

## PERFORMANCE STANDARDS FOR THE ASSESSMENT OF PROPOSED SIMILAR OR MODIFIED

### *In Vitro Macromolecular Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage*

This document contains Performance Standards (PS) for determining the reliability and relevance of similar and modified test methods for ocular hazard that are structurally and mechanistically similar to the *in vitro* macromolecular test method described in TG XXX (OECD 20xx) and the DB-ALM protocol n. 157 (1), in accordance with the principles of Guidance Document No. 34 (2). These PS include Essential Test Methods Components by which to evaluate the structural, mechanistic and procedural similarity of a new similar or modified proposed test method, as well as a list of 30 Reference Chemicals by which to evaluate the minimum reproducibility and predictive capacity necessary for the test method to be considered comparable to the validated reference method (VRM), in this PS the Ocular Irritation<sup>®</sup> assay.

## 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 be deemed to be structurally and mechanistically similar to a Validated Reference Method (VRM) and demonstrate to have sufficient reliability and relevance for specific testing purposes, in accordance with the principles of Guidance Document No. 34 (2). The PS, based on valid and accepted test method(s), can be used to evaluate the reliability and relevance of other similar 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 (2). On the other hand, modified test methods, which propose potential improvements to an approved test method, should also 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 the established PS (2).

2. These PS are proposed for evaluating the validity of new or modified *in vitro* macromolecular test method for identifying i) chemicals inducing serious eye damage (UN GHS Category 1) and ii) chemicals not requiring classification for eye irritation or serious eye damage (UN GHS No Category). As defined by the OECD GD 34 (2) the PS consists of: (i) Essential Test Method Components; (ii) Recommended Reference Chemicals, and; (iii) Defined Reliability and Accuracy Values that the proposed similar or modified test method should meet or exceed. The VRM used to develop the present PS is the Ocular Irritation<sup>®</sup> assay, as described in OECD TG XXX (OECD 20XX) and the DB-ALM protocol n. 157 (1). Definitions are provided in Annex I.

3. Similar (me-too) or modified test methods proposed for the use(s) specified in the OECD TG XXX (X) should be evaluated to determine their reliability and relevance using Reference Chemicals (table 2) representing the full range of the OECD TG 405 *in vivo* ocular hazard, *i.e.*, serious eye damage (UN GHS Cat. 1), eye irritation (UN GHS Cat. 2) and non-classified chemicals (UN GHS No Category) (3) (4). The

1 proposed similar or modified test methods should have reproducibility, sensitivity, specificity and accuracy  
2 values which are comparable or better than those derived from the VRM and as described in paragraphs  
3 18 to 26 of these PS (5). The reliability of the similar or modified test method, as well as its ability to  
4 correctly identify UN GHS No Category and UN GHS Cat. 1 chemicals, should be determined prior to its  
5 use for testing chemicals. Where possible, the classes or types of chemicals that are consistently over - or  
6 under-predicted should be defined.

## 7 **ESSENTIAL TEST METHOD COMPONENTS**

8  
9 4. The Essential Test Method Components consist of essential structural, functional, and procedural  
10 elements of the Validated Reference Method (VRM) that should be included in the protocol of a proposed,  
11 mechanistically and functionally similar or modified test method. These components include unique  
12 characteristics of the test method, critical procedural details, and quality control measures. Adherence to  
13 essential test method components will help to assure that a similar or modified proposed test method is  
14 based on the same concepts as the corresponding VRM (2). The essential test method components to be  
15 considered for similar or modified test methods related to the VRM Ocular Irritation<sup>®</sup> assay are described  
16 in detail in the following paragraphs 5 to 15). For specific parameters (e.g., for Table 1) or modified  
17 procedures, adequate values or procedures should be provided for the proposed similar or modified test  
18 method. These specific values or procedures may vary depending on the specific test method and/or its  
19 modification.

### 20 **Test Method Components**

21  
22  
23 5. The the VRM Ocular Irritation<sup>®</sup> assay for identifying i) chemicals inducing serious eye damage and  
24 ii) chemicals not requiring classification for eye irritation or serious eye damage, consists of two essential  
25 components i.e., a macromolecular matrix and a membrane disc for the controlled delivery of the test  
26 chemical to the macromolecular matrix. The macromolecular matrix serves as the target for the test  
27 chemical and is composed of a mixture of proteins, glycoproteins, carbohydrates, lipids and low molecular  
28 weight components. When hydrated, the protein oligomers within the macromolecular matrix tend to self-  
29 associate and form larger fibrils that are held together by non-covalent into a highly ordered and  
30 transparent structure presumably similar to the transparent cornea.

31 6. Test chemicals that can cause serious eye damage/eye irritation should be able to promote protein  
32 denaturation, protein unfolding and changes in conformation which result in the disruption and  
33 disaggregation of the highly organized macromolecular matrix, and produce turbidity. Such phenomena  
34 should be quantified, by e.g. measuring the changes in light scattering. In the case of the the VRM Ocular  
35 Irritation<sup>®</sup> assay, changes in optical density at 405 nm (OD<sub>405</sub>) are measured and used to determine  
36 Irritation Draize Equivalent (IDE) score for each dose of the tested chemical as described in paragraph 11  
37 of this document and in paragraph 19 of the OECD TG XXX (OECD 20XX). The highest IDE score from a  
38 qualified test run, named the Maximal Qualified Score (MQS), is used for identification of the UN GHS  
39 ocular hazard category (3) for the tested chemical based on a prediction model described in paragraph 15.

