DRAFT OECD GUIDELINE FOR THE TESTING OF CHEMICALS

The Cytosensor Microphysiometer Test Method: An *in vitro* Method for Identifying Ocular Corrosive and Severe Irritant Chemicals as well as Chemicals not Classified as Ocular Irritants

INTRODUCTION

1. The Cytosensor Microphysiometer (CM) test method is an *in vitro* test method that can be used under certain circumstances and with specific limitations for hazard classification and labelling of chemicals for eye corrosion/irritation, according to the United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS) (1). For the purpose of this Test Guideline, severe irritants are defined as chemicals that cause tissue damage in the eye following test chemical administration that is not reversible within 21 days or causes serious physical decay of vision, while ocular corrosives are chemicals that cause irreversible tissue damage to the eye. These chemicals are classified as UN GHS Category 1. Chemicals not classified as an eye irritant are defined as those that do not meet the requirements for classification as UN GHS Category 1, 2A or 2B. While the CM test method is not considered valid as a complete replacement for the *in vivo* rabbit eye test, the CM is recommended for use as part of a tiered testing strategy for regulatory classification and labelling. Thus, the CM is recommended as an initial step within a Top-Down approach to identify ocular corrosives/severe irritants (UN GHS Category 1), as well as an initial step within a Bottom-Up approach to identify chemicals that do not require classification for eye corrosion/irritation (UN GHS Not Classified), specifically for limited types of chemicals (*i.e.* water soluble surfactants and surfactant-containing mixtures) (4)(5).

2. It is currently generally accepted that, in the foreseeable future, no single *in vitro* eye irritation test will be able to replace the *in vivo* Draize eye test to predict across the full range of irritation for different chemical classes. However, strategic combinations of several alternative test methods within a (tiered) testing strategy may be able to replace the Draize eye test (5). The Top-Down approach (5) is designed to be used when, based on existing information, a chemical is expected to have high irritancy potential, while the Bottom-Up approach (5) is designed to be used when, based on existing information, a chemical is expected not to cause sufficient eye irritation to require a classification. Based on the prediction model detailed in paragraph 29, the CM test method can identify substances and mixtures within a limited applicability domain as ocular corrosives/severe irritants (UN GHS Category 1) or as not classified for eye corrosion/irritation (UN GHS Not Classified) without any further testing. Therefore, the CM test method may be used as a partial replacement test for *in vivo* acute eye irritation/corrosion testing to determine the eye irritancy/corrosivity of chemicals, following the sequential testing strategy of TG 405 (6). However, a chemical that is not predicted as ocular corrosive/severe irritant or as not classified for eye corrosion/irritation with the CM test method would require additional testing (*in vitro* and/or *in vivo*) to establish a definitive classification. The CM is so far the only validated *in vitro* test method that can be used to identify chemicals not classified as eye irritants but it is not considered adequately valid for the identification of mild or moderate ocular irritants (*i.e.* UN GHS Categories 2A and 2B).

3. The purpose of this Test Guideline is to describe the procedures used to evaluate the potential for ocular corrosivity or irritancy of a test chemical as measured by its ability to induce changes in cellular metabolism which occur after chemical exposure. The CM test method estimates the decrease in metabolic rate of the cells exposed to the test chemical by measuring the rate of change in pH of the medium (acidity) per unit time as compared to the basal metabolic state. The reduction of the metabolic rate of the exposed cells can be used to estimate the ocular toxicity potential of a test chemical. Annex I provides diagrams of the operating components of the CM. The capsule insert

described in the Cytosensor Manual and which can be seen in the second diagram of Annex I must not be used in the assay (8).

4. Performance Standards (provided in Annex III) have been developed to facilitate the validation of new similar test methods and allow for timely amendment of this test guideline so that new similar test methods can be added if the Performance Standards are met.

5. Definitions are provided in Annex II.

INITIAL CONSIDERATIONS AND LIMITATIONS

6. This Test Guideline is based on INVITTOX protocol No. 130 (7) that has been evaluated in an international validation study by the European Centre for the Validation of Alternative Methods (ECVAM) (8), in collaboration with the US Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) (9) and the Japanese Center for the Validation of Alternative Methods (JaCVAM).

7. The CM test method is not recommended for the identification of mild/moderate irritant chemicals (substances and mixtures), *i.e.* GHS Cat. 2A/2B, as demonstrated by the validation study (4, 8).

8. The CM test method is used for the testing of only water soluble chemicals (substances and mixtures) as well as non water-soluble solids, viscous chemicals or suspensions that maintain uniformity during analysis time, i.e. that do not settle or separate into more than one phase. However, different solubility criteria apply depending on the use given to the test method and the classification system used. More specifically, the following solubility criteria apply:

i) in order to identify a chemical as "not classified for eye irritation" (UN GHS Not Classified) the test chemical must form a single phase solution/suspension in Low-Buffered Medium during analysis time at a concentration > 10 mg/mL for the UN GHS classification and labelling systems (C&L) (1 dose above the cut-off of 10 mg/mL), or at a concentration > 80 mg/mL for the U.S. EPA C&L (1 dose above the cut-off of 80 mg/mL). If at the maximum achievable concentration ≥ 10 mg/mL (UN GHS C&L) or ≥ 80 mg/mL (U.S. EPA C&L) the test chemical still induces a relative cellular metabolic rate > 50%, it can still be identified as UN GHS Not Classified (MRD_{50} must be reported as > 10 mg/mL) or as U.S. EPA Category IV (MRD_{50} must be reported as > 80 mg/mL), respectively.

ii) in order to identify a chemical as "inducing serious eye damage" (UN GHS Category 1) or as an "ocular corrosive or severe irritant" (U.S. EPA Category I) the test chemical must form a single phase solution/suspension in Low-Buffered Medium during analysis time at a concentration ≥ 2 mg/mL (1 dose above the cut-off of 2 mg/mL). However, if the test chemical forms a single phase solution/suspension in Low-Buffered Medium during analysis time only at a concentration ≤ 2 mg/mL, but it already induces a relative metabolic rate < 50% at that concentration (MRD_{50} < 2 mg/mL), it can still be classified as UN GHS Category 1 or as U.S. EPA Category I.

iii) if a test chemical only forms a single phase solution/suspension in Low-Buffered Medium during analysis time at a concentration < 10 mg/mL (UN GHS) or < 80 mg/mL (U.S. EPA C&L) and the relative cellular metabolic rate is still > 50% at that

concentration, the test chemical should be declared as "unsuitable for testing" for the specific classification(s) system(s).

