The OECD Program to Validate the Rat Uterotrophic Bioassay to Screen Compounds for *in Vivo* Estrogenic Responses: Phase 1

Jun Kanno,[1] Lesley Onyon,[2] Joseph Haseman,[3] Penelope Fenner-Crisp,[4] John Ashby,[5] and William Owens[6]

[1]National Institute of Health Sciences, Tokyo, Japan; [2]Environment, Health and Safety Division, OECD, Paris, France; [3]National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA; [4]U.S. Environmental Protection Agency, Washington, DC, USA; [5]Syngenta Central Toxicology Laboratory, Macclesfield, Cheshire, UK; [6]Procter & Gamble, Cincinnati, Ohio, USA

The Organisation for Economic Co-operation and Development has completed the first phase of an international validation program for the rodent uterotrophic bioassay. This uterotrophic bioassay is intended to identify the *in vivo* activity of compounds that are suspected agonists or antagonists of estrogen. This information could, for example, be used to help prioritize positive compounds for further testing. Using draft protocols, we tested and compared two model systems, the immature female rat and the adult ovariectomized rat. Data from 19 participating laboratories using a high-potency reference agonist, ethinyl estradiol (EE), and an antagonist, ZM 189,154, indicate no substantive performance differences between models. All laboratories and all protocols successfully detected increases in uterine weights using EE in phase 1. These significant uterine weight increases were achieved under a variety of experimental conditions (e.g., strain, diet, housing protocol, bedding, vehicle). For each protocol, there was generally good agreement among laboratories with regard to the actual EE doses both in producing the first significant increase in uterine weights and achieving the maximum uterine response. Furthermore, the Hill equation appears to model the dose response satisfactorily and indicates general agreement based on calculated effective dose (ED)$_{10}$ and ED$_{50}$ within and among laboratories. The feasibility of an antagonist assay was also successfully demonstrated. Therefore, both models appear robust, reproducible, and transferable across laboratories for high-potency estrogen agonists such as EE. For the next phase of the OECD validation program, both models will be tested against a battery of weak, partial estrogen agonists. *Key words*: endocrine disruption, estrogen, rat uterus, uterotrophic. *Environ Health Perspect* 109:785–794 (2001). [Online 3 August 2001]
http://ehpnet1.niehs.nih.gov/docs/2001/109p785-794kanno/abstract.html

Concern has been raised that ambient environmental levels of chemicals called environmental estrogens may be causing adverse effects in both humans and wildlife through 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 have noted several cases where local, high-level exposures have produced effects in wildlife (*2–4*).

To address this concern, the Organisation for Economic Co-operation and Development (OECD) initiated a high-priority activity in 1997 to *a*) provide information on testing and assessment activities, particularly at the national regulatory level, and coordinate these activities among member countries as appropriate; *b*) revise existing guidelines and develop new guidelines for screening and testing potential endocrine disrupters; and *c*) harmonize hazard and risk assessment approaches internationally (*5*). The advantage of the OECD activity is that it would produce a set of internationally recognized and harmonized screening and testing guidelines and strategies that would avoid duplication of testing resources, including animals.

The OECD activity is managed by the Task Force on Endocrine Disrupters Testing and Assessment (EDTA), the membership of which includes experts nominated by OECD member countries' regulatory authorities, international organizations, nongovernmental organizations, and industry associations. The activity is part of the OECD Test Guidelines Programme, so overall responsibility of the work lies with the Working Group of National Co-ordinators of the Test Guidelines Programme (WNT).

The OECD conceptual framework identifies short- and long-term assays of increasing complexity and detail to gather information on a chemical. The assays include *a*) structural activity relationships and *in vitro* assays that would identify a chemical based on certain intrinsic characteristics (e.g., estrogen receptor binding affinity); *b*) short-term *in vivo* assays to demonstrate relevant activity in the intact animal (e.g., the uterotrophic assay); and *c*) 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). The OECD strategy aims to develop these assays as multipurpose tools rather than as a rigid scheme. The purpose and use of a bioassay could vary depending on the chemical substance and the available toxicological data on that chemical. An early screen in one case could become a means to determine a chemical's mode of action in another (*5*).

In this article we focus on the OECD validation program for an *in vivo* screen for estrogenic activity. Historically, several candidate systems are available: a vaginal cornification and keratinization response (*6*), a water imbibition response of the uterus after a single dose of the test compound (*7*), and a uterine tissue weight increase after several doses of the test compound (*8–10*). The EDTA reached consensus to select the latter assay, called the uterotrophic assay, for further development and validation. The uterotrophic response has been employed to evaluate estrogenic activity using a number of mammalian and avian species, although primarily laboratory rodents. Because the rat has become the preferred species for reproductive and developmental toxicity testing,

we chose it as the test species for further standardization and development of the uterotrophic assay.

Two possible uterotrophic models are based on the need to have a nonfunctional hypothalamic-pituitary-gonadal axis to ensure a sensitive and consistent uterine response both to administered estrogens alone and to administered antiestrogens in combination with a reference estrogen. One model uses the immature female before significant ovarian estrogen synthesis and regulation by the hypothalamic-pituitary-gonadal axis begins; the other model uses the ovariectomized (OVX) adult female, removing the primary source of estrogen synthesis. An extensive comparison of these models across several laboratories has never been performed. However, data in the literature and available data from laboratories participating in the OECD program suggest that the two models may be equivalent.

The objective of the OECD work on the uterotrophic assay is to develop a new, validated test guideline and clearly define its purpose. OECD member countries formally agreed in a workshop held in Solna, Sweden, in 1996 on the principles of validation and criteria for the acceptability of new and revised test guidelines (animal or nonanimal). These are now commonly known as the "Solna principles and criteria" and follow extensive work by national and regional authorities on the validation and acceptability of alternative test methods, including definitions of key terms (*11*). Validation is defined as the process by which the reliability and the relevance of a procedure are established for a particular purpose. Reliability is defined as the reproducibility of results from an assay within and between laboratories. Relevance describes whether a test is meaningful and useful for a particular purpose. The Solna principles and criteria were originally developed by the European Centre for the Validation of Alternative Methods (ECVAM) and the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and can be summarized as follows: *a*) the test method rationale should be stated, including the scientific need and regulatory purpose; *b*) the relationship of the end point(s) determined by the test method to the *in vivo* biologic effect and to the toxicity of interest must be stated; *c*) the limitations of a method must be described (e.g., metabolic capability); *d*) a detailed protocol must be readily available with sufficient detail to enable the user to adhere to it, including data analysis and decision criteria; *e*) test methods and results should be publicly available and should have been subjected to independent scientific review; *f*) intratest variability, repeatability,

and reproducibility of the test method within and among laboratories should have been demonstrated, including a description of variability with time; *g*) the test method's performance must have been demonstrated using a series of reference chemicals, preferably coded to exclude bias; *h*) the performance of test methods should have been evaluated in relation to existing relevant toxicity data; *i*) all data supporting the assessment of the validity of the test methods including the full data set collected in the validation study should have undergone scientific review; and *j*) these data should have been obtained in accordance with the OECD Principles of Good Laboratory Practice (GLP) (*11*). In 1998, OECD member countries reconfirmed their commitments to these validation principles and criteria and further clarified that the process of validation should allow for flexibility. However, to be able to justify a certain flexibility, a transparent standard procedure should be available that allows for the assessment of the need for and extent of the required level of validation (*12*).

