ANNEX 1

Definitions and Abbreviations

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

Accuracy (concordance): (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with "concordance" to mean the proportion of correct outcomes of a test method.

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

Antagonist: A type of receptor ligand or chemical that does not provoke a biological response itself upon binding to a receptor, but blocks or dampens agonist-mediated responses.

Anti-estrogenic activity, the capability of a chemical to suppress the action of 17β-estradiol mediated through estrogen receptors.

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

BG-1Luc4E2: The BG-1Luc4E2 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.

Cell morphology: The shape and appearance of cells grown in a monolayer in a single well of a tissue culture plate. Cells that are dying often exhibit abnormal cellular morphology.

CF: The OECD Conceptual Framework for the Screening and Testing of Endocrine Disrupting Chemicals.

Charcoal/dextran treatment: Treatment of serum used in cell culture. Treatment with charcoal/dextran (often referred to as "stripping") removes endogenous hormones and hormone-binding proteins.

Cytotoxicity: the harmful effects to cell structure or function ultimately causing cell death and can be a result of 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.

CV: Coefficient of variation

DCC-FBS: Dextran-coated charcoal treated fetal bovine serum.

DMEM: Dulbecco's Modification of Eagle's Medium

DMSO: Dimethyl sulfoxide

E2: 17β-estradiol

EC_{50}: The half maximal effective concentration of a test substance.

ED: Endocrine Disruption

EE: 17α-ethynyl estradiol

EFM: Estrogen-free medium. Dulbecco's Modification of Eagle's Medium (DMEM) supplemented with 4.5% charcoal/dextran-treated FBS, 1.9% L-glutamine, and 0.9% Pen-Strep.

ER: Estrogen receptor
ERE: Estrogen response element

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

FBS: Fetal bovine serum

HeLa: An immortal human cervical cell line

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

hERα: Human estrogen receptor alpha

hERß: Human estrogen receptor beta

LEC: Lowest effective concentration is the lowest concentration of test substance that produces a threshold response (*i.e.* the lowest test substance concentration at which the fold induction is statistically different from the concurrent vehicle control).
IC50: The half maximal effective concentration of an inhibitory test substance.
ICCVAM: The Interagency Coordinating Committee on the Validation of Alternative Methods.

Interlaboratory 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. Interlaboratory reproducibility is determined during the prevalidation and validation processes, and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory reproducibility: The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Me-too test: A colloquial expression for a test methods that is structurally and functionally similar to a validated and accepted reference test method. Such a test method would be a candidate for catch-up validation. Interchangeably used with similar test method.

MT: Metallothionein

MMTV: Mouse Mammary Tumor Virus

OHT: 4-Hydroxytamoxifen

PBTG: Performance-Based Test Guideline.

PC: Positive control (1 nM of E2)

PC_{10}: the concentration of a test chemical at which the measured activity in an agonist assay is 10% of the maximum activity induced by the PC (E2 at 1nM for the STTA assay) in each plate.

PC_{50}: the concentration of a test chemical at which the measured activity in an agonist assay is 50% of the maximum activity induced by the PC (E2 at the reference concentration specified in the test method) in each plate.

PC_{Max}: the concentration of a test chemical inducing the RPCMax

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 substances 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.

Proficiency chemicals (substances): A subset of the 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 twenty two chemicals to be used to demonstrate the ability of a new test method to meet the acceptability criteria demonstrated by the validated reference test methods. These chemicals representative the classes of chemicals for which ER agonism is commonly observed, and represents the full range of potencies (e.g., EC_{50}, PC_{50}) that may be expected for ER agonists (e.g, strong to weak) along with negatives.

Reference estrogen (Positive control, PC): The reference estrogen, 17ß-estradiol (E2, CAS 50-28-2). (single concentration for PC10/50 test chemical)

Reference standard: a reference substance used to demonstrate the adequacy of a test method. 17β-estradiol is the estrogenic reference standard for the BG1Luc ER TA.

Reference test method: The test methods upon which this PBTG is based.

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.

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.

RLU: Relative Light Units

RNA: Ribonucleic Acid

RPCMax: maximum level of response induced by a test chemical, expressed as a percentage of the response induced by 1 nM E2 on the same plate

RPMI: RPMI 1640 medium supplemented with 0.9% Pen-Strep and 8.0% fetal bovine serum (FBS)

RT PCR: Real Time polymerase chain reaction

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.

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.

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 Transcriptional Activation Assay, the ERα transcriptional activation assay using the HeLA 9903 Cell Line.

Substance: Used in the context of the UN GHS (1) 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: Transcriptional activation.

Threshold response: The lowest level of reporter response that is statistically different from that of the concurrent vehicle control (*i.e.* the response that corresponds to the LEC).

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.

Validated test method: An accepted test method for which validation studies have been completed to determine the accuracy and reliability of the method for a specific proposed use.

Validation, a process based on scientifically sound principles by which the reliability and relevance of a particular test, approach, method, or process are established for a specific purpose. Reliability is defined as the extent of reproducibility of results from a test within and among laboratories over time, when performed using the same standardised protocol. The relevance of a test method describes the relationship between the test and the effect in the target species and whether the test method is meaningful and useful for a defined purpose, with the limitations identified. In brief, it is the extent to which the test method correctly measures or predicts the (biological) effect of interest, as appropriate (16).

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

Stably Transfected Human Estrogen Receptor-α Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals using the hERα-HeLa-9903 cell line

INITIAL CONSIDERATIONS AND LIMITATIONS (See also General Introduction)

1. This transcriptional activation (TA) assay uses the hERα-HeLa-9903 cell line to detect estrogenic agonist activity mediated through human estrogen receptor alpha (hERα). The validation study of the Stably Transfected Transactivation Assay (STTA) by the Japanese Chemicals Evaluation and Research Institute (CERI) using the hERα-HeLa-9903 cell line to detect estrogenic agonist activity mediated through human estrogen receptor alpha (hERα) demonstrated the relevance and reliability of the assay for its intended purpose (1).