9. Other identified limitations of the test method are based on false negative and false positive rates. When used as an initial step within a Top-Down approach to identify ocular corrosives/severe irritants (UN GHS Category 1), the CM test method is considered suitable for water soluble chemicals (substances and mixtures). When testing water soluble chemicals the false positive rate for the CM test method ranged from 2% (1/48; UN GHS) to 8.5% (4/47; U.S. EPA) and the false negative rate ranged from 20.5% (7/34; UN GHS) to 26.66% (8/30; U.S. EPA) when compared to *in vivo* results.

10. When used as an initial step within a Bottom-Up approach to identify chemicals that do not require classification for eye corrosion/irritation (UN GHS Not Classified;U.S. EPA Category IV), the CM test method is considered suitable only for water soluble surfactants and surfactant containing mixtures. When testing water soluble surfactants and surfactant containing mixtures, the false negative rate for the CM test method ranged from 0% (0/28; UN GHS) to 2% (1/46; U.S. EPA) and the false positive rate ranged from 50% (3/6; U.S. EPA) to 68% (17/25; UN GHS) when compared to *in vivo* results. Because of high false negative and false positive rates for the CM test method when testing water soluble non-surfactant chemicals, it is not recommended for identifying these types of chemicals as not requiring ocular hazard classification. Surfactant-containing pesticide mixtures were not included in the test method validation chemical set and therefore negative results (UN GHS Not Classified, U.S. EPA Category IV) obtained with this type of mixtures should be considered with caution and substantiated by further supporting information.

11. The current applicability domain might be increased in some cases, but only after analyzing an expanded data set of studied test chemicals, preferably acquired through testing (4). This Test Guideline will be updated accordingly as new information and data are considered.

PRINCIPLE OF THE TEST

12. The CM test method is a cytotoxicity and cell-function based *in vitro* assay that is performed on a sub-confluent monolayer of adherent mouse L929 fibroblasts cultured in a sensor chamber using a pH-meter to detect changes in acidity. The L929 cells were selected because they readily attach to the polycarbonate filter membrane of the transwell and are easy to grow in continuous culture. Although data from the use of other cell types, e.g. normal human keratinocytes, has been reported, virtually all safety studies conducted with the CM instrument since the early 1990's have used L929 cells. Use of normal human keratinocytes with the transwell is not recommended. Mechanistically, the CM test method is intended to model the cytotoxic action of an irritant chemical on the cell membranes of the corneal and conjunctival epithelium where the test chemical would reside in an *in vivo* exposure (10).

13. The CM estimates the metabolic rate of a population of cells maintained in low volume flow-through chambers by measuring the rate of excretion of acid by-products and the resulting decrease in pH of the surrounding medium. The metabolic rate is determined indirectly by the number of protons excreted into Low-Buffered Medium (for composition see paragraph 17) (change in pH) per unit time. The pH-meter forms the bottom of the flow through chamber and serves as a very sensitive and stable pH meter.

14. During the course of an experiment, test samples, prepared as dilution series of a test chemical, are introduced in order of increasing concentration to flow-through chambers containing the cells. Therefore, in the CM test method the same cell population is exposed progressively to

increasing concentrations of the test chemical. The cells cultured in the chamber are exposed to the test chemical for a short period of time, followed by a rinse step with Low-Buffered Medium (for composition see paragraph 18) to remove the test chemical. Finally the flow is stopped and the change in pH is measured. All rate of acidification measurements are made on washed cells. These three steps are repeated with increasing concentrations of the test chemical until either the highest testable concentration has been used or until the population of cells is severely damaged and the metabolic rate has declined to effectively zero.

15. The rate of change in pH per unit time becomes the metabolic rate of the population of cells. If a test chemical causes cytotoxicity to this population of cells it is assumed that the metabolic rate will fall. A transient up-regulation of glucose metabolism can occur if the cells need energy to maintain their integrity in the face of a mild biochemical insult, but it soon falls below the basal level if exposure to the cytotoxic chemical is prolonged or intensified (higher concentration). The concentration of test chemical that leads to a 50% decline in the basal metabolic rate of the population (MRD_{50}; metabolic rate decrement of 50%) is the parameter used to measure the cytotoxic effect of the test chemical on the test system (L929 mouse fibroblast cells). The MRD_{50} value (mg/mL) for each test chemical is calculated from a concentration response curve (see paragraph 28), and is used to provide a measure of the ocular irritancy potential of the test chemical.

16. Recovery is an important part of a test chemical's toxicity profile that is also assessed by the *in vivo* ocular irritation test. The CM test method is non-invasive, thus it could also be used for the determination of recovery of the cells from toxic insult. Additional data, preferably acquired by further testing, would be required to confirm this usefulness (8). This Test Guideline will be updated accordingly as new information and data are considered.

PROCEDURE

Cell maintenance and preparation of the cells for the assay

17. Stock cultures of L929 mouse fibroblasts, grown routinely in cell culture flasks, should be maintained and passaged in Dulbecco's Modified Eagle's Medium with 1.0 mM sodium pyruvate (DMEM) containing 10% Fetal Bovine Serum and 2.0 mM L-glutamine (Growth Medium), under normal growth conditions (see paragraph 18). When preparing the cells for the assay, L929 cells at or near confluency are trypsinised, centrifuged and an appropriate cell suspension is prepared and seeded in DMEM containing 1% Fetal Bovine Serum, 50 µg/mL gentamicin and 2.0 mM L-glutamine (Seeding Medium). At the time of the assay itself the cells should be 70 to 80 % confluent in the sensor chamber to allow for accurate pH readings (8). With the CM instrument, this can be achieved by seeding the cells in the capsule cups at a density of ~6 x 10^5 cells/cup (~5.36 x 10^5 cells/cm^2), and incubating them for 16-32 hours under normal growth conditions (see paragraph 18), before use. Prior to the start of the assay, the medium in sensor chambers containing the cultured L929 cells is changed to serum-free, $NaHCO_3$-free DMEM supplemented to contain 2.0 mM L-glutamine, 50 µg/mL gentamicin, and additional NaCl to preserve osmolarity (NaCl concentration will be increased from 110 mM NaCl to a final concentration of 154 mM NaCl, to substitute for the omission of $NaHCO_3$) (Low-Buffered Medium).
The data provided by the instrument are based on time-dependent changes in pH which occur as a result of cellular metabolism. Use of fully buffered medium would essentially eliminate the ability to detect the necessary level of pH changes.

18. The L929 cell cultures should be kept in incubators in a humidified atmosphere, at 5 ± 1% CO_2 and 37 ± 1ºC. The cells should be free of contamination by bacteria, viruses, mycoplasma and fungi.

