For Official Use **ENV/JM/MONO(2012)19/REV1**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **21-Aug-2015**

English - Or. English

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

PERFORMANCE STANDARDS FOR STABLY TRANSFECTED TRANSACTIVATION IN VITRO ASSAYS TO DETECT ESTROGEN RECEPTOR ANTAGONISTS

Series on Testing & Assessment
No. 174 – Second edition

JT03380919

ENV/JM/MONO(2012)19/REV1
For Official Use

English - Or. English

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 174 ***(Second edition)***

PERFORMANCE STANDARDS FOR STABLY TRANSFECTED TRANSACTIVATION *IN VITRO* ASSAYS TO DETECT ESTROGEN RECEPTOR ANTAGONISTS

(Intended for the developers of new or modified similar test methods)

IOMC
INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS
A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT
Paris 2015

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

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

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

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

This publication is available electronically, at no charge.

Also published in the Series on Testing and Assessment link

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

or contact:

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

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

FOREWORD

This document includes the Performance Standards (PS) for stably transfected transactivation *in vitro* assays to detect estrogen receptor (ER) antagonists. These PS accompany the Performance-Based Test Guideline (PBTG) for Transfected Transactivation *In Vitro* Assays to Detect Estrogen Agonists and Antagonists (TG 455). The PS are intended for the developers of new or modified test methods, similar to the validated reference methods.

PS to detect ER agonists have been developed for TG 455 and are available in the Series on Testing and Assessment as No. 173.

PS to detect ER antagonists have initially been developed by the US for a single test method, the BG1Luc ER transactivation *in vitro* assay to detect ER antagonists. This test method, described in TG 457, was moved to TG 455 when a second method for the detection of ER antagonists became available in 2015. TG 457 was subsequently deleted; in parallel, the PS were updated to take into account this second validated reference method.

The update was performed by the lead countries of the two reference methods (Japan and US) and comments from the Working Group of National Coordinators of the Test Guidelines Programme (WNT) were requested on the draft updated PBTG and on the draft updated PS to detect ER antagonists, in July 2014. Only limited comments were received; they were addressed by Japan and the US in advance of the meeting of the VMG NA held in December 2014, which agreed to WNT submission for approval.

The present document was approved by the WNT in April 2015, declassified and published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides, and Biotechnology on 10 July 2015.

INTRODUCTION

1. The following Performance Standards (PS) accompanies the Performance Based Test Guideline for Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor (ER) Agonists and Antagonist (TG 455) (1). The PS presented in this document are specific to the antagonist part of TG 455. PS for detecting ER agonists are also available (see PS for Stably Transfected Transactivation *In Vitro* Assays to Detect Estrogen Agonists (ER TA Methods) accompanying the Performance Based Test Guideline for TG 455 (1) (13)). Prior to the acceptance of a new test method for regulatory testing applications, validation studies are conducted to assess its reliability (i.e. the extent of intra- and inter-laboratory reproducibility over time when performed using the standardized protocol) and its relevance (i.e. the ability of the test method to correctly predict or measure the biological effect of interest) (3) (4) (5) (6). The purpose of performance standards is to communicate the basis by which new proprietary (i.e. copyrighted, trademarked, registered) and non-proprietary test methods have been determined to have sufficient accuracy (i.e. agreement between a test method result and an accepted reference value) and reliability (i.e. extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol) for a specific testing purpose. New test methods (i.e. "me-too" tests) can be added to TG 455. The Mutual Acceptance of Data will only be guaranteed if any proposed new or updated similar test method, developed according to these Performance Standards (for estrogen antagonist), and to the Performance Standards developed for TG 455 (for estrogen agonist) (13), has been reviewed and adopted by the OECD.

2. Performance standards are based on an adequately validated test method(s) and provide a basis for evaluating the comparability of a proposed test method that is functionally and mechanistically similar (3) (4). The three elements of performance standards are:

- Essential test method components: These consist of essential structural, functional, and procedural elements of a validated test method. They should be included in the protocol of a proposed test method that is functionally and mechanistically similar to the validated method. Essential test method components include unique characteristics of the test method, critical procedural details, and quality control measures.

- A minimum list of reference substances: reference substances are used to assess the accuracy and reliability of a proposed functionally and mechanistically similar test method. These substances are a representative subset of those used to demonstrate the accuracy and reliability of the validated test method, and are the minimum number that should be used to evaluate the performance of a proposed mechanistically and functionally similar test method.

- Test method performance and reliability values: These are the standards for performance (i.e. accuracy, sensitivity, specificity, positive/negative predictivity) and reliability (i.e. degree to which the test method can be performed reproducibly within and among laboratories over time) that the proposed test method should meet or exceed when evaluated using the minimum list of reference substances.

3. The fully validated reference test methods that provide the basis for this PS are:
   - The Stably Transfected TA assay (STTA) using the human (h) ERα-HeLa-9903 cell line (1) and
   - The BG1Luc ER TA assay (2) using the BG1Luc4E2 cell line which predominately expresses hERα with some contribution from hERβ (7) (8).

ESSENTIAL TEST METHOD COMPONENTS AND OTHER VALIDATION CONSIDERATIONS

4. The primary objective of this test method is to provide a qualitative assessment of *in vitro* anti-estrogenic activity (i.e. whether a substance is positive or negative for anti-estrogenic activity). Quantitative analysis is also performed to provide additional information on the potency of test substances. For example, quantitative analysis can determine the half-maximal inhibitory concentration (IC_{50}).

5. Certain principles are important in delineating the essential test method components that determine whether a modified or new test method is functionally and mechanistically similar. *In vitro* ER TA assays are designed to identify substances that might interfere with ER-mediated cellular processes *in vivo*. The interaction of estrogens with cellular ERs initiates a cascade of events leading to the expression of specific genes in multiple target tissues.

