistar Han | Peanut Oil | Altromin 1324 | Tapvei | 3 |
| 18 | Korea – Park | SD CRI CD | Corn Oil | PMI CRD | Autoclaved elm tree | 3 |
| 19 | Huntingdon | SD HSD | Corn Oil | SDS RMI ( e) | None | 3 |

SD : Sprague Dawley
PD : Pelleted diet ; SDS Special Diet Services; CRC; Certified Rodent Chow; CRD Certified Rodent Diet
GKD Ground Klieba Diet

**Analysis of Data**

64. Participating Laboratories recorded the raw experimental data from their phase one work on a standardised excel spreadsheet (Annex 2) developed specifically for this validation study. Participating laboratories also submitted a written report of their work. Both the raw experimental data and the individual reports submitted by each participating laboratory are available on request from the OECD Secretariat.

65. An independent statistician, Dr Joe Haseman, of the US National Institute for Environmental Health Sciences undertook the overall statistical analysis and comparison of laboratory results. A task group of the VMG met with Dr Haseman in October 1999 to discuss the statistical approach to be used. Before Dr Haseman began the statistical analysis, the VMG discussed and agreed Dr Haseman's suggested statistical approach. After the statistical analysis had been carried out the VMG discussed and agreed on the overall interpretation of the results.

66. The ability of each individual laboratory to detect increased uterine weights at various doses of EE was evaluated by an analysis of variance approach, which included body weight as a covariable. Some laboratories used multiple protocols, and these were each evaluated separately. Both wet and blotted uterine weights were also analyzed. Since the variability in uterine weight tended to increase in direct proportion to the increase in mean uterus weight, the variance-stabilizing logarithmic transformation was carried out prior to the data analysis.

67. The primary method of statistical analysis for making pairwise comparisons of each dosed group to vehicle controls was Dunnett's test. Dixon's outlier test was used to detect possible outliers in the data, and Bartlett's test was used to assess homogeneity of variances. If significant heterogeneity was detected, the nonparametric Mann-Whitney U test was used to compare dosed groups with the vehicle control group. This method of analysis makes no distributional assumptions, but does not readily allow for a quantitative comparison of uterine weight responses among laboratories or the adjustment for possible confounding factors such as body weight.

68. Similar analyses were carried out to determine the ability of 0.1 and 1.0 mg/kg ZM used in combination with a specified reference dose of EE to reduce uterine weight relative to the response produced by the reference dose of EE alone.

69. Additionally, for a given protocol the dose-response patterns of the various laboratories were compared by Analysis of Variance techniques to assess the consistency of the uterine weight response across laboratories following the same protocol. The relative sensitivities of the four protocols (A, B, C, and C') were also compared to see if any obvious differences emerged, especially those that may be consistent from laboratory to laboratory.

70. Other factors that were evaluated included:

(i) the within-group variability in uterine weights, both across laboratories and across protocols;
(ii) the strength of the correlation between body weight and uterine weight;
(iii) the relative sensitivities of using uterine wet weight vs. blotted weight;
(iv) the effect of EE on body weight; and
(v) the "effective dose" (ED) of EE that is estimated to produce a specified increase in blotted uterine weight relative to the vehicle control group.

71. In addition, the Hill equation was used to see if the dose-response data fitted a mathematical model.

## RESULTS

### Overall findings

72. In general, the participating laboratories found that each of the four protocol options were straightforward to carry out and that the level of detail given in the standardised protocol was generally sufficient. There were only minor deviations from the standardised protocol. These included:

- one laboratory did not have an untreated control group;
- two laboratories were missing low dose EE groups;
- in a few instances, more or less scattered across laboratories, one or two animals had missing wet and/or blotted uterine weights;
- two laboratories reported uterine weights to the nearest 0.1g rather than 0.1 mg;
- some laboratories misunderstood the dose of EE to use for the combination EE/ZM experimental group; and,
- some laboratories had difficulty in following the protocol guidance for controlling the body weights of the rats within each experimental group.

73. The details of the rat strains, vehicle, animal diets, bedding and number of animals per cage for each of the participating laboratories is given in Table 6.

74. There were no treatment related mortalities recorded in the experiments but unexpectedly there was an overall decrease in body weights reported some laboratories following the protocols which used the adult animal model [Protocol C (OVX, 3 days exposure) and Protocol C' (OVX, 7 days exposure)], and a temporary decrease in some laboratories using the immature protocols (A and B).

75. The inclusion of body weight as a mandatory endpoint in the protocol served three major purposes: 1) as a criterion to minimise the chance of exceeding the restricted age-range for immature animals (Protocol A and Protocol B); 2) as an general index of the animal's general well-being during the duration of the experiment and 3) to test whether there was any correlation between uterine weight and body weight. This latter point could be important if the variability of the assay was influenced by high or low for age uterine weights or certain animal strains.

76. The body weight data of the five highest doses for Protocol A, B, C and C' are shown in Table 7. There were a number of relevant findings in relation to body weight. Firstly some laboratories had difficulty in following the protocol guidance for minimising the body weights ranges in the experiment and commented that to strictly follow the protocol would have required purchasing more animals than needed in the experiment.

77. Secondly, except in the adult OVX protocols there were no general trends of the reduction in body weight associated with the administration of the reference oestrogen EE. In the adult animal protocols (C and C') there was a statistically significant decrease in body weight at the doses of 1, 3, and 10 µg/kg/day (Protocol C) and 0.1, 0.3, 1.3, and 10 µg/kg/day (Protocol C'). In the immature animal experiments there were no similar overall trends, however, some temporary decreases were recorded (Table 7).

**Table 7: Body Weight Data of Five Highest EE Doses for Protocol A, B, C and C'**

| Lab Number | Lab Name | Vehicle | | 0.1 µg/kg/d EE | | 0.3 µg/kg/d EE | | 1.0 µg/kg/d EE | | 3.0 µg/kg/d EE | | 10.0 µg/kg/d EE | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| **Protocol A** | | | | | | | | | | | | | |
| 1 | Japan Bioassay | 58.92 | 3.48 | 57.78 | 2.95 | 59.13 | 3.54 | 58.62 | 3.34 | 59.78 | 1.33 | 56.15 | 3.45 |
| 2 | Sumitomo | 58.75 | 5.19 | 57.88 | 2.80 | 58.07 | 4.16 | 59.02 | 3.85 | 58.97 | 3.20 | 55.87 | 1.86 |
| 3 | Mitsubishi | 58.67 | 4.22 | 58.95 | 2.99 | 57.10 | 1.55 | 59.40 | 2.39 | 57.78 | 1.06 | 56.75 | 4.08 |
| 4 | BASF | 57.55 | 4.90 | 54.90 | 2.10 | 57.97 | 2.91 | 58.23 | 2.76 | 55.90 | 1.69 | 54.28 | 1.34 |
| 5 | Exxon Biosciences | 55.23 | 4.11 | 56.67 | 4.50 | 56.20 | 2.81 | 56.68 | 3.05 | 57.95 | 3.81 | 55.22 | 3.83 |
| 6 | CIT-France | 55.15 | 4.01 | 56.68 | 5.43 | 56.13 | 6.52 | 56.27 | 4.84 | 54.05 | 6.18 | 55.68 | 5.25 |
| 7 | FDSC – Japan | 54.48 | 3.78 | 54.32 | 2.53 | 54.80 | 3.10 | 54.78 | 2.21 | 53.40 | 3.32 | 55.03 | 2.02 |
| 8 | AstraZeneca/Syngenta | 53.88 | 5.41 | 54.68 | 4.67 | 55.75 | 4.71 | 55.57 | 6.93 | 55.45 | 6.32 | 55.07 | 6.09 |
| 9 | CITI-Japan/CERI | 51.63 | 2.75 | 51.53 | 3.44 | 52.83 | 3.23 | 51.58 | 2.28 | 51.10 | 3.04 | 51.80 | 2.49 |
| 10 | Chung Korea | 49.43 | 6.82 | 50.42 | 3.12 | 46.92 | 5.12 | 47.58 | 4.02 | 51.72 | 3.06 | 47.50 | 3.66 |
| 11 | IET-Japan | 48.28 | 1.86 | 49.20 | 2.62 | 48.52 | 2.61 | 48.18 | 3.06 | 47.37 | 2.79 | 48.00 | 2.83 |
| 12 | WIL Research | 47.23 | 3.92 | 48.05 | 3.81 | 46.95 | 3.35 | 48.38 | 4.11 | 46.98 | 3.20 | 46.72 | 3.61 |
| 13 | Bayer | 42.60 | 3.66 | 40.13 | 1.76 | 39.48 | 3.10 | 37.97 | 5.91 | 38.73 | 2.60 | 41.42 | 4.65 |
| 14 | Rhone Poulenc/Aventis | 41.17 | 2.68 | 42.88 | 2.06 | 42.00 | 2.01 | 44.10 | 3.72 | 40.15 | 5.55 | 43.18 | 5.66 |
| 15 | TNO | 38.78 | 1.49 | 39.73 | 1.72 | 38.32 | 1.48 | 37.42 | 1.34 | 38.33 | 1.73 | 39.83 | 2.70 |
| 16 | Berlin | 33.02 | 1.23 | 34.95 | 3.38 | 35.02 | 4.01 | 33.05 | 2.82 | 33.17 | 2.28 | 37.20* | 2.07 |
| **Protocol B** | | | | | | | | | | | | | |
| 1 | Japan Bioassay | 61.48 | 3.05 | 60.00 | 2.47 | 60.45 | 4.41 | 59.32 | 3.77 | 60.28 | 3.89 | 57.37 | 2.77 |
| 2 | Sumitomo | 58.07 | 4.03 | 56.20 | 2.99 | 58.35 | 3.37 | 57.52 | 4.69 | 57.87 | 1.51 | 54.02 | 2.52 |
| 3 | Mitsubishi | 57.85 | 3.57 | 56.70 | 4.08 | 55.98 | 3.49 | 55.62 | 4.64 | 55.03 | 4.63 | 56.12 | 1.56 |
| 7 | FDSC-Japan | 53.08 | 2.37 | 52.32 | 3.12 | 53.03 | 3.51 | 51.98 | 3.49 | 52.77 | 3.48 | 51.76 | 2.19 |
| 8 | AstraZeneca/Syngenta | 54.85 | 4.50 | 55.82 | 4.18 | 54.43 | 5.04 | 54.73 | 6.23 | 56.13 | 5.93 | 54.72 | 5.39 |
| 9 | CITI-Japan/CERI | 52.93 | 3.30 | 52.35 | 4.74 | 51.57 | 4.01 | 50.95 | 3.07 | 51.45 | 1.98 | 50.70 | 2.44 |
| 10 | Chung Korea | 48.00 | 2.95 | 47.32 | 5.30 | 46.08 | 6.87 | 46.85 | 4.03 | 46.62 | 4.31 | 45.53 | 5.84 |
| 11 | IET-Japan | 49.60 | 1.60 | 50.23 | 2.21 | 49.23 | 1.51 | 48.50 | 1.89 | 48.32 | 1.06 | 45.55* | 2.29 |
| 12 | WIL Research | 46.53 | 2.52 | 47.22 | 2.56 | 48.75 | 5.21 | 47.72 | 4.04 | 48.30 | 2.84 | 47.42 | 3.18 |
| 15 | TNO | 42.93 | 1.21 | 41.30 | 2.55 | 41.80 | 2.24 | 42.33 | 1.90 | 42.18 | 2.84 | 42.97 | 1.65 |
| 17 | Denmark | 48.88 | 2.33 | 47.38 | 3.70 | 49.17 | 5.01 | 45.65 | 4.91 | 49.12 | 4.48 | 46.57 | 2.43 |
| 18 | Korea Park | 38.32 | 4.35 | 39.47 | 2.29 | 36.87 | 2.66 | 40.35 | 2.17 | 38.10 | 3.60 | 41.50 | 3.46 |

* $p<0.05$ vs. vehicle

a Denotes the presence of an outlier in the data

| Lab Number | Lab Name | Vehicle | | 0.1 µg/kg/d EE | | 0.3 µg/kg/d EE | | 1.0 µg/kg/d EE | | 3.0 µg/kg/d EE | | 10.0 µg/kg/d EE | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| **Protocol C** | | | | | | | | | | | | | |
| 1 | Japan Bioassay | 210.42 | 9.35 | 203.53 | 11.68 | 201.53 | 9.68 | 202.00 | 12.18 | 194.67* | 13.43 | 192.88* | 10.87 |
| 2 | Sumitomo | 216.55 | 17.91 | 213.38 | 13.73 | 210.38 | 8.00 | 208.08 | 7.98 | 201.08 | 11.29 | 193.58* | 7.75 |
| 3 | Mitsubishi | 235.20 | 12.80 | 228.10 | 11.00 | 228.92 | 12.77 | 218.93* | 12.19 | 211.65* | 10.90 | 212.38* | 10.66 |
| 7 | FDSC – Japan | 239.87 | 16.03 | 232.03 | 21.47 | 234.90 | 15.41 | 227.70 | 19.66 | 224.48 | 15.70 | 219.65 | 12.62 |
| 8 | AstraZeneca/Syngenta | 295.00 | 30.19 | 291.00 | 36.13 | 293.33 | 31.48 | 287.83 | 29.55 | 280.33 | 20.57 | 278.83 | 24.32 |
| 9 | CITI-Japan/CERI | 224.02 | 13.55 | 217.57 | 6.40 | 219.57 | 9.47 | 211.02* | 7.91 | 204.65* | 5.75 | 200.22* | 5.57 |
| 11 | IET-Japan | 169.62 | 7.14 | 169.13 | 6.61 | 165.48 | 7.09 | 162.73 | 4.68 | 159.03* | 4.63 | 156.83* | 5.62 |
| 18 | Korea Park | 215.45 | 3.89 | 210.93 | 7.46 | 212.55 | 6.09 | 209.27 | 5.21 | 205.38* | 4.88 | 195.52* | 5.62 |
| 19 | Huntingdon | 290.63 | 12.84 | 298.60 | 11.89 | 301.75 | 7.84 | 287.18 | 27.35a | 294.57 | 9.68 | 282.43 | 9.62 |
| **Protocol C'** | | | | | | | | | | | | | |
| 1 | Japan Bioassay | 236.20 | 13.13 | 222.38 | 15.31 | 215.03* | 12.42 | 205.43* | 12.85 | 198.63* | 11.29 | 201.32* | 8.84 |
| 3 | Mitsubishi | 256.80 | 8.18 | 242.13 | 12.84 | 243.25 | 8.23 | 223.75* | 11.60 | 218.40* | 10.20 | 210.52* | 10.53 |
| 7 | FDSC-Japan | 252.73 | 15.65 | 238.77 | 11.93 | 240.83 | 16.64 | 227.30* | 12.91 | 225.70* | 13.92 | 221.45* | 14.07 |
| 11 | IET-Japan | 202.63 | 4.45 | 196.32 | 6.70 | 192.00* | 7.37 | 185.63* | 3.29 | 176.27* | 10.05 | 168.87* | 6.23 |

* $p<0.05$ vs. vehicle

a Denotes the presence of an outlier in the data

78. As for the third issue there seemed to be no overall correlation between uterine weight and body weights.

79. Both wet and blotted uterine weights were measured. Two or three laboratories undertook additional histopathology, including BrdU labelling of cell division and others undertook food consumption measurements. No results of the additional histopathology have been provided at he time of preparing this report.

80. The feasibility, reliability and transferability of the standard protocol for detecting both agonists and antagonists of oestrogens were assessed as good. All sixteen laboratories were successful in demonstrating a dose-response increase in both wet and blotted uterine weight at the doses of EE used, by all four protocols. The dose-response curves for each of the laboratories is shown graphically in Figure 3. For each protocol there was generally good agreement among laboratories with regard to the actual EE doses that produced increased uterine weights. The shapes of the dose-response curves were not identical for a given protocol across all laboratories. However, there were a significant number of laboratories that produced parallel dose response curves (e.g., half the laboratories participating in Protocol A; almost half the laboratories in following Protocol B, two thirds of the laboratories following Protocol C and all laboratories following the extended Protocol C'). This indicates that a significant number of laboratories were able to consistently produce approximately the same percent increase in uterine weight at equivalent dose of EE.

81. Blotted weights showed statistical significance at slightly lower EE concentrations than did wet uterine weights. For example for Protocol A significance as generally first achieved at 1.0 µg/kg EE while for Protocol B and C, significance was generally first achieved at the next lower dose of 0.3 µg/kg. Three of the four laboratories conducting the extended Protocol C' first found significance at the 0.1 µg/kg dose. In the higher dosed groups used in the extended Protocol C', blotted uterine weights were reduced substantially relative to wet uterine weights apparently due to a reduction in the luminal fluid content in the uterine tissue between Days 3 and 7. The coefficient of variation for blotted uterine weight was generally lower than for wet uterine weight, however for Protocol A using immature animals the coefficients of variation generally exceeded those of the other protocols. This shows that blotted uterine weight was the most appropriate endpoint in the majority of cases.

82. All laboratories were successful, with the exception of one laboratory, in showing a reduction in uterine weight increase when the reference oestrogen antagonist ZM was used, compared to the uterine weight increase shown with using EE alone. The magnitude of the reduction in uterine weight was similar across all laboratories.

**Detailed statistical findings**

**Validity of underlying statistical assumptions**

83. The logarithmic transformation was generally successful in eliminating the heterogeneity of variances that was clearly evident in the "untransformed" uterine weights. For example, 76% (31/41) of the laboratory and protocol-specific evaluations of blotted uterine weight had no significant heterogeneity (indicated by a $p<0.05$ level Bartlett's test) after a logarithmic transformation. In some cases in which heterogeneity was evident, it was due primarily to the presence of a single outlier in the data (detected by Dixon's test for outliers), some of which are noted in the detailed statistical summary tables (Tables 8-26). In these cases, elimination of the outlier would further reduce or eliminate the heterogeneity. However, the presence of outliers could not account for all cases of significant heterogeneity.

**Figure 3**

**Responses of blotted uterine weight to doses of 17α-ethinyl estradiol (EE) following the four protocols in the validation work.**

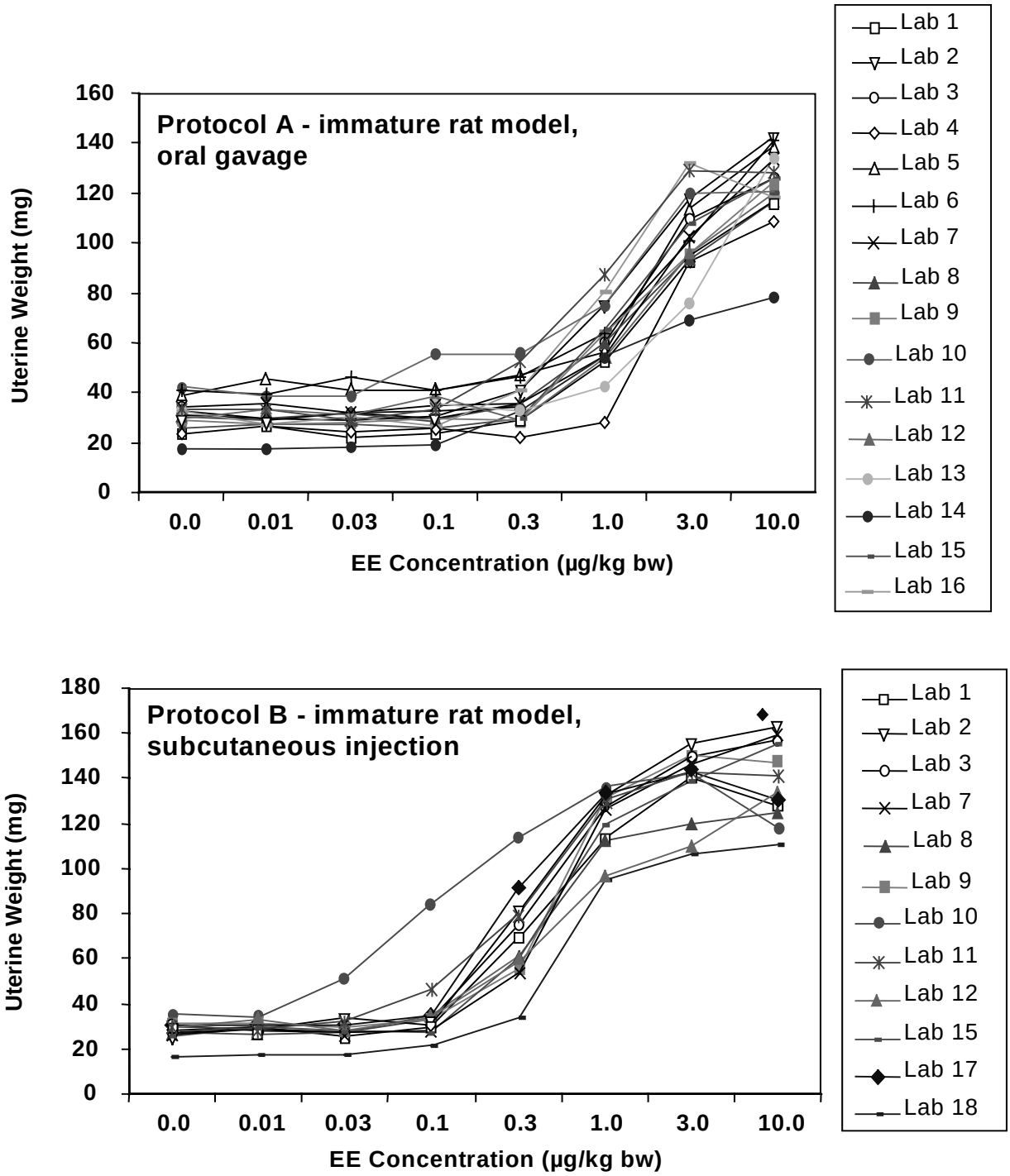
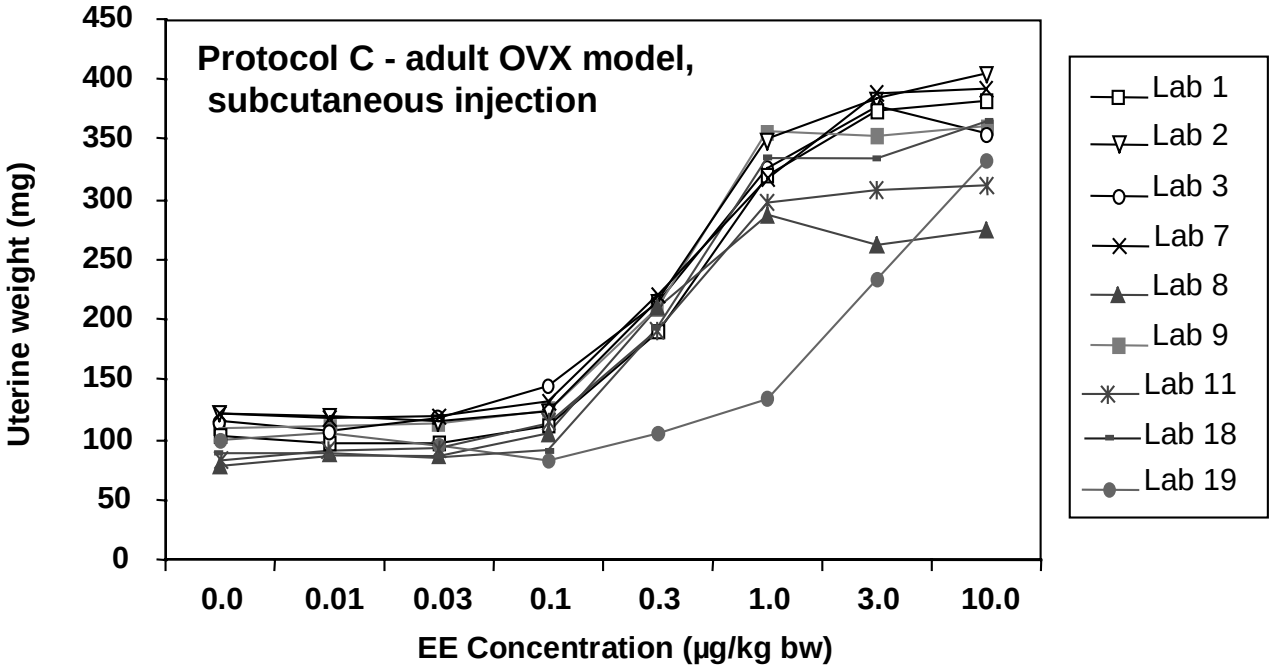
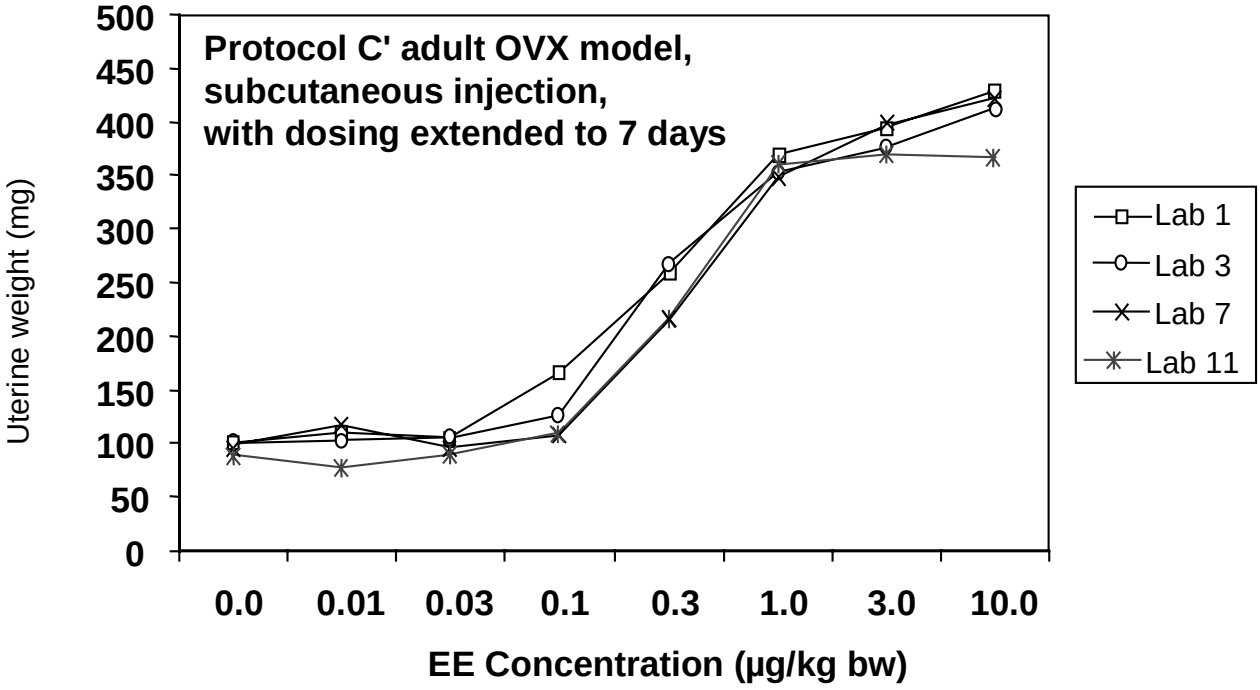


**Figure 3 (continued)**



84. Examples of outliers included a number of cases in which there were blotted or wet uterine weights well outside the range of values for other animals in the group, e.g. a wet weight of 441.8 mg compared to a range of 128.7 mg to 148.6 mg for other animals in the group or a blotted weight of 303.0 mg, compared to a range of 127.4 mg. to 144.5 mg. for other animals in that group. At this stage of the validation work, it is important to identify possible outliers, in order to consider their possible cause as the protocol itself is still being optimized for wider application. Those groups showing large standard deviations (SD's) in uterine weight are often reflecting the presence of an outlier, as indicated in the detailed statistical tables. Outliers occurred more often in the adult animal protocols than in the immature animal protocols (see Table 10).

85. Analysis of Variance procedures are fairly robust against modest departures from normality, and the overall operating characteristics of these procedures (e.g., the type 1 error rate) are not greatly affected by mild to moderate heterogeneity. Nevertheless, in those cases in which heterogeneity was detected following a logarithmic transformation, an alternative nonparametric statistical analysis (Mann-Whitney U test) was carried out to compare dosed groups to controls.