Application of the Test and Control Chemicals

19. A fresh stock solution of test chemical should be prepared for each experimental run and used within 30 minutes of preparation. Test chemicals should be prepared in Low-Buffered Medium.

20. A dose range finding assay is performed to establish an appropriate test chemical dose range for the definitive toxicity test. Solutions at different concentrations are prepared by serial three-fold dilutions in sterile, Low-Buffered Medium that has been left to equilibrate to room temperature overnight. The concentrations to be tested in the dose-range finding assay are as follows: 100 mg/mL; 33.3 mg/mL; 11.1 mg/mL; 3.7 mg/mL; 1.23 mg/mL; 0.412 mg/mL and 0.137 mg/mL.. If possible, the test chemical concentration that results in the reduction of the MRD_{50} value should be calculated from the dose-range finding assay.

21. In the definitive assays seven concentrations are tested. Generally, three concentrations are chosen below the expected MRD_{50} value, one at approximately the MRD_{50} value, and three above the expected MRD_{50}. If the test chemical fails to cause 50% toxicity in the dose range finding assay, the maximum concentration used will generally be 270 mg/mL or less, based on the ability of the test chemical to form stable single phase solution/suspension in the Low-Buffered Medium during analysis time (see paragraph 8).

22. Once a concentration range which includes the MRD_{50} value has been found, the same range of concentrations should be tested at least once more, meaning that the final MRD_{50} value is estimated from the mean of at least two definitive trials A dose range finding trial may be used as a definitive trial if the requirements for a definitive trial (see paragraph 21) are met.

23. The negative control to obtain the basal metabolic rate is Low-Buffered Medium alone. A concurrent positive control should be used in each experimental run. A solvent control is recommended when a solvent other than Low-Buffered Medium is used. The maximum solvent concentration (other than Low-Buffered Medium) should normally be 10%(w/v). The suggested positive control chemical is Sodium lauryl sulphate (CAS No. 151-21-3) prepared as stock solution in deionised water (100 mg/mL), and subsequently diluted in Low-Buffered Medium for testing. A dose range finding assay should be performed once on the positive control to set the appropriate ranges for the subsequent definitive trials. Historical data for the positive control should be established in each user laboratory to ensure that the instrument provides similar readings from day-to-day, and to enable comparing data for different test chemicals tested on different days. These data should not differ substantially from previously established historical ranges. As a reference, the results compiled by the Institute for In Vitro Sciences (IIVS) for 629 assays of the positive control (SLS) conducted over a 12 plus year period as well as for the last 94 of those 629 assays conducted over a period of two years, are summarised below.

Chemical	Dates	No. of Assays	Mean MRD_{50} (mg/mL)	SD	CV (%)
SLS	April, 14 1994 – June 30, 2006	629	0.0799	0.011	14.3

Based on the Positive Control data compiled by IIVS this range is 0.0579 – 0.1019 mg/mL.

Determination of the change in metabolic rate

24. Prior to the start of the assay, the medium in sensor chambers containing the cultured L929 cells is changed to Low-Buffered Medium (see paragraph 17). The medium flow in the instrument is adjusted and obvious bubbles are cleared. While medium is flowing through the chamber, the pH is stable and governed by the medium. When the flow of medium is stopped, the pH begins to drop in a linear fashion over time.

25. It should always be ascertained that the instrument and cells are stable before the experiment can begin, since all subsequent data points are interpreted based on the baseline acidification rate (pH/s). Thus, at the beginning of each assay, at least four to five measurements are taken to assess the basal acidification rate (in the case of the validated CM instrument it is measured in μV/s), which is used as the negative control for each cell culture. For each sensor chamber, these baseline data points should vary from their mean by no more than 10%, and will be determined just prior to introduction of the first sample dilutions. If the baseline data contain one out of five outlying points that can be explained (e.g., caused by a bubble), it is permissible to delete that data point and use only four for calculations. With the validated CM instrument, baseline rates are expected to fall between 50 and 200 μV/s after a stabilization period of approximately 1 hour. The rationale is to verify that the cells in the sensor chamber are sufficient (confluence >70%) to generate a sufficiently high rate to determine the MRD_{50} with a good signal/noise (50 μV/s), and to avoid a situation with a confluence >80% together with a high pH drop (200μV/s). If a sensor chamber with cells fails to achieve these ranges it should be discarded and the cells should be replaced.

26. After the baseline data points have been taken, the cells contained in the chamber undergo cycles of exposure to the test chemical consisting of three phases (exposure, wash-out and measurement). The cycles start from the lowest concentration tested and are repeated for the increasing concentrations of the test chemical in the same cell population. When using the CM instrument, each cycle takes approximately 20 minutes. For standard safety assays the exposure time should be 810 seconds in order to match the experimental conditions for which the main prediction model was established. Longer or shorter exposure times will change the calculated MRD_{50} since toxicity is a function of exposure time. If a different exposure time is used, a conversion algorithm may have to be developed to translate the generated data to validated CM transwell data.

27. The exposure cycle that should be used with the validated CM instrument, and as specified in the validated CM protocol (INVITTOX Protocol No. 130) (7) is as follows. In the first phase of an exposure cycle, the test chemical is introduced into the sensor for 810 seconds. The nominal medium flow rate is 100 μL/min for the first minute and 20 μL/min for the remaining 12 minutes and 30 seconds. During the second phase, which lasts 6 minutes (at a flow rate of 100 μL/min), the test chemical is washed out from the sensor chamber using the Low-Buffered Medium not containing the test chemical. In the third phase the flow is stopped (0 μL/min) for 25 seconds and the rate of pH change is measured. These cycles (exposure, wash-out and measurement phases) are repeated with increasing concentrations of the test chemical until the highest concentration is reached. If a different exposure cycle is used (e.g. due to the use of a different instrument), a conversion algorithm may have to be developed to translate the generated data to validated CM transwell data.

Interpretation of results and Prediction model

28. The acidification rates that occur after exposure to each test chemical concentration are calculated and compared to the mean basal acidification rate of the same cells prior to exposure to the test chemical. The percent of control acidification rate is determined by comparing the dose response acidification rate to the basal acidification rate.

The following equation for the calculation of % control acidification rate should be applied:

$$\% \text{ of control acidification rate} = \frac{\text{acidification rate after exposure to test chemical}}{\text{basal acidification rate}} \times 100$$

The percent of control acidification rates for each concentration are then plotted against the test chemical concentrations. The concentration of the test chemical that results in a 50% reduction in acidification rate is interpolated from the obtained curve and referred to as the MRD_{50}. MRD_{50} data should be expressed in mg/mL.