A Validation Management Group on behalf of the EDTA of the OECD Test Guidelines Programme coordinates the work on human health test methods. The Validation Management Group is composed of experts from eight member countries nominated for their expertise in toxicology, test development and validation, endocrinology, regulatory toxicology, and biostatistics. Experts from ICCVAM and from ECVAM also participate in the Validation Management Group.

Overall Program Design and Objectives

The work on the uterotrophic assay is being performed in phases. Phase 1, now completed, was designed to test, refine, and standardize the immature and the adult OVX rat uterotrophic assays using a high-potency reference agonist compound and to provide data on intra- and interlaboratory variability with this reference compound. In addition, the feasibility of using the protocol for antagonist assays was explored using a reference antagonist. A detailed report of phase 1 of the validation work on the uterotrophic assay, including the rationale for the design of this phase, was submitted to the VMG for final approval in March 2001 (*13*). At this time, further progress on the antagonist portion of the assay awaits synthesis of sufficient quantities of the reference pure antiestrogen. Few pure antagonists such as ZM 189,154 are known (*14*). Most known estrogen antagonists such as tamoxifen will also express low levels of an agonist response and would complicate data interpretation by responding as a

positive in both agonist and antagonist sections of the assay (*15*).

Phase 2, currently underway, is designed to demonstrate the capability of both standardized protocols against a set of test compounds comprising weak estrogen partial agonists and a known negative. Phase 2 is intended to demonstrate the repeatability and variation within and among laboratories for several compounds and over time. To properly investigate intra- and interlaboratory variability, the doses to be used in phase 2 will be specified in all cases. Twenty laboratories are participating in phase 2.

The need for additional work after phase 2 will depend on the outcome of phase 2. The results of the uterotrophic assays along with other relevant biologic and toxicologic data that may exist on the chemicals of interest will be evaluated to demonstrate the reliability and relevance of the uterotrophic screen for its intended use in detecting estrogen agonists *in vivo*.

Design of phase 1. The objectives of the first phase of the OECD validation work were to *a*) demonstrate, in immature and adult OVX rats, the dose–response relationship between uterine weight and a reference estrogen using two possible routes of administration—oral gavage and subcutaneous injection; *b*) investigate intra- and interlaboratory variation and identify any appropriate protocol refinements; *c*) compare the performance of the protocols; and *d*) demonstrate the feasibility of the protocols to identify potential antiestrogenic activity using a pure estrogen antagonist.

Currently, several protocols are in use for the uterotrophic assay. Three principal variables govern their differences: species, age of test animal, and route of administration. The result is eight possible test protocols, each having literature precedents for their use. The literature was reviewed and decisions made on the basis of scientific rationale and practical experience, given the intent to develop an OECD test guideline that could be transferred easily to many laboratories. The rat was chosen over the mouse because it is used most often as the preferred rodent model in toxicology testing paradigms for regulatory purposes. All protocols developed had essentially the same design and differed only in the model used and route of administration. The protocols using oral gavage in immature animals and subcutaneous injection in ovariectomized animals each had large databases of available historical information. The third protocol, subcutaneous injection of immature rat, also had a large database of historical information and was chosen as a link between the other two. A fourth protocol, extending the duration of exposure in ovariectomized animals, was carried out by

some laboratories to explore whether varying this parameter had any effect on the sensitivity of the assay.

The design of phase 1 then consisted of testing four protocols: Protocol A used the immature female rat model, with administration of doses by oral gavage for 3 days at 24-hr intervals followed by humane killing 24 hr after the last administration. Protocol B also used the immature female rat model, with dosing by subcutaneous injection for 3 days at 24-hr intervals followed by humane killing 24 hr after the last administration. Protocol C used the adult OVX rat model, with administration by subcutaneous injection for 3 days at 24-hr intervals followed by humane killing 24 hr after the last administration. Protocol C´ also used the adult OVX model and extended the subcutaneous dosing to 7 days with humane killing 24 hr after the last administration.

The reference estrogen agonist was 17α-ethinyl estradiol (EE; CAS no. 57-63-6), and the reference estrogen antagonist was the pure antagonist ZM 189,154 (ZM; CAS no. 101908-22-9). The same lot of each chemical was distributed from a central repository. These chemicals were gifts of Schering (Kenilworth, NJ, USA) and Astra-Zeneca (Alderley Park, Cheshire, UK), respectively.

The lead laboratory was the National Institute of Health Sciences of Japan. Nineteen laboratories from Denmark, France, Germany, Japan, Korea, the Netherlands, the United Kingdom, and the United States participated in phase 1. Sixteen laboratories from seven nations performed protocol A, 12 laboratories from six nations performed protocol B, nine laboratories from three nations performed protocol C, and four laboratories from one nation performed protocol C´.

Because the uterotrophic assay is intended to be widely practiced, participating laboratories used their traditional rat strain, diet, vehicle, and housing procedures. Animals were to be acquired from standard animal supply sources with general instructions on acclimation and housing (e.g., immature animals transported with litters together accompanied by the dam or a foster dam, or scheduled to arrive as a litter when they are 17 days old; room temperature of 22 ± 3°C and a relative humidity 30–70%; artificial lighting with a 12-hr light and 12-hr dark cycle; feed and tap or filtered drinking water provided *ad libitum*). Each laboratory recorded the specifics, and samples of vehicle and diet were retained. Individual animals were uniquely identified (e.g., by ear tags or tail tattoos), and each group was coded (e.g., by a letter and a color on housing cages). Both an untreated control and a vehicle control were included to allow detection of any significant contamination of the vehicle with phytoestrogen(s). There have been reports both in the older literature and more recently that particular lots of diet, presumably through the presence of phytoestrogens, could influence the baseline uterine weight (*16–20*). If significant variations in the control and vehicle control uterine weights were observed, the contributions of strain, diet, and so on could then be investigated further, if necessary, as retained samples of diet and vehicle were required. Details of these particulars are provided in Table 1.

All protocols were based on a group size of six animals. The total amount for subcutaneous injection per rat per day did not exceed 4 mL/kg, and the total amount for oral gavage per rat per day did not exceed 5 mL/kg. Included were daily measurement of

animal body weights and adjustment of volumes to maintain the specified dose of substance for the allotted period. Body weights starting on the day of administration ranged from 26 to 57 g across laboratories for the immature animals and from 142 to 327 g for the adult OVX animals.

The end points of interest were the wet and blotted uterine weights. The uterine weight increase is a fundamental response of the female to sufficient exposure to estrogen agonists. The response begins with the essential interaction of the estrogen with a high-affinity receptor in uterine tissues that initiates a series of responses culminating in the uterine weight increase. The weight increase is a combination of water imbibition in the tissue and the uterine lumina and a hypertrophic response of the uterine tissues. The estrous cycle in the rat is 4–5 days, so the 3-day administration of a test compound is similar to the response time to endogenous estrogen surges in the intact animal that stimulate the uterine tissue. Thus, estrogen agonists can be identified by a statistically significant increase in uterine weight in treated versus untreated or vehicle control animals. In addition, estrogen antagonists can be identified by blocking or reducing the uterine weight increase of a reference agonist when both are simultaneously administered. The estrogen specificity of the uterine weight increase or decrease can be verified, if necessary, by histologic examination of the uterus and the vagina (*18,21*).