86. Overall results using nonparametric and parametric analyses for these data were very similar. In only two instances did the nonparametric approach reveal significant increases in uterine weight that were not detected by Dunnett's test. Both involved wet weights, both involved cases in which Dunnett's test was very nearly significant, both were cases in which the Mann-Whitney U test was barely significant ($p<0.05$), and both involved instances in which the blotted uterine weight was significantly ($p<0.05$) elevated by Dunnett's test. The nonparametric test results for these two cases are given in Tables 8-10.

87. In addition to possible outliers, other data points of questionable validity were encountered. For example, in some instances the blotted uterine weight for a given animal exceeded it's wet weight. This appeared to be a recurring problem only at one laboratory, for which 11% of the animals (14/126) had blotted weights that reportedly exceeded the wet weights. In fact, for the high dose combination group (ZM + EE; Group 11) in Protocol A, the mean blotted weight (45.50 mg.) actually exceeded the mean wet weight (42.08 mg.; see Table 8). These discrepancies reportedly occurred very early in the study and may well have reflected a need for greater training and/or experience. It is also relevant to note that in several laboratories the extent of GLP procedures was not provided. Many results were given in good faith by the participating laboratories acknowledging that final reports were still in preparation.

88. Most importantly, these discrepancies did not appear to have a major impact on the study results. That is, the elevated uterine weight responses observed in the higher dosed EE groups at this laboratory were so striking, that relatively minor discrepancies in wet weight/blotted weight could not have accounted for these differences.

89. A total of eight animals from five other laboratories (1-2 animals per laboratory) also had blotted weights that reportedly exceeded the wet weights. Again, these minor discrepancies had little impact on overall study results.

**Body weight effects**

90. Body weights varied widely from laboratory to laboratory, no doubt reflecting differences in strain and possibly in age. Mean control body weights are given in Tables 12-14. Although Protocols A and B both used immature animals of similar age, actual body weights varied widely from laboratory to laboratory. For example, in Protocol A, the mean vehicle control body weights ranged from 33.02 g. in one to 58.92 g. in another. The corresponding range in Protocol B was from 38.32 g. to 61.48 g. (Tables 12 and 14).

91. In Protocol C and C' animals aged approximately 6 weeks were used. Here again, the range in body weights was large in Protocol C: from 169.6 g. to 295.0 g., while for Protocol C' the variation was less: from 202.6 g. to 256.8 g. (see Table 14). Yet despite these differences in body weight, generally similar increases in uterine weight were observed at the various laboratories in the different protocols.

92. The experimental design specifications had as one goal the minimization of body weight differences among the animals within a given group. For example, as a guideline it apparently specified that the maximum range of body weights in the immature animal protocols should be approximately 10 g. This standard was met by most, but not all laboratories.

93. For example, the six vehicle control animals used in Protocol A at one laboratory varied in weight from 38.7 g. to 57.7 g., a range approximately double the desired maximum. For two other laboratories, the range of body weights exceeded the desired 10 g. range (46.1 g. to 60.6 g. at one laboratory; 47.9 g. to 61.4 g. at the other).

94. The higher doses of EE tended to reduce body weight in the adult-animal protocols, but not in the immature animal protocols. In Protocol A none of the laboratories showed any significant body weight reductions in the EE groups relative to the controls. The only significant body weight effect (seen at the 10 µg/kg dose at one laboratory) was a significant increase in body weight (from 33.02 g. to 37.20 g.). This 13% increase in body weight was unique to this laboratory and is probably not biologically important. When all laboratories are considered collectively, there is no significant body weight effect (increase or decrease) at any dose in Protocol A.

95. In Protocol B there was only one laboratory that showed a significant body weight change: an 8% body weight reduction for the 10 µg/kg group that was significant ($p<0.05$) relative to the vehicle controls. This decrease is also probably not biologically important. There was no significant reduction in body weight in the 10.0 µg/kg EE group when all laboratories using Protocol B were considered collectively.

96. In Protocol C the 1.0, 3.0, and/or 10 µg/kg EE doses resulted in significantly ($p <0.05$) reduced body weight relative to the vehicle controls for six laboratories (see Table 19). For two other laboratories the high within group variability in body weight prevented the 5-8% body weight reductions seen in the top two dosed groups from being statistically significant. The ninth laboratory showed only a very small (3%) decrease in body weight in the top dose group, and this lack of effect may be related to the weaker uterine weight effect found at this laboratory, as noted above.

97. Despite these differences, the overall results across laboratories were consistent (i.e., there was no significant laboratory by dose interaction for body weight) and Dunnett's test indicated that the body weight reductions in the 1, 3, and 10 µg/kg EE groups (which averaged 4%, 6%, and 8% respectively) were statistically significant ($p<0.05$). These body weight reductions are summarized in Table 19 for each laboratory.

98. In Protocol C there were clear inter-laboratory differences in within-group variability for body weight. In particular, body weights were much more variable for some laboratories than for others (see Table 19). There was an unusually large SD for the 1.0 µg/kg EE group for one laboratory (see Table 19), which was due to a single animal who reportedly lost 11% of her body weight (from 274.6 to 245.6) in a single day.

99. In Protocol C', with extended dosing from 3 to 7 days, the four laboratories showed consistently and significantly reduced body weights in the top 3-4 dosed groups (Table 19). In fact, considered

collectively, Dunnett's test indicated that the body weight reductions at the top five dosed groups (averaging 5%, 6%, 11%, 14%, and 15% for doses of 0.1, 0.3, 1, 3, and 10 µg/kg respectively) were statistically significant ($p<0.05$).

100. These rather large body weight reductions in Protocol C' merit some comment. While Protocol C' appears to have a sensitivity advantage relative to Protocol C, it could be argued that this comes at the "price" of other effects of the test compound. It is beyond the scope of this report to judge whether the 11-15% body weight reductions observed in Protocol C' in the top three EE dosed groups are of sufficient magnitude to "compromise" the biological significance of the uterine weight effects seen at these doses, but clearly this is a matter that requires some consideration.

101. Within a given dosed or control group, there was no consistent correlation between body weight and uterine weight (Table 20). When there was a significant correlation, it tended to be stronger for blotted weight than for wet weight and stronger for Protocols A and B than for Protocols C and C'. For example, half (14/28) of the immature animal experiments (Protocols A and B) showed a significant ($p<0.05$) correlation between body weight and blotted uterine weight, compared with only 15% (2/13) of the adult animal experiments (Protocols C and C'; see Table 20). This correlation (or lack of correlation) is potentially important when considering the need to take body weight differences into account in the evaluation of uterine weights. These data suggest that body weight may be a more important covariable in the immature animal protocols than in the adult animal protocols. This further suggests that it may be more important to equalize body weights across groups (and reduce the variability in body weight) in the immature animal protocols than in the adult animals protocols.

**Laboratory and protocol-specific effects of EE on uterine weight**

102. Qualitatively, there was generally very good agreement among laboratories regarding the EE doses at which significant increases in uterine weight were first detected. For Protocol A, significance was generally first achieved at 1.0 µg/kg EE, while for Protocols B and C, significance was generally first achieved at the next lower dose of 0.3 µg/kg EE. Three of the four laboratories carrying out Protocol C' first found significance at the 0.1 µg/kg EE dose. Tables 8-10 present data summaries for each laboratory highlighting statistical significance at $p< 0.05$ level for each EE dosed group compared to the vehicle control group. These tables also report significant ($p<0.05$) differences between the two antagonist (ZM + EE) groups when compared with the appropriate (3.0 or 0.3 µg/kg) EE groups.

103. The lowest doses giving significant ($p<0.05$) increases in uterine weight for each protocol are summarized in Table 11. This table also shows the number of laboratories where these lowest EE doses were identified.

104. Quantitatively, there were notable differences among laboratories, both in the shape of the dose-response curve and in the magnitude of the uterine weight increases at equivalent doses. Tables 12, 13, and 14 summarize for Protocols A, B and C /C' respectively the percent increase in uterine weight relative to vehicle controls for the four doses at which statistically elevated uterine weights were generally first achieved (depending upon laboratory and protocol): 0.1, 0.3, 1, and 3 µg/kg EE. The actual percentage uterine weight increases observed in these studies is discussed in more detail below.

105. Another question of interest involved the shape of the dose-response curves, and whether this shape was consistent from laboratory to laboratory. One method for assessing this consistency is to evaluate the laboratory by dose interaction in the uterine weight response. A significant interaction implies an inconsistent EE effect across laboratories. Equivalently, it means that the dose-response curves for uterine weight are not parallel.

106. For each of the three major protocols - A, B, and C - analysis of all laboratories revealed a highly significant ($p<0.001$) laboratory by dose interaction, implying that the dose-response curves were not uniformly parallel for all laboratories that carried out a given protocol. However, certain subsets of laboratories did have parallel dose-response curves, as is discussed in more detail below. The most notable departures from parallel dose-response curves included the following:

(i) In Protocol A, only one laboratory that did not find a statistically significant increase in wet uterine weight in the 1.0 µg/kg EE group (see Table 8). In addition, the magnitude of the increase in the 3.0 µg/kg group, while statistically significant, was less than the corresponding increase found at other laboratories (Tables 8 and 12). The blotted uterine weight response at this laboratory was somewhat more consistent with that of other laboratories (Table 12).

A second laboratory found only a modest but statistically significant increase in uterine weight in the 1.0 µg/kg EE group, but a rather dramatic increase in the next higher dosed group. This steep dose-response curve was not seen in other laboratories (Tables 8, 12).

Finally, a third laboratory tended to have lighter control uterine weights than other laboratories (Table 8), and the corresponding percent increase in uterine weight in the 0.3 µg/kg and 1.0 µg/kg EE groups exceeded that of other laboratories (Table 12).

(ii) In Protocol B, one laboratory found significant increases in uterine weight at lower doses and/or greater uterine weight effects at equivalent doses than any other laboratory. For example, this laboratory found a 147% increase in uterine wet weight at the 0.1 µg/kg dose, while the corresponding increase at other laboratories ranged from only 3-46% (Table 13).

(iii) In Protocol C, one laboratory reported only a 4-6% increase in uterine weight at the 0.3 µg/kg dose, while other laboratories were reporting significant uterine weight increases at this dose ranging from 80 to 182% (Table 14). Consequently statistical significance was not achieved until 1.0 µg/kg for both wet and blotted weight (Table 10). Similarly, at the 1.0 µg/kg dose this same laboratory showed only a 35-36% increase in uterine weight, while increases found by the other laboratories at this dose ranged from 161% to 643% (Table 14). This reduced response and the lack of response in the combination groups at this laboratory (discussed below) are arguably the most notable departures from the "norm" seen at other laboratories across all protocols.

107. However, despite the differences noted above, there were groups of laboratories that produced "statistically consistent" responses. That is, there were groups of laboratories for each protocol for which the laboratory by dose interaction (on a log scale) was not statistically significant. This lack of interaction implies that while uterine weights in the various dosed groups may have varied from laboratory to laboratory, the dose-response curves were parallel. This parallelism is arguably the type of reproducibility that is most desirable in studies of this type, i.e., that laboratories can consistently produce approximately the same percent increase in uterine weight at equivalent doses of the test compound.

108. In Protocol A, eight of the sixteen laboratories produced blotted uterine weight effects that were statistically consistent at the various doses evaluated. In Protocol B, five of the twelve laboratories produced uterine weight effects that were statistically consistent at the various doses evaluated. In Protocol C, six of the nine laboratories produced blotted uterine weight effects that were statistically consistent at the various doses evaluated. In Protocol C', all four laboratories produced blotted uterine weights that were statistically consistent (after deleting one outlier).

**Relative sensitivities of the four protocols**

109. Statistical sensitivity can be defined in several ways. One approach is to identify the lowest dose at which statistical significance is achieved. For example, every laboratory that used both Protocols A and B found significance at lower doses (and/or greater percentage increases in uterine weight at equivalent doses) in Protocol B than in Protocol A (see Table 11), so in that sense Protocol B is "more sensitive" than Protocol A. However, for this reference chemical (EE) this "increased sensitivity" is an expected finding due to the differences in the route of administration of EE used in these two protocols.

110. There are no notable differences between Protocols B and C in terms of the dose first producing statistical significance. For example, seven of the eight laboratories that carried out both protocols found significance for uterine wet weights at equivalent doses. The one exception found significance at a lower dose using Protocol B (0.1 µg/kg) than Protocol C (0.3 µg/kg).

111. However, at the higher doses there was a significant difference between the two protocols in terms of the magnitude of the percent increase over controls. For example, for the 12 laboratories carrying out Protocol B, the range of increase in blotted uterine weight over controls was 326-588% for the 1.0 µg/kg EE dose and 370-663% for the 3.0 µg/kg EE dose (Table 13). The corresponding ranges of increase for the nine laboratories carrying out Protocol C were "only" 136-375% for the 1.0 µg/kg dose and 236-375% for the 3.0 µg/kg EE dose (Table 14).

112. Although only four laboratories used Protocols C and C', there is a suggestion that Protocol C' (which extends the dosing period from three to seven days) may be slightly more sensitive than Protocol C. Two of the four laboratories that carried out both protocols found uterine weights significantly elevated at a lower dose in Protocol C' (0.1 µg/kg) then in Protocol C (0.3 µg/kg; see Table 14).

113. Another potential advantage of Protocol C' is that at the higher doses, especially for wet weights, there was much less inter- and intra-laboratory variability in uterine weights in Protocol C' than in Protocol C (see Table 10). Moreover, in the higher dosed groups, wet weights were reduced substantially in Protocol C' relative to Protocol C, whereas the reverse tended to be true for blotted weights (see Table 10). The reduced wet weights in Protocol C' were apparently due to the reduction in luminal fluid content between Days 3 and 7.

114. On the other hand, body weights were more substantially reduced by EE in Protocol C' than in Protocol C, which raises the question of the effects of EE in the higher dosed groups in Protocol C'. This matter is discussed in more detail later.

**Within-group variability in uterine weights**

115. A direct comparison of the standard deviations (SD's) of the uterine weights across laboratories and protocols would be misleading for several reasons. First of all, older animals (Protocols C and C') have uterine weights that are typically 3-4 times the size of the uterine weights in the younger animals (Protocols A and B). Since the variability in uterine weight increases in direct proportion to the mean uterine weight, we would expect to see larger SD's in uterine weight in Protocols C and C' than in Protocols A and B.

116. Similarly, Protocol A tends to have lower uterine weights at equivalent doses than Protocol B, because of the reduced sensitivity of Protocol A relative to Protocol B as noted previously. Thus, we would expect to see larger SD's for uterine weights in Protocol B than in Protocol A for equivalent doses.

117. Two procedures were carried out to permit meaningful comparisons of variability in response. The first was to log-transform the uterine weight data. Then, an analysis of variance was carried out for each laboratory and protocol (using all 11 experimental and control groups) using body weight as a covariable and evaluating the significance of differences among the dosed groups. The error mean square resulting from this analysis can be regarded as a measure of intra-group variability, averaged over doses and corrected for the possible influence of body weight on the observed uterine weight response. These error mean squares are summarized in Table 15.

118. The second procedure was to calculate the coefficient of variation (CV) for each dosed group for each laboratory within each protocol, and then average these values over doses to obtain a "representative" coefficient of variation for each laboratory and protocol. In this report the CV is defined as the ratio of the SD to the mean, multiplied by 100. Thus, for example, a CV of 10.0 implies that the SD is 10% of the mean. These results are summarized in Table 16. Both analyses reveal similar results, namely,

(i) The within-group variability in response was consistently less for blotted weights than for wet weights. This is especially true for Protocols C and C' (Tables 15 and 16).

(ii) Protocol A tended to show more within-group variability in uterine weights than the other three protocols. A comparison of blotted uterine weight CV's by ANOVA (followed up with Dunnett's test) showed the CV's in Protocol A to be significantly ($p<0.05$) greater overall than the corresponding CV's in Protocols B, C, or C', after adjusting for differences due to dose and laboratory (Table 16). This difference is perhaps not unexpected given the different route of administration used in Protocol A relative to the other protocols. The data summarized in Table 16 also suggest that for the three subcutaneous injection protocols, the adult animal protocols (C and C') have slightly lower CV's than the immature animal protocol (B).

Since not all laboratories carried out all protocols, another simple method to illustrate the difference in CV's among the three most widely used protocols (A, B, and C) is to compare for each pair of protocols, the mean CV's (averaged over dosed groups) for those laboratories that carried out both protocols. For example, from Table 16 it can be seen that eight of the ten laboratories that carried out both Protocols A and B had lower CV's on average for blotted uterine weight in Protocol B than in Protocol A. When comparing Protocols A and C, six of the seven laboratories had lower CV's using Protocol C than Protocol A (the other was a tie; see Table 16). Using this method of comparison, there was also a suggestion (6 vs. 2) of a lower CV in Protocol C compared with Protocol B (Table 16).

Table 17 presents the coefficients of variation for the two control groups (untreated and vehicle) for blotted weights. These groups show mean CV's that are similar to those shown in Table 16, which are based on all dosed and control groups. The same general trends discussed in the previous two paragraphs appear to hold for the control groups as well.

(iii) Some laboratories showed more within-group variability in uterine weight than others (see Tables 16 and 17). This high variability often reflected the presence of outliers in the data, as noted previously.

(iv) There was also significant ($p<0.05$) variability in the CV's across doses, although no clear dose-response pattern emerged. Interestingly, the largest CV on average for wet weight

(21.0, averaged over protocols; N=41) was seen in the top dose (10.0 µg/kg) group, whereas for blotted weight, this same group showed the lowest CV (10.8) when compared with the other groups. Moreover, the next highest dose (3.0 µg/kg) had the second smallest mean CV for blotted weights - 11.3. This illustrates the marked reduction in within-group variability achieved among the higher dosed groups by using blotted weights rather than wet weights. The mean vehicle control CVs were 14.1 and 15.2 for wet and blotted uterine weight respectively, somewhat intermediate between the two “extremes” noted above.

**Wet weights vs. Blotted weights**

119. Table 18 compares blotted weights as a function of wet weights for the control groups and for the higher dosed groups. Within control groups, blotted uterine weights for most laboratories were 87-98% of wet weights. There were two exceptions, i.e., there were two laboratories whose control blotted weights averaged only approximately 80% of wet weights (see Table 18). Moreover, this percentage varied widely for individual animals. For example, one of these two laboratories reported individual control animals whose blotted uterine weights were less than 50% of their wet weight. On the other hand, this same laboratory also reported two animals for which blotted uterine weights actually exceeded the wet weights.

120. Importantly, despite some apparent differences in "blotting methodology" from laboratory to laboratory, the resulting blotted weights, especially for the higher dosed groups, still showed much less laboratory-to-laboratory variability than the variability in wet weights (see Tables 8-10). For example, in Protocol C at the top (10.0 µg/kg) dose of EE, the range of mean wet weights across laboratories was 391 mg. to 1227 mg., compared with a range of only 274 mg. to 404 mg. for mean blotted weights (Table 10). Similar results were found for other EE doses and for other protocols.

121. As doses of EE increased, the disparity between blotted and wet weights also increased. These disparities were especially apparent in Protocol C, which in the higher dosed EE groups, blotted weights were typically only 33-45% of wet weights. There were two exceptions, i.e., there were two laboratories in which blotted weights in the top two dosed groups were on average 72% of wet weights (see Table 18).

122. Table 18 also illustrates that in the higher dosed groups, Protocol C' produces blotted weights that are much closer to the wet weights than occurs in Protocol C. For example, blotted weights for the two highest doses in Protocol C' are 76-89% of wet weights, whereas for these same four laboratories the corresponding percentage is only 33-45% in Protocol C. As noted previously, this difference between Protocol C and C' is apparently due to the reduction in luminal fluid content between Days 3 and 7.

**ZM Effects**

123. As noted previously, there was some uncertainty regarding the dose of EE used in the "combination" ZM + EE groups (Groups 10 and 11) in Protocols B, C, and C' in some laboratories. Reconfirmation of the data with the participating laboratories has taken place, but there still remain some ambiguities. The standardized data collection forms provided to the laboratories specified (incorrectly) that a dose of 3.0 µg/kg EE should be used in these combination groups for all protocols. Most laboratories appear to have used the "correct" dose (0.3 µg/kg EE) for protocols B, C and C'. However, two laboratories apparently used the 3.0 µg/kg dose in Protocol B in the two combination groups testing the antagonistic effects of ZM. There is circumstantial evidence (discussed below) that a third laboratory may have done this as well in Protocol B.

124. For the combination groups, the results tended to be consistent qualitatively, but not always quantitatively. For Protocol A, all laboratories found a significantly reduced uterine weight in the top dose

combination group (1.0 mg/kg ZM + 3.0 µg/kg EE) relative to the appropriate (3.0 µg/kg) EE control. Moreover, the magnitude of the reduction was very similar across all laboratories (see Table 21).

125. Four laboratories also showed significantly reduced uterine wet weights in the lower dosed (0.1 mg/kg ZM + 3.0 µg/kg EE) combination group, but only one of these laboratories showed a corresponding significant reduction in uterine blotted weight (see Table 21).

126. For Protocol B, two laboratories apparently used 3.0 µg/kg EE rather than 0.3 µg/kg EE in the combination group, as noted above. Both of these laboratories found a significant reduction in uterine weight in the high dose ZM group, and one also found a significant reduction in the low dose group as well (Table 21).

127. All ten of the remaining laboratories that carried out Protocol B used 0.3 µg/kg EE and found a significant ($p<0.05$) reduction in uterine weight in the top dose combination group. However, the effect at the low dose combination group was more variable. For example, for blotted weights, six laboratories showed a significant reduction in uterine weight (ranging from 65-79% of the uterine weights found in the group receiving 0.3 EE alone), three laboratories showed slight, non-significant decreases (ranging from 85-90% of the 0.3 EE alone response), and one laboratory actually showed a 24% increase in uterine weight (Table 21). Interestingly, in Protocol C, this same laboratory found that this same dose of ZM produced a significant decrease in uterine weight relative to the EE reference dose group, consistent with the results found by other laboratories (Table 22).

128. One possible explanation for the result in Protocol B could have been that this laboratory actually used an EE dose of 3.0 µg/kg rather than 0.3 µg/kg in the combination ZM groups. However, this possibility has been eliminated, since the laboratory in question has confirmed that they used a 0.3 µg/kg dose of EE in the combination groups in Protocol B. Thus, other possible explanations by the laboratory must be explored, and for now this unusual result remains unexplained.

129. In Protocol C, seven of the nine laboratories produced consistent and significant reductions in uterine weight in the top dose group (Table 22). For example, mean blotted uterine weights in these seven laboratories in the top dose group ranged from 50-62% of the uterine weights found in the group receiving 0.3 µg/kg EE alone. Four of these seven laboratories produced significant ($p<0.05$) uterine weight reductions in the low dose group as well, while three did not.

130. There were two laboratories that did not produce a significant reduction in uterine weight at either ZM combination dose in Protocol C. The first laboratory, showed mean responses were similar to the seven laboratories noted above, but had high variability that may have masked the statistical significance of the reduced uterine weight (see Table 10 and earlier discussion of outliers). Thus, it is likely that the effect for this laboratory is "real" but not detected because of the high variability in response.

131. The other laboratory showed no evidence of a reduced uterine weight in either ZM combination group in Protocol C (Table 22). This lack of response may be related to the fact that the reference dose of EE (0.3 µg/kg) used by this laboratory had no impact on uterine weight when given alone, unlike the case for all other laboratories (see Table 10).

132. Interestingly, all eight laboratories that carried out both Protocols B and C had a greater percentage reduction in uterine weight in the top dose (1.0 mg/kg ZM + 0.3 µg/kg EE) combination group relative to vehicle controls in Protocol B when compared to Protocol C (Tables 21 and 22). This statistically significant ($p<0.05$) difference may indicate some increased sensitivity for Protocol B for the detection of strong oestrogen antagonists. However, this matter requires further study, especially since there was no consistent difference in sensitivity between Protocols B and C for the low dose combination

(0.1 mg/kg ZM + 0.3 µg/kg EE) group.

133. In Protocol C', all four laboratories produced a significant ($p<0.05$) decrease in uterine weight in the top dosed combination group and three produced significant reductions in the low dose group as well (Table 22). The one "non-significant" reduction was consistent with the other "significant" responses - one laboratory reported a low dose combination group uterine weight response that was 80-81% of the vehicle controls (Table 22).

134. In Protocol A, thirteen of the nineteen laboratories produced blotted uterine weight decreases in the ZM/EE combination groups that were statistically consistent. In Protocol B, among the nine laboratories that used the 0.3 µg/kg dose of EE, six laboratories produced blotted uterine weight decreases in the ZM/EE combination groups that were statistically consistent. For Protocol C, all laboratories but the one noted above that produced no uterine effects, produced reduced uterine wet effects that were statistically consistent in the ZM/EE combination groups. For Protocol C', all four laboratories produced reduced uterine wet effects that were statistically consistent in the ZM/EE combination groups.

**Other questions addressed in the statistical analysis**

***Was there a vehicle effect on uterine weight?***

135. This is a potentially important issue, since there are literature reports that certain plant oil vehicles have oestrogenic effects that result in increased uterine weights over controls. The only significant ($p<0.05$) vehicle effect occurred for one laboratory in Protocol A, for which the vehicle control group did in fact have significantly ($p<0.05$) increased wet (35.73 mg. vs. 22.80 mg.) and blotted (32.63 mg. vs. 19.38 mg.) uterine weights relative to the untreated control group (see Table 8). The vehicle used by this laboratory was peanut oil (Table 6), but two other laboratories used this same vehicle with no corresponding increase in uterine weight (see Tables 8-10). Thus, it is unclear if this one increase in uterine weight is biologically significant.

***Was a stable maximum response achieved?***

136. A stable maximum completes the dose-response curve and is helpful in selecting the proper reference dose for use with ZM189.154 in the antagonist assays.

137. For Protocol A, no stable maximum was achieved. Although there was highly significant ($p<0.001$) variability among laboratories in the magnitude of the uterine weights found at the two top doses, most laboratories continued to show an increase in uterine weight at the 10 µg/kg dose of EE relative to the 3 µg/kg dose (see Table 8). There were only two exceptions (see Table 8). Excluding these two laboratories, the magnitude of the increase from the 3 µg/kg to the 10 µg/kg dose of EE was relatively consistent across laboratories. The lack of a stable maximum response is consistent with the fact that significantly elevated uterine weights first occurred at higher doses in this protocol compared with the other protocols.

138. For Protocol B, there was no consistent difference in response between the top two EE dosed groups, when the data are considered collectively. For blotted weight there was also a highly significant laboratory by dose interaction, confirming that there is no consistent pattern of response for these two doses, with seven laboratories showing an increase in the 10.0 µg/kg EE group relative to the 3.0 µg/kg EE group, and five laboratories showing a decrease. Analysis of variance indicated that there was no significant overall difference in blotted uterine weight between the 3.0 µg/kg and 10.0 µg/kg EE dosed groups in Protocol B, suggesting (together with the significant interaction) that a stable maximum response had been achieved for most, but not all laboratories.

139. For Protocol C, there was a significant ($p<0.001$) inconsistency in response, due entirely to one laboratory, who showed a marked increase in uterine weight at the 10.0 µg/kg EE dose relative to the 3 µg/kg EE dose that was not seen at other laboratories (see Table 10). If this laboratory is excluded, the remaining laboratories show a relatively consistent and statistically insignificant difference in uterine weight between these two doses, suggesting that a stable maximum was achieved (or very nearly achieved).

140. For Protocol C', the wet weights were clearly ($p<0.001$) continuing to rise in the 10.0 µg/kg EE group relative to the 3.0 µg/kg group (see Table 10). For blotted weights, this increase was less evident, but still significant ($p<0.05$) and relatively consistent (i.e., no significant laboratory by dose interaction). Thus, a stable maximum response was not achieved in Protocol C'.

***What was the impact on study results of the factors that were not standardized and thus varied from laboratory to laboratory?***

141. The uterotrophic assay is intended to be a widely used screen, and although the overall experimental design was standardized as far as possible, several aspects of the protocol design were considered largely impractical to standardize in the absence of specific information showing a relationship between the variable in question and oestrogenic effects. Factors that varied from laboratory to laboratory included strain, diet, housing protocol, bedding, and vehicle. These factors are summarized in Table 6 for each laboratory.