29. The cut-off values of MRD_{50} for predicting chemicals as not classified as irritant, or as ocular corrosives/severe irritants are given below:

Top Down approach: Identification of **severe irritants** (for water soluble chemicals (substances and mixtures))

MRD_{50} (mg/mL)	**UN GHS C&L[1]**	**U.S. EPA C&L[3]**
> 2 mg/mL	No prediction can be made	No prediction can be made
≤ 2 mg/mL	Category 1	Category I

Bottom up approach: Identification of **non irritants** (for water soluble surfactants and surfactant containing mixtures)

MRD_{50} (mg/mL)	**UN GHS C&L[1]**	**U.S. EPA C&L[3]**
> 80 mg/mL	N/A	Category IV
≤ 80 mg/mL	N/A	No prediction can be made
> 10 mg/mL	Not Classified	N/A
≤ 10 mg/mL	No prediction can be made	N/A

C&L: classification and labeling; N/A: Not applicable for the particular classification and labelling system

The CM test method is recommended only for the identification of ocular corrosives and severe irritants (UN GHS Category 1) and for the identification of chemicals not classified as irritant (UN GHS Not Classified), within a previously stated applicability domain (see paragraphs 1, 8, 9 and 10).

In this context, the MRD_{50} cut-off value ≤ 2 mg/mL should be used for the identification of ocular corrosives and severe irritants (for UN GHS) for water soluble chemicals (substances and mixtures) (4, 7, 8), while the MRD_{50} cut-off value > 10 mg/mL (for UN GHS) or > 80 mg/mL (for U.S. EPA C&L) should be used for the identification of chemicals not classified as irritant for water soluble surfactants and water soluble surfactant-containing mixtures (4, 7, 8).

Acceptance of results

30. The test acceptance criteria are as follows:

i) acceptance of test results from a trial are dependent on acceptance of results for the positive control within the trial, which should fall within an acceptable range. A test is considered acceptable if the MRD_{50} of the positive control falls within 2 standard deviations of the historical mean. The positive control historical mean should be established for each participating laboratory based on at least a minimum number of experiments (each consisting of three definitive trials) statistically defined based on the variability of the positive control. As an example, to establish such historical mean in the VRM, the Institute for In Vitro Sciences (IIVS) compiled results from 629 assays of the positive control (SLS) conducted over a 12 plus year period and basedon those (see paragraph 23) this range is 0.0579 – 0.1019 mg/mL.

DATA AND REPORTING

Data

31. For each run, data from individual replicate measurements (*e.g.* basal acidification rate, acidification rate after exposure to a concentration of the test chemical and calculated % of control acidification rate) should be reported in tabular form. In addition means ± SD of individual replicate measurements in each run should be reported.

Test Report

32. The test report should include the following information:

Test and Control Chemicals

- Chemical name(s) such as the structural name used by the Chemical Abstracts Service (CAS), followed by other names, if known;
- Chemical CAS number, if known;
- Purity and composition of the chemical or mixture (in percentage(s) by weight), to the extent this information is available;
- Physical-chemical properties relevant to the conduct of the study (*e.g.* physical state, volatility, pH, stability, water solubility, chemical class);
- Treatment of the test/control chemical prior to testing, if applicable (*e.g.* warming, grinding);
- Storage conditions;

Justification of the Test Method and Protocol Used

- Should include considerations regarding applicability domain and limitations of the test method;

Test Conditions

- Description of cell system used, including certificate of authenticity and the mycoplasma status of the cell line;
- Details of test procedure used;
- Test chemical concentration(s) used;
- Duration of exposure to the test chemical;
- Description of any modifications of the test procedure;
- Description of evaluation criteria used;
- Reference to historical data of the model (*e.g.* negative and positive controls, solvent control, benchmark chemicals, if applicable);

- Information on the technical proficiency demonstrated by the laboratory;

Results

- Tabulation of data from individual baseline measurements, test chemicals and positive control for each trial (including individual results, means and SDs);
- The derived classification(s) with reference to the prediction model and/or decision criteria used;
- Description of other effects observed;

Discussion of the Results

- Should include considerations regarding a non-conclusive outcome (paragraph 28: "No prediction can be made") and further testing;