### 40 **Procedural Conditions**

#### 41 *Applicability of the test method*

42 7. The applicability of the macromolecular test method to specific chemical classes and physico-  
43 chemical properties should be well characterized, as well as the exposure procedures for specific test  
44 chemicals. In the case of the VRM Ocular Irritation<sup>®</sup> assay, most critical physico-chemical property is the  
45 pH and only test chemicals with  $4 \leq \text{pH} \leq 9$  fall within its applicability domain. Some chemicals may cause

1 interference with the Ocular Irritection® test system. However, these can be identified by the quality  
2 controls and acceptance criteria inbuilt in the VRM Ocular Irritection® assay, described in detail in  
3 paragraph 20 to 21 of the OECD TG XXX (OECD, 201X) and in paragraph 14 of this document.

#### 4 5 *Reagent preparation and activation*

6 8. The preparation of the macromolecular matrix and the necessary quality controls should be well  
7 defined. In the case of the VRM Ocular Irritection® assay, the macromolecular matrix powder is hydrated  
8 and activated with a pH lowering activating reagent to form the macromolecular matrix. The pH of the  
9 macromolecular matrix pre and post activation should fall within pre-established ranges of 7.9-8.2 and 6.4-  
10 6.7, respectively. Activator solution is also added to the blanking buffer used as a control for each test  
11 chemical dose.

#### 12 *Application of Test Chemicals*

13 9. An appropriate number of doses/concentrations of the test chemical should be tested according to  
14 the pre-defined conditions. In the case of the VRM Ocular Irritection® assay, solids and non-surfactant  
15 liquids are applied over a cellulose membrane placed on top of the macromolecular matrix, and surfactants  
16 and waxy solids are applied directly to the macromolecular matrix. A series of five doses of each test  
17 chemical are applied, as specified in OECD TG XXXX and DB-ALM protocol (1).

18 10. The macromolecular matrix should be exposed to the test chemicals and concurrent controls for a  
19 well defined duration and temperature conditions. In the case of Ocular Irritection® incubation is performed  
20 at 24.0±0.5 hours at 25±1°C in an incubator.

#### 21 *Determination of Irritation score*

22 11. Determination of an appropriate irritation score for use in a prediction model for identification of  
23 UN GHS Classification categories should be defined. In the case of the VRM Ocular Irritection® assay  
24 Irritation Draize Equivalent (IDE) score is determined for each tested dose/concentration based on the  
25 analysis of the OD<sub>405</sub> measured for the test chemical and analysed against the standard curve established  
26 using a set of 4 calibrating chemicals with well defined ranges of OD<sub>405</sub> response (paragraph 12) tested in  
27 parallel. The IDE scores are calculated by a software incorporating formulas described in the OECD TG  
28 xxx (OECD 201X). In case of the VRM OI® the highest IDE Score of five doses/concentrations obtained for  
29 the test chemical in a qualified test run (see paragraph 14), namely the Maximal Qualified Score (MQS), is  
30 selected for determination of the UN GHS ocular hazard category (3) based on pre-defined cut-off values  
31 described in Table 1.

#### 33 **Control Substances**

34 12. Appropriate controls should be tested in parallel to the test chemical, and should comprise a series  
35 of calibrators and quality control chemicals. In the case of the VRM Ocular Irritection®, these encompass  
36 four calibrating chemicals (Cal<sub>0-3</sub>) included in the assay kit covering the range of OD responses (Cal<sub>0</sub>:  
37 0.062 - 0.262; Cal<sub>1</sub>: 0.089 – 0.315; Cal<sub>2</sub>: 0.351 - 0.945; Cal<sub>3</sub>: 1.277 – 2.127) for derivation of the standard  
38 curve. Furthermore, the VRM Ocular Irritection® commercial kit includes two quality control chemicals (QC)  
39 that should result in well defined ranges of IDE scores within the lower (7.2-20.8) and mid-upper (23.6-  
40 35.6) IDE range of the VRM test.

13. In case that modified or me-too macromolecular assays involve the use of a vehicle or solvent other than distilled water with the test chemical, the vehicle or solvent should fall within the applicability domain of the macromolecular test method, and should not alter the eye irritation potential of the test chemical. When applicable, solvent (or vehicle) controls should be tested concurrently with the test chemical to demonstrate the compatibility of the solvent with the macromolecular matrix system. In the case of the VRM Ocular Irritection<sup>®</sup>, only distilled water is used as solvent for test chemicals with surfactant properties.

### Study Acceptance Criteria

14. The conditions upon which the test result is determined to be acceptable or unacceptable should be clearly defined. In the case of the VRM Ocular Irritection<sup>®</sup> assay the set of acceptance criteria are analysed automatically by the software included in the assay kit and they include: (A) predefined OD ranges for the calibrating chemicals and IDE scores for the of the Quality controls (see paragraph 12); (B) and (C) optimal Net OD for the test chemicals in relation to a minimal value (should be > -0.015), and in relation to the Cal<sub>2</sub> value (Net OD for a test chemical lower than OD<sub>Cal2</sub> triggers a check for interference with the proper response of the macromolecular matrix using an inhibition check solution; (D) Blank OD value higher than 1.2, is an indicator of interference by intensely coloured test chemicals, and (E) fitness of the dose response curves for the test chemicals in relation to the optimal dose response curve and typical dose response curves for known irritant and non-irritant chemicals. These criteria and their application for determination of a qualified result is described in more detail in paragraph 20 and 21 of the OECD TG xxx (OECD 20XXX).