2. This test method is specifically designed to detect hERα-mediated TA by measuring chemiluminescence as the endpoint. However, non-receptor-mediated luminescence signals have been reported at phytoestrogen concentrations higher than 1 μM due to the over-activation of the luciferase reporter gene (2) (3). While the dose-response curve indicates that true activation of the ER system occurs at lower concentrations, luciferase expression obtained at high concentrations of phytoestrogens or similar compounds suspected of producing phytoestrogen-like over-activation of the luciferase reporter gene needs to be examined carefully in stably transfected ER TA assay systems (Appendix 1).

3. The general introduction, performance results from the validation of the TA assays and the common elements for all test methods should be read before using this test method for regulatory purposes. Definitions and abbreviations used in this TG are described in Annex 1.

PRINCIPLE OF THE TEST METHOD (See also General Introduction)

4. The assay is used to signal binding of the estrogen receptor with a ligand. Following ligand binding, the receptor-ligand complex translocates to the nucleus where it binds specific DNA response elements and transactivates a firefly luciferase reporter gene, resulting in increased cellular expression of luciferase enzyme. Luciferin is a substrate that is transformed by the luciferase enzyme to a bioluminescence product that can be quantitatively measured with a luminometer. Luciferase activity can be evaluated quickly and inexpensively with a number of commercially available test kits.

5. The test system utilises the hERα-HeLa-9903 cell line, which is derived from a human cervical tumor, with two stably inserted constructs: (i) the hERα expression construct (encoding the full-length human receptor), and (ii) a firefly luciferase reporter construct bearing five tandem repeats of a vitellogenin Estrogen-Responsive Element (ERE) driven by a mouse metallothionein (MT) promoter TATA element. The mouse MT TATA gene construct has been shown to have the best performance, and so is commonly used. Consequently this hERα-HeLa-9903 cell line can measure the ability of a test chemical to induce hERα-mediated transactivation of luciferase gene expression.

6. Data interpretation for this assay is based upon whether or not the maximum response level induced by a test chemical equals or exceeds an agonist response equal to 10% of that induced by a maximally

inducing (1 nM) concentration of the positive control (PC) 17β estradiol (E2) (*i.e.* the PC10). Data analysis and interpretation are discussed in greater detail in paragraphs 30- 40.

PROCEDURE

Cell Lines

7. The stably transfected hERα-HeLa-9903 cell line should be used for the assay. The cell line can be obtained from the Japanese Collection of Research Bioresources (JCRB) Cell Bank[1], upon signing a Material Transfer Agreement (MTA).

8. Only cells characterised as mycoplasma-free should be used in testing. RT PCR (Real Time Polymerase Chain Reaction) is the method of choice for a sensitive detection of mycoplasma infection (4) (5) (6).

Stability of the cell line

9. To monitor the stability of the cell line, E2, 17α-estradiol, 17α-methyltestosterone, and corticosterone should be used as the reference chemicals and a complete concentration-response curve in the test concentration range provided in Table 1 should be measured at least once each time the assay is performed, and the results should be in agreement with the results provided in Table 1.

Cell Culture and Plating Conditions

10. Cells should be maintained in Eagle's Minimum Essential Medium (EMEM) without phenol red, supplemented with 60 mg/L of antibiotic Kanamycine and 10% dextran-coated-charcoal-treated fetal bovine serum (DCC-FBS), in a CO_2 incubator (5% CO_2) at 37±1°C. Upon reaching 75-90% confluency, cells can be subcultured at 10 mL of $0.4 \times 10^5 - 1 \times 10^5$ cells/mL for 100 mm cell culture dish. Cells should be suspended with 10% FBS-EMEM (which is the same as EMEM with DCC-FBS) and then plated into wells of a microplate at a density of 1×10^4 cells/100 μL/well. Next, the cells should be pre-incubated in a 5% CO_2 incubator at 37°±1°C for 3 hours before the chemical exposure. The plastic-ware should be free of estrogenic activity.

11. To maintain the integrity of the response, the cells should be grown for more than one passage from the frozen stock in the conditioned media and should not be cultured for more than 40 passages. For the hERα-HeLa-9903 cell line, this will be less than three months.

12. The DCC-FBS can be prepared as described in Appendix 2, or obtained from commercial sources.

Acceptability Criteria

Positive and Negative Reference Chemicals

13. Prior to and during the study, the responsiveness of the test system should be verified using the appropriate concentrations of a strong estrogen: E2, a weak estrogen (17α-estradiol), a very weak agonist (17α-methyltestosterone) and a negative compound (corticosterone). Acceptable range values derived from the validation study are given in Table 1 (1). These 4 concurrent reference chemicals should be included with each experiment and the results should fall within the given acceptable limits. If this is not the case, the cause for

[1] JCRB Cell Bank : National Institute of Biomedical Innovation, 7-6-8 Asagi Saito, Ibaraki-shi, Osaka 567-0085, Japan Fax: +81-72-641-9812

the failure to meet the acceptability criteria should be determined (*e.g.* cell handling, and serum and antibiotics for quality and concentration) and the assay repeated. Once the acceptability criteria have been achieved, to ensure minimum variability of EC_{50}, PC_{50} and PC_{10} values, consistent use of materials for cell culturing is essential. The four concurrent reference chemicals, which should be included in each experiment (conducted under the same conditions including the materials, passage level of cells and technicians), can ensure the sensitivity of the assay because the PC10s of the three positive reference chemicals should fall within the acceptable range, as should the PC_{50}s and EC_{50}s where they can be calculated (see Table 1).