Conclusions

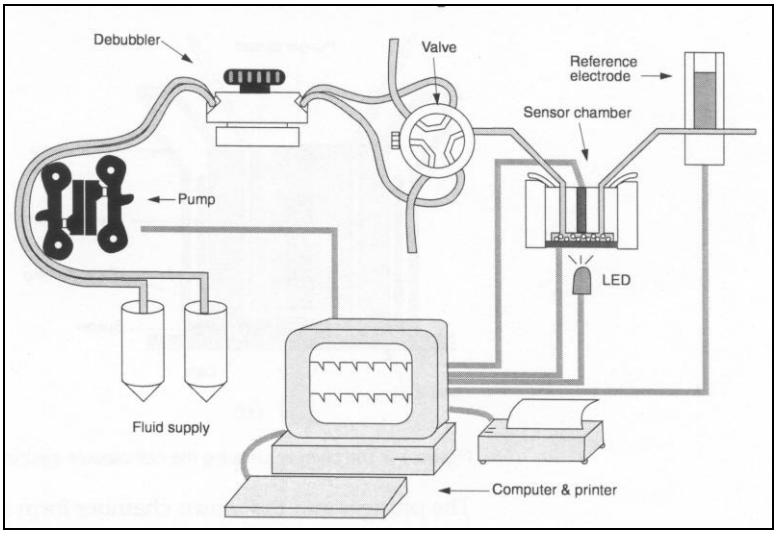
LITERATURE

1. UN (2009), United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Third revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev03/03files_e.html]
2. EC (2008), REGULATION (EC) No 1272/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Official Journal of the European Union L353, 1-1355.
3. U.S. EPA (1996), Label Review Manual: 2nd Edition. EPA737-B-96-001. Washington, DC: U.S. Environmental Protection Agency.
4. EC-ECVAM (2009), Statement on the scientific validity of cytotoxicity/cell-function based *in vitro* assays for eye irritation testing. Available under *Publications* at: [http://ecvam.jrc.it/index.htm]
5. Scott, L., Eskes, C., Hoffmann, S., Adriaens, E., Alepée, N., Bufo, M., Clothier, R., Facchini, D., Faller, C., Guest, R., Harbell, J., Hartung, T., Kamp, H., Varlet, B.L., Meloni, M., McNamee, P., Osborne, R., Pape, W., Pfannenbecker, U., Prinsen, M., Seaman, C., Spielmann, H., Stokes, W., Trouba, K., Berghe, C.V., Goethem, F.V., Vassallo, M., Vinardell, P. and Zuang, V. (2010), A proposed eye irritation testing strategy to reduce and replace *in vivo* studies using Bottom-Up and Top-Down approaches. *Toxicol. In Vitro* 24, 1-9.
6. OECD (2002), *Acute Eye Irritation/Corrosion*, OECD Guideline for the Testing of Chemicals No. 405, OECD, Paris. Available at: [http://www.oecd.org/env/testguidelines]
7. EC-ECVAM (2010), INVITTOX Protocol 130: The Cytosensor Microphysiometer (CM) Toxicity Test. Ispra, Italy: European Centre for the Validation of Alternative Methods (ECVAM). Available at: [http://ecvam-dbalm.jrc.ec.europa.eu]
8. EC-ECVAM (2008), Background Review Document of Existing Methods for Eye Irritation Testing: Silicon Microphysiometer and Cytosensor Microphysiometer. Available under *Validation Study Documents*, Section *Eye Irritation* at: [http://ecvam.jrc.it/index.htm]
9. ICCVAM (2010), ICCVAM Test Method Evaluation report: Current Validation Status of *In Vitro* Test Methods Proposed for Identifying Eye Injury Hazard Potential of Chemicals and products. NIH Publication No. 10-7553. Research Triangle Park, NC; National Institute of Environmental Health Sciences. Available at: [http://iccvam.niehs.nih.gov/methods/ocutox/MildMod-TMER.htm]
10. Harbell, J., S. W. Koontz, R. W. Lewis, D. Lovell and D. Acosta (1997), "IRAG Working Group 4: Cell Cytotoxicity Assays." *Food and Chemical Toxicology* 35, 79-126.
11. OECD (2005), Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, OECD Series on Testing and Assessment No. 34. OECD, Paris. Available at: [http://www.oecd.org/env/testguidelines]
12. EC-ECVAM (2012), Performance Standards for *in vitro* eye irritation test methods based on Cytosensor Microphysiometer (CM). Available at: [http://ecvam.jrc.ec.europa.eu/]

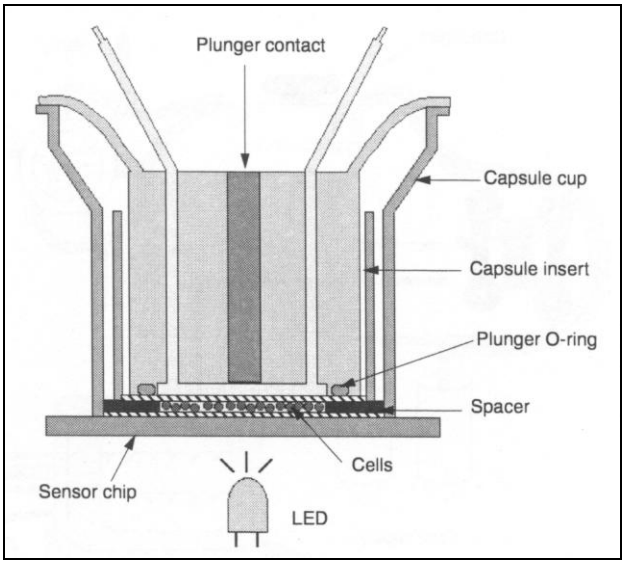
ANNEX I

DIAGRAMS OF THE OPERATING COMPONENTS OF THE CM

The operating components of the CM instrument (diagram taken from the Cytosensor Manual).



The low volume sensor chamber with the transwell in place (diagram taken from the Cytosensor Manual). The capsule insert described in the Cytosensor Manual and which can be seen in this diagram must not be used in the assay (8).



The Cytosensor uses a low volume flow through chamber and a pH-meter to measure the metabolic rate of a cell population. The pH-meter forms the bottom of the flow through chamber and serves as a very sensitive and stable pH meter. The cells are grown on a transwell membrane, which is placed into the sensor chamber and a plunger (with a spacer) pressed down on the membrane to seal it. There is a small medium-filled space between the sensor chip and the bottom of the transwell. The cells are attached to the top of the membrane so

that the acid metabolites should pass through the membrane pores to reach the space in the lower part of the chamber. The medium is passed over the cells on the upper side of the membrane.

ANNEX II

DEFINITIONS

Accuracy: The closeness of agreement between test method results and accepted reference values. 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.

EPA Category I: Chemicals that produce corrosive (irreversible destruction of ocular tissue) or corneal involvement or irritation persisting for more than 21 days (3).

False negative rate: The proportion of all positive chemicals falsely identified by a test method as negative. It is one indicator of test method performance.

False positive rate: The proportion of all negative chemicals that are falsely identified by a test method as positive. It is one indicator of test method performance.

GHS (Globally Harmonized System of Classification and Labelling of Chemicals by the United Nation (UN)): A system proposing the classification of chemicals (substances and mixtures) according to standardized types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (1).

GHS Category 1: Production of tissue damage in the eye, or serious physical decay of vision, following application of a test chemical to the anterior surface of the eye, which is not fully reversible within 21 days of application (1).

Hazard: Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

Me-too test: A colloquial expression for a test method 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.

Mixture: Used in the context of the UN GHS (1) as a mixture or solution composed of two or more substances in which they do not react.

MRD_{50}: Metabolic rate decrement of 50%. The concentration of test chemical as weight/volume %, required to reduce the acidification rate by 50%.

Negative control: An untreated replicate containing all components of a test system. This sample is processed before exposure to the test chemical, to assess the basal acidification rate.

Not-classified: Chemicals that are not classified as UN GHS Categories 1, 2A, or 2B or U.S. EPA Categories I, II, or III ocular irritants (1)(3).

Ocular corrosive: (a) A chemical that causes irreversible tissue damage to the eye. (b) Chemicals that are classified as UN GHS Category 1; or U.S. EPA Category I ocular irritants (1)(3).

Ocular irritant: (a) A chemical that produces a reversible change in the eye following application to the anterior surface of the eye; (b) Chemicals that are classified as UN GHS Categories 2A, or 2B; or U.S. EPA Categories II or III ocular irritants (1)(3).

Ocular severe irritant: (a) A chemical that causes tissue damage in the eye following application to the anterior surface of the eye that is not reversible within 21 days of application or causes serious physical decay of vision. (b) Chemicals that are classified as UN GHS Category 1; or U.S. EPA Category I ocular irritants (1)(3).