### Interpretation of Results and Prediction Model

15. The methodology for interpretation of the results and the prediction model should be clearly defined. In the case of the VRM Ocular Irritection<sup>®</sup> assay, the highest IDE score obtained in a Qualified test run, named the Maximal Qualified Score (MQS) is used to predict the ocular hazard potential of the test chemical according to the UN GHS classification system (3), based on the Prediction Model described in table 1.

Table 1. Ocular Irritection<sup>®</sup> prediction model

Maximal Qualified Score (MQS)	Predicted UN GHS classification
0 – 12.5	No Category
> 12.5 – 30.0	No Prediction Can be Made
> 30.0	Category 1

### MINIMUM LIST OF REFERENCE CHEMICALS

16. 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 *in vitro* macromolecular test method VRM, or representing a minor modification of the VRM, are comparable or better to those of the VRM (5). The 30 Reference Chemicals listed in Table 2 include chemicals

1 representing different chemical classes of interest and are representative of the full range of TG 405 *in vivo*  
 2 ocular hazard, *i.e.*, serious eye damage (UN GHS Cat. 1), eye irritation (UN GHS Cat. 2) and non-  
 3 classified chemicals (UN GHS No Category) (3) (4). The distribution of chemicals in this list comprise 10  
 4 UN GHS Cat. 1 chemicals, 10 Cat. 2 chemicals and 10 No Category test chemicals. The Reference  
 5 Chemicals were selected from the test chemicals used in the validation study of the VRM (5) using the  
 6 selection criteria as described in Table 2 (foot-note A).

8 17. The 30 Reference Chemicals listed in Table 2 represent the minimum number of chemicals that  
 9 should be used to evaluate the reliability and relevance of a proposed similar or modified *in vitro*  
 10 macromolecular test method for identifying i) chemicals inducing serious eye damage (Category 1) and ii)  
 11 chemicals not requiring classification for eye irritation or serious eye damage (UN GHS No Category), in  
 12 accordance with the UN GHS (3). The use of these Reference Chemicals for the development/optimization  
 13 of new similar test methods should be avoided to the extent possible. In situations where a listed chemical  
 14 is unavailable, or where justifiable, another chemicals fulfilling the selection criteria as described in Table 2  
 15 (foot-note A) and for which adequate *in vivo* reference data are available, could be used, e.g. primarily  
 16 from the test chemicals used in the validation study of the VRM (5). Additional chemicals representing  
 17 other chemical or product classes and for which adequate *in vivo* reference data are available are  
 18 recommended to be tested in addition to the minimum list of Reference Chemicals to further evaluate the  
 19 accuracy of the proposed test method.

21 **Table 2. Minimum list of Reference Chemicals for determination of Reproducibility and Predictive**  
 22 **Capacity of similar or modified *In Vitro* macromolecular test method for identifying i) chemicals**  
 23 **inducing serious eye damage (UN GHS Category 1) and ii) chemicals not requiring classification for**  
 24 **eye irritation or serious eye damage (UN GHS No Category)**

Chemical name <sup>A</sup>	CASRN	<i>In vivo</i> UN GHS	Physical state	pH <sup>B</sup>	Organic functional group <sup>C</sup>	VRM Prediction (5)
2-benzyl-4-chlorophenol	120-32-1	Category 1	Solid	5.9	Aryl halide, Benzyl, Phenol	NPCM
2-methylresorcinol	608-25-3	Category 1	Solid	5.8	Benzyl, Phenol	Cat. 1
4-(1,1,3,3-tetramethylbutyl)phenol	140-66-9	Category 1	Solid	5.2	Alkane branched with quaternary carbon, Phenol	NPCM
4-tert-butylcatechol	98-29-3	Category 1	Solid	5.5	Phenol	Cat. 1
Benzalkonium chloride (5%)	63449-41-2	Category 1	Liquid	6.5	Ammonium quaternary (salt), Benzyl	Cat. 1*
Cetylpyridinium bromide (6%)	140-72-7	Category 1	Liquid	4.4	Heterocyclic fragment, Pyridine	Cat. 1*
Lauric acid	143-07-7	Category 1	Solid	4.5	Carboxylic acid	NPCM*
Promethazine hydrochloride	58-33-3	Category 1	Solid	4.5	Aliphatic Amine, Arene, Heterocyclic fragment, Sulfide	Cat. 1
p-tert-butylphenol	98-54-4	Category 1	Solid	7.7	Phenol	NPCM
Sodium oxalate	62-76-0	Category 1	Solid	7.0	Carboxylic acid	NPCM
2,4,11,13-tetraazatetradecanediamine, N,N"-bis(4-chlorophenyl)-3,12-	18472-51-0	Category 2A	Liquid	6.3	Alcohol, Aliphatic Amine secondary, Amidine, Arene, Aryl halide, Carboxylic acid, Imidine (substituted)	Cat. 1