6. The following test method components may vary, so this PS applies to test methods that may differ in:

- cell type (e.g. mammalian, fish, yeast)
- cell line (tissue type)
- characteristics of the cell line including presence of other receptors and metabolism
- culture conditions
- plating density
- plate layout (including how controls are incorporated)
- ERα characteristics (full length or partial, species of origin); if other ER proteins are present, ERα should predominate and the relative expression of each receptor should be known
- reporter gene construct (promoter, receptor binding elements, reporter)
- method of determining cytotoxicity

These elements should be clearly described in the test method, and may be helpful for explaining any possible deviations from the defined reliability and accuracy performance values (see below).

7. Considering that the principal of the method is based on co-exposure of the inducer (E2) and the antiestrogen, essential test method components for *in vitro* ER TA (antagonist) protocols should include:

- The use of a strong reference antagonist (e.g. raloxifene HCl, 4OH-tamoxifen) to demonstrate the adequacy of the method for detecting ER antagonists.

- A weak positive antagonist control (e.g. tamoxifen) that has an IC50 slightly below 10 µM should be included to provide another quality control measure by which to judge the acceptability of the method for detecting a weak agonist, and by which to evaluate the reproducibility of the test method.

- A full concentration response curves should be run on one positive and one negative reference substance in each experiment to demonstrate the adequacy of the test method for detecting ER antagonists.

- In addition, ER TA antagonist studies should include a concurrent control using the reference estrogen (e.g. E2) to establish a baseline level of induction (~80% of E2 maximum) against which the antagonistic activity of test substances can be assessed.

- A vehicle control (e.g. DMSO, EtOH, or H_2O) that is miscible with cell culture media at concentrations that are not cytotoxic and do not otherwise interfere with the test system.

- For initial range-finding, at least six concentrations spaced at decadic logarithmic (log10) intervals should be tested up to the maximum concentration (see below). Based on these range-finding experiments, a suitable concentration range should then be used for testing the chemical in view of generating data on the possible potency of the substance and to derive categorical predictions (e.g. Positive, Negative).

- In the absence of cytotoxicity restraints, the maximum concentration may be the concentration defined in the protocol of each method or even up to the limit of solubility, if appropriate.[1]

- A qualitative or quantitative evaluation of cytotoxicity (potential confounder) and how it is applied to the test method should be included in each study. Concentrations of test substances that reduce viability should not be considered in the analysis of the data.

- All concentrations of the controls (e.g. vehicle, weak positive(s), or negative(s)), the reference estrogen, the reference anti-estrogen, and the test substance should be tested at least in triplicate.

8. No standardized statistical methods for analyzing data obtained from *in vitro* ER TA antagonist assays have been developed. Each test method should establish a well-defined method for classifying a positive and a negative response. Positive results should be characterized by both the magnitude of the effect and the concentration at which the effect occurs (e.g. an IC_{50}, % max, etc.) when possible.

9. To ensure that a proposed *in vitro* ER TA test method possesses characteristics similar to other validated test methods, the reference substances for testing ER antagonists listed in Table 1 should be used to demonstrate the reliability and accuracy of the new test method. The 10 recommended reference substances, representing chemical classes commonly associated with ER activity, have been classified as ER antagonists or negatives based upon published reports, including *in vitro* assays for ER binding and TA (10) (11) (12). If a reference substance is no longer commercially available, a substance with the same classification and comparable potency, mode of action, and chemical class can be used. Supplementary information including the full listings of substances tested in both the STTA and the BG1Luc ER TAs, as well as additional substances tested in each test method during the respective validation studies, is provided in Annex 2. Additional chemicals not included in the reference substances list may be used to demonstrate an improvement (e.g. improved reproducibility and/or accuracy with regard to accepted reference data) of the new test method as compared with the fully validated test methods.

[1] It is necessary to be careful evaluation of positive results at the higher concentration than 10μM, because the non-specific inhibition of luciferase activity due to the cytotoxicity were indicated for some chemicals at the validation study.

Table 1: Reference substances (10) for the Evaluation of Test Method Performance and Reliability for *In vitro* ER TA Assays to Detect ER Antagonists

	Substance[a]	CASRN	ER STTA assay[1]			BG1Luc ER TA assay[2]			ER STTA[1] candidate effects	ICCVAM[5] Consensus Classification	MeSH[6] Chemical Class	Product Class[7]
			ER TA Activity	IC50 Value (M)	Test concentration range(M)	ER TA Activity	IC50 Value[3] (M)	Highest Concentration for Range Finder (M)[4]				
1	4-hydroxytamoxifen	68047-06-3	POS	3.97×10^{-9}	$10^{-12} – 10^{-7}$	POS	2.08×10^{-7}	2.58×10^{-4}	moderate POS	POS	Hydrocarbon (Cyclic)	Pharmaceutical
2	Raloxifene HCl	82640-04-8	POS	7.86×10^{-10}	$10^{-12} – 10^{-7}$	POS	1.19×10^{-9}	1.96×10^{-4}	moderate POS	POS	Hydrocarbon (Cyclic)	Pharmaceutical
3	Tamoxifen	10540-29-1	POS	4.91×10^{-7}	$10^{-10} – 10^{-5}$	POS	8.17×10^{-7}	2.69×10^{-4}	POS	POS	Hydrocarbon (Cyclic)	Pharmaceutical
4	17β estradiol	50-28-2	NEG	-	$10^{-9} – 10^{-4}$	NEG	-	3.67×10^{-3}	to be negative	PN	Steroid	Pharmaceutical, Veterinary Agent
5	Apigenin	520-36-5	NEG	-	$10^{-9} – 10^{-4}$	NEG	-	3.70×10^{-4}	NEG	NEG	Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate
6	Di-n-butyl phthalate	84-74-2	NEG	-	$10^{-8} – 10^{-3}$	NEG	-	3.59×10^{-3}	NEG	NEG	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer
7	Flavone	525-82-6	NEG	-	$10^{-8} – 10^{-3}$	NEG	-	4.50×10^{-4}	to be negative	PN	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical
8	Genistein	446-72-0	NEG	-	$10^{-9} – 10^{-4}$	NEG	-	3.70×10^{-4}	to be negative	NEG	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical

9	p-n-nonylphenol	104-40-5	NEG	-	10^{-9} – 10^{-4}	NEG	-	4.54×10^{-4}	not tested	NEG	Phenol	Chemical Intermediate
10	Resveratrol	501-36-0	NEG	-	10^{-8} – 10^{-3}	NEG	-	4.38×10^{-4}	to be negative	NEG	Hydrocarbon (Cyclic)	Natural Product

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; M = molar; IC50 = half maximal inhibitory concentration of test substance; NEG = negative; PN = presumed negative; POS = positive; PP = presumed positive; IC30 (and IC50) = the concentration of a test substance at which the response is 30% (or 50 % for PC50) inhibition of the response induced by the spike control in each plate.