Historically, most published uterotrophic results have described uterine weights after careful blotting of the uterus after its wall was nicked or split to allow the luminal contents to drain out. The rationale given for measuring blotted uterine weight usually is that the wet weights are more variable, and the variability is increased by the possible loss of luminal

Table 1. Rat strains, vehicle, animal diets, cage bedding, and housing used by participating laboratories.

Laboratory	Rat strain	Vehicle used	Animal diet	Cage bedding	No. per cage
1	SD Crj CD(SD) IGS	Corn oil	CRF-1 PD, Oriental Yeast Co., Ltd., Japan	None	3
2	SD Crj:CD(SD) IGS	Corn oil	CRF-1 PD, Oriental Yeast Co., Ltd., Japan	None	2
3	SD Crj:CD(SD) IGS	Sesame oil	MF PD, Oriental Yeast Co., Ltd., Japan	Autoclaved hardwood	3
4	Wistar Rj WI (SpF) Han	Olive oil	GKD, Provimi Kliba SA, SW	SNIFF (type 3/4)	3
5	SD CD(SD)IG SBR	Corn oil	PMI CRD, W.F. Fisher & Son, USA	None	2
6	SD CD(SD)GS BR	Corn oil	AO4C PD, Usine d'Alimentation Rationelle, France	Autoclaved sawdust	2
7	SD Crj:CD(SD) IGS	Sesame oil	CE2, CLEA, Japan	None	3
8	SD Alpk ApfSD	Peanut oil	SDS RM1, Special Diet Services Ltd., UK	Shredded paper	3
9	SD CD(SD) IGS	Olive oil	MF PD, Oriental Yeast Co., Ltd., Japan	None	1
10	SD CRJ CD	Corn oil	Cheil CRC, Cheil Feed Co., Korea	None	2
11	Jcl:Wistar	Corn oil	MF PD, Oriental Yeast Co., Ltd., Japan	None	3
12	SD Crl CD(SD) IGS BR	Corn oil	PMI CRD 5002, Cincinnati Lab Supply, USA	Ground corn cobs	3
13	Wistar HSB/CpbWU	Corn oil	Altromin 1324, Altromin GmbH, Germany	Low-dust wood granules	3
14	SD ICO:OFA SD(IOPS) caw	Corn oil	AO4C CRC, Usine d'Alimentation Rationelle, France	None	2
15	Wistar Crl (Wi) WU BR	Corn oil	SDS RM3, Special Diet Services Ltd., UK	None	6
16	Wistar HsdCpd:Wu	Peanut oil	Altromin 1324 FORTI, Altromin GmbH, Germany	Wood chip	3
17	Wistar (mol:Wistar/Han)	Peanut oil	Altromin 1324, Brogaarden, Denmark	Tapvei	3
18	SD CRI CD	Corn oil	PM1 CRD, PMI Nutrition International, USA	Autoclaved elm tree	3
19	SD Hsd	Corn oil	SDS RM1 (E) SQC, Special Diet Services Ltd., UK	None	3

Abbreviations: CRC, Certified Rodent Chow; CRD, Certified Rodent Diet; CRF, Certified Rodent Formula; GKD, Ground Klieba Diet; MF, Maintenance Formula; PD, pelleted diet; PMI, Purina Mills, Inc.; RM, rat and mouse; SD, Sprague-Dawley; SDS, special diet services.

fluid during dissection and tissue handling. For test optimization and validation, it was decided to include both wet and blotted uterine weights and to establish their variability in the models among different laboratories using standardized procedures. 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. Because several laboratories were performing the assay for the first time, and to standardize procedures, a videotape of procedures for ovariectomy and uterine dissection and preparation was prepared and distributed to the participating laboratories.

Precaution was taken to specify the age of the animals so that treatment could commence at 19–20 days of age (day of birth counted as day 1) and to limit body weight variability. Limiting the weight variability was thought essential to limit the chances of older animals being inadvertently included in the study. Older animals could enter puberty, leading to an increase in control uterine weight and thereby adding to the variability of the results (*19,22*). For the adult OVX animals, ovariectomy occurred at 6 weeks of age or later, with a minimum period of 1 week after surgery before administration of the reference compounds. In all protocols, groups were randomized according to body weight.

The doses of EE and ZM 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

Table 2. Number of laboratories observing a lowest observed effect level (LOEL) at a given EE dosing based on the first observed significant (*p* < 0.05) increase in uterine weight.

Protocol	Dose of EE (µg/kg/day)				
	0.03	0.1	0.3	1.0	3.0
Wet weight					
A	0	0	4	11	1
B	1	3	8	0	0
C	0	2	6	1	0
C´	0	3	1	0	0
Blotted weight					
A	0	0	4	12	0
B	2	3	7	0	0
C	0	2	6	1	0
C´	0	3	1	0	0

Table 3. Uterine and body weight data of the five highest EE doses (µg/kg/day) for protocol A.