Table 1. Acceptable range values of the 4 reference chemicals for the STTA assay (means ± 2 standard deviations)(SD).

Name	$logPC_{50}$	$logPC_{10}$	$logEC_{50}$	Hill slope	Test range
17β-Estradiol (E2) CAS No: 50-28-2	-11.4 ~ -10.1	<-11	-11.3 ~ -10.1	0.7 ~ 1.5	10^{-14} ~ 10^{-8} M
17α-Estradiol CAS No: 57-91-0	-9.6 ~ -8.1	-10.7 ~ -9.3	-9.6 ~ -8.4	0.9 ~ 2.0	10^{-12} ~ 10^{-6} M
Corticosterone CAS No: 50-22-6	–	–	–	–	10^{-10} ~ 10^{-4}M
17α-Methyltestosterone CAS No: 58-18-4	-6.0 ~ -5.1	-8.0 ~ -6.2	–	–	10^{-11} ~ 10^{-5} M

Positive and Vehicle Controls

14. The positive control (PC) (1 nM of E2) should be tested at least in triplicate in each plate. The vehicle that is used to dissolve a test chemical should be tested as a vehicle control (VC) at least in triplicate in each plate. In addition to this VC, if the PC uses a different vehicle than the test chemical, another VC should be tested at least in triplicate on the same plate with the PC.

Fold-induction

15. The mean luciferase activity of the PC (1 nM E2) should be at least 4-fold that of the mean VC on each plate. This criterion is established based on the reliability of the endpoint values from the validation study (historically between four- and 30-fold).

16. With respect to the quality control of the assay, the fold-induction corresponding to the PC10 value of the concurrent PC (1 nM E2) should be greater than 1+2SD of the fold-induction value (=1) of the concurrent VC. For prioritisation purposes, the PC10 value can be useful to simplify the data analysis required compared to a statistical analysis. Although a statistical analysis provides information on significance, such an analysis is not a quantitative parameter with respect to concentration-based potential, and so is less useful for prioritisation purposes.

Chemicals to Demonstrate Laboratory Proficiency

Vehicle

17. Dimethyl sulfoxide (DMSO), or appropriate solvent, at the same concentration used for the different positive and negative controls and the test chemicals should be used as the concurrent VC. Test substances

should be dissolved in a solvent that solubilizes that test substance and is miscible with the cell medium. Water, ethanol (95% to 100% purity) and DMSO are suitable vehicles. If DMSO is used, the level should not exceed 0.1% (v/v). For any vehicle, it should be demonstrated that the maximum volume used is not cytotoxic and does not interfere with assay performance.

Preparation of Test Chemicals

18. Generally, the test chemicals should be dissolved in DMSO or other suitable solvent, and serially diluted with the same solvent at a common ratio of 1:10 in order to prepare solutions for dilution with media.

Solubility and Cytotoxicity: Considerations for Range Finding.

19. A preliminary test should be carried out to determine the appropriate concentration range of chemical to be tested, and to ascertain whether the test chemical may have any solubility and cytotoxicity problems. Initially, chemicals are tested up to the maximum concentration of 1 µl/mL, 1 mg/mL, or 1 mM, whichever is the lowest. Based on the extent of cytotoxicity or lack of solubility observed in the preliminary test, the first definite run should test the chemical at log-serial dilutions starting at the maximum acceptable concentration (*e.g.* 1 mM, 100 µM, 10 µM, etc.) and the presence of cloudiness or precipitate or cytotoxicity noted. Concentrations in the second, and if necessary third run should be adjusted as appropriate to better characterise the concentration-response curve and to avoid concentrations which are found to be insoluble or to induce excessive cytotoxicity.

20. For ER agonists, the presence of increasing levels of cytotoxicity can significantly alter or eliminate the typical sigmoidal response and should be considered when interpreting the data. Cytotoxicity testing methods that can provide information regarding 80% cell viability should be used, utilising an appropriate assay based upon laboratory experience.

21. Should the results of the cytotoxicity test show that the concentration of the test substance has reduced the cell number by 20% or more, this concentration is regarded as cytotoxic, and the concentrations at or above the cytotoxic concentration should be excluded from the evaluation.

Chemical Exposure and Assay Plate Organisation

22. The procedure for chemical dilutions (Steps-1 and 2) and exposure to cells (Step-3) can be conducted as follows:

Step-1: Each test chemical should be serially diluted in DMSO, or appropriate solvent, and added to the wells of a microtitre plate to achieve final serial concentrations as determined by the preliminary range finding test (typically in a series of, for example 1 mM, 100 µM, 10 µM, 1 µM, 100 nM, 10 nM, 1 nM, 100 pM, and 10 pM (10^{-3}-10^{-11} M)) for triplicate testing.

Step-2: Chemical dilution: First dilute 1.5 µL of the test chemical in the solvent to a concentration of 500 µL of media.

Step-3: Chemical exposure of the cells: Add 50 µL of dilution with media (prepared in Step-2) to an assay well containing 10^4 cells/100 µL/well.

The recommended final volume of media required for each well is 150 µL.

Test samples and reference chemicals can be assigned as shown in Table 3.