142. The original data analysis plan included a formal evaluation of whether or not the factors that varied from laboratory to laboratory could account for any major discrepant uterine weight results observed across laboratories. However, such an analysis was not carried out, primarily because there were very few major "discrepant results" that required explaining. Moreover, an inspection of Table 6 did not reveal an obvious link between the few discrepant results that did occur and any variable in this table. Thus, it is likely that such discrepant results are due to other factors not given in Table 6.

143. The overall consistency of experimental results across many laboratories and protocols is quite encouraging. Because of this consistency, a logical conclusion is that the factors summarized in Table 6 do not have a major impact on study results, at least not for EE.

***Are six animals sufficient for detecting significant increases in uterine weight in this type of study?***

144. One objective of validation is to determine the number of animals that are needed to detect increases in uterine weight if they exist. There is no easy answer to this sample size question, since the required number of animals depends upon the underlying variability and the magnitude of the uterine weight increase above controls. Table 23 compares the power of detecting various increases in uterine weight in the top dose group (by Dunnett's test) as a function of the magnitude of the response (from a 25% to 40% increase in uterine weight), the number of animals per group (6 or 10) and the underlying CV (from 10.0 to 25.0).

145. This is intended to be a rough guideline only, since, for example, it could be argued that the primary method of statistical analysis should be a trend test, which uses all the data and hence would have more power than a procedure that focuses on a specific pairwise comparison such as Dunnett's test. However, time constraints made it impossible to pursue power calculations for more complex statistical procedures such as trend tests.

146. The analysis summarized in Table 23 illustrates the importance of minimizing the coefficient of

variation in this type of study. Six animals per group appear to be sufficient for detecting a 25-35% increase in uterine weight with reasonable power if the coefficient of variation can be kept relatively low (e.g., in the general range of 10.0 to 15.0).

147. Some laboratories appeared to satisfy this criterion, while others did not (see Table 14). Moreover, those uterine weight increases in the 25% to 35% range that were found to be statistically significant relative to vehicle controls (see Tables 12-14) were generally found at those laboratories with relatively low CV's (see Table 16). While larger CV's may be acceptable scientifically, it must be recognized that underlying uterine weight increases in the 25-35% range could go undetected for those laboratories having large CV's in uterine weight.

***Can a mathematical model be found that provides a reasonably good fit to the uterine weight dose-response data?***

148. There are a number of potential advantages in deriving a mathematical model that describes accurately the increase in uterine weight as a function of dose. For example, a laboratory may wish to compare its results quantitatively to those from other laboratories and to choose a reference dose for antagonist studies.

149. Although a number of models could perhaps have been selected to describe these sigmoidal shaped uterine weight dose-response curves, one widely used model that generally provided a good fit to the various data sets was the Hill equation model, a model suggested by Eric Vindimian of INERIS, France. This model has the functional form

$$E(y) = b + (vd^n/(k^n + d^n))$$

where y=uterine weight, E(y) = expected value of uterine weight, d is the dose level, b is the intercept term (estimates the background response), k is the estimated ED50, n is the Hill exponent and is called the shape parameter because it characterizes the curvature of the dose-response, and v is the maximum response above background.

150. This model was applied to the 41 individual experiments by Richard Morris of Analytical Sciences, Inc, and the results are summarized in Tables 24-26. For most datasets the model provided a very good fit to the data.

151. This model and its associated parameter estimates should be used with some caution. For example, it is unclear if the underlying biology would logically point to this specific mathematical model as the appropriate model to explain changes in uterine weight as a function of EE dose. Nevertheless, the good fits obtained are re-assuring. Moreover, this particular model has the advantage of permitting an estimate of the maximum response as well as an estimate of the "effective dose" (ED) of EE that produces a specified response. Tables 24-26 include for each laboratory and protocol the estimated ED10, ED50, and ED90.

152. The concept of "effective dose" perhaps requires some discussion. For example, the ED90 can be regarded as the dose that is estimated to produce a uterine weight that is the sum of the estimated control uterine weight plus 90% of the difference between the estimated maximum uterine weight response and the estimated control uterine weight. Such calculations may have value in the decision of selecting a reference dose of EE for any future "combination group" experiment when the intent is to show a reduction in the increase in uterine weight caused by EE alone.

153. For example, one objective may be to pick an ED90 (suggested by some investigators) as the

reference dose for EE. In the context of these results, perhaps 1.0 µg/kg is a better choice than 0.3 µg/kg for Protocols B, C, and C'. That is, the estimated ED90's in Tables 24-26 are much closer to 1.0 µg/kg than to 0.3 µg/kg.

154. The models summarized in Tables 24-26 support the results previously reported based on other types of statistical analyses. For example, note that in Protocol C, the estimated ED50 for one laboratory (2.60 µg/kg) falls well outside the range of values obtained from the other eight laboratories (0.23 µg/kg to 0.42 µg/kg; see Table 26), confirming that this laboratory had a pattern of uterine weight response that was different from that seen at the other laboratories. In particular, the estimated ED10, ED50, and ED90 for this laboratory are approximately 8-fold higher that seen at other laboratories.

155. In Protocol B note that another laboratory had an estimated ED50 (0.09 µg/kg) that also fell well outside the range of the other 11 laboratories (0.33 to 0.55 µg/kg; see Table 24), confirming the result noted previously for this laboratory.

156. Other parameter estimates are helpful as well. For example, the large estimated slope parameter for one of the laboratories in Protocol A (4.00; see Table 24) confirms that this laboratory had a steeper dose-response curve when compared to other laboratories (range of 1.27 to 2.60), as noted previously. Table 24 also confirms the previously reported low control (and estimated maximum) uterine weight response for another laboratory in Protocol A. Finally, it was noted previously that in Protocol A there was one laboratory that did not to find a significantly elevated uterine weight at the 1.0 µg/kg dose. From a modelling standpoint, this translates to an estimated ED50 (4.08 µg/kg) that is greater than that observed at other laboratories (range 0.68 to 3.09 µg/kg; see Table 24).

157. Comparisons can also be made of sensitivities across protocols. For example, all laboratories that carried out both Protocols A and B had lower estimated ED50's in Protocol B relative to Protocol A, confirming the “reduced sensitivity” in Protocol A noted previously. Conversely, there was no significant difference between the ED50's from Protocols B and C (Tables 24-26). Many other similar comparisons of interest can be made from these tables.

## Table 8

## UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL A

### Wet weights (in mg)

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 34.28 | 7.34 | 34.53 | 2.40 | 36.63 | 3.29 | 63.67* | 9.61 | 137.17* | 27.70 | 194.32* | 49.75 |
| JapanBioassay | 24.52 | 5.53 | 24.43 | 1.35 | 30.33 | 7.69 | 54.37* | 14.25 | 113.62* | 29.26 | 179.98* | 40.94 |
| AstraZeneca | 32.68 | 8.02 | 31.43 | 6.95 | 38.28 | 10.70 | 65.70* | 25.44 | 110.20* | 26.57 | 162.58* | 42.00 |
| Citi-Japan | 35.22 | 4.24 | 35.93 | 2.16 | 36.20 | 4.64 | 55.88* | 10.15 | 120.58* | 22.99 | 202.18* | 30.11 |
| Sumitomo | 32.02 | 2.77 | 31.22 | 1.92 | 42.52* | 9.49 | 78.07* | 24.25 | 153.25* | 40.67 | 273.47* | 46.83 |
| RhonePoulenc | 21.23 | 2.17 | 21.53 | 3.59 | 36.23* | 6.74 | 67.55* | 14.69 | 109.92* | 31.44 | 119.73* | 48.74 |
| CIT-France | 42.25 | 8.04 | 42.51 | 8.45 | 48.04 | 15.78 | 64.90* | 5.01 | 108.66* | 19.58 | 179.38* | 32.27 |
| TNO | 32.00 | 4.74 | 36.83 | 7.39 | 35.17 | 6.97 | 77.00* | 15.17 | 173.00* | 38.33 | 202.67* | 59.35 |
| Berlin | 35.73 | 10.14 | 29.03 | 5.45 | 44.18 | 19.93 | 93.83* | 24.71 | 213.73* | 34.79 | 233.55* | 60.90 |
| BASF | 24.33 | 1.21 | 27.83 | 4.22 | 23.00 | 2.45 | 29.50* | 3.62 | 112.33* | 19.13 | 157.33* | 38.96 |
| Exxon | 41.22 | 9.40 | 44.36 | 6.45 | 51.70 | 11.80 | 57.65* | 9.59 | 132.20* | 34.76 | 201.72* | 5.51 |
| Chung-Korea | 44.70 | 13.01 | 58.17 | 12.02 | 57.43+ | 4.40 | 77.87* | 8.29 | 148.53* | 36.30 | 195.72* | 23.42 |
| FDSC-Japan | 30.08 | 2.72 | 30.37 | 3.75 | 30.23 | 1.20 | 66.52* | 20.70 | 104.82* | 12.29 | 167.32* | 28.54 |
| IET-Japan | 36.85 | 3.21 | 36.60 | 3.73 | 55.60* | 6.17 | 97.93* | 5.91 | 204.47* | 31.87 | 190.55* | 42.60 |
| Bayer | 37.33 | 5.68 | 37.33 | 5.09 | 34.33 | 5.24 | 44.83 | 4.36 | 83.83* | 9.30 | 241.17* | 31.54 |
| WILResearch | 31.65 | 11.97 | 41.00 | 13.61 | 37.23 | 6.57 | 57.70* | 13.04 | 117.25* | 19.37 | 174.73* | 28.47 |

**Table 8 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL A**

**Wet weights (mg)**

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM 189.154 3.0 + 0.1 | | EE + ZM 189.154 3.0 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 31.60 | 4.74 | 31.62 | 2.92 | 33.27 | 5.54 | 101.63# | 30.08 | 45.27# | 8.95 |
| JapanBioassay | 28.50 | 1.67 | 27.98 | 3.72 | 23.38 | 2.28 | 97.47 | 34.27 | 40.15# | 6.19 |
| AstraZeneca | 30.50 | 4.78 | 31.43 | 6.79 | 34.37 | 8.20 | 121.53 | 21.00 | 40.73# | 8.21 |
| Citi-Japan | 39.18 | 5.20 | 35.92 | 7.25 | 32.62 | 3.19 | 106.03 | 26.91 | 49.10# | 6.30 |
| Sumitomo | 31.88 | 3.64 | 30.00 | 3.88 | 30.07 | 1.88 | 144.08 | 21.70 | 54.07# | 7.06 |
| RhonePoulenc | 20.72 | 3.55 | 19.50 | 2.57 | 20.18 | 2.25 | 117.07 | 28.56 | 34.08# | 5.86 |
| CIT-France | 46.82 | 9.24 | 41.52 | 5.98 | 48.07 | 11.02 | 92.11 | 17.85 | 46.88# | 6.43 |
| TNO | 29.67 | 6.44 | 32.50 | 4.64 | 35.00 | 6.07 | 177.33 | 46.59 | 63.83# | 16.73 |
| Berlin | 22.80 | 5.92 | 29.58 | 7.80 | 33.90 | 9.05 | 175.43 | 27.82 | 68.10# | 22.80 |
| BASF | 23.40 | 2.41 | - | - | 25.50 | 5.01 | 50.33# | 8.16 | 25.00# | 4.60 |
| Exxon | - | - | 48.13 | 13.84 | 42.20 | 2.89 | 96.82 | 18.15 | 43.45# | 9.10 |
| Chung-Korea | 46.83 | 14.17 | 41.38 | 14.48 | 40.60 | 8.77 | 170.20 | 42.34 | 83.92# | 11.74 |
| FDSC-Japan | 31.22 | 3.04 | 28.15 | 2.87 | 29.57 | 3.02 | 82.58# | 28.27 | 52.63# | 13.26 |
| IET-Japan | 36.17 | 5.83 | 36.20 | 4.69 | 33.50 | 3.56 | 155.47# | 42.77 | 57.85# | 11.25 |
| Bayer | 37.17 | 9.91 | - | - | - | - | 102.67 | 30.42 | 44.00# | 4.38 |
| WILResearch | 27.72 | 7.20 | 38.17 | 10.57 | 31.75 | 7.10 | 103.72 | 13.15 | 42.08# | 3.08 |

*** p<0.05 vs. vehicle**
**# p<0.05 vs. 3.0 group**
**+ Significant heterogeneity among groups; significant (p<0.05) vs. vehicle controls by a Mann-Whitney U test.**

**Table 8 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL A**

**Blotted weights (mg)**

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 32.50 | 6.75 | 32.92 | 2.44 | 34.38 | 3.12 | 60.28* | 9.27 | 108.85* | 13.60 | 125.68* | 14.05 |
| JapanBioassay | 23.28 | 5.58 | 23.25 | 1.08 | 29.12 | 7.71 | 52.65* | 13.96 | 93.05* | 12.82 | 116.40* | 16.66 |
| AstraZeneca | 30.27 | 8.59 | 29.42 | 6.55 | 35.48 | 9.98 | 59.95* | 21.08 | 94.80* | 19.96 | 117.12* | 15.46 |
| Citi-Japan | 34.28 | 4.12 | 35.02 | 2.67 | 35.33 | 4.90 | 54.75* | 10.20 | 102.53* | 12.57 | 133.48* | 12.97 |
| Sumitomo | 31.45 | 3.00 | 30.57 | 1.89 | 41.23* | 9.60 | 74.70* | 21.81 | 117.18* | 19.61 | 142.35* | 13.11 |
| Rhone Poulenc | 17.63 | 2.32 | 18.65 | 3.91 | 32.35* | 6.92 | 54.62* | 5.59 | 69.20* | 10.95 | 78.07* | 9.51 |
| CIT-France | 41.22 | 7.93 | 41.16 | 7.91 | 46.39 | 15.07 | 63.45* | 5.34 | 100.78* | 12.33 | 141.04* | 17.69 |
| TNO | 25.80 | 5.50 | 26.00 | 7.04 | 29.33 | 4.41 | 65.50* | 12.82 | 108.00* | 7.48 | 126.00* | 13.61 |
| Berlin | 32.63 | 9.47 | 26.20 | 4.74 | 40.67 | 18.57 | 80.47* | 18.34 | 131.58* | 5.62 | 118.07* | 18.67 |
| BASF | 23.83 | 1.60 | 25.83 | 3.06 | 22.17 | 2.32 | 28.33* | 3.08 | 92.50* | 7.42 | 108.67* | 13.22 |
| Exxon | 38.78 | 8.06 | 41.30 | 6.08 | 47.12 | 12.23 | 55.75* | 9.33 | 113.85* | 23.58 | 139.10* | 9.72 |
| Chung-Korea | 42.27 | 12.72 | 55.72 | 12.25 | 55.27* | 4.31 | 74.92* | 7.96 | 119.83* | 27.52 | 120.88* | 16.47 |
| FDSC-Japan | 28.58 | 3.11 | 29.57 | 3.66 | 28.80 | 1.95 | 63.77* | 19.88 | 95.67* | 7.86 | 124.60* | 14.66 |
| IET-Japan | 33.47 | 2.74 | 33.50 | 3.57 | 52.12* | 6.33 | 86.87* | 3.44 | 129.10* | 6.28 | 127.90* | 8.93 |
| Bayer | 34.83 | 5.78 | 36.00 | 4.29 | 32.83 | 5.60 | 42.33* | 3.98 | 76.17* | 4.17 | 134.33* | 12.24 |
| WIL Research | 29.27 | 9.88 | 38.67 | 13.79 | 28.78 | 8.74 | 53.63* | 9.36 | 95.33* | 5.04 | 119.47* | 9.47 |

**Table 8 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL A**

**Blotted weights (mg)**

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 3.0 + 0.1 | | EE + ZM189.154 3.0 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 30.22 | 4.41 | 29.72 | 2.78 | 31.58 | 5.49 | 91.83 | 16.94 | 43.12# | 8.50 |
| JapanBioassay | 26.92 | 2.54 | 26.70 | 4.22 | 22.10 | 2.27 | 78.47 | 16.22 | 38.25# | 6.06 |
| AstraZeneca | 27.93 | 4.26 | 28.80 | 6.66 | 31.78 | 8.05 | 99.93 | 12.62 | 38.12# | 7.43 |
| Citi-Japan | 38.33 | 5.16 | 35.45 | 7.22 | 31.88 | 3.03 | 96.33 | 18.78 | 47.77# | 6.34 |
| Sumitomo | 31.12 | 3.42 | 29.33 | 3.81 | 29.17 | 1.47 | 116.50 | 6.06 | 52.32# | 7.02 |
| RhonePoulenc | 18.55 | 3.16 | 17.33 | 2.62 | 18.15 | 3.03 | 74.07 | 6.34 | 30.40# | 4.95 |
| CIT-France | 45.12 | 8.62 | 39.39 | 6.21 | 46.38 | 10.06 | 86.97 | 14.76 | 45.10# | 6.20 |
| TNO | 24.83 | 3.49 | 27.33 | 4.63 | 27.17 | 6.18 | 107.33 | 10.80 | 53.67# | 11.98 |
| Berlin | 19.38 | 3.80 | 27.10 | 7.37 | 30.02 | 8.92 | 115.12 | 13.88 | 60.93# | 17.04 |
| BASF | 22.80 | 2.17 | - | - | 24.50 | 4.89 | 48.33# | 7.87 | 24.00# | 4.15 |
| Exxon | - | - | 45.85 | 13.59 | 40.68 | 4.65 | 92.40 | 17.40 | 39.92# | 7.41 |
| Chung-Korea | 44.07 | 13.84 | 38.83 | 14.09 | 38.73 | 8.42 | 119.75 | 13.93 | 83.92# | 11.74 |
| FDSC-Japan | 30.00 | 3.37 | 27.13 | 2.57 | 28.32 | 2.85 | 75.48 | 25.52 | 50.87# | 12.95 |
| IET-Japan | 32.75 | 5.62 | 33.12 | 4.18 | 29.10 | 3.33 | 111.05 | 12.20 | 56.08# | 10.93 |
| Bayer | 35.83 | 8.75 | - | - | - | - | 83.17 | 18.20 | 40.50# | 4.18 |
| WILResearch | 26.27 | 3.70 | 33.08 | 11.50 | 31.38 | 4.60 | 93.50 | 13.90 | 45.50# | 3.97 |

*** p<0.05 vs. vehicle**
**# p<0.01 vs. 3.0 group**

## Table 9

### UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL B

### Wet weights (mg)

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 31.98 | 3.30 | 35.12 | 4.19 | 79.28* | 14.38 | 155.33* | 22.93 | 267.58* | 37.50 | 277.45* | 48.56 |
| JapanBioassay | 29.28 | 3.17 | 31.03 | 4.86 | 71.95* | 15.88 | 136.97* | 18.14 | 255.22* | 59.27 | 254.13* | 39.02 |
| AstraZeneca | 29.83 | 2.64 | 37.87+ | 7.20 | 64.57* | 3.78 | 140.70* | 24.61 | 175.58* | 51.89 | 203.13* | 53.79 |
| Citi-Japan | 32.62 | 3.97 | 34.03 | 3.53 | 56.68* | 14.82 | 177.98* | 43.27 | 259.13* | 42.19 | 252.95* | 37.30 |
| Sumitomo | 27.92 | 2.07 | 31.07 | 4.50 | 81.77* | 15.97 | 182.00* | 20.26 | 243.45* | 47.28 | 293.08* | 75.86 |
| Denmark | 33.15 | 4.64 | 38.53 | 5.35 | 108.52* | 29.15 | 240.32* | 66.39 | 284.13* | 42.98 | 242.93* | 24.75 |
| TNO | 36.50 | 7.84 | 37.67 | 7.03 | 78.50* | 6.92 | 181.50* | 55.88 | 238.67* | 52.59 | 304.00* | 75.92 |
| Korea-Park | 21.72 | 2.89 | 28.13* | 4.17 | 44.08* | 6.18 | 142.92* | 18.01 | 201.70* | 24.29 | 257.75* | 50.41 |
| Chung-Korea | 39.20 | 5.88 | 96.72* | 28.89 | 137.25* | 32.79 | 226.13* | 75.76 | 272.18* | 23.16 | 256.43* | 81.95 |
| FDSC-Japan | 27.52 | 0.95 | 29.18 | 2.56 | 54.73* | 18.31 | 157.57* | 23.78 | 205.92* | 27.19 | 220.48* | 58.45 |
| IET-Japan | 33.97 | 4.12 | 49.63* | 8.87 | 91.15* | 12.73 | 202.78* | 50.74 | 278.76* | 46.76 | 307.53* | 46.78 |
| WILResearch | 34.82 | 3.93 | 37.80 | 9.08 | 64.07* | 21.85 | 118.38* | 51.49 | 186.40* | 70.98 | 249.40* | 71.37 |

**Table 9 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL B**

**Wet weights (mg)**

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 0.3 + 0.1 | | EE + ZM189.154 0.3 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 34.02 | 3.70 | 29.37 | 3.44 | 28.67 | 3.14 | 55.08# | 4.60 | 26.67# | 2.08 |
| JapanBioassay | 29.68 | 4.16 | 31.52 | 6.35 | 26.23 | 4.23 | 54.37# | 10.31 | 23.97# | 3.94 |
| AstraZeneca | 25.68 | 4.72 | 28.52 | 2.99 | 29.30 | 2.41 | 57.40 | 6.92 | 23.08# | 2.99 |
| Citi-Japan | 31.15 | 5.44 | 32.00 | 3.37 | 30.78 | 2.27 | 47.97 | 13.19 | 26.40# | 2.80 |
| Sumitomo | 30.78 | 2.02 | 29.53 | 4.50 | 33.95 | 3.20 | 111.12 | 26.21 | 28.00# | 2.99 |
| Denmark | 32.25 | 3.99 | 33.47 | 4.38 | 35.05 | 3.37 | 66.25# | 10.68 | 31.37# | 3.14 |
| TNO | 36.67 | 5.28 | 37.33 | 3.33 | 38.00 | 5.18 | 276.00@ | 85.06 | 57.33@ | 9.22 |
| Korea-Park | 23.48 | 4.33 | 21.92 | 4.45 | 23.45 | 3.20 | 38.07 | 6.46 | 24.22# | 3.17 |
| Chung-Korea | 33.68 | 7.29 | 36.27 | 8.72 | 54.18 | 8.43 | 142.15@ | 47.90 | 56.55@ | 4.15 |
| FDSC-Japan | 31.12 | 4.85 | 30.52 | 3.72 | 28.88 | 2.89 | 40.93# | 4.92 | 26.33# | 1.43 |
| IET-Japan | 32.58 | 1.58 | 32.83 | 6.23 | 37.17 | 3.26 | 69.67# | 12.21 | 34.22# | 2.64 |
| WILResearch | 27.72 | 7.20 | 36.08 | 7.09 | 33.58 | 7.79 | 45.17 | 8.09 | 31.95# | 5.47 |

* $p<0.05$ vs. vehicle
# $p<0.05$ vs. 0.3 group
+ Significant heterogeneity among groups; significant ($p<0.05$) vs. vehicle controls by a Mann-Whitney U test.
@ TNO and Chung-Korea reportedly used 3.0 EE rather than 0.3 in the two combination groups; For TNO, only the top (3.0 + 1.0) dose is significant relative to the 3.0 EE (alone) "control" group. For Korea Chung, both combination doses show significant ($p<0.05$) reductions relative to the 3.0 EE control group.

**Table 9 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL B**

**Blotted weights (mg)**

**Top five doses and vehicle controls**

| LAB | Vehicle | | .01 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 30.72 | 3.45 | 32.82 | 4.02 | 75.37* | 13.73 | 126.95* | 15.08 | 149.67* | 18.81 | 156.95* | 7.98 |
| JapanBioassay | 28.27 | 3.03 | 29.85 | 4.75 | 69.22* | 14.02 | 112.73* | 10.93 | 140.33* | 16.15 | 128.28* | 4.56 |
| AstraZeneca | 27.20 | 2.79 | 34.62* | 6.44 | 60.87* | 3.69 | 112.50* | 12.74 | 119.57* | 15.91 | 125.05* | 16.63 |
| Citi-Japan | 31.68 | 3.94 | 33.15 | 3.24 | 55.68* | 14.63 | 131.67* | 20.53 | 150.05* | 18.88 | 146.67* | 12.65 |
| Sumitomo | 27.30 | 2.11 | 30.62 | 4.49 | 80.70* | 15.93 | 131.77* | 10.98 | 155.57* | 11.64 | 162.68* | 11.19 |
| Denmark | 30.32 | 4.53 | 34.68 | 5.22 | 91.52* | 17.00 | 132.98* | 22.68 | 142.53* | 14.40 | 130.60* | 11.86 |
| TNO | 25.83 | 7.19 | 27.33 | 6.02 | 60.17* | 7.03 | 119.33* | 17.27 | 138.33* | 18.00 | 155.50* | 9.16 |
| Korea-Park | 16.12 | 4.05 | 21.48* | 5.13 | 33.70* | 7.30 | 94.73* | 7.82 | 106.90* | 13.77 | 110.65* | 11.24 |
| Chung-Korea | 35.77 | 6.10 | 84.30* | 17.90 | 113.67* | 26.62 | 136.60* | 19.85 | 141.83* | 20.22 | 118.07* | 21.04 |
| FDSC-Japan | 26.05 | 1.09 | 28.02 | 2.57 | 53.38* | 17.97 | 125.95* | 17.22 | 146.33* | 7.28 | 159.30* | 51.50a |
| IET-Japan | 29.67 | 4.36 | 46.53* | 8.87 | 80.18* | 8.80 | 130.85* | 14.75 | 143.10* | 12.86 | 141.27* | 12.81 |
| WILResearch | 29.67 | 4.24 | 33.60 | 7.51 | 58.78* | 14.70 | 96.63* | 29.29 | 109.73* | 18.98 | 133.95* | 9.13 |

**Table 9 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOL B**

**Blotted weights (mg)**

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 0.3 + 0.1 | | EE + ZM189.154 0.3 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 32.53 | 3.55 | 27.98 | 2.94 | 27.22 | 2.74 | 53.03# | 4.43 | 25.20# | 1.69 |
| JapanBioassay | 29.22 | 3.48 | 30.02 | 5.64 | 25.32 | 3.93 | 53.32# | 9.70 | 22.40# | 3.99 |
| AstraZeneca | 23.58 | 4.26 | 26.40 | 2.80 | 27.18 | 2.34 | 54.30 | 6.36 | 21.28# | 2.96 |
| Citi-Japan | 30.38 | 4.94 | 31.27 | 3.11 | 29.98 | 2.59 | 47.32 | 12.95 | 25.35# | 2.28 |
| Sumitomo | 29.95 | 4.16 | 28.58 | 4.61 | 33.62* | 3.07 | 99.72 | 15.87 | 27.23# | 3.06 |
| Denmark | 29.65 | 4.11 | 30.77 | 4.63 | 30.53 | 3.97 | 60.37# | 8.44 | 28.32# | 3.43 |
| TNO | 26.50 | 7.04 | 29.33 | 2.07 | 28.33 | 5.20 | 144.33@ | 19.40 | 47.00@ | 6.87 |
| Korea-Park | 17.70 | 3.00 | 17.43 | 2.64 | 17.18 | 2.39 | 30.30 | 6.10 | 20.15# | 2.41 |
| Chung-Korea | 31.38 | 6.87 | 33.45 | 10.36 | 50.95* | 8.34 | 107.55@ | 29.93 | 52.50@ | 3.41 |
| FDSC-Japan | 29.25 | 5.07 | 28.72 | 3.98 | 27.60 | 2.79 | 39.77# | 5.10 | 24.73# | 1.47 |
| IET-Japan | 29.48 | 2.18 | 29.13 | 5.80 | 32.18 | 3.06 | 63.58# | 10.74 | 31.98# | 2.17 |
| WILResearch | 26.27 | 3.70 | 32.63 | 6.75 | 28.25 | 9.13 | 38.43# | 6.49 | 28.23# | 4.15 |

*** p<0.05 vs. vehicle**

**# p<0.05 vs. 0.3 group**

**@ TNO and Chung-Korea reportedly used 3.0 EE rather than 0.3 in the two combination groups; For TNO, only the top (3.0 + 1.0) dose is significant relative to the 3.0 EE (alone) "control" group. For Korea Chung, both combination doses show significant (p<0.05) reductions relative to the 3.0 EE control group.**