Laboratory, measurement	Vehicle		0.1		0.3		1.0		3.0		10.0	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1 Wet uterine weight (mg)	24.52	5.53	24.43	1.35	30.33	7.69	54.37*	14.25	113.62*	29.26	179.98*	40.94
Blotted uterine weight (mg)	23.28	5.58	23.25	1.08	29.12	7.71	52.65*	13.96	93.05*	12.82	116.40*	16.66
Body weight (g)	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 Wet uterine weight (mg)	32.02	2.77	31.22	1.92	42.52*	9.49	78.07*	24.25	153.25*	40.67	273.47*	46.83
Blotted uterine weight (mg)	31.45	3.00	30.57	1.89	41.23*	9.60	74.70*	21.81	117.18*	19.61	142.35*	13.11
Body weight (g)	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 Wet uterine weight (mg)	34.28	7.34	34.53	2.40	36.63	3.29	63.67*	9.61	137.17*	27.70	194.32*	49.75
Blotted uterine weight (mg)	32.50	6.75	32.92	2.44	34.38	3.12	60.28*	9.27	108.85*	13.60	125.68*	14.05
Body weight (g)	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 Wet uterine weight (mg)	24.33	1.21	27.83	4.22	23.00	2.45	29.50*	3.62	112.33*	19.13	157.33*	38.96
Blotted uterine weight (mg)	23.83	1.60	25.83	3.06	22.17	2.32	28.33*	3.08	92.50*	7.42	108.67*	13.22
Body weight (g)	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 Wet uterine weight (mg)	41.22	9.40	44.36	6.45	51.70	11.80	57.65*	9.59	132.20*	34.76	201.72*	5.51
Blotted uterine weight (mg)	38.78	8.06	41.30	6.08	47.12	12.23	55.75*	9.33	113.85*	23.58	139.10*	9.72
Body weight (g)	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 Wet uterine weight (mg)	42.25	8.04	42.51	8.45	48.04	15.78	64.90*	5.01	108.66*	19.58	179.38*	32.27
Blotted uterine weight (mg)	41.22	7.93	41.16	7.91	46.39	15.07	63.45*	5.34	100.78*	12.33	141.04*	17.69
Body weight (g)	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 Wet uterine weight (mg)	30.08	2.72	30.37	3.75	30.23	1.20	66.52*	20.70	104.82*	12.29	167.32*	28.54
Blotted uterine weight (mg)	28.58	3.11	29.57	3.66	28.80	1.95	63.77*	19.88	95.67*	7.86	124.60*	14.66
Body weight (g)	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 Wet uterine weight (mg)	32.68	8.02	31.43	6.95	38.28	10.70	65.70*	25.44	110.20*	26.57	162.58*	42.00
Blotted uterine weight (mg)	30.27	8.59	29.42	6.55	35.48	9.98	59.95*	21.08	94.80*	19.96	117.12*	15.46
Body weight (g)	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 Wet uterine weight (mg)	35.22	4.24	35.93	2.16	36.20	4.64	55.88*	10.15	120.58*	22.99	202.18*	30.11
Blotted uterine weight (mg)	34.28	4.12	35.02	2.67	35.33	4.90	54.75*	10.20	102.53*	12.57	133.48*	12.97
Body weight (g)	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 Wet uterine weight (mg)	44.70	13.01	58.17	12.02	57.43*	4.40	77.87*	8.29	148.53*	36.30	195.72*	23.42
Blotted uterine weight (mg)	42.27	12.72	55.72	12.25	55.27*	4.31	74.92*	7.96	119.83*	27.52	120.88*	16.47
Body weight (g)	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 Wet uterine weight (mg)	36.85	3.21	36.60	3.73	55.60*	6.17	97.93*	5.91	204.47*	31.87	190.55*	42.60
Blotted uterine weight (mg)	33.47	2.74	33.50	3.57	52.12*	6.33	86.87*	3.44	129.10*	6.28	127.90*	8.93
Body weight (g)	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 Wet uterine weight (mg)	31.65	11.97	41.00	13.61	37.23	6.57	57.70*	13.04	117.25*	19.37	174.73*	28.47
Blotted uterine weight (mg)	29.27	9.88	38.67	13.79	28.78	8.74	53.63*	9.36	95.33*	5.04	119.47*	9.47
Body weight (g)	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 Wet uterine weight (mg)	37.33	5.68	37.33	5.09	34.33	5.24	44.83	4.36	83.83*	9.30	241.17*	31.54
Blotted uterine weight (mg)	34.83	5.78	36.00	4.29	32.83	5.60	42.33*	3.98	76.17*	4.17	134.33*	12.24
Body weight (g)	42.60	3.66	40.13	1.76	39.48	3.10	37.97	4.09	38.73	2.60	41.42	4.65
14 Wet uterine weight (mg)	21.23	2.17	21.53	3.59	36.23*	6.74	67.55*	14.69	109.92*	31.44	119.73*	48.74
Blotted uterine weight (mg)	17.63	2.32	18.65	3.91	32.35*	6.92	54.62*	5.59	69.20*	10.95	78.07*	9.51
Body weight (g)	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 Wet uterine weight (mg)	32.00	4.74	36.83	7.39	35.17	6.97	77.00*	15.17	173.00*	38.33	202.67*	59.35
Blotted uterine weight (mg)	25.80	5.50	26.00	7.04	29.33	4.41	65.50*	12.82	108.00*	7.48	126.00*	13.61
Body weight (g)	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 Wet uterine weight (mg)	35.73	10.14	29.03	5.45	44.18	19.93	93.83*	24.71	213.73*	34.79	233.55*	60.90
Blotted uterine weight (mg)	32.63	9.47	26.20	4.74	40.67	18.57	80.47*	18.34	131.58*	5.62	118.07*	18.67
Body weight (g)	33.02	1.23	34.95	3.38	35.02	4.01	33.05	2.82	33.17	2.28	37.20*	2.07

*p < 0.05 versus vehicle.

µg/kg/day was specified for both oral gavage and subcutaneous administration. 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 coadministered. EE and ZM 189,154 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 ZM189,154, gentle heating up to 60°C was needed for dissolution. Other protocol details are omitted here for brevity.

Participating laboratories submitted the raw data for central, independent statistical analysis. The ability to detect increased uterine weights was evaluated by an analysis of variance (ANOVA) approach, including body weight as a covariable. A variance-stabilizing logarithmic transformation was performed before the data analysis. Dunnett's test was used for making pairwise comparisons of each dosed group to vehicle controls. Dixon's outlier test was used to detect possible outliers, and Bartlett's test was used to assess homogeneity of variances. If significant heterogeneity was detected, the nonparametric Mann-Whitney U-test was used. For these data, parametric and nonparametric analyses produced similar results.

Results

All participating laboratories confirmed that the protocol was straightforward to perform. Suggested protocol refinements included additional guidance to reduce organ weight variation such as that caused by different prosectors; improved procedures for controlling body weight (in immature animals) and increasing the immature age for administration, because some laboratories encountered weight loss in the early weanlings.

All laboratories and all protocols were successful in detecting increases in uterine weights in the higher dosed EE groups. Within each protocol, there was good agreement among laboratories in the dose–response uterine weight increases for the reference EE. This included the EE doses identified as lowest observed effect levels (LOELs)—that is, the doses at which significant increases in uterine weight were first detected. The number of laboratories that observed a LOEL at a given dose for each protocol are summarized in Table 2. Blotted

weights showed statistical significance at slightly lower EE concentrations than did the wet uterine weights. For protocol A, significance was generally first achieved at 1.0 µg/kg EE. The data for the wet and blotted uterine weights as well as body weights for the five highest dose groups are shown for Protocols A, B, C, and C´ in Tables 3, 4, 5, and 6, respectively. For protocols B and C, significance was generally first achieved at the next lower dose of 0.3 µg/kg EE. This difference was expected the different route of administration and previously published data for EE (18). Three of the four laboratories carrying out protocol C´ first found significant increases in uterine weight at the 0.1 µg/kg EE dose. In the higher dose groups, wet uterine weights were reduced substantially in protocol C´ relative to protocol C, whereas the reverse tended to be true for blotted weights. The reduced wet weights in protocol C´ were apparently caused by the reduction in luminal fluid content between days 3 and 7.

The consistency of dose–response results among laboratories was also evaluated: that is, did laboratories consistently produce a dose–response curve of approximately the

Table 4. Uterine and body weight data of the five highest EE doses (µg/kg/day) for protocol B.