Performance standards (PS): 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; (i) essential test method components; (ii) a minimum list of Reference Chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (iii) 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.

Positive control: A replicate containing all components of a test system and treated with a chemical known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

Reference chemicals: Chemicals selected for use in the validation process, for which responses in the *in vitro* or *in vivo* reference test system or the species of interest are already known. 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. Different sets of reference chemicals may be required for the different stages of the validation process, and for different test methods and test uses (11).

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 (11).

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 and intra-laboratory repeatability.

Replacement test: A test which is designed to substitute for a test that is in routine use and accepted for hazard identification and/or risk assessment, and which has been determined to provide equivalent or improved protection of human or animal health or the environment, as applicable, compared to the accepted test, for all possible testing situations and chemicals.

Solvent control: An untreated sample containing all components of a test system, including the solvent that is processed with the test chemical-treated and other control samples to establish the baseline response for the samples treated with the test chemical dissolved in the same solvent. When tested with a concurrent negative control, this sample also demonstrates whether the solvent interacts with the test system.

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.

Surfactant: Also called surface-active agent, this is a substance, such as a detergent, that can reduce the surface tension of a liquid and thus allow it to foam or penetrate solids; it is also known as a wetting agent.

Surfactant-based formulations: In the context of this Test Guideline, it is a formulation containing one or more surfactants at a final concentration of >5%.

Tiered testing strategy: A stepwise testing strategy where all existing information on a test chemical is reviewed, in a specified order, using a weight of evidence process at each tier to determine if sufficient information is available for a hazard classification decision, prior to progression to the next tier. If the irritancy potential of a test chemical can be assigned based on the existing information, no additional testing is required. If the irritancy potential of a test chemical cannot be assigned based on the existing information, a step-wise sequential animal testing procedure is performed until an unequivocal classification can be made.

Validated test method: A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose (11).

Weight-of-evidence: The process of considering the strengths and weaknesses of various pieces of information in reaching and supporting a conclusion concerning the hazard potential of a chemical.

ANNEX III

PERFORMANCE STANDARDS FOR ASSESSMENT OF PROPOSED SIMILAR OR MODIFIED CYTOSENSOR MICROPHYSIOMETER (CM) TEST METHODS FOR EYE IRRITATION

INTRODUCTION

1. The purpose of Performance Standards (PS) is to provide the basis by which new or modified test methods, both proprietary (*i.e.* copyrighted, trademarked, registered) and non-proprietary can demonstrate to have sufficient reliability and relevance for specific testing purposes. The PS, based on validated and accepted test methods, can be used to evaluate the reliability and relevance of other analogous test methods (colloquially referred to as "me-too" test methods) that are based on similar scientific principles and measure or predict the same biological or toxic effect (9). On the other hand, modified test methods, which propose potential improvements to an approved test method, should be evaluated to determine the effect of the proposed changes on the test method's performance and the extent to which such changes affect the information available for the other components of the validation process. Depending on the number and nature of the proposed changes, the generated data and supporting documentation for those changes, they should either be subjected to the same validation process as described for a new test method, or, if appropriate, to a limited assessment of reliability and relevance using established PS (11).

2. Similar (me-too) or modified test methods proposed for use under this Test Guideline should be evaluated to determine their reliability and relevance using Reference Chemicals (Table 1). The proposed similar or modified test methods should have reliability, sensitivity, specificity and accuracy values which are comparable or better than those derived from the VRM (CM) and as described in paragraphs 8 to 10 of this Annex (Tables 2 and 3). The reliability of the new or modified test method, as well as its ability to correctly identify non-irritant and irritant chemicals, should be determined prior to its use for testing new chemicals.

3. These PS are based on the ECVAM PS (12) for evaluating the validity of new or modified CM test methods. The PS consists of (11): (i) essential test method components; (ii) recommended reference chemicals (including a set of proposed proficiency chemicals), and; (iii) defined reliability and accuracy values that the proposed test method should meet or exceed.

I. ESSENTIAL TEST METHOD COMPONENTS

4. These consist of essential structural, functional, and procedural elements of a validated test method that should be included in the protocol of a proposed, mechanistically and functionally similar or modified test method. These components include unique characteristics of the test method, critical procedural details, and quality control measures. Adherence to essential test method components will help to assure that a similar or modified proposed test method is based on the same concepts as the corresponding VRM. The essential test method components are described in detail in paragraphs 17 to 30 of the Test Guideline:

- Principle of the test (paragraphs 12 to 16)
- Cell maintenance and preparation of the cells for the assay (paragraphs 17-18)
- Application of the Test and Control Chemicals (paragraphs 19-23)
- Determination of the change in metabolic rate (paragraphs 24-27)
- Interpretation of results and Prediction model (paragraphs 28-29)
- Acceptance of results (paragraph 30)

For specific parameters (e.g. for Table 3 and 4), adequate values should be provided for any new similar or modified test method; these specific values may vary depending on the specific test method..

II. MINIMUM LIST OF REFERENCE CHEMICALS

5. Reference Chemicals are used to determine if the reliability and relevance of a proposed similar or modified test method, proven to be structurally and functionally sufficiently similar to the VRM, or representing a minor modification of the VRM, are comparable or better than the VRM. The 30 recommended Reference Chemicals listed in Table 1 include substances representing different chemical classes (*i.e.* chemical categories based on functional groups), while a separate table of only the subset of the surfactants has been prepared (Table 2). The substances included in this list comprise 11 UN GHS Category 1, 9 UN GHS Category 2A/2B and 10 "No-Category" chemicals. The substances listed in Table 1 are selected from the substances used in the validation study of the VRM, with regard to chemical functionality and physical state (8). These Reference Chemicals represent the minimum number of chemicals that should be used to evaluate the reliability and relevance of a proposed similar or modified test method able to identify Category 1 and "No-category" chemicals in accordance with the UN GHS (1). The use of these Reference Chemicals for the development/optimization of new similar test methods should be avoided to the extent possible. In situations where a listed substance is unavailable, other substances for which adequate *in vivo* reference data are available could be used, primarily from the substances used in the validation study of the VRM. If desired, additional substances representing other chemical classes and for which adequate *in vivo* reference data are available may be added to the minimum list of Reference Chemicals to further evaluate the accuracy of the proposed test method.

Table 1: Minimum list of Reference Chemicals (to be used in the Top-down approach to identify chemicals as "inducing serious eye damage" (UN GHS Category 1) or as "ocular corrosives or severe irritants" (U.S. EPA Category I),) for determination of reliability, sensitivity, specificity and accuracy balues for similar or modified CM-based eye irritation test methods*.