## Table 10

## UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOLS C AND C'

### Summary for Protocol C

### Wet weights (mg)

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 121.58 | 6.49 | 151.13 | 55.68a | 231.32* | 21.61 | 656.50* | 177.40 | 1001.52* | 146.81 | 899.68* | 322.24 |
| JapanBioassay | 104.35 | 13.53 | 114.48 | 4.82 | 197.22* | 35.69 | 685.00* | 167.66 | 1052.42* | 43.53 | 1227.00* | 262.58 |
| AstraZeneca | 83.73 | 10.42 | 110.17* | 11.09 | 236.13* | 25.52 | 406.43* | 65.18 | 351.82* | 53.27 | 391.27* | 101.56 |
| Citi-Japan | 110.75 | 13.75 | 125.35 | 17.63 | 219.13* | 30.01 | 717.53* | 180.63 | 859.05* | 164.29 | 866.62* | 182.03 |
| Sumitomo | 125.12 | 19.47 | 128.75 | 14.29 | 225.18* | 32.82 | 697.13* | 89.87 | 886.05* | 104.73 | 1197.35* | 134.96 |
| Huntingdon | 104.17 | 10.76 | 85.33 | 8.78 | 108.33 | 13.47 | 140.50* | 26.57 | 269.17* | 78.70 | 588.33* | 127.61 |
| Korea-Park | 107.02 | 7.43 | 138.60 | 73.48a | 241.33* | 96.54 | 795.48* | 225.16 | 930.65* | 110.47 | 1104.27* | 210.47 |
| FDSC-Japan | 123.28 | 12.15 | 133.80 | 12.01 | 225.77* | 33.61 | 522.88* | 319.22a | 820.57* | 175.01 | 906.77* | 301.50 |
| IET-Japan | 86.00 | 19.07 | 118.42* | 9.09 | 213.22* | 12.32 | 613.82* | 84.49 | 682.80* | 86.85 | 714.18* | 174.83 |

**Table 10 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOLS C AND C'**

**Summary for Protocol C**

**Wet weights (mg)**

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 0.3 + 0.1 | | EE + ZM189.154 0.3 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 122.33 | 5.47 | 110.53 | 15.47 | 124.47 | 3.78 | 170.33# | 16.81 | 111.20# | 15.98 |
| JapanBioassay | 103.55 | 14.28 | 99.87 | 13.05 | 102.22 | 35.69 | 176.70# | 21.94 | 120.18# | 13.89 |
| AstraZeneca | 88.60 | 7.72 | 92.17 | 7.54 | 92.42 | 8.26 | 244.48# | 56.66 | 127.77# | 20.36 |
| Citi-Japan | 121.30 | 21.17 | 113.03 | 14.87 | 116.88 | 18.59 | 223.38# | 44.73 | 129.82# | 13.22 |
| Sumitomo | 108.65 | 11.06 | 123.17 | 7.14 | 118.43 | 11.50 | 185.78# | 16.77 | 128.03# | 14.92 |
| Huntingdon | 94.17 | 13.56 | 111.00 | 19.66 | 101.50 | 18.83 | 107.83 | 26.44 | 112.33 | 19.59 |
| Korea-Park | 109.53 | 9.43 | 105.72 | 12.33 | 106.65 | 19.48 | 169.98 | 19.54 | 160.15 | 101.88a |
| FDSC-Japan | 121.95 | 10.49 | 118.98 | 10.00 | 121.50 | 15.09 | 173.22# | 15.17 | 123.63# | 20.84 |
| IET-Japan | 99.47 | 9.29 | 95.25 | 4.56 | 98.50 | 9.41 | 151.77# | 22.36 | 100.40# | 16.13 |

**Table 10 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOLS C AND C'**

**Summary for Protocol C**

**Blotted weights (mg)**

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 115.92 | 5.27 | 144.42 | 54.14a | 213.95* | 13.02 | 326.07* | 60.22 | 378.37* | 20.79 | 354.37* | 54.84 |
| JapanBioassay | 102.35 | 12.85 | 112.22 | 4.42 | 190.45* | 32.71 | 319.78* | 57.47 | 373.72* | 24.51 | 382.00* | 35.21 |
| AstraZeneca | 79.22 | 10.36 | 105.08* | 10.33 | 211.13* | 13.31 | 287.68* | 23.33 | 262.20* | 25.77 | 273.73* | 42.91 |
| Citi-Japan | 108.47 | 13.23 | 123.60 | 16.68 | 211.37* | 26.45 | 357.57* | 47.52 | 353.82* | 34.38 | 362.05* | 41.95 |
| Sumitomo | 120.82 | 18.43 | 123.47 | 12.71 | 217.48* | 28.73 | 351.32* | 27.51 | 384.72* | 25.94 | 404.32* | 44.63 |
| Huntingdon | 99.17 | 10.17 | 83.17 | 7.70 | 104.67 | 12.09 | 135.17* | 24.19 | 234.17* | 51.43 | 332.67* | 27.34 |
| Korea-Park | 89.25 | 10.33 | 91.80 | 11.71 | 193.07* | 64.32 | 334.95* | 54.84 | 334.48* | 44.55 | 366.20* | 32.62 |
| FDSC-Japan | 121.62 | 12.29 | 131.25 | 12.32 | 220.83* | 32.25 | 317.52* | 63.15 | 387.43* | 34.40 | 391.67* | 40.89 |
| IET-Japan | 82.45 | 15.71 | 113.38* | 10.13 | 191.23* | 10.59 | 297.67* | 15.99 | 307.60* | 35.50 | 312.40* | 43.05 |

**Table 10 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOLS C AND C'**

**Summary for Protocol C**

**Blotted weights (mg)**

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 0.3 + 0.1 | | EE + ZM189.154 0.3 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 117.70 | 5.48 | 106.43 | 14.54 | 118.55 | 3.63 | 163.27# | 14.77 | 106.35# | 14.62 |
| JapanBioassay | 101.17 | 13.75 | 97.77 | 12.44 | 96.78 | 5.74 | 172.55# | 19.68 | 118.07# | 13.64 |
| AstraZeneca | 83.95 | 7.71 | 87.72 | 7.41 | 87.05 | 8.31 | 204.02# | 26.46 | 122.08# | 19.40 |
| Citi-Japan | 118.68 | 21.05 | 111.15 | 15.06 | 114.50 | 18.71 | 206.87# | 21.91 | 126.80# | 13.00 |
| Sumitomo | 105.88 | 10.45 | 119.28 | 6.39 | 115.02 | 10.32 | 182.28# | 15.94 | 124.57# | 13.30 |
| Huntingdon | 92.83 | 13.20 | 105.67 | 18.78 | 95.00 | 18.87 | 104.67 | 27.08 | 109.17 | 20.66 |
| Korea-Park | 98.20 | 9.51 | 88.62 | 10.86 | 83.92 | 20.48 | 137.07 | 14.12 | 139.45 | 93.07a |
| FDSC-Japan | 119.97 | 10.72 | 117.17 | 9.79 | 119.57 | 14.61 | 170.02# | 15.13 | 121.30# | 20.48 |
| IET-Japan | 94.78 | 9.17 | 91.12 | 4.13 | 93.23 | 7.46 | 146.98# | 24.30 | 96.37# | 15.37 |

*** p<0.05 vs. vehicle**

**# p<0.05 vs. 0.3 group**

**a Denotes the presence of an outlier in the data**

**Table 10 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOLS C AND C'**

**Summary for Protocol C'**

**Wet weights (mg)**

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 106.58 | 6.49 | 132.02* | 15.93 | 281.23* | 43.07 | 367.55* | 38.37 | 390.43* | 18.59 | 516.85* | 122.81 |
| JapanBioassay | 102.88 | 11.68 | 111.28 | 15.66 | 223.83* | 35.02 | 395.62* | 58.52 | 443.22* | 95.81 | 685.95* | 154.71 |
| FDSC-Japan | 103.62 | 15.39 | 190.40* | 123.39a | 267.23* | 30.01 | 384.82* | 61.81 | 412.50* | 58.84 | 519.38* | 45.55 |
| IET-Japan | 92.83 | 10.14 | 127.90* | 19.77 | 229.27* | 18.18 | 395.58* | 36.30 | 394.58* | 33.26 | 444.58* | 27.18 |

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 0.3 + 0.1 | | EE + ZM189.154 0.3 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 105.72 | 10.60 | 109.48 | 10.22 | 109.92 | 10.62 | 186.53# | 13.89 | 104.43# | 10.57 |
| JapanBioassay | 99.00 | 12.32 | 120.62 | 39.77 | 98.55 | 8.03 | 181.22 | 70.87a | 98.10# | 19.85 |
| FDSC-Japan | 108.20 | 12.89 | 112.10 | 11.24 | 108.67 | 10.47 | 197.47# | 16.65 | 115.95# | 17.83 |
| IET-Japan | 90.20 | 6.89 | 84.08 | 13.48 | 93.65 | 11.59 | 149.65# | 17.48 | 97.40# | 8.18 |

*** p<0.05 vs. vehicle**
**# p<0.05 vs. 0.3 group**
**a Denotes the presence of an outlier in the data**

**Table 10 (Continued)**

**UTERINE WEIGHT SUMMARY STATISTICS FOR GROUPS RECEIVING EE IN PROTOCOLS C AND C'**

**Summary for Protocol C'**

**Blotted weights(mg)**

**Top five doses and vehicle controls**

| LAB | Vehicle | | 0.1 | | 0.3 | | 1.0 | | 3.0 | | 10.0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 101.62 | 2.71 | 126.07* | 15.46 | 267.58* | 36.94 | 353.92* | 35.94 | 376.05* | 17.61 | 412.77* | 38.11 |
| JapanBioassay | 98.50 | 12.82 | 108.10 | 4.42 | 215.58* | 35.90 | 347.95* | 33.95 | 397.43* | 48.61 | 422.18* | 39.39 |
| FDSC-Japan | 100.87 | 15.16 | 164.32* | 68.25a | 259.18* | 29.10 | 368.30* | 49.85 | 393.82* | 51.62 | 429.95* | 36.53 |
| IET-Japan | 89.48 | 9.68 | 110.35* | 20.69 | 217.85* | 19.76 | 359.03* | 31.43 | 368.37* | 31.53 | 366.43* | 15.03 |

**Other dosed groups**

| LAB | Untreated | | 0.01 | | 0.03 | | EE + ZM189.154 0.3 + 0.1 | | EE + ZM189.154 0.3 + 1.0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mitsubishi | 102.25 | 10.88 | 103.95 | 9.58 | 104.48 | 9.25 | 181.07# | 13.57 | 100.90# | 10.54 |
| JapanBioassay | 95.93 | 11.31 | 116.92 | 37.02 | 96.42 | 8.06 | 173.38 | 62.65a | 95.28# | 18.77 |
| FDSC-Japan | 105.42 | 12.65 | 109.23 | 11.53 | 105.45 | 10.28 | 192.97# | 16.65 | 113.47# | 17.16 |
| IET-Japan | 85.42 | 6.25 | 77.97 | 11.87 | 89.45 | 10.18 | 143.17# | 16.51 | 91.28# | 7.29 |

*** p<0.05 vs. vehicle**
**# p<0.05 vs. 0.3 group**
**a Denotes the presence of an outlier in the data**

**Table 11**

**SUMMARY OF LOWEST EE DOSES SHOWING A SIGNIFICANT (P<0.05) INCREASE IN UTERINE WEIGHT**

**I. Wet weight**

| Protocol | Dose of EE | | | | |
|---|---|---|---|---|---|
| | **0.03** | **0.1** | **0.3** | **1.0** | **3.0** |
| Protocol A | 0 | 0 | 4 | 11 | 1 |
| Protocol B | 1 | 3 | 8 | 0 | 0 |
| Protocol C | 0 | 2 | 6 | 1 | 0 |
| Protocol C' | 0 | 3 | 1 | 0 | 0 |

**II. Blotted weight**

| Protocol | Dose of EE | | | | |
|---|---|---|---|---|---|
| | **0.03** | **0.1** | **0.3** | **1.0** | **3.0** |
| Protocol A | 0 | 0 | 4 | 12 | 0 |
| Protocol B | 2 | 3 | 7 | 0 | 0 |
| Protocol C | 0 | 2 | 6 | 1 | 0 |
| Protocol C' | 0 | 3 | 1 | 0 | 0 |

This table indicates that:

(i) Protocol A is less sensitive than the other protocols for detecting significant increases in uterine weight (this is not unexpected given the different route of administration used in Protocol A); and

(ii) blotted weights are slightly more sensitive than wet weights for detecting increases in uterine weight.

**Table 12**

**RATIO OF MEAN UTERUS WEIGHTS IN EE DOSED GROUPS TO THAT OF THE VEHICLE CONTROLS IN PROTOCOL A FOR FOUR CRITICAL DOSED GROUPS[a]**

| LAB | Vehicle control body wt. | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0.1 | 0.3 | 1.0 | 3.0 | 0.1 | 0.3 | 1.0 | 3.0 |
| JapanBioassay | 58.92 | 100 | 124 | 222* | 463* | 100 | 125 | 226* | 400* |
| Sumitomo | 58.75 | 98 | 133* | 244* | 479* | 97 | 131* | 238* | 373* |
| Mitsubishi | 58.67 | 101 | 107 | 186* | 400* | 101 | 106 | 185* | 335* |
| BASF | 57.55 | 114 | 95 | 121* | 462* | 108 | 93 | 119* | 388* |
| Exxon | 55.23 | 108 | 125 | 140* | 321* | 106 | 122 | 144* | 294* |
| CIT-France | 55.15 | 101 | 114 | 154* | 257* | 100 | 113 | 154* | 244* |
| FDSC-Japan | 54.48 | 101 | 101 | 221* | 348* | 103 | 101 | 223* | 335* |
| AstraZeneca | 53.88 | 96 | 117 | 201* | 337* | 97 | 117 | 198* | 313* |
| Citi-Japan | 51.63 | 102 | 103 | 159* | 342* | 102 | 103 | 160* | 299* |
| Chung-Korea | 49.43 | 130 | 128* | 174* | 332* | 132 | 131* | 177* | 283* |
| IET-Japan | 48.28 | 99 | 151* | 266* | 555* | 100 | 156* | 260* | 386* |
| WILResearch | 47.23 | 130 | 118 | 182* | 370* | 132 | 98 | 183* | 326* |
| Bayer | 42.60 | 100 | 92 | 120 | 225* | 103 | 94 | 122* | 219* |
| Rhone-Poul. | 41.17 | 101 | 171* | 318* | 518* | 106 | 183* | 310* | 443* |
| TNO | 38.78 | 115 | 110 | 241* | 541* | 101 | 114 | 254* | 419* |
| Berlin | 33.02 | 81 | 124 | 263* | 598* | 80 | 125 | 247* | 403* |

* $p<0.05$ vs vehicle control (i.e., vs. a value of 100)

[a] 0.1, 0.3, 1.0, and 3.0 µg/kg, the doses at which significant increases in uterine weight were generally first achieved

**Table 13**

**RATIO OF MEAN UTERUS WEIGHTS IN EE DOSED GROUPS TO THAT OF THE VEHICLE CONTROLS IN PROTOCOL B FOR FOUR CRITICAL DOSED GROUPS[a]**

| LAB | Vehicle control body wt. | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0.1 | 0.3 | 1.0 | 3.0 | 0.1 | 0.3 | 1.0 | 3.0 |
| JapanBioassay | 61.48 | 106 | 246* | 468* | 872* | 106 | 245* | 399* | 496* |
| Sumitomo | 58.07 | 111 | 293* | 652* | 872* | 112 | 296* | 483* | 570* |
| Mitsubishi | 57.85 | 110 | 248* | 486* | 837* | 107 | 245* | 413* | 487* |
| AstraZeneca | 54.85 | 127* | 216* | 472* | 589* | 127* | 224* | 414* | 440* |
| FDSC-Japan | 53.08 | 106 | 199* | 573* | 748* | 108 | 205* | 483* | 562* |
| Citi-Japan | 52.93 | 104 | 174* | 546* | 794* | 105 | 176* | 416* | 474* |
| IET-Japan | 49.60 | 146* | 268* | 597* | 821* | 157* | 270* | 441* | 482* |
| Denmark | 48.88 | 116 | 327* | 725* | 857* | 114 | 302* | 439* | 470* |
| Chung-Korea | 48.00 | 247* | 350* | 577* | 694* | 236* | 318* | 382* | 397* |
| WIL Research | 46.53 | 109 | 184* | 340* | 535* | 113 | 198* | 326* | 370* |
| TNO | 42.93 | 103 | 215* | 497* | 654* | 106 | 233* | 462* | 536* |
| Korea-Park | 38.32 | 130* | 203* | 658* | 929* | 133* | 209* | 588* | 663* |

* $p<0.05$ vs vehicle control (i.e., vs. a value of 100)

Two laboratories showed significant effects at the 0.03 µg/kg dose: Sumitomo (123 for blotted weight) and ChungKorea (138 for wet weight; 142 for blotted weight).

[a] 0.1, 0.3, 1.0, and 3.0 µg/kg, the doses at which significant increases in uterine weight were generally first achieved

**Table 14**

**RATIO OF MEAN UTERUS WEIGHTS IN EE DOSED GROUPS TO THAT OF THE VEHICLE CONTROLS IN PROTOCOLS C AND C' FOR FOUR CRITICAL DOSED GROUPS[a]**

**PROTOCOL C**

| LAB | Vehicle control body wt. | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0.1 | 0.3 | 1.0 | 3.0 | 0.1 | 0.3 | 1.0 | 3.0 |
| AstraZeneca | 295.0 | 132* | 282* | 485* | 420* | 133* | 267* | 363* | 331* |
| Huntingdon | 290.6 | 82 | 104 | 135* | 258* | 84 | 106 | 136* | 236* |
| FDSC-Japan | 239.9 | 109 | 183* | 424* | 666* | 108 | 182* | 261* | 319* |
| Mitsubishi | 235.2 | 124 | 190* | 540* | 824* | 125 | 185* | 281* | 326* |
| Citi-Japan | 224.0 | 113 | 198* | 648* | 776* | 114 | 195* | 330* | 326* |
| Korea-Park | 215.5 | 130 | 226* | 743* | 870* | 103 | 216* | 375* | 375* |
| Sumitomo | 213.1 | 103 | 180* | 557* | 708* | 102 | 180* | 291* | 318* |
| JapanBioassay | 210.4 | 110 | 189* | 656* | 1009* | 110 | 186* | 312* | 365* |
| IET-Japan | 169.6 | 138* | 248* | 714* | 794* | 138* | 232* | 361* | 373* |

**PROTOCOL C'**

| LAB | Vehicle control body wt. | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0.1 | 0.3 | 1.0 | 3.0 | 0.1 | 0.3 | 1.0 | 3.0 |
| Mitsubishi | 256.8 | 124* | 264* | 345* | 366* | 124* | 263* | 348* | 370* |
| FDSC-Japan | 252.7 | 184* | 258* | 371* | 398* | 163* | 257* | 365* | 390* |
| JapanBioassay | 236.2 | 108 | 218* | 385* | 431* | 110 | 219* | 353* | 403* |
| IET-Japan | 202.6 | 138* | 247* | 426* | 425* | 123* | 243* | 401* | 412* |

* $p<0.05$ vs vehicle control (i.e., vs. a value of 100)

[a] 0.1, 0.3, 1.0, and 3.0 µg/kg, the doses at which significant increases in uterine weight were generally first achieved

**Table 15**

**COMPARISON OF WITHIN-GROUP VARIABILITY IN UTERINE WEIGHTS[a]**

| LAB | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | C' | A | B | C | C' |
| Citi-Japan | .0254 | .0354 | .0319 | NT | .0198 | .0288 | .0179 | NT |
| Sumitomo | .0310 | .0298 | .0146 | NT | .0221 | .0157 | .0109 | NT |
| JapanBioassay | .0468 | .0288 | .0193 | .0395 | .0336 | .0200 | .0107 | .0301 |
| CIT-France | .0335 | NT | NT | NT | .0288 | NT | NT | NT |
| Huntingdon | NT | NT | .0347 | NT | NT | NT | .0270 | NT |
| BASF | .0236 | NT | NT | NT | .0153 | NT | NT | NT |
| Bayer | .0335 | NT | NT | NT | .0214 | NT | NT | NT |
| Mitsubishi | .0297 | .0149 | .0342 | .0134 | .0196 | .0108 | .0199 | .0090 |
| Exxon | .0411 | NT | NT | NT | .0393 | NT | NT | NT |
| Chung-Korea | .0469 | .0480 | NT | NT | .0406 | .0292 | NT | NT |
| AstraZeneca | .0565 | .0257 | .0212 | NT | .0500 | .0155 | .0115 | NT |
| Denmark | NT | .0282 | NT | NT | NT | .0202 | NT | NT |
| Berlin | .0743 | NT | NT | NT | .0600 | NT | NT | NT |
| Rhone-Poulenc | .0450 | NT | NT | NT | .0247 | NT | NT | NT |
| Korea-Park | NT | .0259 | .0701 | NT | NT | .0324 | .0501 | NT |
| TNO | .0483 | .0451 | NT | NT | .0333 | .0304 | NT | NT |
| FDSC-Japan | .0334 | .0236 | .0570 | .0354 | .0314 | .0270 | .0146 | .0238 |
| IET-Japan | .0265 | .0230 | .0228 | .0106 | .0137 | .0161 | .0137 | .0109 |
| WILResearch | .0619 | .0687 | NT | NT | .0645 | .0366 | NT | NT |

[a] Tabulated values are the error mean square from an ANOVA of the log-transformed uterine weight data after adjustment for dose effects and body weight differences

NT: Not tested

**Table 16**

**COMPARISON OF COEFFICIENTS OF VARIATION (%) IN UTERINE WEIGHTS (AVERAGED OVER GROUPS WITHIN A LABORATORY AND PROTOCOL)**

| LAB | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | C′ | A | B | C | C′ |
| Citi-Japan | 14.9 | 16.1 | 16.6 | NT | 13.7 | 14.2 | 12.8 | NT |
| Sumitomo | 15.5 | 15.4 | 11.2 | NT | 12.9 | 11.9 | 9.8 | NT |
| JapanBioassay | 18.9 | 16.9 | 13.4 | 19.4 | 16.5 | 14.0 | 11.1 | 16.6 |
| CIT-France | 18.7 | NT | NT | NT | 17.4 | NT | NT | NT |
| Huntingdon | NT | NT | 17.8 | NT | NT | NT | 16.0 | NT |
| BASF | 14.9 | NT | NT | NT | 12.3 | NT | NT | NT |
| Bayer | 16.0 | NT | NT | NT | 14.0 | NT | NT | NT |
| Mitsubishi | 17.1 | 12.4 | 15.9 | 10.5 | 14.2 | 10.7 | 12.0 | 9.0 |
| Exxon | 18.1 | NT | NT | NT | 18.4 | NT | NT | NT |
| Chung-Korea | 20.9 | 22.3 | NT | NT | 20.2 | 19.3 | NT | NT |
| AstraZeneca | 23.8 | 15.4 | 14.1 | NT | 22.2 | 12.3 | 10.8 | NT |
| Denmark | NT | 15.4 | NT | NT | NT | 13.9 | NT | NT |
| Berlin | 26.3 | NT | NT | NT | 22.9 | NT | NT | NT |
| Rhone-Poulenc | 20.0 | NT | NT | NT | 15.2 | NT | NT | NT |
| Korea-Park | NT | 15.3 | 24.8 | NT | NT | 16.4 | 20.0 | NT |
| TNO | 21.1 | 19.1 | NT | NT | 17.0 | 15.9 | NT | NT |
| FDSC-Japan | 15.9 | 14.2 | 18.6 | 16.9 | 15.4 | 14.4 | 11.7 | 14.5 |
| IET-Japan | 14.6 | 14.4 | 12.8 | 10.4 | 10.8 | 12.2 | 10.8 | 10.3 |
| WILResearch | 21.8 | 25.8 | NT | NT | 19.8 | 19.5 | NT | NT |
| Overall[1] Mean | 18.7 | 16.9 | 16.1 | 14.3 | 16.5 | 14.6* | 12.8* | 12.6* |
| SD | 8.0 | 7.5 | 10.6 | 10.3 | 7.8 | 6.3 | 7.8 | 7.4 |
| N | 172 | 132 | 99 | 44 | 172 | 132 | 99 | 44 |

[1] Averaged over doses and laboratories

*significant ($p<0.05$) vs. Protocol A by Dunnett's test (after adjusting for differences due to dose and laboratory;

NT: Not tested.