[a] Common substances tested in the STTA and BG1Luc ER TA assays that were designated as ER antagonists or negatives and used to evaluate accuracy in the BG1 Luc ER TA validation study (9)(10).

[1] The Validation Report of the Stably transfected Transcriptional Activation Assay to Detect ER mediated activity, Part B (12)

[2] ICCVAM Test Method Evaluation Report on the LUMI-CELL ER (BG1Luc ER TA) Test Method: An In Vitro Method for Identifying ER Agonists and Antagonists (10).

[3] Mean IC50 values were calculated with values reported by the laboratories of the BG1Luc ER TA validation study (XDS, ECVAM, and Hiyoshi) (10).

[4]Concentrations reported were the highest concentrations tested (range finder) during the validation of the BG1Luc ER TA Assay. If concentrations differed between the laboratories, the highest concentration is reported. See table 4-11 of ICCVAM Test Method Evaluation Report; The LUMI-Cell®ER (BG1Luc ER TA) Test Method: An *In Vitro* Assay for Identifying Human Estrogen Receptor Agonist and Antagonist Activity of Chemicals (10).

[5] Classification as an ER antagonist or negative was based upon information in the ICCVAM Background Review Documents (BRD) for ER Binding and TA test methods (32) as well as information obtained from publications published and reviewed after the completion of the ICCVAM BRDs (9) (10) (14) (15).

[6] Substances were assigned to one or more chemical classes using the U.S. National Library of Medicine's Medical Subject Headings (MeSH), an internationally recognised standardised classification scheme (available at http://www.nlm.nih.gov/mesh).

[7] Substances were assigned to one or more product classes using the U.S. National Library of Medicine's Hazardous Substances Data Bank (available at http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB).

10. New similar test methods should not be developed on only the basis of the 10 reference substances, but rather on a sufficiently larger test development set including the 10 reference substances. Reference substances should be preferentially used to determine equivalence of performance compared to the validated reference test methods.

11. All substances should be tested in a coded/blinded manner. When evaluated using these reference substances, the reliability and test method performance (i.e. sensitivity, specificity, positive/negative predictivity) of the proposed ER TA test method should approximate the following defined reliability and accuracy values.

DEFINED RELIABILITY AND ACCURACY PERFORMANCE VALUES

12. For the purposes of establishing the reliability and accuracy of the proposed test methods when transferred between laboratories, all 10 reference substances (Table 1) should be tested in two or (preferably) three laboratories. In each laboratory, all 10 references substances should be tested in three runs.

Within-laboratory (Intra-laboratory) reproducibility

13. For the assessment of within-laboratory reproducibility, the concordance of classifications (positive/negative) obtained in three independent consecutive test runs should be 100% for each laboratory for each of the 10 reference substances (Table 1). Three independent consecutive runs are required to fulfill the criteria for acceptance. If, for example, runs 2 and 3 are inconsistent with run 1, one additional run (run 4) will be sufficient to show within-lab reproducibility if run 4 is consistent with runs 2 and 3. If run 4 is consistent with run 1 instead, then at least two additional consecutive runs (runs 5 and 6) showing consistency with run 4 will be required to fulfill the requirement for three consecutive independent runs that have 100% concordance of classifications.

Between-laboratory (Intra- laboratory) reproducibility

14. Between-laboratory reproducibility should be assessed using the 10 reference substances (Table 1). The concordance of classifications (positive/negative) in at least two, but preferably three, laboratories for the 10 reference substances should be 100% for the three positive substances, and at least 86% for the seven negative substances.

Predictive capacity

15. The performance of the proposed test method (i.e. accuracy, sensitivity, specificity, positive/negative predictivity) should be comparable to that demonstrated for the fully validated STTA and BG1Luc ER TA (antagonist) test method (9) (10) when evaluating the 10 reference substances. Based upon the performance values of the validated reference method, the accuracy of the proposed ER TA test method should approximate those of the validated ER TA test method and should be at least 90%.

16. Although it is not realistic to expect test methods to perform identically, discordant results should be addressed in terms of the ability of the test method to accurately classify other substances with similar potencies and from similar chemical classes as demonstrated by the fully validated test method (9) (10).

LITERATURE

1. OECD (2015), *Test No. 455: Performance-Based Test Guideline for The Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing.

2. NIH Publication No. 11-7850: ICCVAM Test Method Evaluation Report, The LUMI-CELL® ER (BG1Luc ER TA) Test Method: An In Vitro Assay for Identifying Human Estrogen Receptor Agonist and Antagonist Activity of Chemicals [http://iccvam.niehs.nih.gov/methods/endocrine/ERTA-TMER.htm]. Appendix B, BG1 Antagonist Protocol [*http://iccvam.niehs.nih.gov/docs/endo_docs/ERTA-TMER/AppxB2-AntagProtocol.pdf*]

3. OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, Series on Testing and Assessment No.34, OECD, Paris.

4. ICCVAM (1997), Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), NIH Publication No. 97-3981, National Institute of Environmental Health Sciences, Research Triangle Park, NC.

5. ICCVAM (2003), ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods, NIH Publication No. 03-4508, Research Triangle Park, NC, National Institute of Environmental Health Sciences

6. OECD (1996), Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods, ENV/MC/CHEM/TG(96)9, OECD, Paris. Available at: [http://www.oecd.org/document/24/0,3746,en_2649_37465_47858904_1_1_1_37465,00.html]

7. Pujol, P., *et al., (1998), Differential expression of estrogen receptor-alpha and -beta messenger RNAs as a potential marker of ovarian carcinogenesis. Cancer* Res. 58(23): p. 5367-73.