Table 3.: Example of plate concentration assignment of the reference chemicals in the assay plate

Row	17α-Methyltestosterone			Corticosterone			17α-Estradiol			E2		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 μM)	→	→	100 μM	→	→	1 μM	→	→	10 nM	→	→
B	conc 2 (1 μM)	→	→	10 μM	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	1 μM	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	100 nM	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	10 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	1 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	100 pM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	→	→	PC	→	→	→	→	→

Plate controls = VC: Vehicle control (DMSO); PC: Positive control (1 nM E2)

23. The reference chemicals (E2, 17α-Estradiol, 17α-methyl testosterone and corticosterone) should be tested in every run (Table 3). PC wells treated with 1 nM of E2 that can produce maximum induction of E2 and VC wells treated with DMSO (or appropriate solvent) alone should be included in each test assay plate (Table 4). If cells from different sources (*e.g.* different passage number, different lot, etc.,) are used in the same experiment, the reference chemicals should be tested for each cell source.

Table 4.: Example of plate concentration assignment of test and plate control chemicals in the assay plate

Row	Test Chemical 1			Test Chemical 2			Test Chemical 3			Test Chemical 4		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 μM)	→	→	1 mM	→	→	1 μM	→	→	10 nM	→	→
B	conc 2 (1 μM)	→	→	100 μM	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	10 μM	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	1 μM	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	100 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	10 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	1 nM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	→	→	PC	→	→	→	→	→

24. The lack of edge effects should be confirmed, as appropriate, and if edge effects are suspected, the plate layout should be altered to avoid such effects. For example, a plate layout excluding the edge wells can be employed.

25. After adding the chemicals, the assay plates should be incubated in a 5% CO_2 incubator at 37±1°C for 20-24 hours to induce the reporter gene products.

26. Special considerations will need to be applied to those compounds that are highly volatile. In such cases, nearby control wells may generate false positives, and this should be considered in light of expected and historical control values. In the few cases where volatility may be of concern, the use of "plate sealers" may help to effectively isolate individual wells during testing, and is therefore recommended in such cases.

27. Repeat definitive tests for the same chemical should be conducted on different days, to ensure independence.

Luciferase assay

28. A commercial luciferase assay reagent [*e.g.* Steady-Glo® Luciferase Assay System (Promega, E2510, or equivalents)] or a standard luciferase assay system (Promega, E1500, or equivalents) can be used for the assay, as long as the acceptability criteria is met. The assay reagents should be selected based on the sensitivity of the luminometer to be used. When using the standard luciferase assay system, Cell Culture Lysis Reagent (Promega, E1531, or equivalents) should be used before adding the substrate. The luciferase reagent should be applied following the manufacturers' instructions.

ANALYSIS OF DATA

29. To obtain the relative transcriptional activity to PC (1 nM of E2), the luminescence signals from the same plate can be analysed according to the following steps (other equivalent mathematical processes are also acceptable):

Step 1. Calculate mean value for the VC.

Step 2. Subtract the mean value of the VC from each well value to normalise the data.

Step 3. Calculate the mean for the normalised PC.

Step 4. Divide the normalised value of each well in the plate by the mean value of the normalised PC (PC=100%).

The final value of each well is the relative transcriptional activity for that well compared to the PC response.

Step 5. Calculate the mean value of the relative transcriptional activity for each concentration group of the test chemical. There are two dimensions to the response: the averaged transcriptional activity (response) and the concentration at which the response occurs (see following section).

EC_{50}, PC_{50} and PC_{10} induction considerations

30. The full concentration-response curve is required for the calculation of the EC_{50}, but this may not always be achievable or practical due to limitations of the test concentration range (for example due to cytotoxicity or solubility problems). However, as the EC_{50} and maximum induction level (corresponding to the top value of the Hill-equation) are informative parameters, these parameters should be reported where possible. For the calculation of EC_{50} and maximum induction level, appropriate statistical software should be used (*e.g.* Graphpad Prism statistical software).

31. If the Hill's logistic equation is applicable to the concentration response data, the EC_{50} should be calculated by the following equation (7):

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10 \exp ((\log EC_{50} - X) \times \text{Hill slope}))$$

Where:

X is the logarithm of concentration; and,

Y is the response and Y starts at the Bottom and goes to the Top in a sigmoid curve.

Bottom is fixed at zero in the Hill's logistic equation.

32. For each test chemical, the following should be provided:

(i) The RPCMax which is the maximum level of response induced by a test chemical, expressed as a percentage of the response induced by 1 nM E2 on the same plate, as well as the PCMax (concentration associated with the RPCMax); and

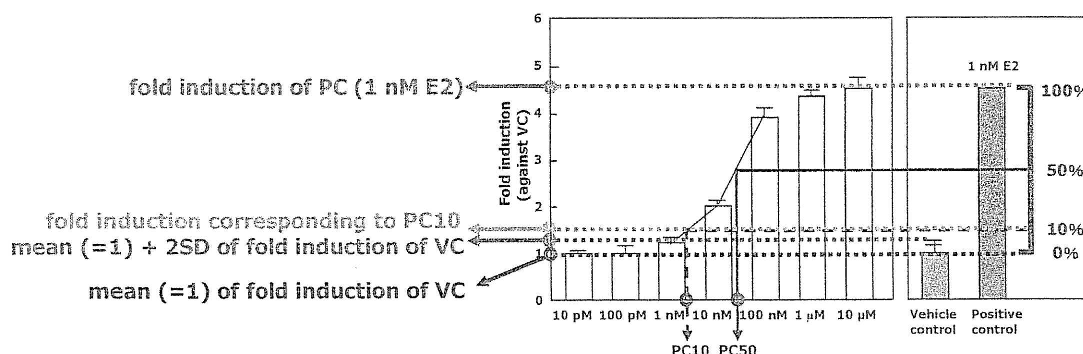
(ii) For positive chemicals, the concentrations that induce the PC10 and, if appropriate, the PC50.