Laboratory, measurement	Vehicle Mean	SD	0.1 Mean	SD	0.3 Mean	SD	1.0 Mean	SD	3.0 Mean	SD	10.0 Mean	SD
1 Wet uterine weight (mg)	29.28	3.17	31.03	4.86	71.95*	15.88	136.97*	18.14	255.22*	59.27	254.13*	39.02
Blotted uterine weight (mg)	28.27	3.03	29.85	4.75	69.22*	14.02	112.73*	10.93	140.33*	16.15	128.28*	4.56
Body weight (g)	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 Wet uterine weight (mg)	27.92	2.07	31.07	4.50	81.77*	15.97	182.00*	20.26	243.45*	47.28	293.08*	75.86
Blotted uterine weight (mg)	27.30	2.11	30.62	4.49	80.70*	15.93	131.77*	10.98	155.57*	11.64	162.68*	11.19
Body weight (g)	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 Wet uterine weight (mg)	31.98	3.30	35.12	4.19	79.28*	14.38	155.33*	22.93	267.58*	37.50	277.45*	48.56
Blotted uterine weight (mg)	30.72	3.45	32.82	4.02	75.37*	13.73	126.95*	15.08	149.67*	18.81	156.95*	7.98
Body weight (g)	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 Wet uterine weight (mg)	27.52	0.95	29.18	2.56	54.73*	18.31	157.57*	23.78	205.92*	27.19	220.48*	58.45
Blotted uterine weight (mg)	26.05	1.09	28.02	2.57	53.38*	17.97	125.95*	17.22	146.33*	7.28	159.30*	51.50a
Body weight (g)	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 Wet uterine weight (mg)	29.83	2.64	37.87#	7.20	64.57*	3.78	140.70*	24.61	175.58*	51.89	203.13*	53.79
Blotted uterine weight (mg)	27.20	2.79	34.62*	6.44	60.87*	3.69	112.50*	12.74	119.57*	15.91	125.05*	16.63
Body weight (g)	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 Wet uterine weight (mg)	32.62	3.97	34.03	3.53	56.68*	14.82	177.98*	43.27	259.13*	42.19	252.95*	37.30
Blotted uterine weight (mg)	31.68	3.94	33.15	3.24	55.68*	14.63	131.67*	20.53	150.05*	18.88	146.67*	12.65
Body weight (g)	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 Wet uterine weight (mg)	39.20	5.88	96.72*	28.89	137.25*	32.79	226.13*	75.76	272.18*	23.16	256.43*	81.95
Blotted uterine weight (mg)	35.77	6.10	84.30*	17.90	113.67*	26.62	136.60*	19.85	141.83*	20.22	118.07*	21.04
Body weight (g)	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 Wet uterine weight (mg)	33.97	4.12	49.63*	8.87	91.15*	12.73	202.78*	50.74	278.76*	46.76	307.53*	46.78
Blotted uterine weight (mg)	29.67	4.36	46.53*	8.87	80.18*	8.80	130.85*	14.75	143.10*	12.86	141.27*	12.81
Body weight (g)	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 Wet uterine weight (mg)	34.82	3.93	37.80	9.08	64.07*	21.85	118.38*	51.49	186.40*	70.98	249.40*	71.37
Blotted uterine weight (mg)	29.67	4.24	33.60	7.51	58.78*	14.70	96.63*	29.29	109.73*	18.98	133.95*	9.13
Body weight (g)	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 Wet uterine weight (mg)	36.50	7.84	37.67	7.03	78.50*	6.92	181.50*	55.88	238.67*	52.59	304.00*	75.92
Blotted uterine weight (mg)	25.83	7.19	27.33	6.02	60.17*	7.03	119.33*	17.27	138.33*	18.00	155.50*	9.16
Body weight (g)	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 Wet uterine weight (mg)	33.15	4.64	38.53	5.35	108.52*	29.15	240.32*	66.39	284.13*	42.98	242.93*	24.75
Blotted uterine weight (mg)	30.32	4.53	34.68	5.22	91.52*	17.00	132.98*	22.68	142.53*	14.40	130.60*	11.86
Body weight (g)	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 Wet uterine weight (mg)	21.72	2.89	28.13*	4.17	44.08*	6.18	142.92*	18.01	201.70*	24.29	257.75*	50.41
Blotted uterine weight (mg)	16.12	4.05	21.48*	5.13	33.70*	7.30	94.73*	7.82	106.90*	13.77	110.65*	11.24
Body weight (g)	38.32	4.35	39.47	2.29	36.87	2.66	40.35	2.17	38.10	3.60	41.50	3.46

aDenotes the presence of an outlier in the data. *$p < 0.05$ versus vehicle. #Significant heterogeneity among groups; significant ($p < 0.05$) versus vehicle controls by a Mann-Whitney U-test.

same shape where the same percentage increase in uterine weight, including the maximal increase, occurred at equivalent doses of the test compound? In protocol A, 8 of the 16 laboratories produced blotted uterine weight responses that were statistically consistent at all doses evaluated. In protocol B, 5 of the 12 laboratories produced blotted uterine weight responses that were statistically consistent at all doses evaluated. In protocol C, six of the nine laboratories produced blotted uterine weight responses that were statistically consistent. In protocol C´, all four laboratories produced blotted uterine weights that were statistically consistent (after deleting one outlier). Dose–response results are shown for all protocols in Figure 1.

The sensitivity of an assay can be defined in several ways. One approach is to identify the lowest dose at which statistical significance

is achieved (see Table 2). Protocol A appeared less sensitive, as expected with the oral route of administration for EE. The data suggest no notable differences between protocols B, C, or C´ in the dose first producing statistical significance. Direct comparisons of performance should be based on data from the same set of laboratories performing both protocols. For example, eight laboratories carried out both protocols B and C. Seven laboratories achieved statistically significant increases at identical doses for the wet uterine weight, and six achieved statistically significant increases in blotted uterine weight.

At the higher EE doses, there was a significant difference between the models in the magnitude of the percentage uterine weight increase over controls. For the 12 laboratories carrying out protocol B, the range of

increased blotted uterine weights over controls was 326–588% for the 1.0 µg/kg EE dose and 370–663% for the 3.0 µg/kg EE dose. For the nine laboratories carrying out protocol C, the range for the increase in blotted weights was 136–375% for the 1.0 µg/kg EE dose and 236–375% for the 3.0 µg/kg EE dose. The responses for all protocols at 3 µg/kg EE dose are shown in Table 7. Protocol B, C, and C´ animals appeared to have reached stable maximal responses in the tested dose range. Protocol A animals did not appear to have reached their stable maximal response even at the 10 µg/kg EE dose relative to the 3 µg/kg dose.

Two procedures were performed to permit meaningful comparisons of variability in response. The first procedure was to log-transform the uterine weight data. This was followed by ANOVA for each laboratory

Table 5. Uterine and body weight data of the five highest EE doses (µg/kg/day) for protocol C.