Chemical name <sup>A</sup>	CASRN	<i>In vivo</i> UN GHS	Physical state	pH <sup>B</sup>	Organic functional group <sup>C</sup>	VRM Prediction (5)
diimino-, di-D-gluconic acid (20% aqueous)						
Ammonium nitrate	6484-52-2	Category 2A	Solid	4.8	n.a.	NPCM
Cetylpyridinium bromide (1%)	140-72-7	Category 2A	Liquid	4.7	Heterocyclic fragment, Pyridine	NPCM*
Methyl acetate	79-20-9	Category 2A	Liquid	6.8	Acetoxy	NPCM
Methyl cyanoacetate	105-34-0	Category 2A	Liquid	5.7	Carboxylic acid ester, Nitrile	Cat. 1*
Naphthalene-1,5-diol	83-56-7	Category 2A	Solid	5.8	Fused polycyclic aromatic, Phenol	No Cat.
Sodium lauryl glucose carboxylate (and) lauryl glucoside	383178-66-3 (110615-47-9)	Category 2A	Liquid	5.7	n.a.	Cat. 1
Propasol solvent P	1569-01-3	Category 2A	Liquid (viscous)	6.2	Alcohol, Ether	Cat. 1
Sodium benzoate	532-32-1	Category 2A	Solid	8.2	Arene, Carboxylic acid	NPCM
Sodium chloroacetate	3926-62-3	Category 2B	Solid	6.1	Alkyl halide, Carboxylic acid	NPCM
1,5-dibromopentane	111-24-0	No category	Liquid	5.7	Alkyl halide	No Cat.
2,2-dimethyl-3-pentanol	3970-62-5	No category	Liquid	5.3	Alcohol, Alkane branched with quaternary carbon	NPCM
2-(2-ethoxyethoxy)ethanol	111-90-0	No category	Liquid (viscous)	5.6	Alcohol, Ether	NPCM
Cetyl pyridinium bromide 0.1%	140-72-7	No category	Liquid	7.1	Heterocyclic fragment, Pyridine	No Cat.
Di-n-propyl disulphide	629-19-6	No category	Liquid	6.1	Disulfide	No Cat.
Diocetyl ether	629-82-3	No category	Liquid	7.1	Ether	No Cat.
Myristyl myristate	3234-85-3	No category	Solid	6.3	Carboxylic acid ester	No Cat.
n,n-dimethylguanidine sulphate	598-65-2	No category	Solid	6.8	Aliphatic Amine, tertiary, Amidine	NPCM
Potassium tetrafluoroborate	14075-53-7	No category	Solid	4.5	n.a.	No Cat.
Sodium lauryl sulfate (3%)	151-21-3	No category	Liquid	6.8	Sulfate	NPCM

1 Abbreviations: CASRN = Chemical Abstracts Service Registry Number; Cat.: category; n.a. = not available;  
2 NPCM: No Prediction Can be Made; UN GHS = United Nations Globally Harmonized System of  
3 Classification and Labelling of Chemicals (3).

4 <sup>A</sup> The 30 Reference Chemicals comprise a representative selection from the 88 chemicals that were used  
5 to validate the reference test method (Ocular Irritation<sup>®</sup>) (5). The goal of the selection process was to  
6 include, to the extent possible, chemicals that: (i) cover the full range of *in vivo* serious eye damage/eye  
7 irritation responses based on the UN GHS classification system (i.e., Categories 1, 2A, 2B or No  
8 Category); (ii) are based on high quality results obtained in the reference *in vivo* rabbit eye test (OECD TG  
9 405) (4) (6); (iii) cover different physical states; (iv) cover a broad range of the chemical classes and  
10 organic functional groups, representative of those used in the validation study (5); (v) reflect the overall  
11 performance characteristics of the reference test method; (vi) cover the full range of *in vitro* responses  
12 based on high quality Ocular Irritation<sup>®</sup> data (0 to 51 MQS); (vii) produced reproducible results in the  
13 VRM; (viii) are commercially available; and (ix) are not associated with prohibitive disposal costs.

1 <sup>B</sup> The pH values are rounded to one decimal point, and values were obtained from the original sources as  
2 indicated in (5).

3 <sup>C</sup> The organic functional groups were characterized using the OECD QSAR toolbox (version 2.3) as  
4 described in (5).

5 \* Test chemicals having limited data in within- and between- laboratory reproducibility but included as  
6 representing relevant chemistries and/or outcome. These chemicals were however not taken into account  
7 for establishing the minimum performance standards for reproducibility as described in paragraphs 23 to  
8 28.

9

## 10 **DEFINED RELIABILITY AND ACCURACY VALUES**

11

12 18. For purposes of establishing the reliability and relevance of proposed similar or modified *in vitro*  
13 macromolecular test methods to be used by several independent laboratories, all 30 Reference Chemicals  
14 listed in Table 2 should be tested in at least three laboratories. However, an assessment of between-  
15 laboratory reproducibility is not essential if the proposed test method is to be used in one laboratory only.  
16 In each laboratory, all Reference Chemicals should be tested in three independent experiments performed  
17 at sufficiently spaced time points. Each experiment should consist of at least five concurrently tested  
18 doses/concentrations for each test chemical, and appropriate controls: blanking samples, four calibrating  
19 chemicals and two Quality Control chemicals.

20

21 19. The calculation of the within-laboratory reproducibility, between-laboratory reproducibility, accuracy,  
22 sensitivity and specificity values of the proposed test method should be done according to the rules  
23 described below to ensure that a predefined and consistent approach is used:

24 - *Within-laboratory reproducibility (WLR)*: for chemicals having MQS within the 0-51 range of responses,  
25 the standard deviation obtained for the three independent experiments should be calculated, and then  
26 the overall mean of the 25 Reference Chemicals having WLR data should be calculated for each  
27 participating laboratory. In addition, the concordance of the predictions of the three independent  
28 experiments of the 25 Reference Chemicals having WLR data should be calculated for both cut-offs  
29 (12.5 and 30.0) and for each participating laboratory (possible non-qualified and excluded test results  
30 should be included in this analysis as considered integral part of the result spectrum of the VRM).