**Table 17**

**COMPARISON OF COEFFICIENTS OF VARIATION (%) IN BLOTTED UTERINE WEIGHTS FOR UNTREATED AND VEHICLE CONTROL GROUPS**

| LAB | Untreated | | | | Vehicle | | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | C′ | A | B | C | C′ |
| Citi-Japan | 13.5 | 16.3 | 17.7 | NT | 12.0 | 12.4 | 12.2 | NT |
| Sumitomo | 11.0 | 3.9 | 9.9 | NT | 9.5 | 7.7 | 15.3 | NT |
| JapanBioassay | 9.4 | 11.9 | 13.6 | 11.8 | 24.0 | 10.7 | 12.6 | 13.0 |
| CIT-France | 19.1 | NT | NT | NT | 19.2 | NT | NT | NT |
| Huntingdon | NT | NT | 14.2 | NT | NT | NT | 10.3 | NT |
| BASF | 9.5 | NT | NT | NT | 6.7 | NT | NT | NT |
| Bayer | 24.4 | NT | NT | NT | 16.6 | NT | NT | NT |
| Mitsubishi | 14.6 | 10.9 | 4.7 | 10.6 | 20.8 | 11.2 | 4.6 | 2.7 |
| Exxon | NT | NT | NT | NT | 20.8 | NT | NT | NT |
| Chung-Korea | 31.4 | 21.9 | NT | NT | 30.1 | 17.1 | NT | NT |
| AstraZeneca | 15.3 | 18.1 | 9.2 | NT | 28.4 | 10.3 | 13.1 | NT |
| Denmark | NT | 13.9 | NT | NT | NT | 14.9 | NT | NT |
| Berlin | 19.6 | NT | NT | NT | 29.0 | NT | NT | NT |
| Rhone-Poulenc | 17.0 | NT | NT | NT | 13.2 | NT | NT | NT |
| Korea-Park | NT | 16.9 | 9.7 | NT | NT | 25.1 | 11.6 | NT |
| TNO | 14.1 | 26.6 | NT | NT | 21.3 | 27.8 | NT | NT |
| FDSC-Japan | 11.2 | 17.3 | 8.9 | 12.0 | 10.9 | 4.2 | 10.1 | 15.0 |
| IET-Japan | 17.2 | 7.4 | 9.7 | 7.3 | 8.2 | 14.7 | 19.1 | 10.8 |
| WILResearch | 14.1 | 14.1 | NT | NT | 33.8 | 14.3 | NT | NT |
| Overall[1] Mean | 16.1 | 15.8 | 10.8 | 10.4 | 19.0 | 14.2 | 12.1 | 10.4 |
| SD | 5.9 | 5.1 | 3.8 | 2.2 | 8.5 | 6.7 | 3.9 | 5.4 |
| N | 15 | 12 | 9 | 4 | 16 | 12 | 9 | 4 |

[1] Averaged over laboratories
NT - not tested

**Table 18**

**MEAN UTERINE BLOTTED WEIGHTS AS A PERCENTAGE OF WET WEIGHTS**

| | Controls[a] | | | | Top two EE Doses[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | **A** | **B** | **C,C′** | **Mean** | **A** | **B** | **C** | **C′** |
| Citi-Japan | 98.0 | 97.5 | 98.0 | 97.8 | 75.5 | 57.9 | 41.5 | NT |
| Sumitomo | 97.9 | 97.3 | 97.0 | 97.4 | 64.3 | 59.7 | 38.6 | NT |
| JapanBioassay | 94.9 | 96.8 | 97.1[c] | 96.5 | 73.3 | 52.7 | 33.3 | 75.6 |
| FDSC-Japan | 95.8 | 94.3 | 98.0[c] | 96.5 | 82.9 | 71.7 | 45.2 | 89.1 |
| CIT-France | 96.3 | NT | NT | 96.3 | 85.7 | NT | NT | NT |
| Huntingdon | NT | NT | 96.3 | 96.3 | NT | NT | 71.8 | NT |
| BASF | 95.8 | NT | NT | 95.8 | 75.7 | NT | NT | NT |
| Mitsubishi | 94.8 | 95.7 | 95.8[c] | 95.5 | 72.0 | 56.2 | 38.6 | 88.1 |
| Bayer | 94.8 | NT | NT | 94.8 | 73.3 | NT | NT | NT |
| Exxon | 94.7 | NT | NT | 94.7 | 77.5 | NT | NT | NT |
| Chung-Korea | 94.2 | 92.2 | NT | 93.2 | 71.2 | 49.1 | NT | NT |
| AstraZeneca | 91.9 | 91.9 | 94.8 | 92.9 | 79.0 | 64.8 | 72.2 | NT |
| IET-Japan | 91.0 | 88.9 | 95.1[c] | 92.5 | 65.1 | 48.6 | 44.4 | 87.9 |
| Denmark | NT | 91.8 | NT | 91.8 | NT | 52.0 | NT | NT |
| WILResearch | 91.3 | 90.1 | NT | 90.7 | 74.8 | 56.3 | NT | NT |
| Berlin | 89.3 | NT | NT | 89.3 | 56.1 | NT | NT | NT |
| Rhone-Poulenc | 87.2 | NT | NT | 87.2 | 64.1 | NT | NT | NT |
| Korea-Park | NT | 76.4 | 85.6 | 81.0 | NT | 48.0 | 34.5 | NT |
| TNO | 82.8 | 73.9 | NT | 78.3 | 62.3 | 54.6 | NT | NT |

[a] average of untreated, vehicle, and 0.01 (another “unaffected” group) EE

[b] average of 3.0 and 10.0 EE

[c] average of C and C' blotted uterine weights as a percentage of wet weights

NT: Not tested

**Table 19**

**EVALUATION OF BODY WEIGHTS IN PROTOCOLS C AND C'**

**PROTOCOL C**

| | Mitsubishi | | JapanBioassay | | Huntingdon | | Sumitomo | | Citi-Japan | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| **Untreated** | 231.68 | 11.40 | 203.87 | 11.04 | 298.25 | 19.08 | 213.13 | 14.62 | 222.35 | 7.97 |
| **Vehicle** | 235.20 | 12.80 | 210.42 | 9.35 | 290.63 | 12.84 | 216.55 | 17.91 | 224.02 | 13.55 |
| **0.01** | 230.83 | 9.31 | 208.73 | 12.71 | 297.90 | 7.22 | 215.38 | 6.13 | 224.57 | 9.92 |
| **0.03** | 232.23 | 8.35 | 212.03 | 11.79 | 298.30 | 16.83 | 215.47 | 10.99 | 221.32 | 9.46 |
| **0.1** | 228.10 | 11.00 | 203.53 | 11.68 | 298.60 | 11.89 | 213.38 | 13.73 | 217.57 | 6.40 |
| **0.3** | 228.92 | 12.77 | 201.53 | 9.68 | 301.75 | 7.84 | 210.38 | 8.00 | 219.57 | 9.47 |
| **1.0** | 218.93* | 12.19 | 202.00 | 12.18 | 287.18 | 27.35a | 208.08 | 7.98 | 211.02* | 7.91 |
| **3.0** | 211.65* | 10.90 | 194.67* | 13.43 | 294.57 | 9.68 | 201.08 | 11.29 | 204.65* | 5.75 |
| **10.0** | 212.38* | 10.66 | 192.88* | 10.87 | 282.43 | 9.62 | 193.58* | 7.75 | 200.22* | 5.57 |

| | AstraZeneca | | FDSC-Japan | | IET-Japan | | Korea-Park | |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| **Untreated** | 292.17 | 24.13 | 241.47 | 18.21 | 167.12 | 5.49 | 211.68 | 2.17 |
| **Vehicle** | 295.00 | 30.19 | 239.87 | 16.03 | 169.62 | 7.14 | 215.45 | 3.89 |
| **0.01** | 299.17 | 15.79 | 236.83 | 21.85 | 168.45 | 5.81 | 213.90 | 1.87 |
| **0.03** | 292.33 | 23.53 | 238.07 | 16.37 | 168.87 | 7.41 | 213.75 | 5.67 |
| **0.1** | 291.00 | 36.13 | 232.03 | 21.47 | 169.13 | 6.61 | 210.93 | 7.46 |
| **0.3** | 293.33 | 31.48 | 234.90 | 15.41 | 165.48 | 7.09 | 212.55 | 6.09 |
| **1.0** | 287.83 | 29.55 | 227.70 | 19.66 | 162.73 | 4.68 | 209.27 | 5.21 |
| **3.0** | 280.33 | 20.57 | 224.48 | 15.70 | 159.03* | 4.63 | 205.38* | 4.88 |
| **10.0** | 278.83 | 24.32 | 219.65 | 12.62 | 156.83* | 5.62 | 195.52* | 5.62 |

**Table 19 (Continued)**

**EVALUATION OF BODY WEIGHTS IN PROTOCOLS C AND C'**

**PROTOCOL C'**

| | Mitsubishi | | JapanBioassay | | FDSC-Japan | | IET-Japan | |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| Untreated | 262.70 | 13.27 | 230.77 | 9.30 | 252.43 | 12.77 | 196.97 | 6.43 |
| Vehicle | 256.80 | 8.18 | 236.20 | 13.13 | 252.73 | 15.65 | 202.63 | 4.45 |
| 0.01 | 261.53 | 10.65 | 228.63 | 11.25 | 251.38 | 15.82 | 202.55 | 6.24 |
| 0.03 | 257.12 | 9.86 | 229.98 | 12.74 | 249.83 | 16.95 | 201.13 | 7.05 |
| 0.1 | 242.13 | 12.84 | 222.38 | 15.31 | 238.77 | 11.93 | 196.32 | 6.70 |
| 0.3 | 243.25 | 8.23 | 215.03* | 12.42 | 240.83 | 16.64 | 192.00* | 7.37 |
| 1.0 | 223.75* | 11.60 | 205.43* | 12.85 | 227.30* | 12.91 | 185.63* | 3.29 |
| 3.0 | 218.40* | 10.20 | 198.63* | 11.29 | 225.70* | 13.92 | 176.27* | 10.05 |
| 10.0 | 210.52* | 10.53 | 201.32* | 8.84 | 221.45* | 14.07 | 168.87* | 6.23 |

*** p<0.05 vs. vehicle (Dunnett's test)**

**a this large SD was due to a single animal who reportedly lost 11% of her body weight (from 274.6 to 245.6) in a single day.**

**Table 20**

**SUMMARY OF THE STRENGTH OF THE CORRELATION BETWEEN UTERINE WEIGHT AND BODY WEIGHT**

| LAB | Wet weight | | | | Blotted weight | | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | C′ | A | B | C | C′ |
| Citi-Japan | - | - | - | NT | + | - | - | NT |
| Sumitomo | - | - | - | NT | - | - | - | NT |
| JapanBioassay | - | - | ++ | - | - | + | ++ | - |
| FDSC-Japan | + | + | - | - | + | - | - | - |
| CIT-France | + | NT | NT | NT | ++ | NT | NT | NT |
| Huntingdon | NT | NT | - | NT | NT | NT | - | NT |
| BASF | + | NT | NT | NT | + | NT | NT | NT |
| Mitsubishi | ++ | ++ | - | - | ++ | ++ | - | - |
| Bayer | - | NT | NT | NT | + | NT | NT | NT |
| Exxon | - | NT | NT | NT | - | NT | NT | NT |
| Chung-Korea | + | ++ | NT | NT | ++ | ++ | NT | NT |
| AstraZeneca | - | - | - | NT | - | + | + | NT |
| IET-Japan | - | - | - | - | - | + | - | - |
| Denmark | NT | - | NT | NT | NT | - | NT | NT |
| WILResearch | - | - | NT | NT | ++ | ++ | NT | NT |
| Berlin | - | NT | NT | NT | - | NT | NT | NT |
| Rhone-Poulenc | - | NT | NT | NT | - | NT | NT | NT |
| Korea-Park | NT | - | - | NT | NT | - | - | NT |
| TNO | - | - | NT | NT | - | - | NT | NT |

-: no correlation; +: weak ($p<0.05$) correlation; ++: strong ($p<0.01$) correlation
NT - not tested

**Table 21**

**RATIO OF MEAN UTERUS WEIGHTS IN TWO GROUPS RECEIVING 0.1 MG/KG OR 1.0 MG/KG ZM 189.154 AND 3.0 µG/KG EE (PROTOCOL A) OR 0.3 µG/KG EE (PROTOCOL B) EE RELATIVE TO THE GROUP RECEIVING THE CORRESPONDING DOSE OF EE ALONE**

| Dose EE | Protocol A | | | | Protocol B | | | |
|---|---|---|---|---|---|---|---|---|
| | Wet weight | | Blotted weight | | Wet weight | | Blotted weight | |
| (µg/kg) | 0.1 | 1.0 | 0.1 | 1.0 | 0.1 | 1.0 | 0.1 | 1.0 |
| Citi-Japan | 88 | 41* | 94 | 47* | 85 | 47* | 85 | 46* |
| Sumitomo | 94 | 35* | 99 | 45* | 136 | 34* | 124 | 34* |
| JapanBioassay | 86 | 35* | 84 | 41* | 76* | 33* | 77* | 32* |
| CIT-France | 85 | 43* | 86 | 45* | NT | NT | NT | NT |
| BASF | 45* | 22* | 52* | 26* | NT | NT | NT | NT |
| Bayer | 122 | 52* | 109 | 53* | NT | NT | NT | NT |
| Mitsubishi | 74* | 33* | 84 | 40* | 69* | 34* | 70* | 33* |
| Exxon | 73 | 33* | 81 | 35* | NT | NT | NT | NT |
| AstraZeneca | 110 | 37* | 105 | 40* | 89 | 36* | 89 | 35* |
| Berlin | 82 | 32* | 87 | 46* | NT | NT | NT | NT |
| Rhone-Poulenc | 107 | 31* | 107 | 44* | NT | NT | NT | NT |
| Denmark | NT | NT | NT | NT | 61* | 29* | 66* | 31* |
| Korea-Park | NT | NT | NT | NT | 86 | 55* | 90 | 60* |
| FDSC-Japan | 79* | 50* | 82 | 53* | 75* | 48* | 74* | 46* |
| IET-Japan | 76* | 28* | 86 | 43* | 76* | 38* | 79* | 40* |
| WILResearch | 88 | 36* | 98 | 48* | 71 | 50* | 65* | 48* |
| Chung-Korea@ | 115 | 57* | 100 | 44* | 52* | 21* | 76* | 37* |
| TNO@ | 103 | 37* | 99 | 50* | 116 | 24* | 104 | 34* |

@ this laboratory used 3.0 µg/kg EE, not 0.3 µg/kg EE, for Protocol B

* $p<0.05$ vs vehicle control (i.e., vs. a value of 100)

NT - Not tested

**Table 22**

**RATIO OF MEAN UTERUS WEIGHTS IN TWO GROUPS RECEIVING 0.1MG/KG OR 1.0MG/KG ZM189.154 AND 0.3 μG/KG EE RELATIVE TO THE GROUP RECEIVING 0.3 μG/KG EE ALONE FOR PROTOCOLS C AND C'**

| LAB | Protocol C | | | | Protocol C′ | | | |
|---|---|---|---|---|---|---|---|---|
| | Wet weight | | Blotted weight | | Wet weight | | Blotted weight | |
| | 0.1 | 1.0 | 0.1 | 1.0 | 0.1 | 1.0 | 0.1 | 1.0 |
| Citi-Japan | 102 | 59* | 98 | 60* | NT | NT | NT | NT |
| Sumitomo | 82* | 57* | 84* | 57* | NT | NT | NT | NT |
| JapanBioassay | 90 | 61* | 91 | 62* | 81 | 44* | 80 | 44* |
| Mitsubishi | 74* | 48* | 76* | 50* | 66* | 37* | 68* | 38* |
| Huntingdon | 100 | 104 | 100 | 104 | NT | NT | NT | NT |
| AstraZeneca | 104 | 54* | 97 | 58* | NT | NT | NT | NT |
| Korea-Park | 70[a] | 66[a] | 71[a] | 72[a] | NT | NT | NT | NT |
| FDSC-Japan | 77* | 55* | 77* | 55* | 74* | 43* | 74* | 44* |
| IET-Japan | 71* | 47* | 77* | 50* | 65* | 42* | 66* | 42* |

*p<0.05

* p<0.05 vs. vehicle control (i.e., vs. a value of 100)

[a] Not significant due to the presence of outliers (see Discussion)
NT - Not tested

**Table 23**

**SUMMARY OF POWER CALCULATIONS**

| Percent increase in uterine weight in the top dose group | Coefficient of variation | Approximate power (%) for detecting top dose effect | |
|---|---|---|---|
| | | N = 6 | N = 10 |
| 25 | 10.0 | 90.5 | 99.4 |
| 25 | 15.0 | 52.2 | 79.4 |
| 25 | 20.0 | 28.2 | 48.5 |
| 25 | 25.0 | 15.4 | 28.7 |
| | | | |
| 30 | 10.0 | 97.5 | 99.96 |
| 30 | 15.0 | 67.4 | 91.4 |
| 30 | 20.0 | 39.6 | 64.7 |
| 30 | 25.0 | 22.4 | 40.2 |
| | | | |
| 35 | 10.0 | 99.5 | 100.0 |
| 35 | 15.0 | 81.1 | 97.1 |
| 35 | 20.0 | 51.0 | 77.6 |
| 35 | 25.0 | 30.6 | 51.6 |
| | | | |
| 40 | 10.0 | 99.9 | 100.0 |
| 40 | 15.0 | 89.7 | 99.1 |
| 40 | 20.0 | 61.5 | 87.1 |
| 40 | 25.0 | 39.0 | 63.3 |

This table presents the approximate power of a design with nine groups of N animals each for detecting (at the top dose) a significant ($p<0.05$) increase in uterine weight by Dunnett's test as a function of the magnitude of the increase in the top dose group and the underlying coefficient of variation. Power calculations based on 5000 simulated studies per condition.

**Table 24**

**HILL EQUATION PARAMETER ESTIMATES AND BENCHMARK DOSE ESTIMATES FOR UTERINE BLOTTED WEIGHT: PROTOCOL A**

| LAB | Estimated benchmark Parameter estimates | | | | dose (μg/kg) | |
|---|---|---|---|---|---|---|
| | b[1] | k[2] | n[3] | (v+b)[4] | ED10 | ED90 |
| Citi-Japan | 33.82 | 2.07 | 1.90 | 137.56 | 0.65 | 6.55 |
| Sumitomo | 29.52 | 1.47 | 1.49 | 147.24 | 0.34 | 6.42 |
| JapanBioassay | 22.25 | 1.52 | 1.66 | 102.52 | 0.40 | 5.71 |
| FDSC-Japan | 27.74 | 1.39 | 2.10 | 118.67 | 0.49 | 3.96 |
| CIT-France | 41.68 | 3.09 | 1.35 | 161.78 | 0.60 | 15.82 |
| BASF | 24.11 | 2.11 | 4.00 | 109.11 | 1.22 | 3.66 |
| Mitsubishi | 31.56 | 1.48 | 2.14 | 126.62 | 0.53 | 4.12 |
| Bayer | 34.51 | 4.08 | 1.95 | 151.92 | 1.32 | 12.62 |
| Exxon | 42.52 | 2.02 | 2.60 | 140.14 | 0.86 | 4.72 |
| Chung-Korea | 42.38 | 1.26 | 1.27 | 132.49 | 0.22 | 7.06 |
| AstraZeneca | 29.53 | 1.42 | 1.63 | 117.28 | 0.37 | 5.49 |
| IET-Japan | 31.53 | 0.81 | 1.57 | 133.93 | 0.20 | 3.29 |
| WILResearch | 31.89 | 1.91 | 1.93 | 122.73 | 0.61 | 5.95 |
| Berlin | 28.75 | 0.86 | 2.03 | 126.96 | 0.29 | 2.55 |
| Rhone-Poulenc | 17.07 | 0.68 | 1.45 | 77.20 | 0.15 | 3.10 |
| TNO | 26.18 | 1.24 | 2.10 | 124.66 | 0.44 | 3.54 |

[1] estimated control value

[2] estimated ED50

[3] estimated shape parameter

[4] estimated maximum response

**Table 25**

**HILL EQUATION PARAMETER ESTIMATES AND BENCHMARK DOSE ESTIMATES FOR UTERINE BLOTTED WEIGHT: PROTOCOL B**

| LAB | Estimated benchmark Parameter estimates | | | | dose (μg/kg) | |
|---|---|---|---|---|---|---|
| | b[1] | k[2] | n[3] | (v+b)[4] | ED10 | ED90 |
| Citi-Japan | 30.81 | 0.47 | 2.58 | 147.09 | 0.20 | 1.11 |
| Sumitomo | 29.11 | 0.34 | 2.76 | 152.30 | 0.15 | 0.76 |
| JapanBioassay | 27.41 | 0.35 | 2.54 | 129.73 | 0.15 | 0.84 |
| FDSC-Japan | 27.02 | 0.55 | 2.42 | 150.09 | 0.22 | 1.36 |
| Mitsubishi | 28.33 | 0.37 | 2.36 | 148.33 | 0.14 | 0.93 |
| Denmark | 30.48 | 0.27 | 3.22 | 135.53 | 0.14 | 0.53 |
| Chung-Korea | 33.78 | 0.09 | 1.52 | 131.88 | 0.02 | 0.40 |
| Korea-Park | 17.43 | 0.55 | 2.38 | 110.46 | 0.22 | 1.39 |
| AstraZeneca | 26.61 | 0.42 | 1.74 | 125.91 | 0.12 | 1.49 |
| IET-Japan | 29.24 | 0.33 | 1.52 | 145.43 | 0.08 | 1.41 |
| WILResearch | 29.58 | 0.49 | 1.72 | 122.30 | 0.14 | 1.78 |
| TNO | 26.85 | 0.52 | 1.98 | 148.19 | 0.17 | 1.58 |

[1] estimated control value

[2] estimated ED50

[3] estimated shape parameter

[4] estimated maximum response

**Table 26**

**HILL EQUATION PARAMETER ESTIMATES AND BENCHMARK DOSE ESTIMATES FOR UTERINE BLOTTED WEIGHT: PROTOCOLS C AND C'**

| LAB: Protocol C | Estimated benchmark Parameter estimates | | | | dose (μg/kg) | |
|---|---|---|---|---|---|---|
| | b[1] | k[2] | n[3] | (v+b)[4] | ED10 | ED90 |
| Citi-Japan | 111.9 | 0.35 | 2.63 | 362.6 | 0.15 | 0.80 |
| Sumitomo | 117.0 | 0.39 | 2.32 | 390.7 | 0.15 | 1.00 |
| JapanBioassay | 98.5 | 0.42 | 2.06 | 372.9 | 0.14 | 1.22 |
| FDSC-Japan | 118.4 | 0.38 | 2.04 | 378.1 | 0.13 | 1.12 |
| Mitsubishi | 111.7 | 0.39 | 1.62 | 368.8 | 0.10 | 1.51 |
| Korea-Park | 86.9 | 0.31 | 3.61 | 342.4 | 0.17 | 0.58 |
| AstraZeneca | 84.4 | 0.23 | 2.54 | 276.3 | 0.10 | 0.55 |
| IET-Japan | 87.8 | 0.33 | 1.77 | 316.9 | 0.10 | 1.15 |
| Huntingdon | 96.2 | 2.60 | 1.79 | 349.2 | 0.76 | 8.85 |

| LAB: Protocol C′ | Estimated benchmark Parameter estimates | | | | dose (μg/kg) | |
|---|---|---|---|---|---|---|
| | b[1] | k[2] | n[3] | (v+b)[4] | ED10 | ED90 |
| IET-Japan | 85.2 | 0.32 | 2.13 | 372.2 | 0.11 | 0.88 |
| JapanBioassay | 101.1 | 0.43 | 1.90 | 410.3 | 0.13 | 1.35 |
| FDSC-Japan | 97.3 | 0.29 | 1.24 | 422.4 | 0.05 | 1.72 |
| Mitsubishi | 102.4 | 0.26 | 2.41 | 384.6 | 0.10 | 0.65 |

[1] estimated control value

[2] estimated ED50

[3] estimated shape parameter

[4] estimated maximum response

## CONCLUSIONS

158. From the results obtained in the first phase of the validation work on the rodent uterotrophic assay, including the overall statistical analysis and the practical experience of the participating laboratories, the following conclusions can be drawn on each of the stated objectives for the work.

**Objective 1: Demonstration of the dose-response relationship between uterine weight in immature rodents and in ovariectomized rodents following oral or subcutaneous injection of the reference oestrogen 17alpha ethinyl oestradiol (EE).**

159. All laboratories were successful in demonstrating a dose-response curve using the strong oestrogen agonist EE. The dose response curve was demonstrated using each of the four protocols (Protocols A, B, C and C'). For each protocol there was generally good agreement between the laboratories.

160. Statistical analysis of wet and blotted uterine weight data from laboratories showed similar results, although blotted weights appeared to be slightly more sensitive. This higher sensitivity was related, in part, to the reduced inter-laboratory and intra-laboratory variability of these measurements.

161. A mathematical model -- the Hill equation, generally provides a good fit to the data and had the additional further advantage of generating further parameters of the dose–response curve that can be used to further identify differences in performance between the laboratories.

**Objective 2: Demonstration of consistency of results between laboratories**.

162. For each protocol there was generally good agreement among laboratories with regard to the actual EE that produced increased uterine weights. Some laboratories showed differences in the shapes of the uterine weight dose-response curve as well in the magnitude of the increases found at equivalent doses. However there did not appear to be any trends associated with these incidental findings nor did they seem significant in the context of the overall inter-laboratory findings.

163. Similar statistically significant increases in uterine weight were observed by the participating laboratories under a variety of different experimental conditions (e.g. strain, diet, housing protocol, bedding, vehicle etc.) suggesting a certain robustness of the protocols.

164. Body and uterine weights were more variable at some laboratories than others. Overall a correlation of uterine weight with body weight could not be demonstrated. Nonetheless, controlling the magnitude of the underlying variability in uterine weights, especially wet weights, is still considered important when evaluating the overall sensitivity of the assay. Whether additional investigation and/or measures are needed on this issue may need to be considered in future phases.

**Objective 3: Demonstration of the robustness of all four protocols and identification of possible differences in sensitivity between the protocols.**

165. Protocol A was less sensitive than the other three protocols in the sense of requiring higher doses to produce equivalent increases in uterine weight. This result could have been expected given that Protocol A was the only protocol in which the EE was administered by oral gavage – the other three protocols (Protocols B, C and C 'were administered by sub-cutaneous injection.

166. For the adult animal protocols, limited data suggest that using a seven-day dosing regimen (Protocol C') may have a sensitivity advantage over the three-day exposure period (Protocol C). Administration of EE in the adult animal protocols led to reductions in body weight that did not occur in

the immature animal protocols (Protocol A and B). In the extended exposure protocol this was particularly marked and raises questions about the effect of EE on the adult animal.

167. Uterine weight increases in the range 25-30% can be detected in all protocols with sufficient power if the underlying coefficient of variation can be maintained at a reasonably low level e.g. in the range 10 .0 to 15.0).

168. The coefficient of variation for blotted uterine weight appears to be somewhat greater in protocol A than in the other three protocols, this may be due to the difference in route of administration and therefore the bioavailability of the reference oestrogen. However it may also be related to the animal model used as the coefficients for blotted uterine weights in the adult animal protocols (Protocols C and C') have slightly lower CV's than the immature animal protocol B which also involved sub-cutaneous injection.

**Objective 4: Selection of appropriate reference dose of ethinyl oestradiol for use in subsequent protocol(s) for investigating chemicals of unknown oestrogenic activity**

169. The first phase study was successful in helping to select the reference dose of EE to use in further validation and possible routine application of the protocol. Based on the findings of the phase one study, two reference doses are proposed.

a) The lowest dose of EE for which laboratories could demonstrate a statistically significant uterine weights – this would provide a test for to ensure the assay continues to be sensitive when testing other test substances.

   - For protocols B, C and C' a dose of 0.3 µg/kg/day gave a low but consistent statistically significant response and

   - For protocol A the dose of 1.0 µg/kg/day gave a low but statistically significant response.

b) A higher dose of EE on the ascending point of the dose-response curve is needed to be able to demonstrate and antagonist effect of a chemical, when co-administered with EE (as in the case of ZM). This dose would also demonstrate a very large uterine increase in any test system. In this situation. For use in combination with possible antagonists the effects that would be shown by EE alone must be reduced when co-administered with the antagonist. Therefore the effect of EE should not be sufficiently large so that the reduction of weight gain by the antagonist is masked by the EE effect - as would be the case if the maximal EE effect had been reached. An effect level of approximately ED 80 may be considered suitable. Using the results of the phase one work the modelled data was used to estimate a dose of EE equivalent to the ED 80.

   - For Protocol B, C and C' a dose of 1.0 µg/kg/day is proposed and,

   - For Protocol A the dose of 3.0 µg/kg/day is proposed.

**Objective 5: Confirmation of the anti-oestrogenic effects of the oestrogen antagonist known as ZM 189.154;**

170. ZM 189.154 co-administered with EE successfully reduced the uterine weights of animals when compared to the increases shown with EE alone. The magnitude of these responses was generally similar across laboratories, especially for the higher dose combination (EE and ZM) group.

**Objective 6: Identification of necessary protocol refinements.**

171. Very few protocol refinements were identified as necessary after considering the overall statistical analysis of the results.

172. Although there were a number of local variations allowed for in the OECD standardised protocol, e.g. laboratories choice of strain of rat historically used, regular housing conditions, diet and vehicle etc, there was very few major discrepant uterine weight results observed across laboratories. Where there were minor discrepancies looking at the listing of diets, strains of animals use and so on did not reveal any obvious need for more detailed analysis or consequent protocol refinement.

173. The fact that the protocol was successfully applied by a wide range of laboratories, increasing research laboratories, chemical testing facilities and governmental laboratories, some with no prior experience of the assay – also strengthens the conclusion that the results from phase one did not lead to any need to refine the protocols.

174. On the other hand, participating laboratories raised a number of issues to clarify and refine the protocol from a practical point of view, for example clarification of the day of birth as day zero rather than day 1 – and improving standardisation of the body weights of animals located to the various experimental groups.

175. Also based on the experience of the laboratories it was proposed to increase the time between ovariectomy of the adult animals and starting the experimental procedure to increase the possibility of detecting animals with incomplete ovariectomy. For the immature animal it was agree to propose increasing the age range of animals included in the study to be between 18 and 20 days at the start of the experiment. The need for good historical control data and good supply arrangements especially for the immature animals was also stressed from a practical point of view.

**RECOMMENDATIONS**

176. After considering the results and conclusions of the phase one work, it was agreed by the VMG during its second meeting in January 2000, that sufficient information had been obtained on the reliability of the different protocol options (i.e. OVX or immature with either s.c. injection or oral gavage) to demonstrate that the assay was robust and transferable for the reference chemical EE.

177. The main outstanding question was whether the same degree of reliability would be demonstrated using test substances of different oestrogenic potencies.

178. In order to answer this question it was agreed that it was unnecessary to further compare protocols which varied only in the route or length of exposure to the test substance. The basic minimum needed was to obtain information on the reliability and sensitivity of the tests by a comparison of the performance of a test in immature rodents and in ovariectomised adult females.

179. The VMG confirmed that an OECD Test Guideline which allowed choice in the use of route of exposure and animal model should be the goal, if possible. For the route of exposure this would mean that the route of exposure chosen would be the most appropriate considering both the proposed use of the chemical and existing toxicological information.

180. The VMG recommended that the subcutaneous route of administration would be the most appropriate one to compare the two model systems – immature and mature. Therefore the VMG recommended that future work should focus on Protocol B (s.c. exposure and immature animal model) and

Protocol C (s.c. exposure and mature animal model) using a number of chemicals of known oestrogenic activity. In addition a chemical not expected to have any oestrogenic activity should be included. If sufficient resources were available the VMG recommended comparing these results with results obtained using Protocol A, (exposure by oral gavage in the immature animal). This would provide additional information given the common use of the oral route of this exposure route in routine toxicological testing.