8. Rogers, J.M. and M.S. Denison (2000), *Recombinant cell bioassays for endocrine disruptors: development of a stably transfected human ovarian cell line for the detection of estrogenic and anti-estrogenic chemicals*. *In Vitro* and Molecular Toxicology: Journal of Basic and Applied Research, 13(1): p. 67-82.

9. OECD (2015), The Validation Report of the Stably transfected Transcriptional Activation Assay to Detect ER mediated activity, Series on Testing and Assessment No.225 OECD, Paris.

10. ICCVAM (2011), ICCVAM Test Method Evaluation Report (TMER) on the LUMI-CELL® ER (BG1Luc ER TA) Test Method an *In Vitro* Method for Identifying ER Agonists and Antagonists, National Institute of Environmental Health Sciences, Research Triangle Park, NC. NIH Publication No. 11-785 Available at: [http://iccvam.niehs.nih.gov/docs/endo_docs/ERTA-TMER/BG1LucER-TA-TMER-Combined.pdf].

11. ICCVAM (2003), ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and

Transcriptional Activation Assays,
Available at [http://iccvam.niehs.nih.gov/docs/endo_docs/edfinalrpt0503/edfinrpt.pdf]

12. ICCVAM (2002), Background Review Document: Estrogen Receptor Binding, Appendiz D, Substances Tested in the ER Binding Assay, NIH Publication Report No. 03-4504. Available at: [http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erbndbrd/ERBd034504.pdf]

13. OECD (2012), *Performance Standards for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Agonists for TG 455*, Series on Testing and Assessment No.173, OECD, Paris.

14. EDSTAC (1998), Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) Final Report. Available at: [http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm]

15. ICCVAM (2002), Background Review Document: Estrogen Receptor Transcriptional Activation (TA) Assay. Appendix D, Substances Tested in the ER TA Assay, NIH Publication Report No. 03-4505. Available at: [http://iccvam.niehs.nih.gov/docs/endo_docs/final1002/erta_brd/ERTA034505.pdf]

ANNEX 1

Definitions and Abbreviations

Acceptability criteria: Minimum standards for the performance of experimental controls and reference standards. All acceptability criteria must be met for an experiment to be considered valid.

Accuracy: The closeness of agreement between a test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance (2). Accuracy is determined by using the number of reference substances as denominator with number of correct responses in numerator normally expressed as a percent.

Agonist: A substance that produces a response, e.g. transcription, when it binds to a specific receptor.

Antagonist: A substance that inhibits an agonist response, e.g. transcription.

Anti-estrogenic activity: The capability of a chemical to inhibit 17β-estradiol or other estrogens in their ability to bind to and activate estrogen receptors. hERα-mediated estrogenic activity can be detected with the PBTG.

BG-1: An immortalized adenocarcinoma cell that endogenously express estrogen receptor.

BG1Luc4E2: The BG1Luc4E2 cell line was derived from BG-1 immortalized human-derived adenocarcinoma cells that endogenously express both forms of the estrogen receptor (ERα and ERβ) and have been stably transfected with the plasmid pGudLuc7.ERE. This plasmid contains four copies of a synthetic oligonucleotide containing the estrogen response element upstream of the mouse mammary tumor viral (MMTV) promoter and the firefly luciferase gene.

Cytotoxicity: Harmful effects to cell structure or function that can ultimately cause cell death and can be reflected by a reduction in the number of cells present in the well at the end of the exposure period or a reduction of the capacity for a measure of cellular function when compared to the concurrent vehicle control.

E2: 17β-estradiol

ER: Estrogen receptor

hERα: Human estrogen receptor alpha

hERß: Human estrogen receptor beta

ERE: Estrogen response element

Estrogenic activity: The capability of a substance to mimic 17β-estradiol in its ability to bind to and activate estrogen receptors. hERα-mediated estrogenic activity can be detected with the PBTG.

HeLa: An immortal human cervical cell line

HeLa9903: A HeLa cell subclone into which hERα and a luciferase reporter gene have been stably transfected

IC_{50}: The half maximal inhibitory concentration of a test substance.

Within-laboratory (Intra-laboratory) reproducibility: A determination of the extent that qualified people within the same laboratory can successfully replicate results using a specific protocol at different times (2).

Between-laboratory (Inter-laboratory) reproducibility: A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is determined during the prevalidation and validation processes, and indicates the extent to which a test method can be successfully transferred between laboratories, also referred to as between-laboratory reproducibility (2).

Me-too test: A colloquial expression for a test method that is structurally and functionally similar to a validated and accepted reference test method. Interchangeably used with similar test method.

PBTG: Performance-Based Test Guideline.

Performance standards: Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are (1) essential test method components; (2) a minimum list of reference chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (3) the comparable levels of accuracy and reliability, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of reference chemicals (2).

Predictivity (negative): The proportion of correct negative responses among substances testing negative by a test method. It is an indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

Predictivity (positive): The proportion of correct positive responses among substances testing positive by a test method. It is one indicator of test method accuracy. Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.

Proficiency chemicals (substances): reference chemicals included in the Performance Standards that can be used by laboratories to demonstrate technical competence with a

standardized test method. Selection criteria for these substances typically include that they represent the range of responses, are commercially available, and have high quality reference data available.

Proficiency: The demonstrated ability to properly conduct a test method prior to testing unknown substances.

Reference Chemicals (substances): A set of chemicals to be used to demonstrate the ability of a new test method to meet the acceptability criteria demonstrated by the validated reference test method(s). These chemicals should be representative of the classes of chemicals for which the test method is expected to be used, and should represent the full range of responses that may be expected from the chemicals for which it may be used, from strong, to weak, to negative.

Reference anti-estrogen: Raloxifene HCl (Ral, CASRN 82640-04-8).

Reference estrogen: 17ß-estradiol (E2, CASRN 50-28-2).

Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (2).

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (2).

SD: Standard deviation

Sensitivity: The proportion of all positive/active substances that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (2).