33. The PCx value can be calculated by interpolating between 2 points on the X-Y coordinate, one immediately above and one immediately below a PCx value. Where the data points lying immediately above and below the PCx value have the coordinates (a,b) and (c,d) respectively, then the PCx value may be calculated using the following equation:

$$\log[PCx] = \log[c]+(x-d)/(d-b)$$

34. Descriptions of PC values are provided in Figure 1 below.

Figure 1: Example of how to derive PC-values. The PC (1 nM of E2) is included on each assay plate



35. The results should be based on two (or three) independent runs. If two runs give comparable and therefore reproducible results, it is not necessary to conduct a third run. To be acceptable, the results should:

- Meet the performance standard requirements:
    - The mean luciferase activity of the PC (1 nM E2) should be at least 4-fold that of the mean VC on each plate
    - The fold induction corresponding to the PC10 value of the concurrent PC (1 nM E2) should be greater than 1+2SD of the fold induction value (=1) of the VC.
    - The results of 4 reference chemicals should be within the acceptable range (Table 1).
- Be reproducible.

Data Interpretation Criteria

Table 5. : Positive and negative decision criteria

Positive	If the RPCMax is obtained that is equal to or exceeds 10% of the response of the positive control in at least two of two or two of three runs.
Negative	If the RPCMax fails to achieve at least 10% of the response of the positive control in two of two or two of three runs.

36. Data interpretation criteria are shown in Table 5. Positive results will be characterised by both the magnitude of the effect and the concentration at which the effect occurs. Expressing results as a concentration at which a 50% (PC50) or 10% (PC10) of PC values are reached accomplishes both of these goals. However, a test chemical is determined to be positive, if the maximum response induced by the test chemical (RPCMax) is equal to or exceeds 10% of the response of the PC in at least two of two or two of three runs, while a test chemical is considered negative if the RPCMax fails to achieve at least 10% of the response of the positive control in two of two or two of three runs.

37. The calculations of PC10, PC50 and PCMax can be made by using a spreadsheet available with the Test Guideline on the OECD public website[2].

38. It should be sufficient to obtain PC10 or PC50 values at least twice. However, should the resulting base-line for data in the same concentration range show variability with an unacceptably high coefficient of variation (CV; %) the data may not be considered reliable and the source of the high variability should be identified. The CV of the raw data triplicates (*i.e.* luminescence intensity data) of the data points that are used for the calculation of PC10 should be less than 20%.

39. Meeting the acceptability criteria indicates the assay system is operating properly, but it does not ensure that any particular run will produce accurate data. Duplicating the results of the first run is the best insurance that accurate data were produced, see paragraphs 41 and 42.

40. Where more information is required in addition to the screening and prioritisation purposes of this TG for positive test compounds, particularly for PC10-PC49 chemicals, as well as chemicals suspected to over-stimulate luciferase, it can be confirmed that the observed luciferase-activity is solely an ERα-specific response, using an ERα antagonist (see Appendix 1).

TEST REPORT

41. See paragraph 20 of the Common Elements to all methods.

[2] [http://www.oecd.org/env/testguidelines]

LITERATURE

1. CERI (2006), Draft validation report of TA assay using HeLa-hER-9903 to detect estrogenic activity. [Available at: http://www.oecd.org/document/62/0,3343,en_2649_34377_2348606_1_1_1_1,00.html]
2. Escande, A., Pillon, A., Servant, N., Cravedi, J.P., Larrea, F., Muhn, P., Nicolas, J.C., Cavailles, V. and Balaguer, P. (2006), Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta. *Biochem. Pharmacol.,* 71, 1459-1469.
3. Kuiper, G.G., Lemmen, J.G., Carlsson, B., Corton, J.C., Safe, S.H., van der Saag, P.T., van der Burg, B. and Gustafsson, J.A. (1998), Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinol.,* 139, 4252-4263.
4. Spaepen, M., Angulo, A.F., Marynen, P. and Cassiman, J.J. (1992), Detection of bacterial and mycoplasma contamination in cell cultures by polymerase chain reaction. *FEMS Microbiol Lett.* 78(1), 89-94.
5. Kobayashi, H., Yamamoto, K., Eguchi, M., Kubo, M., Nakagami, S., Wakisaka, S., Kaizuka, M. and Ishii H (1995), Rapid detection of mycoplasma contamination in cell cultures by enzymatic detection of polymerase chain reaction (PCR) products. *J. Vet. Med. Sci.* 57(4), 769-71.
6. Dussurget, O. and Roulland-Dussoix D. (1994), Rapid, sensitive PCR-based detection of mycoplasmas in simulated samples of animal sera. *Appl. Environ. Microbiol.* 60(3), 953-9.
7. De Lean, A., Munson, P.J. and Rodbard D. (1978), Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose-response curves. *Am. J. Physiol.* 235, E97-E102.

Appendix 1

False positives: Assessment of non-receptor mediated luminescence signals

1. False positives might be generated by non-ER-mediated activation of the luciferase gene, or direct activation of the gene product or unrelated fluorescence. Such effects are indicated by an incomplete or unusual dose-response curve. If such effects are suspected, the effect of an ER antagonist (*e.g.* 4-hydroxytamoxifen (OHT) at non-toxic concentration) on the response should be examined. The pure antagonist ICI 128780 may not be suitable for this purpose as a sufficient concentration of ICI 128780 may decrease the VC value, and this will affect the data analysis.

2. To ensure validity of this approach, the following needs to be tested in the same plate:

- Agonistic activity of the unknown chemical with / without 10 µM of OHT
- VC (in triplicate)
- OHT (in triplicate)
- 1 nM of E2 (in triplicate) as agonist PC
- 1 nM of E2 + OHT (in triplicate)

3. ***Data interpretation criteria***

Note: All wells should be treated with the same concentration of the vehicle.