31 - *Between-laboratory reproducibility (BLR)*: for chemicals having MQS within the 0-51 range of  
32 responses, the mean MQS value is calculated for each laboratory (obtained from the three independent  
33 experiments), the standard deviation obtained for the results from at least three laboratories is then  
34 calculated, and the overall mean SD for the 25 Reference Chemicals having BLR data should be then  
35 calculated. In addition, the concordance of the predictions between at least three laboratories should be  
36 calculated for both cut-offs (12.5 and 30.0) based on the majority laboratory classification (possible non-  
37 qualified and excluded test results should be included in this analysis as considered integral part of the  
38 result spectrum of the VRM).

39 - *Predictive capacity*: should be assessed based on the majority of predictions available per chemical for  
40 all 30 Reference Chemicals. The concordance of these predictions with the expected result as defined  
41 by the *in vivo* UN GHS classifications (dichotomized into UN GHS Category 1 vs. non-Category 1  
42 chemicals and into UN GHS No Category vs. classified chemicals) should then be used to calculate the  
43 specificity, sensitivity and concordance of results.

44

45 *Within-laboratory reproducibility*

46

20. An assessment of within-laboratory reproducibility (WLR) for similar or modified test method proposed should show in terms of MQS, an overall mean standard deviation (SD) for the 25 Reference Chemicals having WLR data from three independent experiments that is smaller or equal ( $\leq$ ) than 3 within each laboratory (actual for Ocular Irritation<sup>®</sup>: 1.7, 2.2, 1.5 in each laboratory for the 25 Reference Chemicals having WLR data, and 2.0, 2.6, 2.2 for the overall validation dataset (5)).

21. In addition, to discriminate UN GHS Category 1 from non-Cat. 1 chemicals, the assessment of within-laboratory reproducibility for similar or modified test method proposed should show in every laboratory, a concordance of predictions (UN GHS Cat. 1 versus non-Cat. 1 test chemicals) obtained for the 25 Reference Chemicals having WLR data from three independent experiments that is equal or higher ( $\geq$ ) than 85% (actual for Ocular Irritation<sup>®</sup>: 96%, 88%, 92% in each laboratory for the 25 Reference Chemicals having WLR data, and 91%, 88%, 84% for the overall validation dataset (5)).

22. Finally, to discriminate UN GHS No Category from UN GHS classified chemicals (but not to categorize classified chemicals), the assessment of within-laboratory reproducibility for similar or modified test method proposed should show in every laboratory, a concordance of predictions (UN GHS No Category versus classified test chemicals) for the 25 Reference Chemicals having WLR Reference Chemicals data obtained from three independent experiments that is equal or higher ( $\geq$ ) than 80% (actual for Ocular Irritation<sup>®</sup>: 96%, 88% and 84% in each laboratory for the 25 Reference Chemicals having WLR data, and 89%, 86%, 78% for the overall validation dataset (5)).

#### *Between-laboratory reproducibility*

23. An assessment of between-laboratory reproducibility (BLR) for similar or modified test method proposed should show in terms of MQS, an overall mean standard deviation (SD) for the 25 Reference Chemicals having BLR data that is smaller or equal ( $\leq$ ) than 3 for studies conducted in three different laboratories actual for Ocular Irritation<sup>®</sup>: 1.9 for the 25 Reference Chemicals having BLR data, and 2.5 for the overall validation dataset (5)).

24. In addition, for similar or modified test methods proposed to discriminate UN GHS Category 1 from non-Cat. 1 chemicals, the between-laboratory concordance of predictions (UN GHS Cat. 1 versus non-Cat. 1 test chemicals) obtained for the 25 Reference Chemicals having BLR data should be equal or higher ( $\geq$ ) than 80% (actual for Ocular Irritation<sup>®</sup>: 84% for the 25 Reference Chemicals having BLR data, and 84% for the overall validation dataset (5)).

25. Finally, for similar or modified test methods proposed to discriminate UN GHS No Category from UN GHS classified chemicals (but not to categorize classified chemicals), the between-laboratory concordance of predictions (UN GHS No Category versus classified test chemicals) obtained for the 25 Reference Chemicals having BLR data should be equal or higher ( $\geq$ ) than 80% (actual for Ocular Irritation<sup>®</sup>: 92% for the 25 Reference Chemicals having BLR data, and 82% for the overall validation dataset (5)).

#### *Predictive capacity*

26. The predictive capacity (sensitivity, specificity, false negative rate, false positive rate, ability to correctly identify UN GHS Cat. 1 and UN GHS No Category chemicals of the proposed similar or modified *in vitro* macromolecular test methods should be comparable or better to that of the validated test method (5) (Table 3).