Specificity: The proportion of all negative/inactive substances that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of attest method (2).

Stable transfection: When DNA is transfected into cultured cells in such a way that it is stably integrated into the cells genome, resulting in the stable expression of transfected genes. Clones of stably transfected cells are selected by stable markers (e.g. resistance to G418).

STTA: Stably Transfected Transactivation Assay, the ERα transactivation assay using the HeLa 9903 Cell Line.

Substance: Used in the context of the UN GHS as chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.

TA: Transactivation.

Transcription: mRNA synthesis

Transcriptional activation: The initiation of mRNA synthesis in response to a specific chemical signal, such as a binding of an estrogen to the estrogen receptor.

Validation: The process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose.

VC: Vehicle control, the solvent that is used to dissolve test and control chemicals is tested solely as vehicle without dissolved chemical.

Weak positive control: A weakly active substance selected from the reference chemicals list that is included in all tests to help ensure proper functioning of the assay.

ANNEX 2

Supplementary Information

for the

Stably Transfected Human Estrogen Receptor-α TransActivation (STTA) Assay for Detection of Estrogenic Antagonist-Activity of Chemicals using the hERα-HeLa-9903 cell line

And

The Estrogen Receptor (BG1Luc ER TA) Transactivation Test Method for Identifying ER Antagonists

Table 1: Comparison of Results from STTA and BG1Luc ER TA Assays for 15 substances Tested in Both Assays and Classified as Positive (POS) or Negative (NEG) for ER Antagonists

			ER STTA assay[1]		BG1Luc ER TA assay[2]					
	Substance[a]	CASRN	ER TA Activity	IC50 Value[b] (M)	ER TA Activity	IC50 Value[b,3] (M)	ER STTA [1] candidate effects	ICCVAM [4] Consensus Classification	MeSH[5] Chemical Class	Product Class[6]
1	4-hydroxytamoxifen	68047-06-3	POS	3.97×10^{-9}	POS	2.08×10^{-7}	moderate POS	POS	Hydrocarbon (Cyclic)	Pharmaceutical
2	Dibenzo[a.h] anthracene	53-70-3	POS	No IC50	POS	No IC50	POS	PP	Polycyclic Compound	Laboratory Chemical, Natural Product
3	Mifepristone	84371-65-3	POS	5.61×10^{-6}	NEG	-	mild POS	NEG	Steroid	Pharmaceutical
4	Raloxifene HCl	82640-04-8	POS	7.86×10^{-10}	POS	1.19×10^{-9}	moderate POS	POS	Hydrocarbon (Cyclic)	Pharmaceutical
5	Tamoxifen	10540-29-1	POS	4.91×10^{-7}	POS	8.17×10^{-7}	POS	POS	Hydrocarbon (Cyclic)	Pharmaceutical
6	17-b estradiol	50-28-2	NEG	-	NEG	-	PN	PN	Steroid	Pharmaceutical, Veterinary Agent
7	Apigenin	520-36-5	NEG	-	NEG	-	NEG	NEG	Heterocyclic Compound	Dye, Natural Product, Pharmaceutical Intermediate
8	Atrazine	1912-24-9	NEG	-	NEG	-	NEG	PN	Heterocyclic Compound	Herbicide
9	Di-n-butyl phthalate	84-74-2	NEG	-	NEG	-	NEG	NEG	Ester, Phthalic Acid	Cosmetic Ingredient, Industrial Chemical, Plasticizer
10	Fenarimol	60168-88-9	NEG	-	NEG	-	not tested	PN	Heterocyclic Compound, Pyrimidine	Fungicide
11	Flavone	525-82-6	NEG	-	NEG	-	PN	PN	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical
12	Flutamide	13311-84-7	NEG	-	NEG	-	NEG	PN	Amide	Pharmaceutical, Veterinary Agent
13	Genistein	446-72-0	NEG	-	NEG	-	PN	NEG	Flavonoid, Heterocyclic Compound	Natural Product, Pharmaceutical
14	p-n-nonylphenol	104-40-5	NEG	-	NEG	-	not tested	NEG	Phenol	Chemical Intermediate
15	Resveratrol	501-36-0	NEG	-	NEG	-	PN	NEG	Hydrocarbon (Cyclic)	Natural Product

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; M = molar; IC50 = half maximal inhibitory concentration of test chemical; NEG = negative; PN = presumed negative; POS = positive; PP = presumed positive; IC30 (and IC50) = the concentration of a test chemical at which the response is 30% (or 50 % for PC50) inhibition of the response induced by the spike control in each plate.

[a] Common substances tested in the STTA and BG1Luc ER TA assays that were designated as ER antagonists or negatives and used to evaluate accuracy in the BG1 Luc ER TA validation study (9)(10).

[b] Maximum concentration tested in the absence of limitations due to cytotoxicity or insolubility was 1 x 10-3 M (STTA Assay) and 1 x 10-5 M (BG1Luc ER TA Assay).

[1] The Validation Report of the Stably transfected Transcriptional Activation Assay to Detect ER mediated activity, Part B (9)

[2] ICCVAM Test Method Evaluation Report on the LUMI-CELL ER (BG1Luc ER TA) Test Method: An In Vitro Method for Identifying ER Agonists and Antagonists (10).

[3] Mean IC50 values were calculated with values reported by the laboratories of the BG1Luc ER TA validation study (XDS, ECVAM, and Hiyoshi) (10).

[4] Classification as an ER antagonist or negative was based upon information in the ICCVAM Background Review Documents (BRD) for ER Binding and TA test methods (10) as well as information obtained from publications published and reviewed after the completion of the ICCVAM BRDs (12).

[5] Substances were assigned to one or more chemical classes using the U.S. National Library of Medicine's Medical Subject Headings (MeSH), an internationally recognized standardized classification scheme (available at http://www.nlm.nih.gov/mesh).

[6] Substances were assigned to one or more product classes using the U.S. National Library of Medicine's Hazardous Substances Data Bank (available at http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB).

Table 2: Substances Tested for ER Antagonist Activity during the STTA ER TA Validation Study (9)

	Substance [1]	CASRN	ER STTA [1] antagonist Classification	ER STTA [1,a] Mean IC50(M)
1	17b-estradiol	50-28-2	NEG	-
2	4,4'-(Hexafluoroisopropylidene)diphenol	1478-61-1	NEG	-
3	4,4'-[1-[4-[1-(4-Hydroxyphenyl)-1-methylethyl]phenyl]ethylidene]bis[phenol]	110726-28-8	POS	2.51×10^{-6}
4	4,4'-Cyclohexylidenebisphenol	843-55-0	NEG	-
5	4-Hydroxytamoxifen	68047-06-3	POS	3.97×10^{-9}
6	Apigenin	520-36-5	NEG	-
7	Atrazine	1912-24-9	NEG	-
8	Clomiphene citrate(cis and trans mixture)	50-41-9	POS	4.26×10^{-7}
9	Dibenzo[a,h]anthracene	53-70-3	POS	No IC50
10	Dibutyl phthalate	84-74-2	NEG	-
11	Fenarimol	60168-88-9	NEG	-
12	Flavone	525-82-6	NEG	-
13	Flutamide	13311-84-7	NEG	-
14	Genistein	446-72-0	NEG	-
15	ICI 182,780	129453-61-8	POS	2.67×10^{-10}
16	Methylpiperdinylpyrazole dihydrochloride	289726-02-9	POS	3.11×10^{-8}
17	Mifepristone(Mifeprex)＝RU-486	84371-65-3	POS	5.61×10^{-6}
18	p-n-nonylphenol	104-40-5	NEG	-
19	Raloxifene HCl	82640-04-8	POS	7.86×10^{-10}
20	Resveratrol	501-36-0	NEG	-
21	Tamoxifen	10540-29-1	POS	4.91×10^{-7}

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; M = molar; IC50 = half maximal inhibitory concentration of test chemical; NEG = negative; POS = positive.

[1]Table is sorted by classification and then alphabetically by chemical name. Only substances for which a definitive POS/NEG call could be made were included in the table.

[1] The Validation Report of the Stably transfected Transcriptional Activation Assay to Detect ER mediated activity, Part B (9)

[a] Maximum concentration tested in the absence of limitations due to cytotoxicity or insolubility was 1 x 10-3 M

Table 3: Substances Tested for ER Antagonist Activity during the BG1Luc ER TA Validation Study (10)

	Substance[1]	CASRN	BG1Luc ER TA Classification[3]	BG1Luc ER TA Mean IC_{50} (M)[4]
1	4-hydroxytamoxifen	68047-06-3	POS	2.08×10^{-7}
2	Actinomycin D[2]	50-76-0	POS	1.43×10^{-7}
3	Apomorphine	58-00-4	POS	NC
4	Cycloheximide[2]	66-81-9	POS	9.67×10^{-7}
5	Dibenzo[*a.h*] anthracene	53-70-3	POS	NC
6	Ketoconazole	65277-42-1	POS	1.23×10^{-6}
7	Medroxy-progesterone acetate	71-58-9	POS	NC
8	Raloxifene HCl	82640-04-8	POS	1.19×10^{-9}
9	Tamoxifen	10540-29-1	POS	8.17×10^{-7}
10	12 – *O* –tetradecanoyl-phorbol-13-acetate	16561-29-8	NEG	-
11	17-α ethinyl estradiol	57-63-6	NEG	-
12	17-β estradiol	50-28-2	NEG	-
13	17β-trenbolone	10161-33-8	NEG	-
14	19-nortestosterone	434-22-0	NEG	-
15	2-*sec*-butylphenol	89-72-5	NEG	-
16	2,4,5-trichlorophenoxy-acetic acid	93-76-5	NEG	-
17	4-androstenedione	63-05-8	NEG	-
18	4-cumylphenol	599-64-4	NEG	-
19	4-hydroxyandrostenedione	566-48-3	NEG	-
20	Apigenin	520-36-5	NEG	-
21	4-*tert*-octylphenol	140-66-9	NEG	-
22	5α-dihydrotestosterone	521-18-6	NEG	-
23	Ammonium perchlorate	7790-98-9	NEG	-
24	Chrysin	480-40-0	NEG	-
25	Atrazine	1912-24-9	NEG	-
26	Bicalutamide	90357-06-5	NEG	-
27	Bisphenol A	80-05-7	NEG	-
28	Bisphenol B	77-40-7	NEG	-
29	Butylbenzyl phthalate	85-68-7	NEG	-
30	Coumestrol	479-13-0	NEG	-
31	Corticosterone	50-22-6	NEG	-

32	Genistein	446-72-0	NEG	-
33	Cyproterone acetate	427-51-0	NEG	-
34	Daidzein	486-66-8	NEG	-
35	Dexamethasone	50-02-2	NEG	-
36	Di-*n*-butyl phthalate	84-74-2	NEG	-
37	Dicofol	115-32-2	NEG	-
38	Diethylhexyl phthalate	117-81-7	NEG	-
39	Diethylstilbestrol	56-53-1	NEG	-
40	Estrone	53-16-7	NEG	-
41	Ethyl paraben	120-47-8	NEG	-
42	Fenarimol	60168-88-9	NEG	-
43	Finasteride	98319-26-7	NEG	-
44	Flavone	525-82-6	NEG	-
45	Fluoranthene	206-44-0	NEG	-
46	Fluoxymestrone	76-43-7	NEG	-
47	Flutamide	13311-84-7	NEG	-
48	Kaempferol	520-18-3	NEG	-
49	Haloperidol	52-86-8	NEG	-
50	Hydroxyflutamide	52806-53-8	NEG	-
51	Resveratrol	501-36-0	NEG	-
52	Kepone	143-50-0	NEG	-
53	L-thyroxine	51-48-9	NEG	-
54	Linuron	330-55-2	NEG	-
55	*meso*-hexestrol	84-16-2	NEG	-
56	Methyl testosterone	58-18-4	NEG	-
57	Mifepristone	84371-65-3	NEG	-
58	Morin	480-16-0	NEG	-
59	Nilutamide	63612-50-0	NEG	-
60	Norethynodrel	68-23-5	NEG	-
61	*o.p'*-DDT	789-02-6	NEG	-
62	Oxazepam	604-75-1	NEG	-
63	*p*-n-nonylphenol	104-40-5	NEG	-
64	*p.p'*-DDE	72-55-9	NEG	-
65	*p.p'*-methoxychlor	72-43-5	NEG	-
66	Phenobarbital	50-06-6	NEG	-