Substance	CAS Number	Physical state	Conc. Tested (%)	In vivo GHS/EPA	In vitro GHS/EPA[4]
Benzalkonium chloride[1,2]	8001-54-5	Liquid	10	Cat.1/ cat.I	Cat1 / CatI
Sodium lauryl sulphate[1,2]	151-21-3	Solid	15	Cat.1/ cat.I	Cat1 / CatI
Dibenzoyl-L-tartaric acid[2,3]	2743-38-6	Solid	100	Cat.1/ cat.II	Cat1 / CatI
Cetylpyridinium bromide[1,2]	140-72-7	Solid	6	Cat.1/SCNM	Cat1 / CatI
Benzalkonium chloride[1,2]	8001-54-5	Liquid	5	Cat.1/ cat.I	Cat1 / CatI
Trichloroacetic acid[2]	76-03-9	Solid	30	Cat.1/ cat.I	Cat1 / CatI
Promethazine HCl[2,3]	58-33-3	Solid	100	Cat.1/ cat.I	Cat1 / CatI
Sodium perborate, $4H_2O$[2]	10486-00-7	Solid	100	Cat.1/ cat.I	Cat1 / CatI
Triton X-100[1,2]	9002-93-1	Liquid	10	Cat.1/ cat.II	Cat1 / CatI
Sodium hydroxide[2]	1310-73-2	Solid	10	Cat.1/ cat.I	NP / NP
Triton X-100[1,2]	9002-93-1	Liquid	5	Cat.2A/ cat.III	NP / NP
Sodium lauryl sulphate[1,2]	151-21-3	Solid	3	No Cat./ cat.III	NP / NP
Benzalkonium chloride[1,2]	8001-54-5	Liquid	1	Cat.1/ cat.I	NP / NP
Tween 20[1,2]	9005-64-5	Liquid	100	No Cat./ cat.III	NP / NP
Sodium hydroxide[2]	1310-73-2	Solid	1	Cat.2B/ cat.III	NP / NP
Ethyl-2-methylacetoacetate[2]	609-14-3	Liquid	100	Cat.2B/ cat.III	NP / NP
Methyl ethyl ketone[2]	78-93-3	Liquid	100	Cat.2A/ cat.III	NP / NP
Methyl acetate[2]	79-20-9	Liquid	100	Cat.2A/ cat.II	NP / NP
Ethanol[2,3]	64-17-5	Liquid	100	Cat.2A/ cat.II-III	NP / NP
Acetone[2]	67-64-1	Liquid	100	Cat.2A/ cat.II	NP / NP
Isobutanol[2]	78-83-1	Liquid	100	Cat.2A/ cat.II	NP / NP
Ammonium nitrate[2]	6484-52-2	Solid	100	Cat.2A/ cat.III	NP / NP
Trichloroacetic acid[2]	76-03-9	Solid	3	No Cat./ cat.III	NP / NP
Triton X-100[1,2]	9002-93-1	Liquid	1	No Cat./ cat.III	No Cat / NP
Methyl isobutyl ketone[2,3]	108-10-1	Liquid	100	No Cat./ cat.III	NP / NP
Ethyl acetate[2]	141-78-6	Liquid	100	No Cat./ cat.III	NP / NP
Cetylpyridinium bromide[1,2]	140-72-7	Solid	0,1	No Cat./ cat.III	No Cat / CatIV
Glycerol[2]	56-81-5	Liquid	100	No Cat./ cat.IV	NP / NP
Propylene glycol[2]	57-55-6	Liquid	100	No Cat./ cat.IV	NP / NP
Polyethylene glycol 400[1,2]	25322-68-3	Liquid	100	No Cat./cat.IV	No Cat / CatIV

[1]Highlighted substances are surfactants which, together with the rest of the chemicals have been used to estimate the predictive capacity in a top-down approach. [2]*In vivo* study from ECETOC. [3]*In vivo* study from Gautheron et al., 1994. [4]Refer to the prediction model

*Number of laboratories where the substances were tested varied from 1-6.

NP: means no prediction could be made according to the prediction model as described in the TG, paragraph 29.

Table 2: Minimum list of Reference Chemicals (to be used in the Bottom-up approach to identify surfactants or surfactant-based formulations as "not classified for eye irritation"- UN GHS Not Classified or U.S. EPA Category IV**) for determination of reliability, sensitivity, specificity and accuracy balues for similar or modified CM-based eye irritation test methods***.

Substance	CAS Number	Physical state	Conc. Tested (%)	In vivo GHS/EPA	In vitro GHS/EPA[4]
Benzalkonium chloride[1,2]	8001-54-5	Liquid	10	Cat.1/ cat.I	NP / NP
Sodium lauryl sulphate[1,2]	151-21-3	Solid	15	Cat.1/ cat.I	NP / NP
Cetylpyridinium bromide[1,2]	140-72-7	Solid	6	Cat.1/SCNM	NP / NP
Benzalkonium chloride[1,2]	8001-54-5	Liquid	5	Cat.1/ cat.I	NP / NP
Triton X-100[1,2]	9002-93-1	Liquid	10	Cat.1/ cat.II	NP / NP
Triton X-100[1,2]	9002-93-1	Liquid	5	Cat.2A/ cat.III	NP / NP
Sodium lauryl sulphate[1,2]	151-21-3	Solid	3	No Cat./ cat.III	NP / NP
Benzalkonium chloride[1,2]	8001-54-5	Liquid	1	Cat.1/ cat.I	NP / NP
Tween 20[1,2]	9005-64-5	Liquid	100	No Cat./ cat.III	NP / NP
Triton X-100[1,2]	9002-93-1	Liquid	1	No Cat./ cat.III	No Cat / NP
Cetylpyridinium bromide[1,2]	140-72-7	Solid	0,1	No Cat./ cat.III	No Cat / CatIV
Polyethylene glycol 400[1,2]	25322-68-3	Liquid	100	No Cat./cat.IV	No Cat / CatIV

[1]Highlighted substances are surfactants used to estimate the predictive capacity in a bottom-up approach. [2]*In vivo* study from ECETOC. [3]*In vivo* study from Gautheron et al., 1994. [4]Refer to the prediction model

*Number of laboratories where the substances were tested varied from 1-6.

NP: means no prediction could be made according to the prediction model as described in the TG, paragraph 29.