181. In summary these recommendations would allow the performance of the protocols to be rigorously compared with the performance in the phase one work with the reference oestrogen – EE. It would also test the judgement of the VMG that flexibility in the use of either animal model was scientifically justified.

182. The VMG recommended that phase two of the validation work would be organised to include two complementary approaches:

- a multi-chemical single dose approach in which a number of test substances (including one not expected to have any oestrogenic effect) are tested blind; and

- a comparative dose-response of test substances, to determine variability within a laboratory at different doses of test substances.

182. If one of the doses tested in the dose-response approach should be same as the multi-chemical approach this would allow an assessment of intra-laboratory variability with the same dose of test substance being tested twice, once in an open and once in a blind or coded manner.

**REFERENCES**


1. Colborn T, Clement C, eds. Chemically induced alterations in sexual and functional development. Princeton, NJ: Princeton Scientific Publishing, 1992.

2. EC/EEA/OECD/WHO. European workshop on the impact of endocrine disrupters on human health and wildlife. European Union Report EUR 17549. Workshop held at Weybridge, UK, Dec. 1996. Brussels:European Union, 1997.

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

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

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

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

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

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

9. Allen E, Doisy EA. An ovarian hormone. Preliminary report on its localization, extraction and partial purification, and action in test animals. J. Am. Med. Assoc. 81:819-821 (1923).

10. Astwood EB. A six-hour assay for the quantitative determination of estrogen. Endocrinologv 23:25-31 (1938).

11. Bülbring E, Burn JH. The estimation of oestrin and of male hormone in oily solution. J. Physiol. 85:320-33 (1935).

12.. Dorfman RL, Gallagher TF, Koch FC. The nature of the estrogenic substance in human male urine and bull testis. Endocrinology 19:33-41 (1936).

13. Lauson HD, Heller CG, Golden JB, Servinghaus EL. The immature rat uterus in the assay of estrogenic substances, and a comparison of estradiol, estrone and estriol. Endocrinology 24:35-44 (1939).

14. Boettger-Tong H, Murphy L, Chiappetta C, Kirkland JL, Goodwin B, Adlercreutz H, Stancel GM, Mäkelä S. A case of a laboratory animal feed with high estrogenic activity and its impact on *in vivo* responses to exogenously administered estrogens. Environ. Health Per. 106:369-373 (1998).

15. Casanova M, You L, Gaido KW, Archibeque-Engle S, Janszen DB, d'A. Heck H. Developmental effects of dietaryphytoestogens in Sprague-Dawley rats and interactions of genistein and daidzein with rat estrogen receptors $\alpha$ and $\beta$ *in vitro*. Toxicol. Sci. 51:236-244 (1999).

16. Odum J, Lefevre PA, Tittensor S, Paton D, Routledge EJ, Beresford NA, Sumpter JP, Ashby J. The rodent uterotrophic assay: Critical protocol features, studies with nonylphenols, and comparison with a yeast estrogenicity assay. Reg. Toxicol. Pharmacol. 25:176-188 (1997).

17. Thigpen JE, Li LA, Richter CB, Lebetkin EH, Jameson CW. The mouse bioassay for the detection of estrogenic activity in rodent diets: I. A standardized method for conducting the mouse bioassay. Lab. Anim. Sci. 37:596-601 (1987).

18 Zarrow MX, Lazo-Wasem EA, Shoger RL. Estrogenic activity in a commercial animal ration. Science 118:650-651 (1953).

19. Wakeling, A.E and J Bowler, 1988. Novel anti-oestrogens without partial agonist activity. J Steroid Biochem 31: 645-653

**ACKNOWLEDGMENTS**

The dedicated efforts and work of the lead laboratory (the National Institute for Health Sciences) and especially the participating laboratories which generated the data are acknowledged, namely the Centre Evaluation and Research Institute, (CERI) Japan; the Food Drug Safety Centre, Japan; the Institute of Environmental Toxicology, Japan; Mitsubishi Chemical Safety Institute, Japan; Japan Bioassay Research Centre, Japan; Sumitomo Chemical Company, Japan; Syngenta Central Toxioclogy Laboratory, UK; TNO, the Netherlands; Research Institute of Chemical Technology, Korea; WIL Research Laboratories, US; BASF, Germany; BAYER-AG, Germany; CIT, France; Aventis CropScience, France; Exxon Biomedical Sciences Inc., US; Free University of Berlin, Germany; National Institute of Toxicological Research, Korea; Huntingdon Life Sciences, UK; Institute of Food Safety and Toxicology. Denmark.

Also acknowledged is Eric Vindimian for suggesting and performing examples of the effective dose calculations based on the Hill equation, Schering and AstraZeneca (now Syngenta) for the donation of the reference chemicals and J William Owens for his outstanding contribution to the writing of the reports and providing generous amounts of time carrying out literature reviews and analysis. Finally we wish to especially thank Joe Hasemen of the US National Institute for Environmental Health Sciences and Richard Morris of Analytical Sciences, Inc. for their outstanding statistical and computational support related to this project. The statistical analyses summarized in this report involved data from 19 different laboratories and four different protocols (41 separate experiments), with data arriving at different times, and in formats that were not easily amenable to immediate statistical analysis.

**ANNEXURES**

**Annex 1**

**Participating Laboratories for OECD Validation of Rodent Uterotrophic Assay**

**First Phase**

**Listing of Participating Laboratories**

**19 Laboratories**

---

## Lead Laboratory

**Dr INOUE, Tohru**
**Chairperson**
**Division of Cellular and Molecular Toxicology**
**Centre for Biological Safety Research**
**National Institute of Health Sciences**
**1-18-1 Kamiyohga**
**Setagaya- ku, Tokyo 158-8501**
**JAPAN**

**Tel :81-3-3700 – 9619**
**Fax :81-3-3700-9647**
**E-mail: tohru@nihs.go.jp**

**Dr. KANNO, Jun**
**Senior Scientist**
**Division of Cellular and Molecular Toxicology**
**National Institute of Health Sciences**
**1-18-1 Kamigyohga**
**Setagaya-ku, Tokyo 158-8501**
**JAPAN**

**Tel: 81-3-3700-9639**
**Fax: 81-3-3700-9647**
**E-mail: kanno@nihs.go.jp**

---

## OECD Secretariat

Ms. ONYON, Lesley
Environmental Health and Safety Division
Environment Directorate
Organisation of Economic Co-operation and Development
2, rue André Pascal
75775 Paris Cedex
FRANCE

Tel: 33-1-4524-9849
Fax: 33-1-4524-1675
E-mail: Lesley.Onyon@oecd.org
Cc: Tiffany.Larsen@oecd.org

**Syngenta**

Dr ASHBY, John
Syngenta
Central Toxicology Laboratory
Alderley Park
Macclesfield
GB Cheshire SK 10 4TJ
UK

Tel: +44- 1625 51 28 33
Fax: +44 – 1625 59 02 49
Email: john.ashby@syngenta.com

Dr ODUM, Jenny
Syngenta
Central Toxicology Laboratory
Alderley Park
Macclesfield
GB Cheshire SK 10 4TJ
UK

Tel: +44- 1625 51 61 82
Fax: +44 – 1625 59 02 49
Email: jenny.odum@syngenta.com

**Aventis CropScience**

Dr. BARS, Remi
Aventis CropScience
355 rue Dostoievski
BP 153
F-06903 Sophia Antipolis
FRANCE

Tel: 33-492-94-34-83
Fax: 33-493-65-41- 04
E-mail: Remi.Bars@aventis.com

Dr. CHUZEL, Franck
Aventis CropScience
355 rue Dostoievski
BP 153
F-06903 Sophia Antipolis
FRANCE

Tel: 33-492-94-34-41
Fax: 33-493-95-84-54
E-mail: Franck.Chuzel@aventis.com

**BASF**

Dr. SCHILLING, Karsten
BASF AG Z470
Abt Toxicologie
Carl-Bosch-Strasse 35
D-67056 Ludwigshafen
GERMANY

Tel: 49-621-60-58114
Fax: 49-621-60-51734
E-mail: Karsten.Schilling@basf-ag.de
(cc Elard.Jacob@basf-ag.de)

**Bayer AG**

Dr. KRÖTLINGER, Franc
Bayer AG
Institute of Toxicology
Pharma Research Centre
D-42096 Wuppertla
GERMANY

Tel: 49-202-368-429
Fax: 49-202-368-704
E-mail: Franc.Kroetlinger.fk@bayer-ag.de

**Chemicals Evaluation and Research Institute, (CERI) Japan**

Dr. YAMASAKI, Kanjii
CERI (Formerly CITI)
Hita Research Laboratories
822 3-chome
Ishii-machi Hita Oita 877-0061
JAPAN

Tel: 81-973-24-7211
Fax: 81-973-23-9800
E-mail: Yamasaki-kanji@hita.cerij.or.jp
Cc takatsuki-mineo@cerij.or.jp
miyachi-shigeki@cerij.or.jp

Dr. SAWAKI, Masakuni
CERI (Formerly CITI)
Hita Research Laboratories
822 3-chome
Ishii-machi Hita Oita 877-0061
JAPAN

Tel: 81-973-24-7211
Fax: 81-973-23-9800
E-mail: sawaki-masakuni@hita.cerij.or.jp

**CIT, France**

Dr. RICHARD, Jacques
Senior Toxicologist
CIT
BP 563 27005
Evreux
FRANCE

Tel: 33-2-32-29-2626
Fax: 33-2-32-67-8705
E-mail: Jacques.Richard@citox.com

Dr. FORSTER, Roy
Scientific Director
CIT
BP 563 27005
Evreux
FRANCE

Tel: 33-2-32 29 26 26
Fax : 33 2-32-67 87 05
Email Roy.Forster@citox.com

**Exxon Biomedical Sciences Inc.**

Dr. FREEMAN, James
Laboratory Operations
Exxon Biomedical Sciences, Inc.
P.O. Box 971
1545 Route 22 East
Annandale, NJ 08801-0971
UNITED STATES

Tel: 1-908-730-1123
Fax: 1-908-730-1197
E-mail: jjfreem@erenj.com

**Food Drug Safety Centre, Japan**

Dr. IMAI, Kiyoshi
Food Drug Safety Center
Hatano Research Institute
729-5 Ochiai Hatano-shi
Kanagawa 257 8523
JAPAN

Tel: 81-463-82-4751
Fax: 81-463-82-9627
E-mail: Kiyoimai@aol.com

**Free University of Berlin, Germany**

Dr. CHAHOUD, Ibrahim
Institute of Clinical Pharmacology and Toxicology
Free University of Berlin
Garystrasse 5
14195 Berlin
GERMANY

Tel: 49-30-8445-1750
Fax: 49-30-8445-1761
E-mail: Chahoud@zedat.fu-berlin.de

Dr TALSNESS, Chris
Institute of Clinical Pharmacology and Toxicology
Free University of Berlin
Garystrasse 5
14195 Berlin
GERMANY

Tel: 49-30-8445-1750
Fax: 49-30-84-45-1671
Email: talsness@flintstone.ukbf.fu-berlin.de

**Huntingdon Life Sciences PLC**

Dr. RUCKMAN, Steve
Huntingdon Life Sciences Ltd.
Woolley Road, Huntingdon
Cambridgeshire PE17 5HS
UNITED KINGDOM

Tel: 44-1480-892-035
Fax: 44-1480-891-685
E-mail: RuckmanS@UKOrg.Huntingdon.com

Dr. BLANCHARD, Emma
Huntingdon Life Sciences Ltd.
Woolley Road
Huntingdon
Cambridgeshire PE17 5HS
UNITED KINGDOM

Tel: 44-1480-892-035
Fax: 44-1480-891-685
E-mail: BlanchaE@UKOrg.Huntingdon.com

**Institute of Environmental Toxicology, Japan**

Dr. AOYAMA, Hiroaki
Institute of Environmental Toxicology
4321 Uchimoriya-cho
Ibaraki 303 0043
JAPAN

Tel: 81-297-27-4532
Fax: 81-297-27-4519
E-mail: Aoyama@iet.or.jp

**Institute of Food Safety and Toxicology, Denmark**

Dr. HOSSAINI, Alireza
Institute of Food Safety and Toxicology
Danish Veterinary and Food Administration
Morkhoj Bygade 19
DK 2860
DENMARK

Tel: 45-3395-6541
Fax: 45-3395-6698
E-mail: Alh@fdir.dk
Cc JJL@VFD.DK

**Japan Bioassay Research Centre**

Dr OKUDA, Hirokazu
Japan Bioassay Research Centre
2445 Hirasawa Hatano-shi
Kanagawa 257 0015
JAPAN

Tel: +81 463 82 3911
Fax: +81 463 32 3860
Email: Hokuda@jisha.or.jp

**Mitsubishi Chemical Safety Institute**

Dr. NAGAI, Kenjii
Mitsubishi Chemical Safety Institute
Kashima Laboratory
14 Sunayama
Hasaki-machi, Kashima-gun
Ibraraki-ken 314 0255
JAPAN

Tel: 81-479-46-5084
Fax: 81-479-46-5097
E-mail: k-nagai@ankaken.co.jp

**National Institute of Toxicological Research, Korea**

Dr. HAN, Soon Young
Head of Endocrine Toxicology Division
Department of Toxicology
National Institute of Toxicological Research
Korea Food and Drug Administration
Nokbundong 5, Eunpyunggu
Seoul 122-704
KOREA

Tel: 82-2-380-1877-9
Fax: 82-2-355-9446
E-mail: soonyoungh@kfda.go.kr

cc:

Dr. PARK, Kui Lea
National Institute of Toxicological Research
Korea Food and Drug Administration
Nokbundong 5, Eunpyunggu
Seoul 122-704
KOREA

Tel: 82-2-380-1789
Fax: 82-2-355-9446
E-mail: Parkkl@kfda.go.kr

**Research Institute of Chemical Technology**
**Korea**

Dr CHUNG, Moon
Toxicology Research Centre
Korea Research Institute of Chemical Technology
P.O Box 107
Yusong
Taejon 305-600
KOREA

Tel: 82-42-860-7476
Fax: 82-42-860-7488
Email: MkChung@pado.krict.re.fr

**Sumitomo Chemical Company, Japan**

Dr. YAMADA, Tomoya
Sumitomo Chemical Company
Environmental Health Science Laboratory
1-98 Kasugade-naka 3 chome
Konohana-ku
Osaka 554 0022
JAPAN

Tel: 81-6-6466-5343
Fax: 81-6-6466-5443
E-mail: Yamadat8@sc.sumitomo-chem.co.jp

**TNO Nutrition and Food Institute**

Dr. PRINSEN, Menk
Nutrition and Food Research Institute
TNO
Toxicology Division
THE NETHERLANDS

Tel: 31-30-69-44-558
Fax: 31-30-69-60-264
E-mail: Prinsen@voeding.tno.nl

**WIL Research Laboratories, US**

Dr VARSHO, Bennett
WIL Research Laboratories
1407 George Road
Ashland, OHIO 44805
USA

Tel: 1-419-289-8700
Fax: 1-419-289-3650
Email Refeathers@wilresearch.com
Bvarsho@wilresearch.com
Dstump@wilresearch.com
(copy rick.becker@cmahq.com )

**Annex 2**

**Standardised Spreadsheet**

**(The standardised spreadsheet which was used by the participating laboratories for recording experimental data can be found in a separate PDF file entitled “StandardisedSpreadsheet.pdf”.)**

## Annex 3

## OECD Protocols for the Phase One Work

# OECD VALIDATION WORK ON IN-VIVO UTEROTROPHIC SCREENING ASSAY

## FINAL PROTOCOL A

# OECD VALIDATION WORK ON IN-VIVO UTEROTROPHIC SCREENING ASSAY

## FINAL PROTOCOL A

## Immature female rats with oral gavage

**Protocol for demonstration of dose-response of the uterotrophic assay using the reference oestrogen ethinyl oestradiol.**

## LEAD LABORATORY

**National Institute of Health Sciences**
**Cellular & Molecular Toxicology Division**
**1-18-1 Kamiyoga, Setagayaku, Tokyo-158-8501**
**Japan**
Lead Investigator: Dr. Tohru INOUE

**TABLE OF CONTENTS**

## RATIONALE

1. The rodent uterotrophic assay evaluates the ability of a chemical to show biological activities consistent with agonism or antagonism of natural oestrogens. *Uterotrophic* is a term used to describe an increased growth of tissue of the uterus. A chemical causing an increase in weight of the uterus thereby indicates that it has activity consistent with natural oestrogens.

2. **Protocol A** will generate data as the first step in an OECD project to validate the rodent uterotrophic assay. In particular it will:

- Demonstrate the **dose-response relationship** between uterine weight in young immature rats following the oral administration of the reference oestrogen 17-ethinyl oestradiol (hereafter referred to as EE) (CAS No. 57-63-6).

- Enable **variation between laboratories to be investigated** and **protocol refinement** to be proposed.

- Enable a **comparison of the results from similar protocols**, i.e. the comparison of this protocol with two similar ones in which immature rats are used with exposure by the sub-cutaneous route (**Protocol B**) and one in which mature ovariectomised rats are used (**Protocol C**) with exposure by the sub-cutaneous route.

- Assist in **selecting the appropriate reference dose of ethinyl oestradiol** to use in a subsequent protocol(s) for investigating chemicals of unknown oestrogenic activity.

3. In addition the protocol is intended to confirm the anti-oestrogenic effects of the compound ZM 189.154 (CAS No. 101908-22-9). This chemical is referred to hereafter as ZM. This is important as both anti-oestrogens and oestrogens will be included in the later steps of the OECD validation work.

4. In this protocol, EE will be administered by gavage to very young, female rats for three consecutive days. Twenty-four hours after the last administration, the rats will be humanely killed. The weight of the uterus (both wet and blotted weights) will be used to detect possible uterotrophic (oestrogenic) activity. Two satellite groups of rats will be used to confirm the anti-oestrogenic activity of ZM.

## TIME SCHEDULE

5. It is anticipated that the test can be conducted within one working week, provided that the animals have been received and sufficiently acclimatised before the experimental start date (see paragraph 12).

## TEST SUBSTANCES

### Characterisation of test substances

6. Characterisation of the test substances is the responsibility of the original chemical suppliers and those managing the chemical repository. It is not the responsibility of the lead or participating laboratories.

7. The test substances (EE and ZM) will be characterised by name, supplier, batch number, purity, appearance, storage conditions and expiry date.

**Non-routine health and safety requirements**

8. The test substances are known as possible reproductive and developmental toxicants therefore appropriate precautions should be taken to protect personnel, e.g. necessary training, labelling and storage procedures, and protective handling procedures during dose preparation and dose administration.

9. In addition, appropriate precautions such as wearing protective gloves, protective clothing and eye protection should be taken when handling the animals, diets, cages, and wastes (e.g. remaining test solutions, faeces, and carcasses). Waste disposal will be in accordance with good practice and existing regulations.

**TEST SYSTEM**

**Characterisation of the test system**

10. The study will be conducted with very young, female laboratory rats that will be obtained from a colony maintained under SPF-conditions. The specific strain of rat will be selected by the participating laboratories based on experience and historical control data under their own operating conditions. The participating laboratories will record the strain used their study report(s).

11. At the commencement of the treatment period, the age of the rats will be 19 to 20 days (day of birth counted as day 1). The exact age of the animals should be specified by the supplier or animal facility. The body weight variation among the animals will not exceed 10g or will be within +/- 5g of the mean weight.

**Animal allocation**

12. In order to reduce stress and improve the acclimatisation of the animals, particular care should be taken with arrangements for supply of animals. If the animals are externally supplied, the supplier should be requested to transport the litter together with the dam or a foster dam when the pups are about two weeks of age. If this is not possible, the immature animals should be scheduled to arrive when they are 17 days old. The pups should be then acclimatised prior to the experimental start date.

13. Upon arrival, the rats will be taken to the room assigned to this study and checked for overt signs of ill health and anomalies. Prior to the experimental start date, the animals will be allocated to the various treatment groups. The procedure shall be in principle by randomisation but should also ensure that all groups of animals have the same mean weight population within ± 5% probability level.

**Identification of the test system**

14. The study will be identified with a unique study number and individual rats will be uniquely identified e.g. by ear tags or tail tattoos. Each group of rats will be coded e.g. by a letter and a colour. Each cage will be labelled to show the laboratory code for the group, the animal identification numbers, the cage number, and the study number. The specific identification system used by the participating laboratory will be recorded and included in the study report.

## EXPERIMENTAL CONDITIONS

### Animal maintenance

15. Appropriate husbandry conditions should be followed. The room will be maintained at a temperature of 22 ± 3 Cº and a relative humidity of between 30% and 70%, other than during room cleaning. Lighting will be artificial with a cycle of 12 hours light and 12 hours dark. Prior to and at the end of the study, the cages and other materials the animals may touch will be cleaned with appropriate agents as specified in the laboratory standard operating procedure. These procedures should be recorded and this information included in the study report.

16. As some bedding materials may contain naturally occurring oestrogenic compounds, the particular bedding used by the participating laboratory should be recorded and details included in the study report.

17. Group housing is recommended because single housing of immature animals may cause considerable stress on the animals by loss of social contact. This stress may interfere with the hormonal control of uterine weight. If the animals are caged in a group, then the group number should not exceed three rats per cage, i.e. two cages of three animals per group.

### Feed and drinking water

18. Feed and drinking water (tap or filtered) will be provided *ad libitum.*

19. The rats will be fed the usual rodent diet used by the participating laboratory. Because of the possibility of dietary phytoestrogens the participating laboratory should record the details of the diet, supplier, and the batch used. This information should be included in the study report. Each batch of diet should be analysed by the supplier for nutrients and contaminants according to the supplier's normal practice. The certificate of analysis for the batch used in the study will be included in the study report. The same diet batch should be used throughout the study for all animals.

20. The participating laboratory should maintain a frozen sample of the rodent diet used so that the diet can be further analysed, if necessary, e.g. for phytoestrogens and isoflavones.

## EXPERIMENTAL PROCEDURES

### Administration of the test substances

21. The oral gavage route will be used to administer all substances.

22. The test substances will be administered once per day on three consecutive days i.e. 19, 20 and 21 days of age or 20, 21 and 22 days of age. The amount administered should be calculated on the body weight of the animal on the day of treatment. Treatment on each consecutive day will be at approximately the same time and sequence for each animal. Test dilutions of the test substance will be prepared daily unless information is available is available which confirms the stability of the test solutions. In the latter case, the dilutions of the test substance can be made before the start of the study consistent with the substance's known stability.

23. The same test vehicle should be used for both EE and ZM. The participating laboratories will record the test vehicle to be used and include this information in their study reports.

24. The EE will be dissolved in a minimal amount of 95% ethanol and diluted to final working concentration by the test vehicle (e.g. corn, arachis, sesame or olive oil). Details of the test vehicle, as well as the vehicle supplier and lot, should be recorded and this information included in the study report. The participating laboratories should preserve a sample of the vehicle, if a further analysis, e.g. of the phytoestrogen content should become necessary.

25. Two satellite groups will be used to confirm the anti-oestrogenic effects of ZM. The ZM189154 will be dissolved in minimal amount of 95% ethanol and diluted to final working concentration by the test vehicle. Gentle heating up to 60 C° may be needed for dissolution.

26. When testing the oestrogen antagonist ZM, it is administered first by gavage and then the EE is administered, also by gavage within a short a time as possible (e.g. 15 minutes).

27. The total amount of gavage per rat per day will not exceed 5mL/kg.

**Experimental groups and dose levels**

28. The study will comprise eleven groups of six females each, as shown in **Table 1** and including seven dose groups of EE, one untested control group, one vehicle control group and two satellite groups of ZM.

**Table 1 - Details of Experimental Groups and Dose Levels**

| Groups | N= | Dose | | Route | Maximum total p.o. volume/day /rat |
|---|---|---|---|---|---|
| | | EE (microgram/kg) | ZM189154 (milligram/kg) | | |
| 1 ( untreated control) | 6 | 0 | 0 | not applicable | Not applicable |
| 2 (vehicle control) | 6 | 0 | 0 | p.o. | 5ml/kg/day |
| 3 | 6 | 0.01 | 0 | p.o. | 5ml/kg/day |
| 4 | 6 | 0.03 | 0 | p.o. | 5ml/kg/day |
| 5 | 6 | 0.10 | 0 | p.o. | 5ml/kg/day |
| 6 | 6 | 0.30 | 0 | p.o. | 5ml/kg/day |
| 7 | 6 | 1.00 | 0 | p.o. | 5ml/kg/day |
| 8 | 6 | 3.00 | 0 | p.o. | 5ml/kg/day |
| 9 | 6 | 10.00 | 0 | p.o. | 5ml/kg/day |
| 10 | 6 | 3.0 | 0.1 | p.o. | 5ml/kg/day |
| 11 | 6 | 3.0 | 1.0 | p.o. | 5ml/kg/day |

**Observations, analyses and measurements**

***Clinical signs***

29. Animal observations will be conducted according to the usual routine of the participating laboratory. On working days, all cages will be checked in the morning and afternoon for dead or moribund animals. On Saturdays and Sundays and other non-working days, a minimum of one check per day will be carried out. All abnormalities will be recorded and included in the study report.

***Body weight and food consumption***

30. The body weight of each rat will be recorded daily to the nearest 0.1 g, starting just prior to initiation of treatment i.e. when the animals are allocated into groups. As an optional measurement, the amount of food consumed during the treatment period may be measured per cage by weighing the feeders. The food consumption results will be expressed in grams per rat per day.

***Measurement of uterus weight***

31. Both wet and blotted uterus weights are the mandatory endpoints of this test protocol. Measurement of the wet weight includes the uterus and its luminal contents. Blotted weight is measured after the luminal contents of the uterus have been expressed and removed.

32. Twenty-four hours after the last treatment, the rats will be humanely killed in the same sequence as the test substance was administered. The method of humane killing will be the one routinely used by the participating laboratory, and this should be recorded and details included in the study report.

33. The uterus (without ovaries) will be carefully dissected and trimmed of fascia and fat to avoid loss of luminal contents. The vagina shall be removed from the uterus at the level of the uterine cervix. Further details for the removal and preparation of uterine tissues for weight measurement are included in the legend to **Figure 1**.

34. The uterus will be transferred to a uniquely marked and weighed container (e.g. a petri-dish) with care to avoid desiccation before weighing. The uterus will be weighed with the luminal contents (wet weight) to the nearest 0.1 mg.

35. Each uterus will then be individually processed to open the uterine wall and carefully blot the excess fluid. For example, both uterine horns may be pierced or cut longitudinally, placed on moistened filter paper (e.g. Whatman No. 3) and gently pressed to absorb the luminal fluid. The procedure used must have good reproducibility within the laboratory and not be too severe to render the tissue unacceptable for histopathological analysis, as this additional investigation will be undertaken by some by some laboratories. A video is available from the lead laboratory, on request, demonstrating the blotting procedure to help improve the level of reproducibility between laboratories.

36. For those laboratories wishing to perform a histopathological examination of the vagina and/or uterus, and ovaries should be fixed in 10% neutral buffered formalin (4% formaldehyde). If histopathology is done, the procedure used must be recorded and included in the study report. As it is known that tissue reactions differ in each portion of the uterus, **Figure 2** shows the points at which histological cross sections should be made.

## STATISTICAL ANALYSIS OF THE RESULTS

37. Each participating laboratory should record and provide the raw data with the items as listed below. A report of this data and an analysis of the results should be made to the lead laboratory. The lead laboratory will be responsible for making an overall assessment and presentation to the Validation Management Group. The raw data will include body weight, clinical status of animals during the test and before necropsy and uterine weight (wet and blotted). The Validation Management Group will determine the statistical procedures to be used in the evaluation of data taking into account dependent statistical advice.

## RETENTION OF RECORDS, SAMPLES AND SPECIMENS

38. A reference sample of EE and ZM will be retained by the chemical repository or chemical supplier until the end of the whole project if its nature allows this. Samples of diet, and test vehicle should be retained by the participating laboratories, so that further analyses can be carried out if needed, Participating laboratories should retain raw data, the master copy of the final report and all other information relevant to the quality and integrity of the study.