- If the agonistic activity of the unknown chemical is NOT affected by the treatment with ER antagonist, it is classified as "Negative".
- If the agonistic activity of the unknown chemical is completely inhibited, apply the decision criteria.
- If the agonistic activity at the lowest concentration is equal to, or is exceeding, PC10 response the unknown chemical is inhibited equal to or exceeding PC10 response. The difference in the responses between the non-treated and treated wells with the ER antagonist is calculated and this difference should be considered as the true response and should be used for the calculation of the appropriate parameters to enable a classification decision to be made.

4. ***Data analysis***

Check the performance standard.
Check the CV between wells treated under the same conditions.

1. Calculate the mean of the VC
2. Subtract the mean of VC from each well value **not** treated with OHT
3. Calculate the mean of OHT
4. Subtract the mean of the VC from each well value treated with OHT
5. Calculate the mean of the PC
6. Calculate the relative transcriptional activity of all other wells relative to the PC.

Appendix 2

Preparation of Serum treated with Dextran Coated Charcoal (DCC)

1. The treatment of serum with dextran-coated charcoal (DCC) is a general method for removal of estrogenic compounds from serum that is added to cell medium, in order to exclude the biased response associated with residual estrogens in serum. 500 mL of fetal bovine serum (FBS) can be treated by this procedure.

Components

2. The following materials and equipment will be required:

Materials

Activated charcoal
Dextran
Magnesium chloride hexahydrate ($MgCl_2 \cdot 6H_2O$)
Sucrose
1 M HEPES buffer solution (pH 7.4)
Ultrapure water produced from a filter system

Equipment

Autoclaved glass container (size should be adjusted as appropriate)
General Laboratory Centrifuge (that can set temperature at 4°C)

Procedure

3. The following procedure is adjusted for the use of 50 mL centrifuge tubes:

[Day-1] Prepare dextran- coated charcoal suspension with 1 L of ultrapure water containing 1.5 mM of $MgCl_2$, 0.25 M sucrose, 2.5 g of charcoal, 0.25 g dextran and 5 mM of HEPES and stir it at 4°C, overnight.

[Day-2] Dispense the suspension in 50 mL centrifuge tubes and centrifuge at 10000 rpm at 4°C for 10 minutes. Remove the supernatant and store half of the charcoal sediment at 4°C for the use on Day-3. Suspend the other half of the charcoal with FBS that has been gently thawed to avoid precipitation, and heat-inactivated at 56°C for 30 minutes, then transfer into an autoclaved glass container such as an Erlenmeyer flask. Stir this suspension gently at 4°C, overnight.

[Day-3] Dispense the suspension with FBS into centrifuge tubes for centrifugation at 10000 rpm at 4°C for 10 minutes. Collect FBS and transfer into the new charcoal sediment prepared and stored on Day-2. Suspend the charcoal sediment and stir this suspension gently in an autoclaved glass container at 4°C, overnight.

[Day-4] Dispense the suspension for centrifugation at 10000 rpm at 4°C for 10 minutes and sterilise the supernatant by filtration through 0.2 μm sterile filter. This DCC treated FBS should be stored at -20°C and can be used for up a year.

Annex 3

BG1Luc Estrogen Receptor Transactivation Test Method for Identifying Estrogen Receptor Agonists and Antagonists

INITIAL CONSIDERATIONS AND LIMITATIONS (See also General Introduction)

1. This assay uses the BG1Luc4E2 cell line. It has been validated by the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) (1). The BG1Luc cell lines endogenously predominantly express ERα and a minor amount of ERβ (2) (3) (4).

2. This assay is applicable to a wide range of substances, provided they can be dissolved in dimethyl sulfoxide (DMSO; CASRN 67-68-5*)*, do not react with DMSO or the cell culture medium, and are not cytotoxic. The demonstrated performance of the BG1Luc ER TA agonist test method suggests that data generated with this test method may inform upon ER mediated mechanisms of action and could be considered for prioritization of substances for further testing.

3. This test method is specifically designed to detect hERα and hERß-mediated TA by measuring chemiluminescence as the endpoint. Chemiluminescence use in bioassays is widespread because luminescence has a high signal-to-background ratio (10). However, the activity of firefly luciferase in cell-based assays can be confounded by compounds that inhibit the luciferase enzyme, causing both apparent inhibition or increased luminescence due to protein stabilization (10). In addition, in some luciferase-based ER reporter gene assays, non-receptor-mediated luminescence signals have been reported at phytoestrogen concentrations higher than 1 µM due to the over-activation of the luciferase reporter gene (9) (11). While the dose-response curve indicates that true activation of the ER system occurs at lower concentrations, luciferase expression obtained at high concentrations of phytoestrogens or similar compounds suspected of producing phytoestrogen-like over-activation of the luciferase reporter gene needs to be examined carefully in stably transfected ER TA assay systems (see Annex 2)

4. The general introduction, performance results from the validation of the TA assays and the common elements for all test methods should be read before using this test method for regulatory purposes. Definitions and abbreviations used in this TG are described in Annex 1.

PRINCIPLE OF THE TEST METHOD (See also General introduction)

5. The assay is used to indicate ER ligand binding, followed by translocation of the receptor-ligand complex to the nucleus. In the nucleus, the receptor-ligand complex binds to specific DNA response elements and transactivates the reporter gene (*luc*), resulting in the production of luciferase and the subsequent emission of light, which can be quantified using a luminometer. Luciferase activity can be quickly and inexpensively evaluated with a number of commercially available kits. The BG1Luc ER TA utilizes an ER responsive human ovarian adenocarcinoma cell line, BG-1, which has been stably transfected with a firefly *luc* reporter construct under control of four estrogen response elements placed upstream of the mouse mammary tumor virus promoter (MMTV), to detect substances with *in vitro* ER agonist activity. This MMTV promoter exhibits only minor cross-reactivity with other steroid and non-steroid hormones (8). Criteria for data interpretation are

described in detail in paragraph 37. Briefly, a positive response is identified by a concentration-response curve containing at least three points with nonoverlapping error bars (mean ± SD), as well as a change in amplitude (normalized relative unit [RLU]) of at least 20% of the maximal value for the reference substance (17β-estradiol [E2; CASRN 50-28-2].