1 **Table 3. Required sensitivity, specificity and accuracy for similar or modified *in vitro***  
2 **macromolecular (based on the predictive capacity obtained with the Reference Test Method)<sup>1</sup>**

Purpose	Chemical set	Accuracy	Sensitivity	Specificity
Identification of UN GHS Cat. 1 chemicals	Ref. Chems	70%	50%	80%
	Validation set (88)	74%	50%	81%
Identification of UN GHS No Cat. chemicals	Ref. Chems.	83%	95%	60%
	Reduced validation set (79)	81%	98%	63%
	Validation set (88)	76%	93%	58%

3 <sup>1</sup> Table 3 provides the accuracy of the reference test method in correctly identifying i) chemicals inducing  
4 serious eye damage (UN GHS Category 1) and ii) chemicals not requiring classification for eye irritation or  
5 serious eye damage (UN GHS No Category) based on the results of the 30 Reference Chemicals (Table  
6 2).

## 1 LITERATURE

- 2 (1) DB-ALM (INVITTOX) (2013) Protocol 157: Ocular Irritection® Assay System, 25pp. Available:  
3 <http://ecvam-dbalm.jrc.ec.europa.eu>.  
4
- 5 (2) OECD (2005). OECD Series on Testing and Assessment No. 34. Guidance Document on the  
6 Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment.  
7 Available at: [[http://www.oecd.org/document/30/0,3343,en\\_2649\\_34377\\_1916638\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/30/0,3343,en_2649_34377_1916638_1_1_1_1,00.html)].  
8
- 9 (3) United Nations (UN) (2017). Globally Harmonized System of Classification and Labelling of  
10 Chemicals (GHS), Seventh revised edition, UN New York and Geneva, 2017. Available at:  
11 [https://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs\\_rev07/English/03e\\_part3](https://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev07/English/03e_part3).  
12
- 13 (4) OECD (2017). Test Guideline 405. OECD Guideline for Testing of Chemicals. Acute eye  
14 irritation/corrosion. Organisation for Economic Cooperation and Development, Paris, France. Available at:  
15 [http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-](http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788)  
16 [effects\\_20745788](http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788).  
17
- 18 (5) Eskes C, Hoffmann S, Facchini D, Ulmer R, Wang A, Flego M, Vassallo M, Bufo M, van Vliet E,  
19 d'Abrosca F, Wilt N (2014). Validation Study on the Ocular Irritection® Assay for Eye Irritation Testing.  
20 *Toxicology In Vitro* 28, 1046-1065.  
21
- 22 (6) Barroso J., Pfannenbecker U., Adriaens E., Alépée N., Cluzel M., De Smedt A., Hibatallah J., Klaric  
23 M., Mewes K.R., Millet M., Templier M., McNamee P. (2017). Cosmetics Europe compilation of historical  
24 serious eye damage/eye irritation *in vivo* data analysed by drivers of classification to support the selection  
25 of chemicals for development and evaluation of alternative methods/strategies: the Draize eye test  
26 Reference Database (DRD). *Archives of Toxicology* 91, 521-547.  
27
- 28 (7) Scott L, Eskes C, Hoffman S, Adriaens E, Alepee N, Bufo M, Clothier R, Facchini D, Faller C, Guest  
29 R, Hamernik K, Harbell J, Hartung T, Kamp H, Le Varlet B, Meloni M, Mcnamee P, Osborn R, Pape W,  
30 Pfannenbecker U, Prinsen M, Seaman C, Spielmann H, Stokes W, Trouba K, Vassallo M, Van den Berghe  
31 C, Van Goethem F, Vinardell P, Zuang V (2010). A proposed Eye Irritation Testing Strategy to Reduce and  
32 Replace *in vivo* Studies Using Bottom-up and Top-down Approaches. *Toxicology In vitro* 24,1-9.  
33
- 34 (8) OECD (2017). Guidance Document on an Integrated Approach on Testing and Assessment for  
35 Serious Eye Damage and Eye Irritation. Series on Testing and Assessment, No. 263. Environment, Health  
36 and Safety Publications, Organisation for Economic Cooperation and Development, Paris, France.  
37 Available at:  
38 [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2017\)15&doclang](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2017)15&doclang)  
39 [uage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2017)15&doclang)

## ANNEX 1

### 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 (2).

**Activator:** Solution employed to initiate formation of the ordered macromolecular matrix when the protein has been rehydrated.

**Benchmark chemical:** A chemical used as a standard for comparison to a test chemical. A benchmark chemical should have the following properties; (i), a consistent and reliable source(s); (ii), structural and functional similarity to the class of chemicals being tested; (iii), known physical/chemical characteristics; (iv) supporting data on known effects; and (v), known potency in the range of the desired response.

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

**Blank qualification:** The blank OD for each sample is checked to fall in an appropriate range (i.e., pre-established minimum and maximum blank OD), and for flatness (i.e., OD variability between two consecutive doses/concentrations, and between the highest and lowest doses/concentrations in a group of 3 doses/concentrations).

**Bottom-Up approach:** Step-wise approach used for a test chemical suspected of not requiring classification and labelling for eye irritation or serious eye damage, which starts with the determination of chemicals not requiring classification and labelling (negative outcome) from other chemicals (positive outcome) (7) (8).

**Calibrators:** defined irritant solutions having well characterized IDE scores. The calibrators are used to derive a standard curve with which the results of the test method are compared to, and ensure optimal performance.

**Cornea:** The transparent part of the front of the eyeball that covers the iris and pupil and admits light to the interior.

**Concordance:** This is a measure of test method performance for test methods that give a categorical result, and is one aspect of relevance. The term is sometimes used interchangeably with accuracy, and is defined as the proportion of all chemicals tested that are correctly classified as positive or negative. Concordance is highly dependent on the prevalence of positives in the types of test chemical being examined (2).