67	Phenolphthalin	81-90-3	NEG	-
68	Pimozide	2062-78-4	NEG	-
69	Procymidone	32809-16-8	NEG	-
70	Progesterone	57-83-0	NEG	-
71	Propylthiouracil	51-52-5	NEG	-
72	Sodium azide	26628-22-8	NEG	-
73	Spironolactone	52-01-7	NEG	-
74	Testosterone	58-22-0	NEG	-
75	Vinclozolin	50471-44-8	NEG	-

[1]Table is sorted by classification and then alphabetically by chemical name. Only substances for which a definitive POS/NEG call could be made were included in the table.

[2]Actinomycin D and cycloheximide, inhibit protein biosynthesis, and should not be considered as antagonists.

[3]Classification based upon results reported in the ICCVAM Test Method Evaluation Report (TMER) on the LUMI-CELL® ER (BG1Luc ER TA) Test Method an *In Vitro* Method for Identifying ER Agonists and Antagonists [10].

[4]NC = An IC_{50} value could not be calculated for this substance.

Summary of the Accuracy and Reliability Values Obtained During the Validation Studies for the STTA and BG1Luc ER TA (Antagonists)

The respective validation study reports for the STTA and BG1Luc ER TAs provide comprehensive descriptions of the data used to develop the reliability and accuracy values for each of antagonist assays (9, 10). The following is a summary of the test method performance and intra- and inter-laboratory reproducibility for the validated test method.

I. **Intra-laboratory (within-laboratory) reproducibility:** The closeness of agreement between test results obtained within a single laboratory when the procedure is performed using the same substance under identical conditions within a given period of time.

a. BG1Luc ER TA validation study: The intra-laboratory reproducibility of the BG1Luc ER TA Antagonist test method was evaluated using 12 substances (2 positive, 10 negative), that were each tested three times on three separate days at each laboratory. There was 100% agreement within each laboratory for each of the three repeat tests of these reference substances (Phase 2 of the antagonist validation study).

Table 4: Intra-laboratory reproducibility for the BG1Luc ER TA Antagonist Assay (10)

Activity per Test	XDS	ECVAM	Hiyoshi
Agreement Within Laboratory	12/12 (100%)	12/12 (100%)	12/12 (100%)
+++	2/12	2/12	2/12
---	10/12	10/12	10/12
Discordance Within Laboratory	0/12 (0%)	0/12 (0%)	0/12 (0%)
++-	0/12	0/12	0/12
+--	0/12	0/12	0/12

Abbreviations: ECVAM = European Centre for the Validation of Alternative Methods; XDS = Xenobiotic Detection Systems, Inc.

\+ Denotes a positive test result.

\- Denotes a negative test result.

+++ Indicates that each of three replicate tests within each laboratory had a classification as positive.

--- Indicates that each of three replicate tests within each laboratory had a classification as negative.

++- Indicates that a test substance was classified as positive in two of three replicate tests. The substance was classified as negative in a third replicate test.

+-- Indicates that the test substance was classified as positive in one of three replicate tests. The substance was classified as negative in the remaining two tests.

b. STTA ER TA validation study: The intra-laboratory reproducibility of the ER STTA Antagonist assay method was evaluated using 20 substances (8 positives, 10 negatives and 2 unknowns), and all or a part of these substances were distributed to each laboratory in coded manner, then three to six sets of data that met the quality criteria were obtained on separate days. There was

67% - 100% agreement within each laboratory for each repeated tests of test substances (Task-3 of the antagonist validation study).

Table 5: Intra-laboratory reproducibility for the ER STTA Antagonist Assay (9)

Activity per Test	CERI	OTSUKA	KANEKA	HIYOSHI
Agreement Within Laboratory	19/20 (95%)	19/19 (100%)	8/12 (67%)	18/20 (90%)
Positive candidate	8/9	9/9	3/5	8/9
Negative candidate	9/9	8/9	6/7	8/9
Unknown	2/2	1/1	0/1	2/2
Discordance Within Laboratory	1/20 (5%)	0/19 (0%)	4/12 (33%)	2/20 (10%)
Positive candidate (Results*)	1/9 (NNP)*	0/9	2/5 (PPPPNP) (PPPNP)	1/9 (NPP)
Negative candidate (Results)	0/9	1/9 (PNN)	1/6 (NNNPN)	1/9 (NPP)
Unknown (Results)	0/2	0/1	1/1 (PNNNNN)	0/2

*: Deatails of discordant results, (NNP) means two negatives and one positive.

II. **Inter-laboratory (Between-laboratory) reproducibility:** A measure of the extent to which different qualified laboratories using the same protocol and testing the same substances can produce qualitatively and quantitatively similar results. Inter- laboratory reproducibility is determined during the validation process and indicates the extent to which a test method can be transferred successfully among laboratories.

a. BG1Luc ER TA validation study: Inter-laboratory reproducibility was assessed using 53 substances that were tested at least once in each of 3 laboratories. There was 94% (50/53) agreement on the classifications for these substances among the laboratories. Two substances (2/53) had inadequate overall classifications (i.e. 1 positive, 1 negative and 1 inadequate call). The agreement between the laboratories is shown in Table 6.