PROCEDURE

Cell Line

6. The stably transfected BG1Luc4E2 Cell line should be used for the assay. The cell line is available with a technical licensing agreement from the University of California, Davis, California, USA[3], and from Xenobiotic Detection Systems Inc., Durham, North Carolina, USA[4].

Stability of the Cell Line

7. To maintain the stability and integrity of the cell line, the cells should be grown for more than one passage from the frozen stock in cell maintenance media (see paragraph 9). Cells should not be cultured for more than 30 passages. For the BG1Luc4E2 cell line, 30 passages will be approximately three months.

Cell Culture and Plating Conditions

8. Procedures specified in the Guidance on Good Cell Culture Practice (5) (6) should be followed to assure the quality of all materials and methods in order to maintain the integrity, validity, and reproducibility of any work conducted.

9. BG1Luc4E2 Cells are maintained in RPMI 1640 medium supplemented with 0.9% Pen-Strep and 8.0% fetal bovine serum (FBS) in a dedicated tissue culture incubator at 37°C ± 1°C, 90% ± 5% humidity, and 5.0% ± 1% CO_2/air.

10. Upon reaching ~80% confluence, BG1Luc4E2 cells are subcultured and conditioned to an estrogen-free environment for 48 hours prior to plating the cells in 96-well plates for exposure to test substances and analysis of estrogen dependent induction of luciferase activity. The estrogen-free medium (EFM) contains Dulbecco's Modification of Eagle's Medium (DMEM) without phenol red, supplemented with 4.5% charcoal/dextran-treated FBS, 1.9% L-glutamine, and 0.9% Pen-Strep. All plasticware should be free of estrogenic activity [see detailed protocol (7)].

Acceptability Criteria

11. Acceptance or rejection of a test is based on the evaluation of reference standard and control results from each experiment conducted on a 96-well plate. Each reference standard is tested in multiple concentrations and there are multiple samples of each reference and control concentration. Results are compared to quality controls (QC) for these parameters that were derived from the historical databases generated by each laboratory during the demonstration of proficiency. The historical databases are updated with reference standard and control values on a continuous basis. Changes in equipment or laboratory conditions may necessitate generation of updated historical databases.

[3] UC Davis, Office of Research Technology Transfer Services 1850 Research Park Drive, Suite 100 Davis, CA 95618. (530) 754-8649 e-mail address to be added

[4] Xenobiotic Detection Systems, Inc. 1601 E Geer St # S Durham, NC 27704. (919) 688-4804

Range Finder Test

12. Acceptability criteria for the range finder test are as follows:

- Induction: Plate induction should be measured by dividing the average highest E2 reference standard relative light unit (RLU) value by the average DMSO control RLU value. Five-fold induction is usually achieved, but for purpose of acceptance, induction should be greater than or equal to four-fold.
- DMSO control results: Solvent control RLU values should be within 2.5 times the standard deviation of the historical solvent control mean RLU value.
- An experiment that fails either acceptance criterion should be discarded and repeated.

Comprehensive Test

13. It should include acceptability criteria from the range finder test and the following:

- Reference standard results: The E2 reference standard concentration-response curve should be sigmoidal in shape and have at least three values within the linear portion of the concentration-response curve.
- Positive control results: Methoxychlor control RLU values should be greater than the DMSO mean plus three times the standard deviation from the DMSO mean.
- An experiment that fails any single acceptance criterion will be discarded and repeated.

Reference Standards, Positive, and Vehicle Controls

Vehicle Control

14. The vehicle that is used to dissolve the test substances should be tested as a vehicle control. The vehicle used during the validation of the BG1Luc method was 1% (v/v) dimethylsulfoxide (DMSO, CASRN 67-68-5) (see paragraph 21). If a vehicle other than DMSO is used, all reference standards, controls, and test substances should be tested in the same vehicle. .

Reference Standard (Range Finder)

15. The reference standard is E2 (CASRN 50-28-2). For range finder testing, the reference standard is comprised of a serial dilution of four concentrations of E2 (1.84×10^{-10}, 4.59×10^{-11} , 1.15×10^{-11} and 2.87×10^{-12} M), with each concentration tested in duplicate wells.

Reference Standard (Comprehensive)

16. E2 for comprehensive testing is comprised of a 1:2 serial dilution consisting of 11 concentrations (ranging from 3.67×10^{-10} to 3.59×10^{-13} M) of E2 in duplicate wells.

Weak Positive Control

17. The weak positive control is 9.06×10^{-6} M *p,p'*-methoxychlor (methoxychlor; CASRN 72-43-5) in EFM.

Fold-Induction

18. The induction of luciferase activity of the reference standard (E2) is measured by dividing the average highest E2 reference standard RLU value by the average DMSO control RLU value, and the result should be greater than four-fold.

Demonstration of Laboratory Proficiency

19. To demonstrate proficiency with the BG1Luc ER TA test method, a laboratory should compile historical databases with reference standard and control data generated from at least 10 independant experiments, conducted on different days. These experiments are the foundation for reference standards and the historical controls. Future acceptable results should be added to enlarge the database. A successful demonstration of proficiency will be achieved by producing values that are no more than 2.5 standard deviations of the historical controls (see paragraph 11).