**CV:** Coefficient of Variation.

1 **Eye irritation:** Production of changes in the eye, which are fully reversible, occurring after the exposure of  
2 the eye to a substance or mixture. Interchangeable with "Reversible effects on the Eye" and with "UN GHS  
3 Category 2" (3).

4  
5 **False negative rate:** The proportion of all positive chemicals falsely identified by a test method as  
6 negative. It is one indicator of test method performance.

7  
8 **False positive rate:** The proportion of all negative (non-active) chemicals that are falsely identified as  
9 positive. It is one indicator of test performance.

10  
11 **Foam test:** employed to determine whether the unknown substance should be tested utilizing  
12 surfactant or non-surfactant application procedure (8).

13  
14 **Hazard:** Inherent property of an agent or situation having the potential to cause adverse effects when an  
15 organism, system or (sub) population is exposed to that agent.

16  
17 **Hydrating Solution:** Solution employed to rehydrate the reagent powder and facilitate formation of the  
18 ordered protein matrix.

19  
20 **IATA:** Integrated Approach on Testing and Assessment (8).

21  
22 **Inhibition check solution:** An irritating substance known to quickly react with the macromolecular reagent  
23 and produce evident turbidity, which can be employed to verify the functionality of macromolecular reagent  
24 when the OD readings of qualified test chemical doses/concentrations are less than Calibrator 2. The  
25 inhibition check solution verifies that the macromolecular reagent in those wells is still able to produce  
26 evident turbidity (e.g., > OD Calibrator 2) and identifies inaccurate low OD reading (or inaccurate non-  
27 irritant) results when the turbidity is less than OD Calibrator 2.

28  
29 **Irreversible effects on the eye:** See "Serious eye damage" and "UN GHS Category 1".

30  
31 **Irritation Draize Equivalent (IDE) Score:** A numerical score derived from the optical density  
32 measurement of the Ocular Irritation<sup>®</sup> assay for a tested dose/concentration when compared to the curve  
33 obtained with the calibrators.

34  
35 **Maximal Qualified Score (MQS):** Represents the highest IDE Score obtained from the different tested  
36 doses/concentrations of a test chemical. Ranging from 0 to 51 it is used to predict the irritation potential of  
37 the test chemical.

38  
39 **Membrane Discs:** A semi-permeable membrane that facilitates controlled delivery of the test chemical into  
40 the protein reagent.

41  
42 **Me-too test:** A colloquial expression for a test method that is structurally and functionally similar to a  
43 validated and accepted reference test method. Such a test method would be a candidate for catch-up  
44 validation (2). The term is interchangeably used with similar test method.

45  
46 **Mixture:** a mixture or solution composed of two or more substances in which they do not react (3).

1 **Net Optical Density Check:** Controls the net optical density by measuring the OD of the activated protein  
2 reagent and subtracting the OD of the activated blanking buffer. The Net OD (OD<sub>reagent</sub> – OD<sub>blank</sub> =  
3 OD<sub>Net</sub>) should be > -15.  
4

5 **Not Classified:** Test chemicals that are not classified for eye irritation (UN GHS Category 2, 2A, or 2B) or  
6 serious eye damage (UN GHS Category 1). Interchangeable with “UN GHS No Category”.  
7

8 **Performance standards (PS):** Standards, based on a validated test method, that provide a basis for  
9 evaluating the comparability of a proposed test method that is mechanistically and functionally similar.  
10 Included are; (i) essential test method components; (ii) a minimum list of Reference Chemicals selected  
11 from among the chemicals used to demonstrate the acceptable performance of the validated test method;  
12 and (iii) the similar levels of reliability and accuracy, based on what was obtained for the validated test  
13 method, that the proposed test method should demonstrate when evaluated using the minimum list of  
14 Reference Chemicals (2).  
15

16 **Prediction Model:** a formula or algorithm (e.g., formula, rule or set of rules) used to convert the results  
17 generated by a test method into a prediction of the (toxic) effect of interest. Also referred to as decision  
18 criteria. A prediction model contains four elements: (i) a definition of the specific purpose(s) for which the  
19 test method is to be used; (ii) specifications of all possible results that may be obtained, (iii) an algorithm  
20 that converts each study result into a prediction of the (toxic) effect of interest, and (iv) specifications as to  
21 the accuracy of the prediction model (e.g., sensitivity, specificity, and false positive and false negative  
22 rates). Prediction models are generally not used in *in vivo* ecotoxicological tests (2).  
23

24 **Predictive Capacity:** The predictive capacity reflects the test method performance in terms of correct and  
25 incorrect predictions in comparison to reference data. It gives quantitative information (e.g. correct  
26 prediction rate) on the relevance of the test method. It comprises amongst others, the sensitivity and  
27 specificity of the test method.  
28

29 **Quality Control chemicals:** Two defined irritant solutions (QC1 and QC2) with well characterized IDE  
30 scores within the lower (7.2-20.8) and mid-upper range (23.6-35.6) of the Ocular Irritation® test method.  
31 The quality control check verifies that the method is functioning properly and can correctly detect eye  
32 irritation potency in the lower and mid/upper IDE ranges.  
33

34 **Reagent Powder:** Consists of a mixture of proteins, glycoproteins, carbohydrates, lipids and low molecular  
35 weight components. When hydrated, the reagent powder forms a solution containing an ordered  
36 macromolecular matrix. Proteins in this solution undergo changes in conformation when exposed to an  
37 irritant test chemical.  
38