		Vehicle		0.1		0.3		1.0		3.0		10.0	
Laboratory, measurement		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	Wet uterine weight (mg)	104.35	13.53	114.48	4.82	197.22*	35.69	685.00*	167.66	1052.42*	43.53	1227.00*	262.581
	Blotted uterine weight (mg)	102.35	12.85	112.22	4.42	190.45*	32.71	319.78*	57.47	373.72*	24.51	382.00*	35.21
	Body weight (g)	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	Wet uterine weight (mg)	125.12	19.47	128.75	14.29	225.18*	32.82	697.13*	89.87	886.05*	104.73	1197.35*	134.96
	Blotted uterine weight (mg)	120.82	18.43	123.47	12.71	217.48*	28.73	351.32*	27.51	384.72*	25.94	404.32*	44.63
	Body weight (g)	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	Wet uterine weight (mg)	121.58	6.49	151.13	55.68[a]	231.32*	21.61	656.50*	177.40	1001.52*	146.81	899.68*	322.24
	Blotted uterine weight (mg)	115.92	5.27	144.42	54.14[a]	213.95*	13.02	326.07*	60.22	378.37*	20.79	354.37*	54.84
	Body weight (g)	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	Wet uterine weight (mg)	123.28	12.15	133.80	12.01	225.77*	33.61	522.88*	319.22[a]	820.57*	175.01	906.77*	301.50
	Blotted uterine weight (mg)	121.62	12.29	131.25	12.32	220.83*	32.25	317.52*	63.15	387.43*	34.40	391.67*	40.89
	Body weight (g)	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	Wet uterine weight (mg)	83.73	10.42	110.17*	11.09	236.13*	25.52	406.43*	65.18	351.82*	53.27	391.27*	101.56
	Blotted uterine weight (mg)	79.22	10.36	105.08*	10.33	211.13*	13.31	287.68*	23.33	262.20*	25.77	273.73*	42.91
	Body weight (g)	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	Wet uterine weight (mg)	110.75	13.75	125.35	17.63	219.13*	30.01	717.53*	180.63	859.05*	164.29	866.62*	182.03
	Blotted uterine weight (mg)	108.47	13.23	123.60	16.68	211.37*	26.45	357.57*	47.52	353.82*	34.38	362.05*	41.95
	Body weight (g)	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	Wet uterine weight (mg)	86.00	19.07	118.42*	9.09	213.22*	12.32	613.82*	84.49	682.80*	86.85	714.18*	174.83
	Blotted uterine weight (mg)	82.45	15.71	113.38*	10.13	191.23*	10.59	297.67*	15.99	307.60*	35.50	312.40*	43.05
	Body weight (g)	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	Wet uterine weight (mg)	107.02	7.43	138.60	73.48[a]	241.33*	96.54	795.48*	225.16	930.65*	110.47	1104.27*	210.47
	Blotted uterine weight (mg)	89.25	10.33	91.80	11.71	193.07*	64.32	334.95*	54.84	334.48*	44.55	366.20*	32.62
	Body weight (g)	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	Wet uterine weight (mg)	104.17	10.76	85.33	8.78	108.33	13.47	140.50*	26.57	269.17*	78.70	588.33*	127.61
	Blotted uterine weight (mg)	99.17	10.17	83.17	7.70	104.67	12.09	135.17*	24.19	234.17*	51.43	332.67*	27.34
	Body weight (g)	290.63	12.84	298.60	11.89	301.75	7.84	287.18	27.35[a]	294.57	9.68	282.43	9.62

[a]Denotes the presence of an outlier in the data. *$p < 0.05$ versus vehicle.

Table 6. Uterine and body weight data of the five highest EE doses (µg/kg/day) for protocol C´.

		Vehicle		0.1		0.3		1.0		3.0		10.0	
Laboratory, measurement		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	Wet uterine weight (mg)	102.88	11.68	111.28	15.66	223.83*	35.02	395.62*	58.52	443.22*	95.81	685.95*	154.71
	Blotted uterine weight (mg)	98.50	12.82	108.10	4.42	215.58*	35.90	347.95*	33.95	397.43*	48.61	422.18*	39.39
	Body weight (g)	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	Wet uterine weight (mg)	106.58	6.49	132.02*	15.93	281.23*	43.07	367.55*	38.37	390.43*	18.59	516.85*	122.81
	Blotted uterine weight (mg)	101.62	2.71	126.07*	15.46	267.58*	36.94	353.92*	35.94	376.05*	17.61	412.77*	38.11
	Body weight (g)	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	Wet uterine weight (mg)	103.62	15.39	190.40*	123.39[a]	267.23*	30.01	384.82*	61.81	412.50*	58.84	519.38*	45.55
	Blotted uterine weight (mg)	100.87	15.16	164.32*	68.25[a]	259.18*	29.10	368.30*	49.85	393.82*	51.62	429.95*	36.53
	Body weight (g)	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	Wet uterine weight (mg)	92.83	10.14	127.90*	19.77	229.27*	18.18	395.58*	36.30	394.58*	33.26	444.58*	27.18
	Blotted uterine weight (mg)	89.48	9.68	110.35*	20.69	217.85*	19.76	359.03*	31.43	368.37*	31.53	366.43*	15.03
	Body weight (g)	202.63	4.45	196.32	6.70	192.00*	7.37	185.63*	3.29	176.27*	10.05	168.87*	6.23

[a]Denotes the presence of an outlier in the data. *$p < 0.05$ versus vehicle.

and protocol, using body weight as a covariable. The error mean square resulting from this analysis can be regarded as a measure of intragroup variability, averaged over doses and corrected for the possible influence of body weight on the observed uterine weight response. The second procedure was to calculate the coefficient of variation (CV) in uterine weight for each dosed group for each laboratory within each protocol. The CVs were averaged over doses to obtain a representative value for each laboratory and protocol. Each procedure produced similar findings. The results for the second procedure are summarized in Table 8. These analyses revealed that *a*) within-group variability in response was consistently less for blotted weights than for wet weights; *b*) protocol A tended to show more within-group variability in both the wet and blotted measures of uterine weights, which was not unexpected given the oral route of administration; and *c*) the adult OVX subcutaneous protocols (C and C′) have slightly lower CVs than the immature animal subcutaneous protocol (B). Note from Table 8 that some laboratories have consistently different (lower or higher) CVs across all protocols. This suggests an important role

for laboratory technique in controlling variability in both the wet and blotted uterine weight response measurements.

Although body weights were controlled tightly within a laboratory, animal body weights varied widely in both the immature and the adult OVX protocols from laboratory to laboratory. Yet despite these differences in body weights, generally similar relative increases in both wet and blotted uterine weight were observed at the various laboratories for all of the protocols. High EE doses reduced body weight in the adult OVX protocols, but not in the immature animal protocols. In protocols A and B, only one laboratory recorded a significantly ($p < 0.05$) reduced body weight at the 10 μg/kg EE dose relative to the vehicle, while one laboratory recorded a significant increase. In protocol C, the 1.0, 3.0, and/or 10 μg/kg EE doses significantly reduced body weight relative to the vehicle controls for six laboratories. In protocol C′, with extended dosing from 3 to 7 days, the four laboratories showed consistently and significantly reduced body weights in the high EE dose groups. This weight loss is characteristic of potent estrogens such as EE or diethylstilbestrol. Loss of

body weight and adverse effects including mortality require further consideration when determining the doses to be tested in any proposed OECD Test Guideline. These issues will be considered further in dose selection of test substances in phase 2 of the validation program.

Body weight and uterine weight showed no consistent correlation, with less than half of the studies showing a significant ($p < 0.05$) correlation between these two variables. Significant associations were found more often in the immature animal protocols than in the adult animal protocols. These findings, coupled with the lack of significant body weight effects in the immature animal protocols, meant that the body weight adjustment had relatively little impact on the evaluation of uterine weights in these studies.

In the ZM antagonism dose groups, most laboratories found blotted uterine weight decreases in the ZM/EE combination groups that were statistically consistent, with the magnitude of the reduction similar across all laboratories. Interestingly, in all eight laboratories that carried out both protocols B and C, protocol B had a greater percentage reduction in uterine weight at the
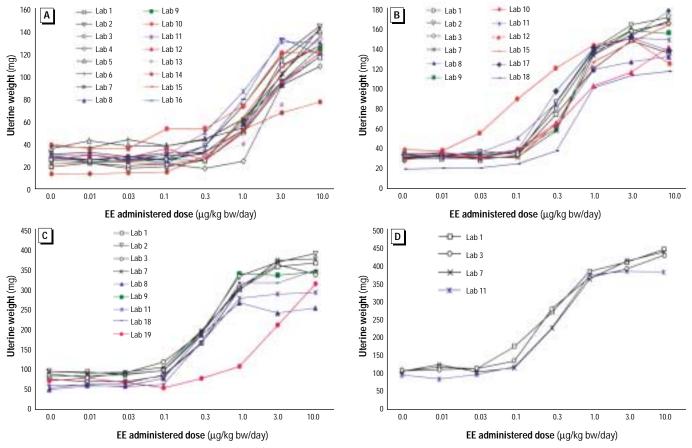

Figure 1. Response of blotted uterine weight to doses of EE. (*A*) Participating laboratory results for protocol A using immature female rats, dosing by oral gavage for 3 consecutive days. (*B*) Participating laboratory results for protocol B using immature female rats, dosing by subcutaneous injection for 3 consecutive days. (*C*) Participating laboratory results for protocol C using adult OVX rats, dosing by subcutaneous injection for 3 consecutive days. (*D*) Participating laboratory results for protocol C′ using extended subcutaneous injection dosing to 7 days. In all cases, animals were humanely sacrificed 24 hr after the last dose administration.

top dose (1.0 mg/kg ZM + 0.3 μg/kg EE). However, 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. One laboratory could not demonstrate a decreased uterine weight in the antagonist coadministration experiments. This was apparently related to an inability to induce a maximum uterine response at an EE-alone dose so that a statistically significant reduction could not be observed (see Figure 1, protocol C, lab 19, and Tables 7 and 10).

The original data analysis plan included a formal evaluation of whether the factors that varied from laboratory to laboratory could introduce variability in uterine weight response. Factors that varied from laboratory to laboratory included strain, diet, housing protocol, bedding, and vehicle. However, because of the overall consistency of the uterine weight responses across laboratories when these factors were reviewed (Table 1), a formal analysis of this aspect was judged unnecessary at this stage.

The analysis summarized in Table 9 illustrates the importance of minimizing the coefficient of variation in this type of study. The power of detecting various increases in uterine weight in the top dose group (by Dunnett's test) is analyzed 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). Six animals per group appear to be sufficient for detecting a 25–35% increase in uterine weight with reasonable power if the CV can be kept relatively low (e.g., in the general range of 10.0–15.0).

A widely used mathematical model, the Hill equation model, generally provided a good fit to the various data sets. The Hill model was applied to the 41 individual experiments. This permits an estimate of the effective doses at various levels such as 10%, 50%, and 90% of the maximum, the ED_{10}, ED_{50}, and ED_{90}, respectively. The calculated results for the ED_{10} and the ED_{90} are summarized in Table 10. The model calculations support the results previously reported based on other types of statistical analyses. For example, the estimated ED_{10} values in protocols B, C, and C´ are lower than those in protocol A, and no significant difference was found between the ED_{10} values from protocols B, C, and C´.

Discussion and Conclusions

All laboratories and all protocols were successful in phase 1 of the OECD validation program in detecting increases in uterine weights using EE as the reference agonist. These significant uterine weight increases were achieved in both the immature and the adult OVX models under a variety of different experimental conditions besides route of administration (e.g., strain, diet, housing protocol, bedding, vehicle). This suggests a certain robustness of the protocols, at least for the reference EE. The consistency of the results also suggests that no further specification of the strain of rat, diet, and so on is necessary for this screening assay to detect potent estrogen agonists. For each protocol, there was generally good agreement among laboratories with regard to the actual EE doses that produced increased uterine weights and the maximum response observed. The shapes of the uterine weight dose–response curves, although similar for many labs, did show some variation. The feasibility of an antagonist assay was also successfully demonstrated, but in less detail because availability of the ZM reference compound was limited.

At this time, no substantive difference or advantage in model—immature versus adult OVX—has been found. Differences in response to EE and ZM caused by the route of administration were expected and did occur. Nonetheless, these results reinforce the sense of robustness of the protocols. For example, there is a relatively consistent half-log difference between subcutaneous and

Table 7. Ratio of mean uterine weights in the 3.0 μg/kg/day EE group to that of the vehicle control.[a]

Laboratory	Wet weight				Blotted weight			
	A	B	C	C´	A	B	C	C´
1	463	872	1,009	431	400	496	365	403
2	479	872	708	NT	373	570	318	NT
3	400	837	824	366	335	487	326	370
4	462	NT	NT	NT	388	NT	NT	NT
5	321	NT	NT	NT	294	NT	NT	NT
6	257	NT	NT	NT	244	NT	NT	NT
7	348	748	666	398	335	562	319	390
8	337	589	420	NT	313	440	331	NT
9	342	794	776	NT	299	474	326	NT
10	332	694	NT	NT	283	397	NT	NT
11	555	821	794	425	386	482	373	412
12	370	535	NT	NT	326	370	NT	NT
13	225	NT	NT	NT	219	NT	NT	NT
14	518	NT	NT	NT	443	NT	NT	NT
15	541	654	NT	NT	419	536	NT	NT
16	598	NT	NT	NT	403	NT	NT	NT
17	NT	857	NT	NT	NT	470	NT	NT
18	NT	929	870	NT	NT	663	375	NT
19	NT	NT	258	NT	NT	NT	236	NT

NT, not tested.
[a]All responses are significant: $p < 0.05$ versus vehicle control (stated as a value of 100).

Table 8. Comparison of coefficients of variation (%) in uterine weights (averaged over dose groups within a laboratory and protocol).

Laboratory	Wet weight				Blotted weight			
	A	B	C	C´	A	B	C	C´
1	18.9	16.9	13.4	19.4	16.5	14.0	11.1	16.6
2	15.5	15.4	11.2	NT	12.9	11.9	9.8	NT
3	17.1	12.4	15.9	10.5	14.2	10.7	12.0	9.0
4	14.9	NT	NT	NT	12.3	NT	NT	NT
5	18.1	NT	NT	NT	18.4	NT	NT	NT
6	18.7	NT	NT	NT	17.4	NT	NT	NT
7	15.9	14.2	18.6	16.9	15.4	14.4	11.7	14.5
8	23.8	15.4	14.1	NT	22.2	12.3	10.8	NT
9	14.9	16.1	16.6	NT	13.7	14.2	12.8	NT
10	20.9	22.3	NT	NT	20.2	19.3	NT	NT
11	14.6	14.4	12.8	10.4	10.8	12.2	10.8	10.3
12	21.8	25.8	NT	NT	19.8	19.5	NT	NT
13	16.0	NT	NT	NT	14.0	NT	NT	NT
14	20.0	NT	NT	NT	15.2	NT	NT	NT
15	21.1	19.1	NT	NT	17.0	15.9	NT	NT
16	26.3	NT	NT	NT	22.9	NT	NT	NT
17	NT	15.4	NT	NT	NT	13.9	NT	NT
18	NT	15.3	24.8	NT	NT	16.4	20.0	NT
19	NT	NT	17.8	NT	NT	NT	16.0	NT
Overall mean[a]	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

NT, not tested.
[a]Averaged over doses and all laboratories. *Significant ($p < 0.05$) versus protocol A by Dunnett's test (after adjusting for differences due to dose and laboratory).

oral gavage administration in observed LOEL doses and the calculated ED_{10} doses across protocols (see Tables 2 and 6 and Figure 1). Group sizes of six animals appear to be sufficient to detect modest percentage increases (25–35%) in uterine weight that have been observed for weak partial estrogen agonists ([18,23]). Both the wet and blotted uterine weight end points were sensitive in all protocols. The blotted weight was less variable and, qualified by the use of high-potency reference compounds, was first in a few instances to indicate a statistically significant difference at lowest doses. Furthermore, the Hill equation appears to model satisfactorily the dose response to provide additional perspective on ED_{10}, ED_{50}, and maximal dose responses within and among laboratories.

Minor protocol refinements were identified. For example, protocol A was amended to allow a wider variation in body weights so that unnecessary animal use could be avoided. A body weight variation of ± 20% of the mean body weight (e.g., 35 g ± 7 g) was proposed as sufficient for the next phase of the study. Randomization among the groups will be maintained. Additionally, the age range of immature animals at first administration was expanded to 18–20 days. Protocol C was amended to lengthen the postoperative acclimatization period to 14 days. This allows further time for uterine regression, the use of the vaginal smears to confirm complete removal of the ovaries, and greater flexibility for laboratories in timing their experiments.

The current intent of the OECD program is to proceed with both the immature and the adult OVX models unless a substantive difference in the ability to detect estrogenic responses of the uterus is found. If confirmed as equivalent, they may both be considered for adoption as OECD Test Guidelines. The OECD is now implementing phase 2 of the program, which will entail demonstration of the protocols using weak estrogen partial agonists, such as genistein, o,p'-DDT, methoxychlor, nonylphenol, and bisphenol A. Phase 2 will continue to examine the performance of both immature and adult OVX animal models.

Table 9. Variation in statistical power and animal usage with CV.[a]

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

[a]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 CV. Power calculations were based on 5,000 simulated studies per condition.

Table 10. Hill equation dose estimates (EE µg/kg/day) for uterine blotted weights.

Laboratory	Protocol A ED_{10}	Protocol A ED_{90}	Protocol B ED_{10}	Protocol B ED_{90}	Protocol C ED_{10}	Protocol C ED_{90}	Protocol C′ ED_{10}	Protocol C′ ED_{90}
1	0.40	5.71	0.15	0.84	0.14	1.22	0.13	1.35
2	0.34	6.42	0.15	0.76	0.15	1.00	NT	NT
3	0.53	4.12	0.14	0.93	0.10	1.51	0.10	0.65
4	1.22	3.66	NT	NT	NT	NT	NT	NT
5	0.86	4.72	NT	NT	NT	NT	NT	NT
6	0.60	15.82	NT	NT	NT	NT	NT	NT
7	0.49	3.96	0.22	1.36	0.13	1.12	0.05	1.72
8	0.37	5.49	0.12	1.49	0.10	0.55	NT	NT
9	0.65	6.55	0.20	1.11	0.15	0.80	NT	NT
10	0.22	7.06	0.02	0.40	NT	NT	NT	NT
11	0.20	3.29	0.08	1.41	0.10	1.15	0.11	0.88
12	0.61	5.95	0.14	1.78	NT	NT	NT	NT
13	1.32	12.62	NT	NT	NT	NT	NT	NT
14	0.15	3.10	NT	NT	NT	NT	NT	NT
15	0.44	3.54	0.17	1.58	NT	NT	NT	NT
16	0.29	2.55	NT	NT	NT	NT	NT	NT
17	NT	NT	0.14	0.53	NT	NT	NT	NT
18	NT	NT	0.22	1.39	0.17	0.58	NT	NT
19	NT	NT	NT	NT	0.76	8.85	NT	NT

NT, not tested.

REFERENCES AND NOTES

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. 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, et al. Environmental endocrine disruption: an effects assessment and analysis. Environ Health Perspect 106(suppl 1):11–56 (1998).
4. National Research Council. Hormonally Active Agents in the Environment. Washington, DC:National Academy Press, 1999.
5. OECD. Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Working Group, 10–11 March 1998. ENV/MC/CHEM/RA(98)5. Paris:Organisation for Economic Co-operation and Development, 1998.
6. Allen E, Doisy EA. An ovarian hormone. Preliminary report on its localization, extraction and partial purification, and action in test animals. JAMA 81:819–821 (1923).
7. Astwood EB. A six-hour assay for the quantitative determination of estrogen. Endocrinology 23:25–31 (1938).
8. Bülbring E, Burn JH. The estimation of oestrin and of male hormone in oily solution. J Physiol 85:320–333 (1935).
9. Dorfman RL, Gallagher TF, Koch FC. The nature of the estrogenic substance in human male urine and bull testis. Endocrinology 19:33–41 (1936).
10. 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).
11. OECD. Final Report of the OECD Workshop on the Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods. Seventh Meeting of the National Co-ordinators of the Test Guidelines Programme, 18–19 September 1996. ENV/MC/CHEM/TG(96)9. Paris:Organisation for Economic Co-operation and Development, 1996.
12. OECD. The Validation of Test Methods Considered for Adoption as OECD Test Guidelines. ENV/MC/CHEM(98)6. Paris:Organisation for Economic Co-operation and Development, 1998.
13. OECD Final Report of the Phase 1 of the Validation Study of the Uterotrophic Assay. ENV/JM/TG/EDTA(2001)1. Paris:Organisation for Economic Co-operation and Development, March 2001.
14. Dukes M, Chester R, Yarwood L, Wakeling AE. Effects of a non-steroidal pure antioestrogen, ZM 189,154, on oestrogen target organs of the rat including bones. J Endocrinol 141:335–341 (1994).
15. Wakeling AE, Slater SR. Estrogen-receptor binding and biologic activity of tamoxifen and its metabolites. Cancer Treat Rep 64:741–744 (1980).
16. 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 Perspect 106:369–373 (1998).
17. Casanova M, You L, Gaido KW, Archibeque-Engle S, Janszen DB, d'A Heck H. Developmental effects of dietary phytoestogens in Sprague-Dawley rats and interactions of genistein and daidzein with rat estrogen

receptors α and β *in vitro*. Toxicol Sci 51:236–244 (1999).

18. 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. Regul Toxicol Pharmacol 25:176–188 (1997).

19. 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 con-

ducting the mouse bioassay. Lab Anim Sci 37:596–601 (1987).

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

21. Reel JR, Lamb JC IV, Neal BH. Survey and assessment of mammalian estrogen biological assays for hazard characterization. Fundam Appl Toxicol 34:288–305 (1996).

22. Christian MS, Hoberman AM, Bachmann S, Hellwig J.

Variability in the uterotrophic response assay (an *in vivo* estrogenic response assay) in untreated control and positive control (DES-DP, 2.5 µg/kg, BID) Wistar and Sprague-Dawley rats. Drug Chem Toxicol 21(suppl 1):51–100 (1998).

23. Ashby J, Tinwell H. Uterotrophic activity of bisphenol A in the immature rat. Environ Health Perspect 106:719–720 (1998).