Table 6: Inter-laboratory reproducibility for the BG1Luc ER TA Antagonist Assay (10)

Results Among Laboratories	Percent Agreement
Agreement Among Laboratories	50/53 (94%)
+++	4/53 (8%)
---	43/53 (81%)
++I	1/53 (2%)
--I	2/53 (4%)
Discordance Among Laboratories	3/53 (6%)
++-	0/53 (0%)
+--	1/53 (2%)
+-I	2/53 (4%)

Abbreviations: I = inadequate data (i.e. Data are classified as inadequate if, because of major qualitative or quantitative limitations, they cannot be interpreted as valid for showing either the presence or absence of activity. Inadequate data typically result from some type of systemic error, such as high background across the test plate or failure of a multi-tip pipette to dispense liquid in numerous wells.
[a]Only those substances that produced a definitive result in at least two of the three laboratories were used in this evaluation.
[b]Substances that produced an inadequate result in two laboratories during agonist testing were not included in this table.
+ Denotes a positive test result.
- Denotes a negative test result.
+++ Indicates that the substance was classified as positive at all three laboratories.
--- Indicates that the substance was classified as negative at all three laboratories.
++I Indicates that the substance was classified as positive at two of three laboratories but had inadequate data in the third.
--I Indicates that the substance was classified as negative at two of three laboratories but had inadequate data in the third.
+-I Indicates that the substance was classified as positive at one laboratory, negative at one laboratory, and inadequate at the third laboratory

b. STTA ER TA validation study: Inter-laboratory reproducibility was assessed using 12 substances that were tested in four labs and 8 substances that were tested in two or three labs. There was 67%-100% agreement for each substance among the laboratories. Then 18 out of 20 substances gave perfect matching results (100%) in all labs. The agreement between the laboratories is shown in Table 7.

Table 7: Inter-laboratory reproducibility for the ER STTA Antagonist Assay (9)

Chemical name	Candidate effect	CERI	OTSUKA	KANEKA	HIYOSHI	Concordance between labs
ICI 182,780	P	P	P	P	P	100%
Mifepristone(Mifeprex)=RU-486	P	P	P	P	P	100%
4,4'-(Hexafluoroisopropylidene)diphenol	N	N	N	N	N	100%
Methylpiperdinylpyrazole dihydrochloride	P	P	P		P	100%
4-Hydroxytamoxifen	P	P	P	P	P	100%
Raloxifene HCl	P	P	P		P	100%
Clomiphene citrate(cis and trans mixture)	P	P	P	P	P	100%
Dibutyl phthalate	N	N	N	N	N	100%
Atrazine	N	N	N		P	67%
Flutamide	N	N	N	N	N	100%
4,4'-Cyclohexylidenebisphenol	N	N	N	N	N	100%
4,4'-[1-[4-[1-(4-Hydroxyphenyl)-1 -methylethyl]phenyl]ethylidene]bis[phenol]	P	P	P		P	100%
Apigenin	N	N	N		N	100%
Genistein	N	N	N	N	N	100%

Dibenzo[a,h]anthracene	P	N	P		P	67%
p-n-nonylphenol	Unknown	N	N	N	N	100%
Flavone	N	N	N		N	100%
Resveratrol	N	N	N	N	N	100%
Fenarimol	Unknown	N			N	100%
17b-estradiol	N	N	N	N	N	100%
P: Positive, N: Negative.	Accuracy:	94%	100%	100%	94%	100% (97%)*
	Sensitivity:	88%	100%	100%	100%	100% (97%)*
	Specificity :	100%	100%	100%	88%	100% (97%)*

*: Values in parenthesis are calculated with all individual data derived in four laboratories (n=65).

Table 8: Summary of Inter-laboratory reproducibility for the ER STTA Antagonist Assay (9)

Results Among Laboratories	**Percent Agreement**
Agreement Among Laboratories	18/20 (90%)
PPPP	4/20 (20%)
PPP	3/20 (15%)
NNNN	8/20 (40%)
NNN	2/20 (10%)
NN	1/20 (5%)
Discordance Among Laboratories	2/20 (10%)
PPPN	1/20 (5%)
NNNP	1/20 (5%)

III. **Predictive Capacity**: Measures of test method performance (i.e. accuracy, sensitivity, specificity, positive and negative predictivity), and overall accuracy provide a quantitative assessment of the closeness of agreement (e.g. the proportion of correct outcomes) between test method results and the values obtained from reference substances.

a. BG1Luc ER TA validation study: The predictive capacity was assessed using 25 reference substances (3 positive, 22 negative) that produced definitive results in the BG1Luc ER TA assay for antagonist activity (See Section 3.4 in reference10).

Table 9: Predictive Capacity for the BG1 Luc ER TA (Antagonist Assay) (10)

		BG1Luc ER TA		
		Positive	Negative	Total
ICCVAM Classification	Positive	3	0	3
	Negative	0	22	22
	Total	3	22	25

Accuracy	100%	25/25
Sensitivity	100%	3/3
Specificity	100%	22/22
Positive Predictivity	100%	3/3
Negative Predictivity	100%	22/22

b. STTA ER TA validation study: The predictive capacity was assessed using 18 test substances (8 positives and 10 negatives) possessing their candidate effects for antagonist activity. The accuracy of the ER STTA antagonist assay was calculated as

94%-100% in each lab, and the overall accuracy for the total decisions (n=18) and all data (n=65) were calculated as 100% and 97%, respectively (9).

Table 10: Predictive Capacity for the STTA ER TA (Antagonist Assay) (9)

		STTA ER TA (IC30 based)		
		Positive	Negative	Total
Candidate effect	Positive	8	0	8
	Negative	0	10	10
	Total	8	10	18

Accuracy	100%	18/18
Sensitivity	100%	8/8
Specificity	100%	10/10
Positive Predictivity	100%	8/8
Negative Predictivity	100%	10/10

Table 11: Template for Accuracy Analysis

		New Test Outcome		
		Positive	Negative	Total
Reference Test Classification	Positive	a	c	a + c
	Negative	b	d	b + d
	Total	a + b	c + d	a+b+c+d

a = positive in both new assay and by reference test classification

b = positive in new assay and negative by reference test classification c = negative in new assay and positive by reference test classification d = negative in both new assay and by reference test classification Accuracy = ([a+d]/[a+b+c+d])

Sensitivity = (a/[a+c]) Specificity = (d/[b+d])

Positive Predictivity = (a/[a+b]) Negative Predictivity = (d/[c+d])