## REPORTING REQUIREMENTS

39. A final report will be prepared for each experiment conducted by each participating laboratory including details of:

**Laboratory Protocol**:

- Including date and approval

**Testing facility**:

- Address details
- Responsible personnel and their study responsibilities

**Test Substance**:

- Characterisation of ZM and EE (to be provided by chemical supplier/repository)
- Method and frequency of preparation of dilutions

**Vehicle**:

- Characterisation of test vehicle (nature, supplier and lot)

**Test animals:**

- Strain
- Supplier and specific supplier facility
- Age on supply with birth date
- Whether or not supplied with dam or foster dam
- If supplied with dam or foster dam and information is available, process and date of weaning
- Details of acclimatisation procedure
- Number of animals per cage
- Detail and method of individual animal and group identification.

**Test Conditions:**

- Details of randomisation process (i.e. method used)
- Record of cage location in laboratory racks

- Diet ( name, type, supplier, content)
- Water source (e.g. tap water or filtered water) and supply (by tubing from a large container, in bottles etc)
- Bedding
- Record of lighting interval
- Record of air conditioning (filter maintenance)
- Record of room clean up
- Description of blotting procedure details
- Details of histopathological procedures (including copy of standard operating procedures)

**Results**

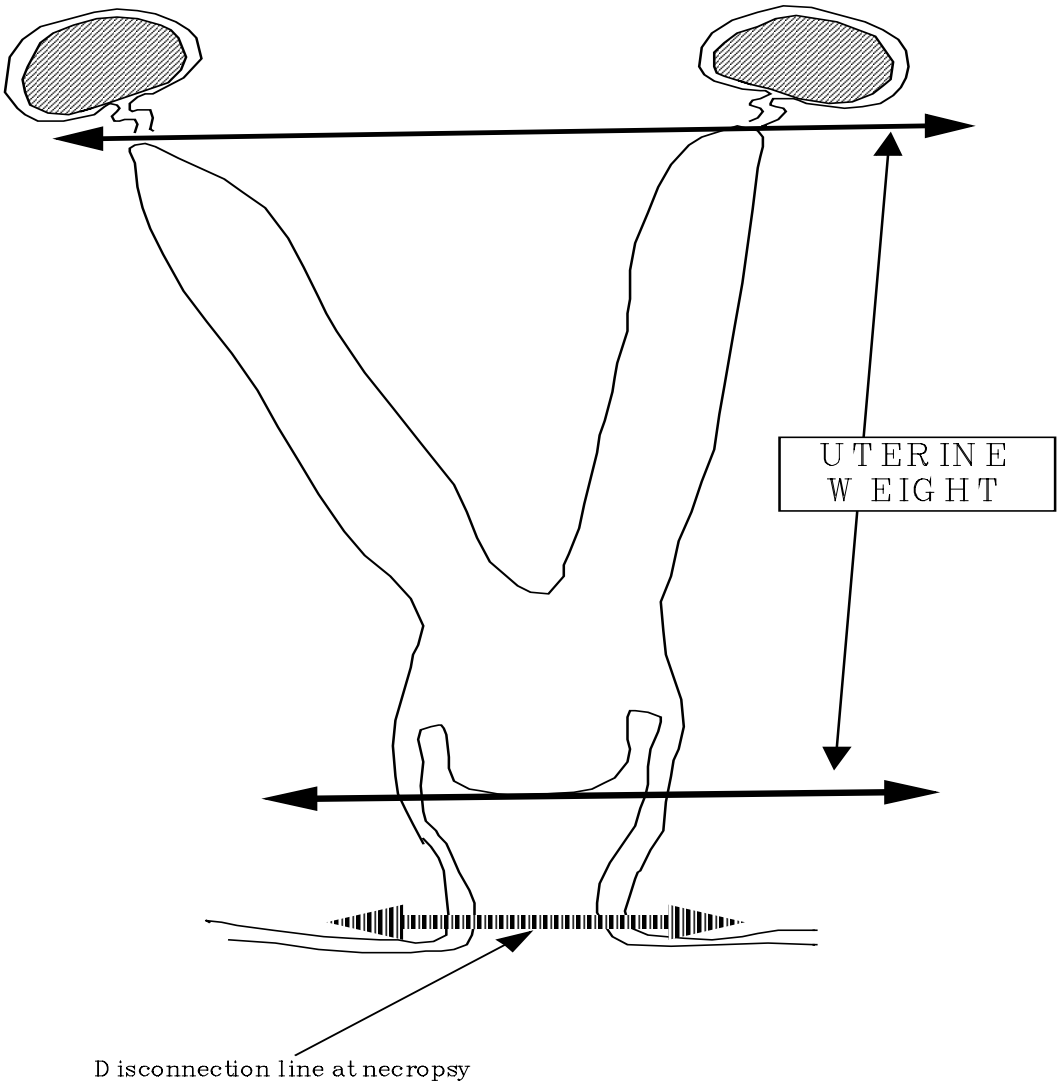
**For individual animals:**

- Daily body weight - from the day the animals are allocated into groups to the day of necropsy
- Age of each animal (in days counting birth date as day 1) when administration of test compound begins
- Date and time of each gavage administration
- Calculated amount of each gavage and any observations on losses during or after administration
- Daily record of status of animal, including relevant symptoms and observations
- Suspected cause of death (if found during study in moribund state or dead)
- Date and time of humane killing
- Approximate time interval in hours between last test substance administration and humane killing
- Organ weight at necropsy
- Wet uterine weight per animal and any observations on loss of lamina fluid during dissection and preparation for weighing (to the nearest 0.1 mg)
- Blotted uterine weight per animal (to the nearest 0.1 mg)
- If undertaken, histopathogical report of uterus, vagina and ovary.

***For each group of animals***

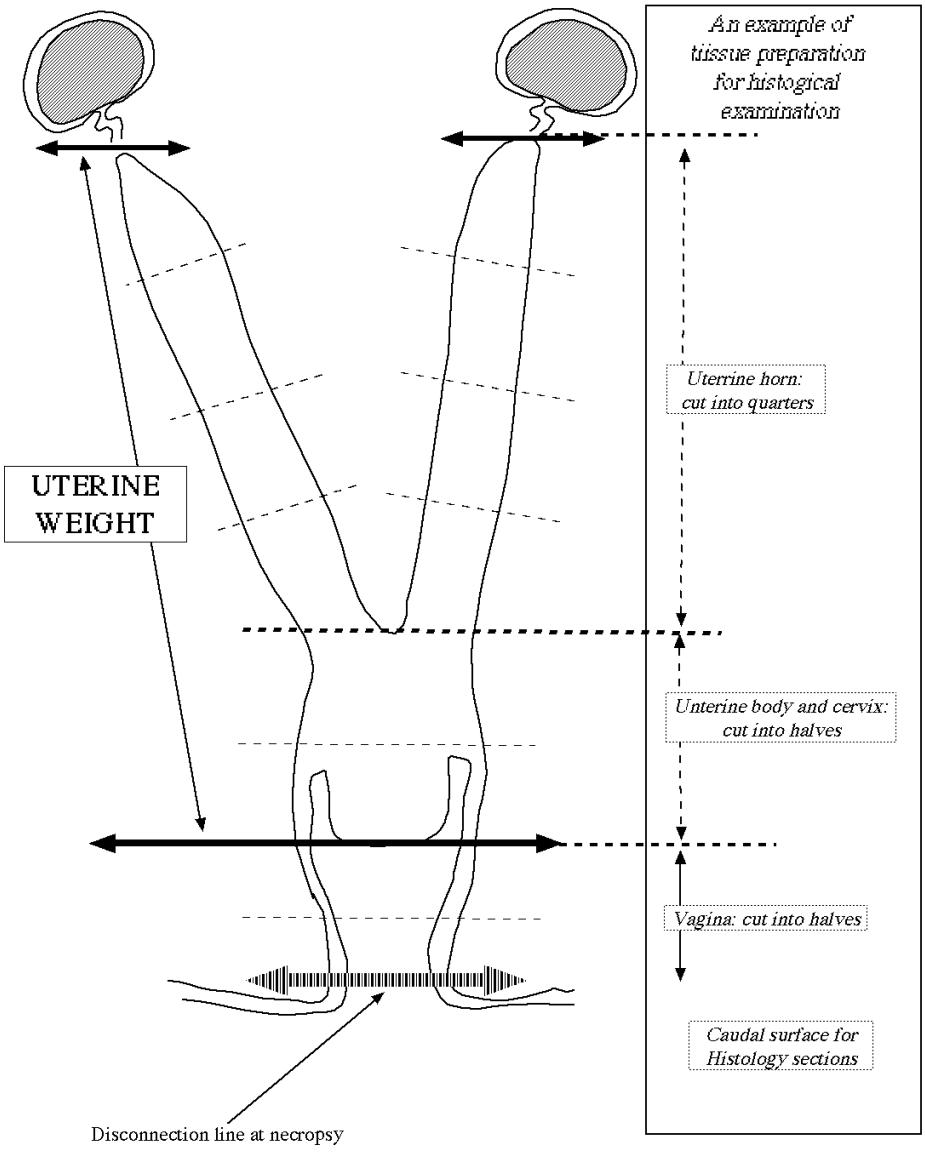
- Daily body weights (from day of allocation into groups to the day of necropsy)
- Uterine weights (both wet and dry ) per dose given
- If measured , daily food consumption

**Figure 1: The removal and preparation of the uterine tissues for weight measurement.**




In detail the procedure is to open the pubic symphysis. Then, each ovary and uterine horn is detached from the dorsal abdominal wall. Urinary bladder and ureters are removed from the ventral and lateral side of uterus and vagina. Fibrous adhesion between the rectum and the vagina is detached until the junction of vaginal orifice and perineal skin is identified. The uterus and vagina is detached from the body by incising the vaginal wall just above the junction between perineal skin as shown in the figure. The excess fat and connective tissue is trimmed away. The vagina is removed from the uterus as shown in the figure for uterus weight measurement. Weight with luminal fluid (wet weight) and without the luminal fluid (blotted weight) are measured.

**Figure 2: One example for the preparation of uterus and vagina for optional histopathological examination**.




As it is known that the tissue reaction differs in each portion of uterus. It is recommended to prepare cross sections from different portions of this hollow organ, to observe cell proliferation (for example BrdU labelling) as well as histological changes of the uterine components.

# OECD VALIDATION WORK ON IN-VIVO UTEROTROPHIC SCREENING ASSAY
## FINAL PROTOCOL B

# OECD VALIDATION WORK ON IN-VIVO UTEROTROPHIC SCREENING ASSAY

## FINAL PROTOCOL B

## Immature female rats with SUB-CUTANEOUS administration

(14/07/99 10:29)

**Protocol for demonstration of dose-response of the uterotrophic assay using the reference oestrogen ethinyl oestradiol.**

## LEAD LABORATORY

**National Institute of Health Sciences**
**Cellular & Molecular Toxicology Division**
**1-18-1 Kamiyoga, Setagayaku, Tokyo-158-8501**
**Japan**
Lead Investigator: Dr. Tohru INOUE

**TABLE OF CONTENTS**

## RATIONALE

1. The rodent uterotrophic assay evaluates the ability of a chemical to show biological activities consistent with agonism or antagonism of natural oestrogens. *Uterotrophic* is a term used to describe an increased growth of tissue of the uterus. A chemical causing an increase in weight of the uterus thereby indicates that it has activity consistent with natural oestrogens.

2. **Protocol B** will generate data as the first step in an OECD project to validate the rodent uterotrophic assay. In particular it will:

- Demonstrate the **dose-response relationship** between uterine weight in young immature rats following sub-cutaneous injection of the reference oestrogen 17-ethinyl oestradiol (hereafter referred to as EE) (CAS No. 57-63-6).

- Enable **variation between laboratories to be investigated** and **protocol refinement** to be proposed.

- Enable a **comparison of the results from similar protocols**, i.e. the comparison of this protocol with two similar ones in which immature rats are used with exposure by oral gavage (**Protocol A**) and one in which mature ovariectomised rats are used (**Protocol C**) with exposure by the sub-cutaneous route.

- Assist in **selecting the appropriate reference dose of ethinyl oestradiol** to use in a subsequent protocol(s) for investigating chemicals of unknown oestrogenic activity.

3. In addition the protocol is intended to confirm the anti-oestrogenic effects of the compound ZM 189.154 (CAS No. 101908-22-9). This chemical is referred to hereafter as ZM. This is important as both anti-oestrogens and oestrogens will be included in the later steps of the OECD validation work.

4. In this protocol, EE will be administered by s.c injection to very young, female rats for three consecutive days. Twenty-four hours after the last administration, the rats will be humanely killed. The weight of the uterus (both wet and blotted weights) will be used to detect possible uterotrophic (oestrogenic) activity. Two satellite groups of rats will be used to confirm the anti-oestrogenic activity of ZM.

## TIME SCHEDULE

5. It is anticipated that the test can be conducted within one working week, provided that the animals have been received and sufficiently acclimatised before the experimental start date (see paragraph 12).

## TEST SUBSTANCES

### Characterisation of test substances

6. Characterisation of the test substances is the responsibility of the original chemical suppliers and those managing the chemical repository. It is not the responsibility of the lead or participating laboratories.

7. The test substances (EE and ZM) will be characterised by name, supplier, batch number, purity, appearance, storage conditions and expiry date.

**Non-routine health and safety requirements**

8. The test substances are known as possible reproductive and developmental toxicants therefore appropriate precautions should be taken to protect personnel, e.g. necessary training, labelling and storage procedures, and protective handling procedures during dose preparation and dose administration.

9. In addition, appropriate precautions such as wearing protective gloves, protective clothing and eye protection should be taken when handling the animals, diets, cages, and wastes (e.g. remaining test solutions, faeces, and carcasses). Waste disposal will be in accordance with good practice and existing regulations.

**TEST SYSTEM**

**Characterisation of the test system**

10. The study will be conducted with very young, female laboratory rats that will be obtained from a colony maintained under SPF-conditions. The specific strain of rat will be selected by the participating laboratories based on experience and historical control data under their own operating conditions. The participating laboratories will record the strain used their study report(s).

11. At the commencement of the treatment period, the age of the rats will be 19 to 20 days (day of birth counted as day 1). The exact age of the animals should be specified by the supplier or animal facility. The body weight variation among the animals will not exceed 10g or will be within +/- 5g of the mean weight.

**Animal allocation**

12. In order to reduce stress and improve the acclimatisation of the animals, particular care should be taken with arrangements for supply of animals. If the animals are externally supplied, the supplier should be requested to transport the litter together with the dam or a foster dam when the pups are about two weeks of age. If this is not possible, the immature animals should be scheduled to arrive when they are 17 days old. The pups should be then acclimatised prior to the experimental start date.

13. Upon arrival, the rats will be taken to the room assigned to this study and checked for overt signs of ill health and anomalies. Prior to the experimental start date, the animals will be allocated to the various treatment groups. The procedure shall be in principle by randomisation but should also ensure that all groups of animals have the same mean weight population within ± 5% probability level.

**Identification of the test system**

14. The study will be identified with a unique study number and individual rats will be uniquely identified e.g. by ear tags or tail tattoos. Each group of rats will be coded e.g. by a letter and a colour. Each cage will be labelled to show the laboratory code for the group, the animal identification numbers, the cage number, and the study number. The specific identification system used by the participating laboratory will be recorded and included in the study report.

## EXPERIMENTAL CONDITIONS

### Animal maintenance

15. Appropriate husbandry conditions should be followed. The room will be maintained at a temperature of 22 ± 3 Cº and a relative humidity of between 30% and 70%, other than during room cleaning. Lighting will be artificial with a cycle of 12 hours light and 12 hours dark. Prior to and at the end of the study, the cages and other materials the animals may touch will be cleaned with appropriate agents as specified in the laboratory standard operating procedure. These procedures should be recorded and this information included in the study report.

16. As some bedding materials may contain naturally occurring oestrogenic compounds, the particular bedding used by the participating laboratory should be recorded and details included in the study report.

17. Group housing is recommended because single housing of immature animals may cause considerable stress on the animals by loss of social contact. This stress may interfere with the hormonal control of uterine weight. If the animals are caged in a group, then the group number should not exceed three rats per cage i.e. two cages of three animals per group.

### Feed and drinking water

18. Feed and drinking water (tap or filtered) will be provided *ad libitum.*

19. The rats will be fed the usual rodent diet used by the participating laboratory. Because of the possibility of dietary phytoestrogens the participating laboratory should record the details of the diet, supplier, and the batch used. This information should be included in the study report. Each batch of diet should be analysed by the supplier for nutrients and contaminants according to the supplier's normal practice. The certificate of analysis for the batch used in the study will be included in the study report. The same diet batch should be used throughout the study for all animals.

20. The participating laboratory should maintain a frozen sample of the rodent diet used so that the diet can be further analysed, if necessary, e.g. for phytoestrogens and isoflavones.

## EXPERIMENTAL PROCEDURES

### Administration of the test substances

21. Sub-cutaneous injection will be used to administer all substances.

22. The test substances will be administered once per day on three consecutive days i.e. 19, 20 and 21 days of age or 20, 21 and 22 days of age. The amount administered should be calculated on the body weight of the animal on the day of treatment. Treatment will be for three consecutive days at approximately the same time and sequence for each animal. Test dilutions of the test substance will be prepared daily unless information is available is available which confirms the stability of the test solutions. In the latter case, the dilutions of the test substance can be made before the start of the study consistent with the substance's known stability.

23. The same test vehicle should be used for both EE and ZM. The participating laboratories will record the test vehicle to be used and include this information in their study reports.

24. The EE will be dissolved in a minimal amount of 95% ethanol and diluted to final working concentration by the test vehicle (e.g. corn, arachis, sesame or olive oil). Details of the test vehicle, as well as the vehicle supplier and lot, should be recorded and this information included in the study report. The participating laboratories should preserve a sample of the vehicle, if a further analysis, e.g. of the phytoestrogen content should become necessary.

25. Two satellite groups will be used to confirm the anti-oestrogenic effects of ZM. The ZM189154 will be dissolved in minimal amount of 95% ethanol and diluted to final working concentration by the test vehicle. Gentle heating up to 60 C° may be needed for dissolution.

26. When testing the oestrogen antagonist ZM, it is administered first by s.c. injection and then the EE is administered, also by s.c. injection within a short a time as possible (e.g. 15 minutes).

27. The total amount of s.c. injection per rat per day will not exceed 4mL/kg.

**Experimental groups and dose levels**

28. The study will comprise eleven groups of six females each, as shown in **Table 1** including 7 doses of EE, one untreated control group, one vehicle control group and two satellite groups for ZM.

**Table 1 - Details of Experimental Groups and Dose Levels**

| Groups | N= | Dose | | Route | Maximum total s.c. volume/day/rat |
|---|---|---|---|---|---|
| | | EE (microgram/kg) | ZM189154 (milligram/kg) | | |
| 1 ( untreated control) | 6 | 0 | 0 | not applicable | Not applicable |
| 2 (vehicle control) | 6 | 0 | 0 | s.c. | 4ml/kg/day |
| 3 | 6 | 0.01 | 0 | s.c. | 4ml/kg/day |
| 4 | 6 | 0.03 | 0 | s.c. | 4ml/kg/day |
| 5 | 6 | 0.10 | 0 | s.c. | 4ml/kg/day |
| 6 | 6 | 0.30 | 0 | s.c. | 4ml/kg/day |
| 7 | 6 | 1.00 | 0 | s.c. | 4ml/kg/day |
| 8 | 6 | 3.00 | 0 | s.c. | 4ml/kg/day |
| 9 | 6 | 10.00 | 0 | s.c. | 4ml/kg/day |
| 10 | 6 | 0.3 | 0.1 | s.c. | 4ml/kg/day |
| 11 | 6 | 0.3 | 1.0 | s.c. | 4ml/kg/day |

**Observations, analyses and measurements**

***Clinical signs***

29. Animal observations will be conducted according to the usual routine of the participating laboratory. On working days, all cages will be checked in the morning and afternoon for dead or moribund animals. On Saturdays and Sundays and other non-working days, a minimum of one check per day will be carried out. All abnormalities will be recorded and included in the study report.

***Body weight and food consumption***

30. The body weight of each rat will be recorded daily to the nearest 0.1 g, starting just prior to initiation of treatment i.e. when the animals are allocated into groups. As an optional measurement, the amount of food consumed during the treatment period may be measured per cage by weighing the feeders. The food consumption results will be expressed in grams per rat per day.

***Measurement of uterus weight***

31. Both wet and blotted uterus weights are the mandatory endpoints of this test protocol. Measurement of the wet weight includes the uterus and its luminal contents. Blotted weight is measured after the luminal contents of the uterus have been expressed and removed.

32. Twenty-four hours after the last treatment, the rats will be humanely killed in the same sequence as the test substance was administered. The method of humane killing will be the one routinely used by the participating laboratory and this should be recorded and details included in the study report.

33. The uterus (without ovaries) will be carefully dissected and trimmed of fascia and fat to avoid loss of luminal contents. The vagina shall be removed from the uterus at the level of the uterine cervix. Further details for the removal and preparation of uterine tissues for weight measurement are included in the legend to **Figure 1**.

34. The uterus will be transferred to a uniquely marked and weighed container (e.g. a petri-dish) with care to avoid desiccation before weighing. The uterus will be weighed with the luminal contents (wet weight) to the nearest 0.1 mg.

35. Each uterus will then be individually processed to open the uterine wall and carefully blot the excess fluid. For example, both uterine horns may be pierced or cut longitudinally, placed on moistened filter paper (e.g. Whatman No. 3) and gently pressed to absorb the luminal fluid. The procedure used must have good reproducibility within the laboratory and not be too severe to render the tissue unacceptable for histopathological analysis, as this additional investigation will be undertaken by some by some laboratories. A video is available from the lead laboratory, on request, demonstrating the blotting procedure to help improve the level of reproducibility between laboratories.

36. For those laboratories wishing to perform a histopathological examination, the uterus, vagina and ovaries should be fixed in 10% neutral buffered formalin (4% formaldehyde). If histopathology is done, the procedure used must be recorded and included in the study report. As it is known that tissue reactions differ in each portion of the uterus, **Figure 2** shows the points at which histological cross sections should be made.

## STATISTICAL ANALYSIS OF THE RESULTS

37. Each participating laboratory should record and provide the raw data with the items as listed below. A report of this data and an analysis of the results should be made to the lead laboratory. The lead laboratory will be responsible for making an overall assessment and presentation to the Validation Management Group. The raw data will include body weight, clinical status of animals during the test and before necropsy and uterine weight (wet and blotted). The Validation Management Group will determine the statistical procedures to be used in the evaluation of data taking into account dependent statistical advice.

## RETENTION OF RECORDS, SAMPLES AND SPECIMENS

38. A reference sample of EE and ZM will be retained by the chemical repository or chemical supplier until the end of the whole project if its nature allows this. Samples of diet, and test vehicle should be retained by the participating laboratories, so that further analyses can be carried out if needed. Participating laboratories should retain raw data, the master copy of the final report and all other information relevant to the quality and integrity of the study.

## REPORTING REQUIREMENTS

39. A final report will be prepared for each experiment conducted by each participating laboratory including details of:

**Laboratory Protocol**:

- Including date and approval

**Testing facility**:

- Address details
- Responsible personnel and their study responsibilities

**Test Substance**:

- Characterisation of ZM and EE ( to be provided by chemical supplier/repository)
- Method and frequency of preparation of dilutions

**Vehicle**:

- Characterisation of test vehicle (nature, supplier and lot)

**Test animals:**

- Strain
- Supplier and specific supplier facility
- Age on supply with birth date
- Whether or not supplied with dam or foster dam
- If supplied with dam or foster dam and information is available, process and date of weaning
- Details of acclimatisation procedure

- Number of animals per cage
- Detail and method of individual animal and group identification.

**Test Conditions:**

- Details of randomisation process (i.e. method used)
- Record of cage location in laboratory racks
- Diet (name, type, supplier, content)
- Water source (e.g. tap water or filtered water) and supply (by tubing from a large container, in bottles, etc.)
- Bedding
- Record of lighting interval
- Record of air conditioning (filter maintenance)
- Record of room clean up
- Description of blotting procedure
- Details of histopathological procedures (including copy of standard operating procedures)

**Results**

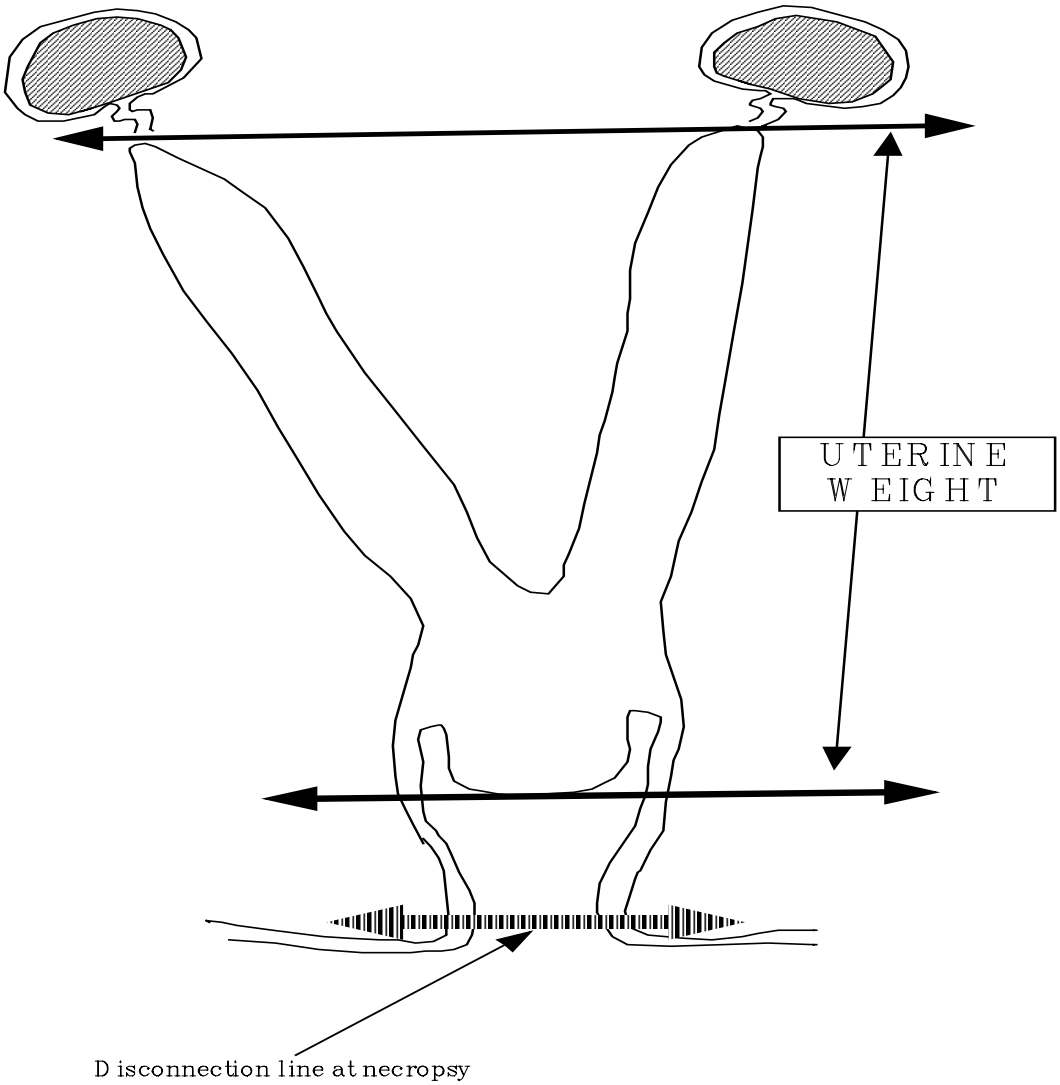
**For individual animals:**

- Daily body weight - from the day the animals are allocated into groups to the day of necropsy
- Age of each animal ( in days counting birth date as day 1) when administration of test compound begins
- Date and time of each s.c. administration
- Calculated amount of s.c. injection and any observations on losses during or after administration
- Daily record of status of animal, including relevant symptoms and observations
- Suspected cause of death ( if found during study in moribund state or dead)
- Date and time of humane killing
- Approximate time interval in hours between last test substance administration and humane killing
- Organ weight at necropsy
- Wet uterine weight per animal and any observations on loss of luminal fluid during dissection and preparation for weighing (to the nearest 0.1 mg)
- Blotted uterine weight per animal (to the nearest 0.1 mg)
- If undertaken, report of histopathogical analysis of uterus, vagina and ovaries.