39 **Reference Chemicals:** Chemicals selected for use in the validation process, for which responses in the *in*  
40 *vitro* or *in vivo* reference test system or the species of interest are already known. These chemicals should  
41 be representative of the classes of chemicals for which the test method is expected to be used, and should  
42 represent the full range of responses that may be expected from the chemicals for which it may be used,  
43 from strong, to weak, to negative. Different sets of reference chemicals may be required for the different  
44 stages of the validation process, and for different test methods and test uses (2).  
45

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

- 1  
2 **Reliability:** Measures of the extent that a test method can be performed reproducibly within and between  
3 laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and  
4 inter-laboratory reproducibility (2).  
5
- 6 **Reproducibility:** The agreement among results obtained from testing the same substance using the same  
7 test protocol (2).  
8
- 9 **Reversible effects on the eye:** See "Eye irritation" and "UN GHS Category 2".  
10
- 11 **SD:** Standard Deviation.  
12
- 13 **Sensitivity:** The proportion of all positive/active chemicals that are correctly classified by the test method.  
14 It is a measure of accuracy for a test method that produces categorical results, and is an important  
15 consideration in assessing the relevance of a test method (2).  
16
- 17 **Serious eye damage:** Production of tissue damage in the eye, or serious physical decay of vision, which  
18 is not fully reversible occurring after exposure of the eye to a substance or mixture. Interchangeable with  
19 "Irreversible effects on the eye" and with "UN GHS Category 1" (3).  
20
- 21 **Solvent/vehicle control:** An untreated sample containing all components of a test system, including the  
22 solvent or vehicle that is processed with the test chemical-treated and other control samples to establish  
23 the baseline response for the samples treated with the test chemical dissolved in the same solvent or  
24 vehicle. When tested with a concurrent negative control, this sample also demonstrates whether the  
25 solvent or vehicle interacts with the test system.  
26
- 27 **Specificity:** The proportion of all negative/inactive chemicals that are correctly classified by the test  
28 method. It is a measure of accuracy for a test method that produces categorical results and is an important  
29 consideration in assessing the relevance of a test method (2).  
30
- 31 **Standard Operating Procedures (SOP):** Formal, written procedures that describe in detail how specific  
32 routine, and test-specific, laboratory operations should be performed. They are required by GLP.  
33
- 34 **Substance:** means chemical elements and their compounds in the natural state or obtained by any  
35 production process, including any additive necessary to preserve the stability of the product and any  
36 impurities deriving from the process used, but excluding any solvent which may be separated without  
37 affecting the stability of the substance or changing its composition (4).  
38
- 39 **Surfactants:** Also called surface-active agent, this is a substance and/or its dilution (in an appropriate  
40 solvent/vehicle), which consists of one or more hydrophilic and one or more hydrophobic groups, that is  
41 capable of reducing the surface tension of a liquid and of forming spreading or adsorption monolayers at  
42 the water-air interface, and/or of forming emulsions and/or microemulsions and/or micelles, and/or of  
43 adsorption at water-solid interfaces.  
44
- 45 **Test chemical:** The term "test chemical" is used to refer to what is being tested.  
46
- 47 **Tiered testing strategy:** A stepwise testing strategy, which uses test methods in a sequential manner. All  
48 existing information on a test chemical is reviewed at each tier, using a weight-of-evidence process, to  
49 determine if sufficient information is available for a hazard classification decision, prior to progression to

1 the next tier in the strategy. If the hazard potential/potency of a test chemical can be assigned based on  
2 the existing information at a given tier, no additional testing is required (7) (8).

3  
4 **Top-Down approach:** Step-wise approach used for a chemical suspected of causing serious eye damage,  
5 which starts with the determination of chemicals inducing serious eye damage (positive outcome) from  
6 other chemicals (negative outcome) (7) (8).

7  
8 **UN GHS (United Nations Globally Harmonized System of Classification and Labelling of**  
9 **Chemicals):** A system proposing the classification of chemicals (substances and mixtures) according to  
10 standardized types and levels of physical, health and environmental hazards, and addressing  
11 corresponding communication elements, such as pictograms, signal words, hazard statements,  
12 precautionary statements and safety data sheets, so that to convey information on their adverse effects  
13 with a view to protect people (including employers, workers, transporters, consumers and emergency  
14 responders) and the environment (3).

15  
16 **UN GHS Category 1:** See "Serious eye damage" and/or "Irreversible effects on the eye".

17  
18 **UN GHS Category 2:** See "Eye irritation" and/or "Reversible effects to the eye".

19  
20 **UN GHS No Category:** Chemicals that do not meet the requirements for classification as UN GHS  
21 Category 1 or 2 (2A or 2B). Interchangeable with "Not Classified".

22  
23 **Validated Reference Method(s) (VRM(s)):** one (or more) test method(s) that was(were) used to develop  
24 the related official Test Guidelines and Performance Standards (PS). The VRM(s) is(are) considered the  
25 reference test method(s) to compare new proposed similar or modified test methods in the framework of a  
26 PS-based validation study.

27  
28 **Within-laboratory reproducibility:** determination of the extent that qualified people within the same  
29 laboratory can successfully replicate results using a specific protocol at different times, also referred to as  
30 intra-laboratory reproducibility (2).