III. DEFINED RELIABILITY AND ACCURACY VALUES

6. For purposes of establishing the reliability and relevance of proposed similar or modified CM-based test methods to be used by several independent laboratories, all 30 Reference Chemicals listed in Table 1 should be tested in at least three laboratories. It is however essential that all PS-based validation studies are independently assessed by internationally recognized validation bodies, in agreement with international guidelines (11). In each laboratory, three independent classifications should be obtained for each of the reference chemicals for validation purposes. Each classification is obtained from the combination of results of two to three independent and qualified trials, performed with different cell batches and at sufficiently spaced time points. A **qualified trial** consists of a test that meets the test acceptance criteria for the PC, as defined in the corresponding SOP. Otherwise, the trial is considered as non-qualified. The number of qualified trials recommended per tested chemical may be reduced from three to two if statistically/scientifically justified.

7. The calculation of the reliability, sensitivity, specificity and accuracy values of the proposed test method should be done according to the rules described below to ensure that a predefined and consistent approach is used:

1. The within-laboratory reproducibility (WLR) expresses the extent of concordance of the three **classifications** obtained for each Reference Chemical, within each participating laboratory. Each classification is decided using only two/three qualified trials. The number and identity of the Reference Chemicals which have less than three classifications should be reported.

2. The between-laboratory reproducibility (BLR) expresses the extent of concordance among participating laboratories of **final classifications** for each Reference Chemical. The final classification for each Reference Chemical in each participating laboratory is decided based on whether the first two classifications agree with each other. If not, then a third one is obtained which defines the final classification in each laboratory. The number and identity of the Reference Chemicals which have less than three classifications per laboratory should be reported.

3. The calculation of the sensitivity, specificity and accuracy values should be done using all classifications obtained for each Reference Chemical in each laboratory. The calculations should be based on the individual predictions of each classification for each Reference Chemical in each laboratory.

Each laboratory should not produce more than three qualified trials per test chemical for one classification. Excess production of data and subsequent data selection are regarded as not appropriate.

Within-laboratory reproducibility

8. Depending on the regulatory framework and the classification system applied where this method is considered valid, the test method should meet the requirements of either the UN GHS or the U.S. EPA or both classification systems, as described below:

i) When used in the Bottom-Up Approach (5) to identify surfactants or surfactant-based formulations as "not classified for eye irritation" (UN GHS Not Classified or U.S. EPA Category IV), the concordance of final classifications for the 12 RC (Table 2) obtained in different, independent

(minimum of three) experiments within a single participating laboratory must be equal or higher (≥) than 90%.

ii) When used in the Top-Down Approach (5) to identify chemicals as "inducing serious eye damage" (UN GHS Category 1), the concordance of final classifications for the 30 RC (Table 1) obtained in different, independent (minimum of three) experiments within a single participating laboratory must be equal or higher (≥) than 80%

Between-laboratory reproducibility

9. Depending on the regulatory framework and the classification system applied where this method is considered valid, the test method should meet the requirements of either the UN GHS or the U.S. EPA or both classification systems, as described below:

i) When used in the Bottom-Up Approach (5) to identify surfactants or surfactant-based formulations as "not classified for eye irritation" (UN GHS Not Classified or U.S. EPA Category IV), the concordance of final classifications for the 12 RC (Table 2) obtained between the different participating laboratories (minimum of three) must be equal or higher (≥) than 90%.

[The actual values for the validated method were: 100% (UN GHS) and 94.44% (U.S. EPA)].

ii) When used in the Top-Down Approach (5) to identify chemicals as "inducing serious eye damage" (UN GHS Category 1), the concordance of final classifications for the 30 RC (Table 1) obtained between the different participating laboratories (minimum of three) must be equal or higher (≥) than 80%

[The actual value for the validated method was: 87.62% (UN GHS)].

Predictive capacity (accuracy)

10. The calculation of the accuracy values of the test method should be done on the basis of the RC tested during the validation study. The accuracy values (sensitivity, specificity, false negative rate, false positive rate and overall accuracy) of the proposed similar or modified test method should be done using all qualifying trials obtained for each RC in each laboratory. They should be comparable to those derived from the VRM, taking into consideration additional information relating to relevance in the species of interest. Depending on the regulatory framework and the classification system applied where this method is considered valid, the test method should meet the requirements of either the UN GHS/EU CLP or the U.S. EPA or both classification systems, as described below:

i) When used in the Bottom-Up Approach (5) to identify surfactants or surfactant-based formulations as "not classified for eye irritation" (UN GHS Not Classified or U.S. EPA Category IV), the test method should comply with the following capacity predictions (Table 2):

The **sensitivity** should be equal to 100% (UN GHS).
The **sensitivity** should be equal or higher (≥) than 94% (U.S. EPA).
The **specificity** should be equal or higher (≥) than 60% (UN GHS).
The **specificity** should be equal to 100% (U.S. EPA).
Overall accuracy should be equal or higher (≥) than 83.3% (UN GHS)
Overall accuracy should be equal or higher (≥) than 94.4% (U.S. EPA).

Table 3: Required accuracy values for any similar or modified test method to be considered scientifically valid.

PREDICTIVE CAPACITY – Bottom Up Approach for surfactants					
Sensitivity		**Specificity**		**Overall Accuracy**	
GHS	EPA	GHS	EPA	GHS	EPA
100%	94%	60%	100%	83.3%	94.4%
FNR		**FPR**			
GHS	EPA	GHS	EPA		
0%	6%	40%	0%		

ii) When used in the Top-Down Approach (5) to identify chemicals as "inducing serious eye damage" (UN GHS/EU CLP Category 1) or as an "ocular corrosive or severe irritant" (U.S. EPA Category I), the test method should comply with the following capacity predictions (Table 3):

The **sensitivity** should be equal or higher (≥) than 76% (UN GHS)
The **sensitivity** should be equal or higher (≥) than 72.1% (U.S. EPA).
The **specificity** should be equal or higher (≥) than 96.8% (UN GHS)
The **specificity** should be equal or higher (≥) than 89.5% (U.S. EPA).
Overall accuracy should be equal or higher (≥) than 89.2% (UN GHS)
Overall accuracy should be equal or higher (≥) than 84.7% (U.S. EPA).

Table 4: Required accuracy values for similar or modified CM-based test method to be considered valid.

PREDICTIVE CAPACITY – Top Down Approach for all chemicals					
Sensitivity		**Specificity**		**Overall Accuracy**	
GHS	EPA	GHS	EPA	GHS	EPA
76%	72.1%	96.8%	89.5%	89.2%	84.7%
FNR		**FPR**			
GHS	EPA	GHS	EPA		
24%	27.9%	3.2%	10.5%		