***For each group of animals***

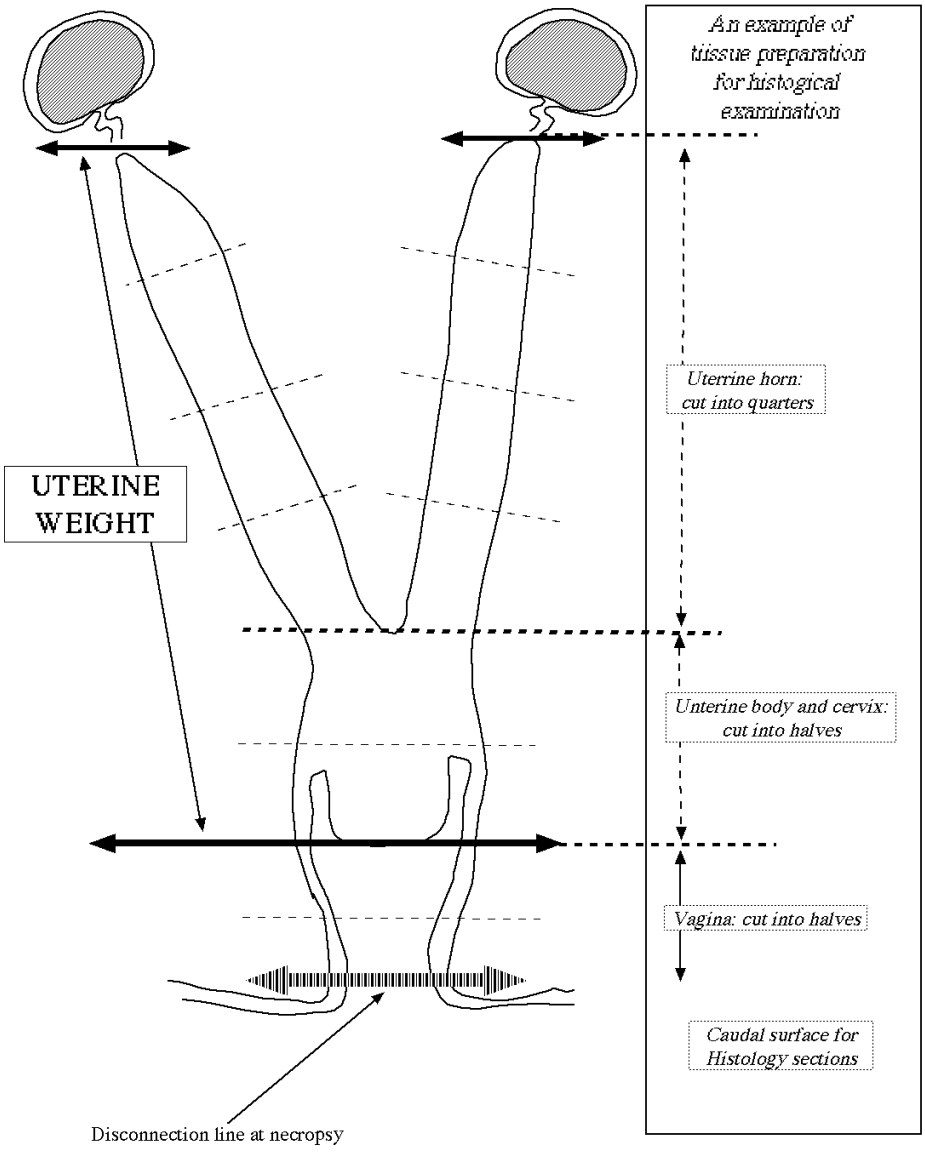
- Daily body weights (from day of allocation into groups to the day of necropsy)
- Uterine weights (both wet and dry ) per dose given
- If measured , daily food consumption

**Figure 1: The removal and preparation of the uterine tissues for weight measurement.**




In detail the procedure is to open the pubic symphysis. Then, each ovary and uterine horn is detached from the dorsal abdominal wall. Urinary bladder and ureters are removed from the ventral and lateral side of uterus and vagina. Fibrous adhesion between the rectum and the vagina is detached until the junction of vaginal orifice and perineal skin is identified. The uterus and vagina is detached from the body by incising the vaginal wall just above the junction between perineal skin as shown in the figure. The excess fat and connective tissue is trimmed away. The vagina is removed from the uterus as shown in the figure for uterus weight measurement. Weight with luminal fluid (wet weight) and without the luminal fluid (blotted weight) are measured.

**Figure 2: One example for the preparation of the uterus and vagina for optional histopathological examinations**.




As it is known that the tissue reaction differs in each portion of uterus. It is recommended to prepare cross sections from different portions of this hollow organ, to observe cell proliferation (for example BrdU labelling) as well as histological changes of the uterine components.

# OECD VALIDATION WORK ON IN-VIVO UTEROTROPHIC SCREENING ASSAY

## FINAL PROTOCOL C

## Including Variations for Protocol C'

# OECD VALIDATION WORK ON IN-VIVO UTEROTROPHIC SCREENING ASSAY

## FINAL PROTOCOL C (including variations for Protocol C')

## mature OVARIECTOMISED rats with SUB-CUTANEOUS ADMINISTRATION

(14/07/99 14:27)

**Protocol for demonstration of dose-response of the uterotrophic assay using the reference oestrogen ethinyl oestradiol.**

## LEAD LABORATORY

**National Institute of Health Sciences**
**Cellular & Molecular Toxicology Division**
**1-18-1 Kamiyoga, Setagayaku, Tokyo-158-8501**
**Japan**
Lead Investigator: Dr. Tohru INOUE

**TABLE OF CONTENTS**

## RATIONALE

1. The rodent uterotrophic assay evaluates the ability of a chemical to show biological activities consistent with agonism or antagonism of natural oestrogens. *Uterotrophic* is a term used to describe an increased growth of tissue of the uterus. A chemical causing an increase in weight of the uterus thereby indicates that it has activity consistent with natural oestrogens.

2. **Protocol C** will generate data as the first step in an OECD project to validate the rodent uterotrophic assay. In particular it will:

- **Demonstrate** the dose-response relationship **between uterine weight in adult female ovariectomised rats following sub-cutaneous injection of the reference oestrogen 17-ethinyl oestradiol (hereafter referred to as EE) (CAS No. 57-63-6)**.

- **Enable** variation between laboratories **to be investigated and** protocol refinement to be proposed.

- **Enable** a comparison of the results from similar protocols, i**.e. the comparison of this protocol with two similar ones in which immature female rats are used with exposure by oral gavage (Protocol A) and one in which immature female rats are used with exposure by sub-cutaneous injection (Protocol B).**

- **Assist in** selecting the appropriate reference dose of ethinyl oestradiol **to use in a subsequent protocol(s) for investigating chemicals of unknown oestrogenic activity.**

3. In addition, the protocol is intended to confirm the anti-oestrogenic effects of the compound ZM 189.154 (CAS No. 101908-22-9). This chemical is referred to hereafter as ZM. This is important as both anti-oestrogens and oestrogens will be included in the later steps of the OECD validation work.

4. In this protocol, EE will be administered by s.c. injection to adult female ovariectomised rats for three consecutive days. As an option, an additional study may be undertaken by the participating laboratory in which a longer period of administration of the test substance(s) is undertaken e.g. for 5 or 7 days. Two satellite groups of rats will be used to confirm the anti-oestrogenic activity of ZM.

5. Twenty-four hours after the last administration, the rats will be humanely killed. The weight of the uterus (both wet and blotted weights) will be used to detect possible uterotrophic (oestrogenic) activity.

## TIME SCHEDULE

6. The time needed for this assay will depend on whether the animals are supplied following ovariectomy or whether this procedure is carried out by the participating laboratory. It will also depend on the length of the exposure period.

## TEST SUBSTANCES

### Characterisation of test substances

7. Characterisation of the test substances is the responsibility of the original chemical suppliers and those managing the chemical repository. It is not the responsibility of the lead or participating laboratories.

8. The test substances (EE and ZM) will be characterised by name, supplier, batch number, purity, appearance, storage conditions and expiry date.

**Non-routine health and safety requirements**

9. The test substances are known as possible reproductive and developmental toxicants therefore appropriate precautions should be taken to protect personnel, e.g. necessary training, labelling and storage procedures, and protective handling procedures during dose preparation and dose administration.

10. In addition, appropriate precautions such as wearing protective gloves, protective clothing and eye protection should be taken when handling the animals, diets, cages, and wastes (e.g. remaining test solutions, faeces, and carcasses). Waste disposal will be in accordance with good practice and existing regulations.

**TEST SYSTEM**

**Characterisation of the test system**

11. The study will be conducted with adult female ovariectomised laboratory rats. These rats will be obtained from a colony maintained under SPF-conditions. The specific strain of rat will be selected by the participating laboratories based on experience and historical control data under their own operating conditions. The participating laboratories will record the strain used their study report(s).

12. At the commencement of the ovariectomy the rats will be 6 weeks old and over.

**Animal allocation**

13. Upon arrival, the rats will be taken to the room assigned to this study and checked for overt signs of ill health and anomalies. The adult females will be acclimatised to laboratory conditions for at least 7 days. Prior to the experimental start date, the animals will be allocated to the various treatment groups. The procedure shall be in principle by randomisation but should also ensure that all groups of animals have the same mean weight population within ± 5% probability level.

**Identification of the test system**

14. The study will be identified with a unique study number and individual rats will be uniquely identified e.g. by ear tags or tail tattoos. Each group of rats will be coded e.g. by a letter and a colour. Each cage will be labelled to show the laboratory code for the group, the animal identification numbers, the cage number, and the study number. The specific identification system used by the participating laboratory will be recorded and included in the study report.

**EXPERIMENTAL CONDITIONS**

**Animal maintenance**

15. Appropriate husbandry conditions should be followed. The room will be maintained at a temperature of 22 ± 3 C° and a relative humidity of between 30% and 70%, other than during room cleaning. Lighting will be artificial with a cycle of 12 hours light and 12 hours dark. Prior to and at the end of the study, the cages and other materials the animals may touch will be cleaned with appropriate

agents as specified in the laboratory standard operating procedure. These procedures should be recorded and this information included in the study report.

16. As some bedding materials may contain naturally occurring oestrogenic compounds, the particular bedding used by the participating laboratory should be recorded and details included in the study report.

17. Animals may be caged singly or in groups. In the case of group housing then the group number should not exceed three rats per cage i.e., two cages of three animals per group.

**Feed and drinking water**

18. Feed and drinking water (tap or filtered) will be provided *ad libitum.*

19. The rats will be fed the usual rodent diet used by the participating laboratory. Because of the possibility of dietary phytoestrogens the participating laboratory should record the details of the diet, supplier, and the batch used. This information should be included in the study report. Each batch of diet should be analysed by the supplier for nutrients and contaminants according to the supplier's normal practice. The certificate of analysis for the batch used in the study will be included in the study report. The same diet batch should be used throughout the study for all animals.

20. The participating laboratory should maintain a frozen sample of the rodent diet used so that the diet can be further analysed, if necessary, e.g. for phytoestrogens and isoflavones.

**EXPERIMENTAL PROCEDURES**

**Procedure for Ovariectomy**

21. A video is available, on request from the lead laboratory which shows the surgical procedure for ovariectomy. Essentially the procedure is as follows. The dorso-lateral abdominal wall should be cut 1 cm lengthways at the mid point between the costal inferior border and the illiac crest, and a few millimetres lateral to the lateral margin of the lumbar muscle. The ovary should be pulled out and disconnected at the junction of the oviduct and the uterine body. After confirming that no massive bleeding is occurring, the abdominal wall should be closed by one suture and the skin closed by autoclips. The ligation points are shown schematically in **Figure 1**.

**Administration of the test substances**

22. S.c. injection will be used to administer all substances.

23. The test substances will be administered once per day on three consecutive days. If the participating laboratory is additionally running a comparative study with longer exposure periods, the test substances will be administered for the longer period as appropriate, e.g. 7 days.

24. The amount administered should be calculated on the body weight of the animal on the day of treatment. Treatment will be at approximately the same time and sequence for each animal. Test dilutions of the test substance will be prepared daily unless information is available which confirms the stability of the test solutions. In the latter case, the dilutions of the test substance can be made before the start of the study consistent with the substance's known stability.

25. The same test vehicle should be used for both EE and ZM. The participating laboratories will record the test vehicle to be used and include this information in their study reports.

26. The EE will be dissolved in a minimal amount of 95% ethanol and diluted to final working concentration by the test vehicle (e.g. corn, arachis, sesame or olive oil). Details of the test vehicle, as well as the vehicle supplier and lot, should be recorded and this information included in the study report. The participating laboratories should preserve a sample of the vehicle, if a further analysis, e.g. of the phytoestrogen content should become necessary.

27. Two satellite groups will be used to confirm the anti-oestrogenic effects of ZM. The ZM189154 will be dissolved in minimal amount of 95% ethanol and diluted to final working concentration by the test vehicle. Gentle heating up to 60 C° may be needed for dissolution.

28. When testing the oestrogen antagonist ZM, it is administered first by s.c. injection and then the EE is administered, also by s.c. injection within a short a time as possible (e.g. 15 minutes)

29. The total amount of s.c. injection per rat per day will not exceed 4mL/kg.

**Experimental groups and dose levels**

30. The study will comprise eleven groups of six females each, as shown in **Table 1** and including the dose groups for EE, one untreated control group, one vehicle control group and two satellite groups for ZM.

**Table 1 - Details of Experimental Groups and Dose Levels**

| Groups | N= | Dose | | Route | Maximum total s.c. volume/day/rat |
|---|---|---|---|---|---|
| | | EE (microgram/kg) | ZM189154 (milligram/kg) | | |
| 1 ( untreated control) | 6 | 0 | 0 | Not applicable | Not applicable |
| 2 (vehicle control) | 6 | 0 | 0 | s.c. | 4ml/kg/day |
| 3 | 6 | 0.01 | 0 | s.c. | 4ml/kg/day |
| 4 | 6 | 0.03 | 0 | s.c. | 4ml/kg/day |
| 5 | 6 | 0.10 | 0 | s.c. | 4ml/kg/day |
| 6 | 6 | 0.30 | 0 | s.c. | 4ml/kg/day |
| 7 | 6 | 1.00 | 0 | s.c. | 4ml/kg/day |
| 8 | 6 | 3.00 | 0 | s.c. | 4ml/kg/day |
| 9 | 6 | 10.00 | 0 | s.c. | 4ml/kg/day |
| 10 | 6 | 0.3 | 0.1 | s.c | 4ml/kg/day |
| 11 | 6 | 0.3 | 1.0 | s.c. | 4ml/kg/day |

**Observations, analyses and measurements**

***Clinical signs***

31. Animal observations will be conducted according to the usual routine of the participating laboratory. On working days, all cages will be checked in the morning and afternoon for dead or moribund animals. On Saturdays and Sundays and other non-working days, a minimum of one check per day will be carried out. All abnormalities will be recorded and included in the study report.

***Body weight and food consumption***

32. The body weight of each rat will be recorded daily to the nearest 0.1 g, starting just prior to initiation of treatment i.e. when the animals are allocated into groups. As an optional measurement, the amount of food consumed during the treatment period may be measured per cage by weighing the feeders. The food consumption results will be expressed in grams per rat per day.

***Measurement of uterus weight***

33. Both wet and blotted uterus weights are the mandatory endpoints of this test protocol. Measurement of the wet weight includes the uterus and its luminal contents. Blotted weight is measured after the luminal contents of the uterus have been expressed and removed.

34. Twenty-four hours after the last treatment, the rats will be humanely killed in the same sequence as the test substance was administered. The method of humane killing will be the one routinely used by the participating laboratory, and this should be recorded and details included in the study report.

35. The uterus will be carefully dissected and trimmed of fascia and fat to avoid loss of luminal contents. The vagina shall be removed from the uterus at the level of the uterine cervix. Further details for the removal and preparation of uterine tissues for weight measurement are included in the legend to **Figure 2**.

36. The uterus will be transferred to a uniquely marked and weighed container (e.g. a petri-dish) with care to avoid desiccation before weighing. The uterus will be weighed with the luminal contents (wet weight) to the nearest 0.1 mg.

37. Each uterus will then be individually processed to open the uterine wall and carefully blot the excess fluid. For example, both uterine horns may be pierced or cut longitudinally, placed on moistened filter paper (e.g. Whatman No. 3) and gently pressed to absorb the luminal fluid. The procedure used must have good reproducibility within the laboratory and not be too severe to render the tissue unacceptable for histopathological analysis, as this additional investigation will be undertaken by some by some laboratories. A video is available from the lead laboratory, on request, demonstrating the blotting procedure to help improve the level of reproducibility between laboratories.

38. For those laboratories wishing to perform a histopathological examination of the vagina and/or uterus, the uterus and vagina should be fixed in 10% neutral buffered formalin (4% formaldehyde). If histopathology is done, the procedure used must be recorded and included in the study report. As it is known that tissue reactions differ in each portion of the uterus, **Figure 3** shows the points at which histological cross sections should be made.

## STATISTICAL ANALYSIS OF THE RESULTS

39. Each participating laboratory should record and provide the raw data with the items as listed below. A report of this data and an analysis of the results should be made to the lead laboratory. The lead laboratory will be responsible for making an overall assessment and presentation to the Validation Management Group. The raw data will include body weight, clinical status of animals during the test and before necropsy and uterine weight (wet and blotted). The Validation Management Group will determine the statistical procedures to be used in the evaluation of data taking into account dependent statistical advice.

## RETENTION OF RECORDS, SAMPLES AND SPECIMENS

40. A reference sample of EE and ZM will be retained by the chemical repository or chemical supplier until the end of the whole project if its nature allows this. Samples of diet, and test vehicle should be retained by the participating laboratories, so that further analyses can be carried out if needed, Participating laboratories should retain raw data, the master copy of the final report and all other information relevant to the quality and integrity of the study.

## REPORTING REQUIREMENTS

41. A final report will be prepared for each experiment conducted by each participating laboratory including details of:

**Laboratory Protocol**:

- Including date and approval

**Testing facility**:

- Address details
- Responsible personnel and their study responsibilities

**Test Substance**:

- Characterisation of ZM and EE (to be provided by chemical supplier/repository)
- Method and frequency of preparation of dilutions

**Vehicle**:

- Characterisation of test vehicle ( nature, supplier and lot)

**Test animals:**

- Strain
- Supplier and specific supplier facility
- Age of animal when ovariectomised
- Age of animal when administration of test substance began
- Details of acclimatisation procedure
- Number of animals per cage
- Detail and method of individual animal and group identification.

**Test Conditions:**

- Details of randomisation process (i.e. method used)
- Record of cage location in laboratory racks
- Diet (name, type, supplier, content)
- Water source (e.g. tap water or filtered water) and supply (by tubing from a large container, in bottles, etc.)
- Bedding
- Record of lighting interval
- Record of air conditioning (filter maintenance)
- Record of room clean up
- Description of blotting procedure details
- Details of histopathological procedures (including copy of standard operating procedures)

**Results**

**For individual animals:**

- Daily body weight - from the day the animals are allocated into groups to the day of necropsy
- Age of each animal (in days counting birth date as day 1) when administration of test compound begins
- Date and time of each s.c. injection
- Calculated amount of each s.c. injection
- Daily record of status of animal, including relevant symptoms and observations
- Suspected cause of death (if found during study in moribund state or dead)
- Date and time of humane killing
- Approximate time interval in hours between last test substance administration and humane killing
- Organ weight at necropsy
- Wet uterine weight per animal and any observations on loss of luminal fluid during dissection and preparation for weighing to the nearest 0.1 mg
- Blotted uterine weight per anima to the nearest 0.1 mg
- If undertaken, histopathogical report of uterus and vagina

**For each group of animals**

- Daily body weights (from day of allocation into groups to the day of necropsy)
- Uterine weights (both wet and dry) per dose given
- If measured, daily food consumption

**Figure 1: Schematic diagram showing points of ligation in surgically removing ovaries**

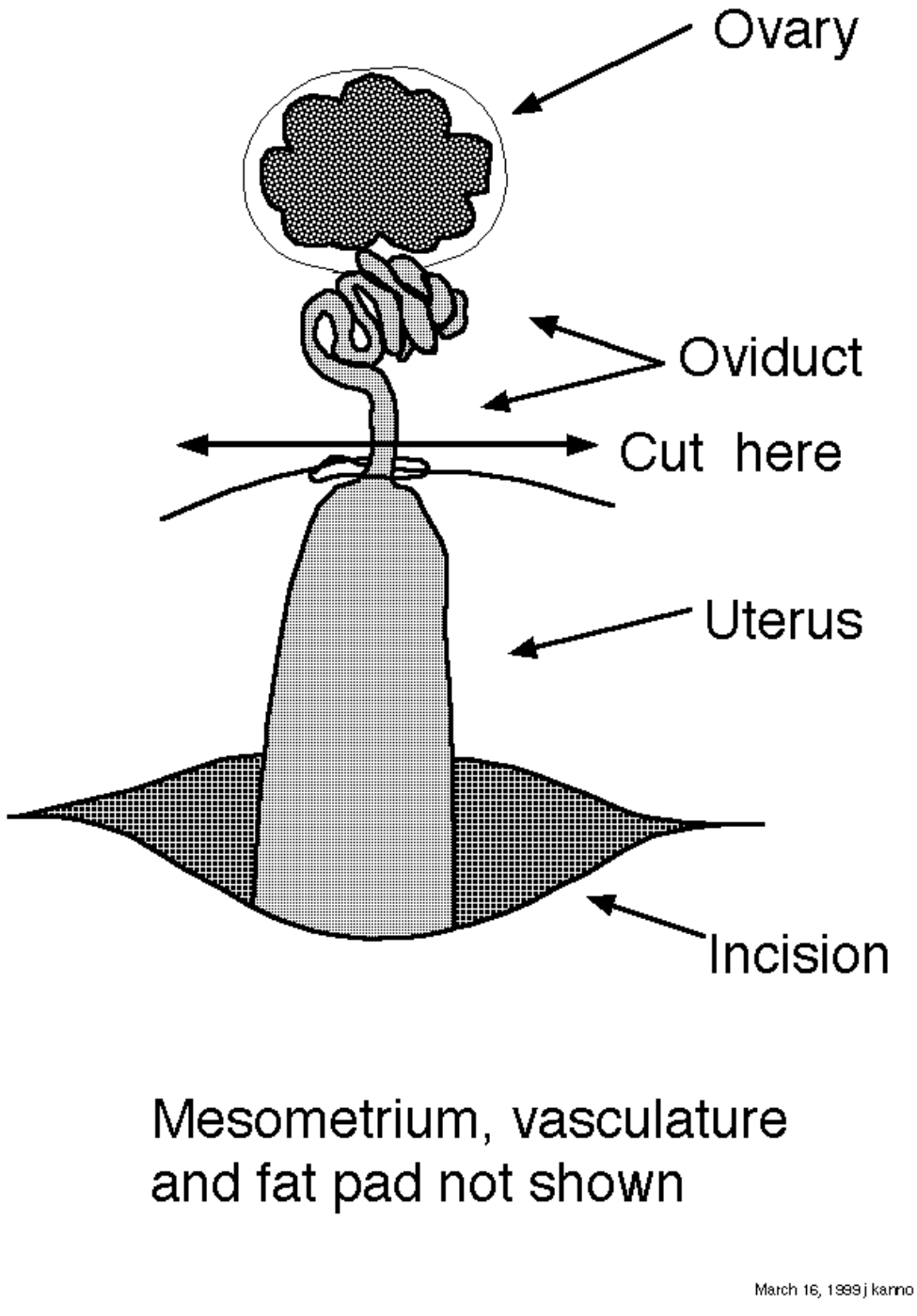
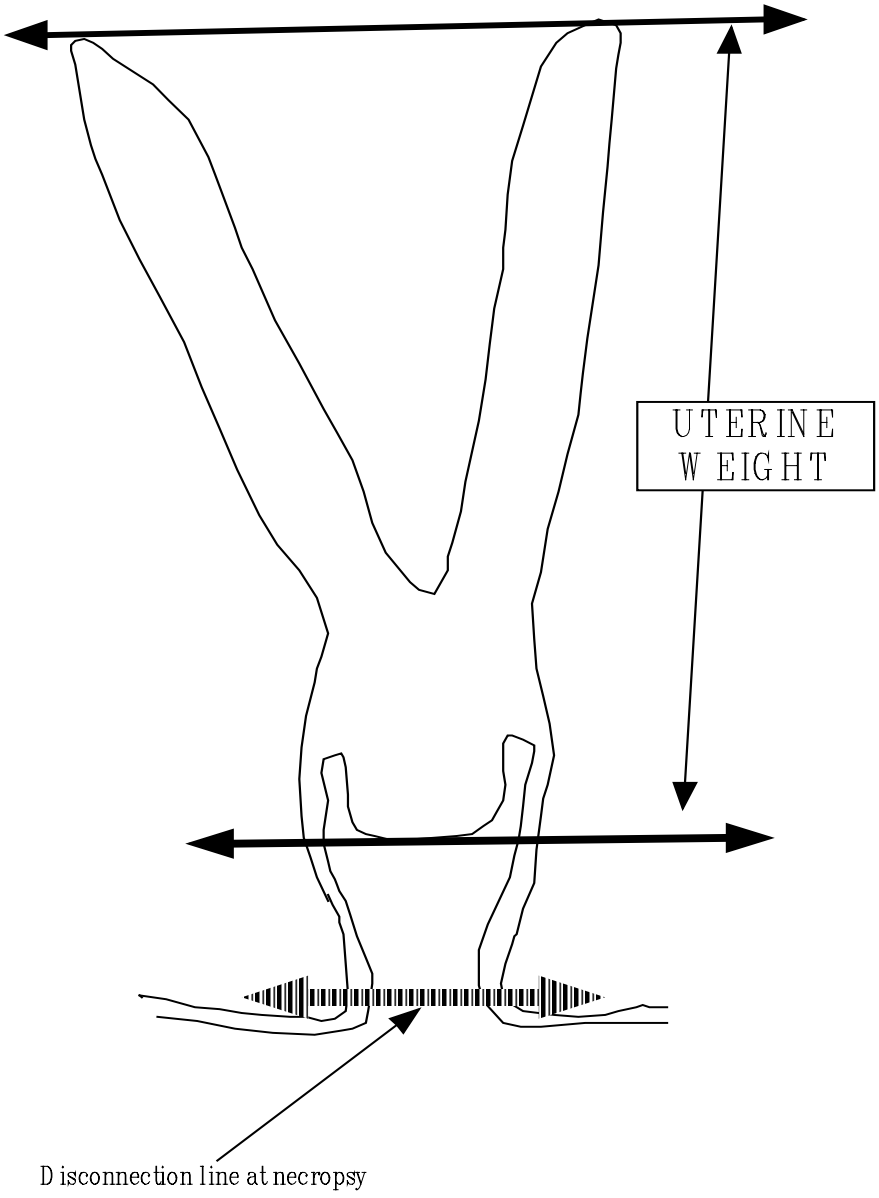


**Figure 2: The removal and preparation of the uterine tissues for weight measurement.**




In detail the procedure is to open the pubic symphysis. Then, each ovary and uterine horn is detached from the dorsal abdominal wall. Urinary bladder and ureters are removed from the ventral and lateral side of uterus and vagina. Fibrous adhesion between the rectum and the vagina is detached until the junction of vaginal orifice and perineal skin is identified. The uterus and vagina is detached from the body by incising the vaginal wall just above the junction between perineal skin as shown in the figure. The excess fat and connective tissue is trimmed away. The vagina is removed from the uterus as shown in the figure for uterus weight measurement. Weight with luminal fluid (wet weight) and without the luminal fluid (blotted weight) are measured.

**Figure 3: One example for the preparation of uterus and vagina for optional histopathological examination**.