20. Once the historical databases are compiled, the proficiency substances listed in Table 2 of the Common Elements for all test methods, should be tested.

Vehicle

21. Test Substances should be dissolved in a solvent that solubilizes the test substance and is miscible with the cell medium. Water, ethanol (95% to 100% purity) and DMSO are suitable vehicles. If DMSO is used, the level should not exceed 1% (v/v). For any vehicle, it should be demonstated that the maximum volume used is not cytotoxic and does not interfere with assay performance. Reference standards and controls are dissolved in 100% solvent and then diluted down to appropriate concentrations in EFM.

Preparation of Test Substances

22. The test substances are dissolved in 100% DMSO (or appropriate solvent), and then diluted down to appropriate concentrations in EFM. All test substances should be allowed to equilibrate to room temperature before being dissolved and diluted. Test substance solutions should be prepared fresh for each experiment.. Solutions should not have noticeable precipitate or cloudiness. Reference standard and control stocks may be prepared in bulk; however, final reference standard, contrl dilutions and test substances should be freshly prepared for each experiment abd used within 24 hours of preparation.

Solubility and Cytotoxicity: Considerations for Range Finding

23. Range finder testing consists of seven point, 1:10 serial dilutions run in duplicate. Initially, test substances are tested up to the maximum concentration of 1 mg/mL (~1 mM). Range finder experiments are used to determine the following:

- Test substance starting concentrations to be used during comprehensive testing
- Test substance dilutions (1:2 or 1:5) to be used during comprehensive testing

24. An assessment of cell viability/cytotoxicity is included in the test method protocol (7) and is incorporated into range finder and comprehensive testing. The cytotoxicity method that was used to assess cell viability during the validation of the BG1Luc ER TA (1) was a scaled qualitative visual observation method, however, a quantitative method for the determination of cytotoxicity can be used (see protocol (7). Data from test substance concentrations that cause more than 20% reduction in viability cannot be used.

Test Substance Exposure and Assay Plate Organization

25. Cells are counted and plated into 96-well tissue culture plates (2×10^5cells per well) in EFM and incubated for 24 hours to allow the cells to attach to the plate. The EFM is removed and replaced with test and

reference chemicals and incubated for 19-24 hours. Special considerations will need to be applied to those compounds that are highly volatile since nearby control wells may generate false positive results. In such cases, "plate sealers" may help to effectively isolate individual wells during testing, and is therefore recommended in such cases.

Range Finder Tests

26. Range finder testing uses all wells of the 96-well plate to test up to six substances as seven point 1:10 serial dilutions in duplicate (see Figure 1). It uses four concentrations of E2 (Table 1) in duplicate as the reference standard and four replicate wells for the DMSO control.

Figure 1 Range Finder Test 96-well Plate Layout

	1	2	3	4	5	6	7	8	9	10	11	12
A	TS1-1	TS1-1	TS2-1	TS2-1	TS3-1	TS3-1	TS4-1	TS4-1	TS5-1	TS5-1	TS6-1	TS6-1
B	TS1-2	TS1-2	TS2-2	TS2-2	TS3-2	TS3-2	TS4-2	TS4-2	TS5-2	TS5-2	TS6-2	TS6-2
C	TS1-3	TS1-3	TS2-3	TS2-3	TS3-3	TS3-3	TS4-3	TS4-3	TS5-3	TS5-3	TS6-3	TS6-3
D	TS1-4	TS1-4	TS2-4	TS2-4	TS3-4	TS3-4	TS4-4	TS4-4	TS5-4	TS5-4	TS6-4	TS6-4
E	TS1-5	TS1-5	TS2-5	TS2-5	TS3-5	TS3-5	TS4-5	TS4-5	TS5-5	TS5-5	TS6-5	TS6-5
F	TS1-6	TS1-6	TS2-6	TS2-6	TS3-6	TS3-6	TS4-6	TS4-6	TS5-6	TS5-6	TS6-6	TS6-6
G	TS1-7	TS1-7	TS2-7	TS2-7	TS3-7	TS3-7	TS4-7	TS4-7	TS5-7	TS5-7	TS6-7	TS6-7
H	E2-1	E2-2	E2-3	E2-4	VC	VC	VC	VC	E2-1	E2-2	E2-3	E2-4

Abbreviations: E2-1 to E2-4 = concentrations of the E2 reference standard (from high to low); TS1-1 to TS1-7 = concentrations (from high to low) of test substance 1 (TS1); TS2-1 to TS2-7 = concentrations (from high to low) of test substance 2 (TS2); TS3-1 to TS3-7 = concentrations (from high to low) of test substance 3 (TS3); TS4-1 to TS4-7 = concentrations (from high to low) of test substance 4 (TS4); TS5-1 to TS5-7 = concentrations (from high to low) of test substance 5 (TS5); TS6-1 to TS6-7 = concentrations (from high to low) of test substance 2 (TS2); VC = vehicle control (DMSO [1% v/v EFM.]).

27. The recommended final volume of media required for each well is 200 µL. Only use test plates in which the cells in all wells give a viability of 80% and above.

28. Determination of starting concentrations for comprehensive testing are described in depth in the protocol (7). Briefly, the following criteria are used:

- If there are no points on the test substance concentration curve that are greater than the mean plus three times the standard deviation of the DMSO control, comprehensive testing will be conducted using an 11-point 1:2 serial dilution starting at the maximum soluble concentration.

- If there are points on the test substance concentration curve that are greater than the mean plus three times the standard deviation of the DMSO control), the starting concentration to be used for the 11-point dilution scheme in comprehensive testing should be one log higher than the concentration giving the highest adjusted RLU value in the range finder. The 11-point dilution scheme will be based on either 1:2 or 1:5 dilutions according to the following criteria: