# OECD/OCDE 431

Adopted:
28 July 2015

# OECD GUIDELINES FOR THE TESTING OF CHEMICALS

## *In vitro* skin corrosion: reconstructed human epidermis (RHE) test method

### INTRODUCTION

1. Skin corrosion refers to the production of irreversible damage to the skin manifested as visible necrosis through the *epidermis* and into the *dermis*, following the application of a test chemical [as defined by the United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS)] (1). This updated Test Guideline 431 provides an *in vitro* procedure allowing the identification of non-corrosive and corrosive substances and mixtures in accordance with UN GHS (1). It also allows a partial sub-categorization of corrosives.

2. The assessment of skin corrosion potential of chemicals has typically involved the use of laboratory animals (OECD Test Guideline 404 (TG 404); originally adopted in 1981 and revised in 1992, 2002 and 2015) (2). In addition to the present TG 431, two other *in vitro* test methods for testing corrosion potential of chemicals have been validated and adopted as OECD Test Guidelines 430 (3) and 435 (4). Furthermore the *in vitro* OECD TG 439 (5) has been adopted for testing skin irritation potential. A document on Integrated Approaches to Testing and Assessment (IATA) for Skin Corrosion and Irritation describes several modules which group information sources and analysis tools, and provides guidance on (i) how to integrate and use existing testing and non-testing data for the assessment of skin irritation and skin corrosion potentials of chemicals and (ii) proposes an approach when further testing is needed (6).

3. This Test Guideline addresses the human health endpoint skin corrosion. It makes use of reconstructed human *epidermis* (RhE) (obtained from human derived non-transformed epidermal keratinocytes) which closely mimics the histological, morphological, biochemical and physiological properties of the upper parts of the human skin, *i.e.* the *epidermis*. This Test Guideline was originally adopted in 2004 and updated in 2013 to include additional test methods using the RhE modelsand the possibility to use the methods to support the sub-categorisation of corrosive chemicals, and updated in 2015 to refer to the IATA guidance document and introduce the use of an alternative procedure to measure viability.

4. Four validated test methods using commercially available RhE models are included in this Test Guideline. Prevalidation studies (7), followed by a formal validation study for assessing skin corrosion (8) (9) (10) have been conducted (11) (12) for two of these commercially available test methods, EpiSkin™ Standard Model (SM) and EpiDerm™ Skin Corrosivity Test (SCT) (EPI-200) (referred to in the following text as the Validated Reference Methods – VRMs). The outcome of these studies led to the recommendation that the two VRMs mentioned above could be used for regulatory purposes for distinguishing corrosive (C) from non-corrosive (NC) substances, and that the EpiSkin™ could moreover be used to support sub-categorization of corrosive substances (13) (14) (15). Two other commercially

available *in vitro* skin corrosion RhE test methods have shown similar results to the EpiDerm™ VRM according to PS-based validation (16) (17) (18). These are the SkinEthic™ RHE[1] and epiCS® (previously named EST-1000) that can also be used for regulatory purposes for distinguishing corrosive from non-corrosive substances (19) (20). Post validation studies performed by the RhE model producers in the years 2012 to 2014 with a refined protocol correcting interferences of unspecific MTT reduction by the test chemicals improved the performance of both discrimination of C/NC as well as supporting sub-categorisation of corrosives (21) (22).

5. Before a proposed similar or modified *in vitro* RhE test method for skin corrosion other than the VRMs can be used for regulatory purposes, its reliability, relevance (accuracy), and limitations for its proposed use should be determined to ensure its similarity to the VRMs, in accordance with the requirements of the Performance Standards (PS) (23) set out in accordance with the principles of Guidance Document No.34 (24). The Mutual Acceptance of Data will only be guaranteed after any proposed new or updated test method following the PS have been reviewed and included in this Test Guideline. The test methods included in this Test Guideline can be used to address countries' requirements for test results on *in vitro* test method for skin corrosion, while benefiting from the Mutual Acceptance of Data.

**DEFINITIONS**

6. Definitions used are provided in Annex 1.

**INITIAL CONSIDERATIONS**

7. This Test Guideline allows the identification of non-corrosive and corrosive substances and mixtures in accordance with the UN GHS (1). This Test Guideline further supports the sub-categorization of corrosive substances and mixtures into optional Sub-category 1A, in accordance with the UN GHS (1), as well as a combination of Sub-categories 1B and 1C (21) (22). A limitation of this Test Guideline is that it does not allow discriminating between skin corrosive Sub-category 1B and Sub-category 1C in accordance with the UN GHS (1) due to the limited set of well-known *in vivo* corrosive Sub-category 1C chemicals. EpiSkin™, EpiDerm™, SkinEthic™ and epiCS® test methods are able to sub-categorize (i.e. 1A versus 1B-and-1C versus NC) but differences are observed between EpiSkin™ and the three other test methods, EpiDerm™, SkinEthic™ and epiCS® in view of their capacity to provide information on sub-categorisation. Results from EpiSkin™ can be used as such; whereas results from EpiDerm™, SkinEthic™ and epiCS® generate high over-classification rates for a combination of Sub-categories 1B-and-1C (see Annex 3). Therefore, for EpiDerm™, SkinEthic™ and epiCS®, chemicals that are classified as 1B-and-1C can be considered as 1B-and-1C, while chemicals for which cell viability at 3 minutes is below 50% should just be considered as Category 1, since the Sub-category 1A predictions of these three test methods contain a high rate of over-predictions of chemicals of Sub-categories 1B-and-1C. The regulatory framework in member countries will decide how this Test Guideline will be used, e.g. acknowledging the significant probability of overclassification, a Sub-category 1A classification may still be accepted or further testing may be conducted to confirm the result.

8. A wide range of chemicals representing mainly individual substances has been tested in the validation supporting the test methods included in this Test Guideline when they are used for identification of non-corrosives and corrosives; the empirical database of the validation study amounted to 60 chemicals covering a wide range of chemical classes (8) (9) (10). Testing to demonstrate sensitivity, specificity, accuracy and within-laboratory-reproducibility of the assay for sub-categorization was performed by the

[1] The abbreviation RhE (=Reconstructed human Epidermis) is used for all models based on RhE technology. The abbreviation RHE as used in conjunction with the SkinEthic™ model means the same, but, as part of the name of this specific test method as marketed, is spelled all in capitals.

test method developers and results were reviewed by the OECD (21) (22). On the basis of the overall data available, the Test Guideline is applicable to a wide range of chemical classes and physical states including liquids, semi-solids, solids and waxes. The liquids may be aqueous or non-aqueous; solids may be soluble or insoluble in water. Whenever possible, solids should be ground to a fine powder before application; no other prior treatment of the sample is required. In cases where evidence can be demonstrated on the non-applicability of test methods included in the Test Guideline to a specific category of test chemicals, these test methods should not be used for that specific category of test chemicals. In addition, this Test Guideline is assumed to be applicable to mixtures as an extension of its applicability to substances. However, due to the fact that mixtures cover a wide spectrum of categories and composition, and that only limited information is currently available on the testing of mixtures, in cases where evidence can be demonstrated on the non-applicability of the Test Guideline to a specific category of mixtures (*e.g.* following a strategy as proposed in (25)), the Test Guideline should not be used for that specific category of mixtures. Before use of the test guideline on a mixture for generating data for an intended regulatory purpose, it should be considered whether, and if so why, it may provide adequate results for that purpose. Such considerations are not needed, when there is a regulatory requirement for testing of the mixture. Gases and aerosols have not been assessed yet in validation studies (8) (9) (10). While it is conceivable that these can be tested using RhE technology, the current Test Guideline does not allow testing of gases and aerosols.

9. Test chemicals absorbing light in the same range as MTT formazan and test chemicals able to directly reduce the vital dye MTT (to MTT formazan) may interfere with the tissue viability measurements and need the use of adapted controls for corrections. The type of adapted controls that may be required will vary depending on the type of interference produced by the test chemical and the procedure used to measure MTT formazan (see paragraphs 25-31).

10. While this Test Guideline does not provide adequate information on skin irritation, it should be noted that OECD TG 439 specifically addresses the health effect skin irritation *in vitro* and is based on the same RhE test system, though using another protocol (5). For a full evaluation of local skin effects after a single dermal exposure, the Guidance Document No. 203 on Integrated Approaches for Testing Assessment should be consulted (6). This IATA approach includes the conduct of *in vitro* tests for skin corrosion (such as described in this Test Guideline) and skin irritation before considering testing in living animals. It is recognized that the use of human skin is subject to national and international ethical considerations and conditions.

**PRINCIPLE OF THE TEST**

11. The test chemical is applied topically to a three-dimensional RhE model, comprised of non-transformed, human-derived epidermal keratinocytes, which have been cultured to form a multi-layered, highly differentiated model of the human *epidermis*. It consists of organized basal, spinous and granular layers, and a multi-layered *stratum corneum* containing intercellular lamellar lipid layers representing main lipid classes analogous to those found *in vivo.*

12. The RhE test method is based on the premise that corrosive chemicals are able to penetrate the *stratum corneum* by diffusion or erosion, and are cytotoxic to the cells in the underlying layers. Cell viability is measured by enzymatic conversion of the vital dye MTT [3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue tetrazolium bromide; CAS number 298-93-1], into a blue formazan salt that is quantitatively measured after extraction from tissues (26). Corrosive chemicals are identified by their ability to decrease cell viability below defined threshold levels (see paragraphs 35 and 36). The RhE-based skin corrosion test methods have shown to be predictive of *in vivo* skin corrosion effects assessed in rabbits according to the OECD guideline 404 (2).

**DEMONSTRATION OF PROFICIENCY**

13. Prior to routine use of any of the four validated RhE test methods that adhere to this Test Guideline, laboratories should demonstrate technical proficiency by correctly classifying the twelve Proficiency Substances listed in Table 1. In case of the use of a method for sub-classification, also the correct sub-categorization should be demonstrated. In situations where a listed substance is unavailable or where justifiable, another substance for which adequate *in vivo* and *in vitro* reference data are available may be used (e.g. from the list of reference chemicals (23)) provided that the same selection criteria as described in Table 1 is applied.

**Table 1: List of Proficiency Substances[1]**

| Substance | CASRN | Chemical Class[2] | UN GHS Cat. Based on *In Vivo* results[3] | VRM Cat. Based on *In Vitro* results[4] | MTT Reducer[5] | Physical State |
|---|---|---|---|---|---|---|
| **Sub-category 1A *In Vivo* Corrosives** | | | | | | |
| Bromoacetic acid | 79-08-3 | Organic acid | 1A | (3) 1A | -- | S |
| Boron trifluoride dihydrate | 13319-75-0 | Inorganic acid | 1A | (3) 1A | -- | L |
| Phenol | 108-95-2 | Phenol | 1A | (3) 1A | -- | S |
| Dichloroacetyl chloride | 79-36-7 | Electrophile | 1A | (3) 1A | -- | L |
| **Combination of sub-categories 1B-and-1C *In Vivo* Corrosives** | | | | | | |
| Glyoxylic acid monohydrate | 563-96-2 | Organic acid | 1B-and-1C | (3) 1B-and-1C | -- | S |
| Lactic acid | 598-82-3 | Organic acid | 1B-and-1C | (3) 1B-and-1C | -- | L |
| Ethanolamine | 141-43-5 | Organic base | 1B | (3) 1B-and-1C | Y | Viscous |
| Hydrochloric acid (14.4%) | 7647-01-0 | Inorganic acid | 1B-and-1C | (3) 1B-and-1C | -- | L |
| ***In Vivo* Non Corrosives** | | | | | | |
| Phenethyl bromide | 103-63-9 | Electrophile | NC | (3) NC | Y | L |
| 4-Amino-1,2,4-triazole | 584-13-4 | Organic base | NC | (3) NC | -- | S |
| 4-(methylthio)-benzaldehyde | 3446-89-7 | Electrophile | NC | (3) NC | Y | L |
| Lauric acid | 143-07-7 | Organic acid | NC | (3) NC | -- | S |

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; UN GHS = United Nations Globally Harmonized System (1); VRM = Validated Reference Method; NC = Not Corrosive

[1]The proficiency substances, sorted first by corrosives versus non-corrosives, then by corrosive sub-category and then by chemical class, were selected from the substances used in the ECVAM validation studies of EpiSkin™ and EpiDerm™ (8) (9) (10) and from post-validation studies based on data provided by EpiSkin™ (22), EpiDerm™, SkinEthic™ and epiCS® developers. Unless otherwise indicated, the substances were tested at the purity level obtained when purchased from a commercial source (8) (10). The selection includes, to the extent possible, substances that: (i) are representative of the range of corrosivity responses (*e.g.* non-corrosives; weak to strong corrosives) that the VRMs are capable of measuring or predicting; (ii) are representative of the chemical

classes used in the validation studies; (iii) have chemical structures that are well-defined; (iv) induce reproducible results in the VRM; (v) induce definitive results in the *in vivo* reference test method; (vi) are commercially available; and (vii) are not associated with prohibitive disposal costs.

[2]Chemical class assigned by Barratt *et al.* (8).

[3]The corresponding UN Packing groups are I, II and III, respectively, for the UN GHS 1A, 1B and 1C.

[4]The VRM *in vitro* predictions reported in this table were obtained with the EpiSkin™ and the EpiDerm™ test methods (VRMs) during post-validation testing performed by the test method developers.

[5]The viability values obtained in the ECVAM Skin Corrosion Validation Studies were not corrected for direct MTT reduction (killed controls were not performed in the validation studies). However, the post-validation data generated by the test method developers that are presented in this table were acquired with adapted controls.

14. As part of the proficiency exercise, it is recommended that the user verifies the barrier properties of the tissues after receipt as specified by the RhE model manufacturer. This is particularly important if tissues are shipped over long distance/time periods. Once a test method has been successfully established and proficiency in its use has been demonstrated, such verification will not be necessary on a routine basis. However, when using a test method routinely, it is recommended to continue to assess the barrier properties in regular intervals.

## PROCEDURE

15. The following is a generic description of the components and procedures of the RhE test methods for skin corrosion assessment covered by this Test Guideline. The RhE models endorsed as scientifically valid for use within this Test Guideline, *i.e.* the EpiSkin™ (SM), EpiDerm™ (EPI-200), SkinEthic™ RHE and epiCS® models (16) (17) (19) (27) (28) (29) (30) (31) (32), can be obtained from commercial sources. Standard Operating Procedures (SOPs) for these four RhE models are available (33) (34) (35) (36), and their main test method components are summarized in Annex 2. It is recommended that the relevant SOP be consulted when implementing and using one of these methods in the laboratory. Testing with the four RhE test methods covered by this Test Guideline should comply with the following:

## RHE TEST METHOD COMPONENTS

### *General Conditions*

16. Non-transformed human keratinocytes should be used to reconstruct the epithelium. Multiple layers of viable epithelial cells (basal layer, *stratum spinosum, stratum granulosum*) should be present under a functional *stratum corneum*. The *stratum corneum* should be multi-layered containing the essential lipid profile to produce a functional barrier with robustness to resist rapid penetration of cytotoxic benchmark chemicals, *e.g.* sodium dodecyl sulphate (SDS) or Triton X-100. The barrier function should be demonstrated and may be assessed either by determination of the concentration at which a benchmark chemical reduces the viability of the tissues by 50% ($IC_{50}$) after a fixed exposure time, or by determination of the exposure time required to reduce cell viability by 50% ($ET_{50}$) upon application of the benchmark chemical at a specified, fixed concentration (see paragraph 18). The containment properties of the RhE model should prevent the passage of material around the *stratum corneum* to the viable tissue, which would lead to poor modelling of skin exposure. The RhE model should be free of contamination by bacteria, viruses, mycoplasma, or fungi.

***Functional Conditions***

*Viability*

17. The assay used for quantifying tissue viability is the MTT-assay (26). The viable cells of the RhE tissue construct reduce the vital dye MTT into a blue MTT formazan precipitate, which is then extracted from the tissue using isopropanol (or a similar solvent). The OD of the extraction solvent alone should be sufficiently small, i.e., OD < 0.1. The extracted MTT formazan may be quantified using either a standard absorbance (OD) measurement or an HPLC/UPLC-spectrophotometry procedure (37). The RhE model users should ensure that each batch of the RhE model used meets defined criteria for the negative control. An acceptability range (upper and lower limit) for the negative control OD values should be established by the RhE model developer/supplier. Acceptability ranges for the negative control OD values for the four validated RhE test methods included in this Test Guideline are given in Table 2. An HPLC/UPLC-Spectrophotometry user should use the negative control OD ranges provided in Table 2 as the acceptance criterion for the negative control. It should be documented that the tissues treated with negative control are stable in culture (provide similar OD measurements) for the duration of the exposure period.

**Table 2: Acceptability ranges for negative control OD values to control batch quality**

| | Lower acceptance limit | Upper acceptance limit |
|---|---|---|
| **EpiSkin™ (SM)** | ≥ 0.6 | ≤ 1.5 |
| **EpiDerm™ SCT (EPI-200)** | ≥ 0.8 | ≤ 2.8 |
| **SkinEthic™ RHE** | ≥ 0.8 | ≤ 3.0 |
| **epiCS®** | ≥ 0.8 | ≤ 2.8 |

*Barrier function*

18. The *stratum corneum* and its lipid composition should be sufficient to resist the rapid penetration of certain cytotoxic benchmark chemicals (*e.g.* SDS or Triton X-100), as estimated by $IC_{50}$ or $ET_{50}$ (Table 3). The barrier function of each batch of the RhE model used should be demonstrated by the RhE model developer/vendor upon supply of the tissues to the end user (see paragraph 21).

*Morphology*

19. Histological examination of the RhE model should be performed demonstrating multi-layered human *epidermis*-like structure containing *stratum basale*, *stratum spinosum*, *stratum granulosum* and *stratum corneum* and exhibits lipid profile similar to lipid profile of human epidermis. Histological examination of each batch of the RhE model used demonstrating appropriate morphology of the tissues should be provided by the RhE model developer/vendor upon supply of the tissues to the end user (see paragraph 21).

*Reproducibility*

20. Test method users should demonstrate reproducibility of the test methods over time with the positive and negative controls. Furthermore, the test method should only be used if the RhE model developer/supplier provides data demonstrating reproducibility over time with corrosive and non-corrosive chemicals from *e.g.* the list of Proficiency Substances (Table 1). In case of the use of a test method for sub-categorization, the reproducibility with respect to sub-categorization should also be demonstrated.

*Quality control (QC)*

21. The RhE model should only be used if the developer/supplier demonstrates that each batch of the RhE model used meets defined production release criteria, among which those for *viability* (paragraph 17), *barrier function* (paragraph 18) and *morphology* (paragraph 19) are the most relevant. These data are provided to the test method users, so that they are able to include this information in the test report. Only results produced with QC accepted tissue batches can be accepted for reliable prediction of corrosive classification. An acceptability range (upper and lower limit) for the $IC_{50}$ or the $ET_{50}$ is established by the RhE model developer/supplier. The acceptability ranges for the four validated test methods are given in Table 3.

**Table 3: QC batch release criteria**

| | Lower acceptance limit | Upper acceptance limit |
|---|---|---|
| **EpiSkin™ (SM)** (18 hours treatment with SDS)(33) | $IC_{50}$ = 1.0 mg/mL | $IC_{50}$ = 3.0 mg/mL |
| **EpiDerm™ SCT (EPI-200)** (1% Triton X-100)(34) | $ET_{50}$ = 4.0 hours | $ET_{50}$ = 8.7 hours |
| **SkinEthic™ RHE** (1% Triton X-100)(35) | $ET_{50}$ = 4.0 hours | $ET_{50}$ = 10.0 hours |
| **epiCS®**(1% Triton X-100)(36) | $ET_{50}$ = 2.0 hours | $ET_{50}$ = 7.0 hours |

***Application of the Test Chemical and Control Substances***

22. At least two tissue replicates should be used for each test chemical and controls for each exposure time. For liquid as well as solid chemicals, sufficient amount of test chemical should be applied to uniformly cover the epidermis surface while avoiding an infinite dose, *i.e.* a minimum of 70 μL/cm$^2$ or 30 mg/cm$^2$ should be used. Depending on the methods, the epidermis surface should be moistened with deionized or distilled water before application of solid chemicals, to improve contact between the test chemical and the epidermis surface (33) (34) (35) (36). Whenever possible, solids should be tested as a fine powder. The application method should be appropriate for the test chemical (see *e.g.* references (33-36). At the end of the exposure period, the test chemical should be carefully washed from the *epidermis* with an aqueous buffer, or 0.9% NaCl. Depending on which of the four validated RhE test methods is used, two or three exposure periods are used per test chemical (for all four valid RhE models: 3 min and 1 hour; for EpiSkin™ an additional exposure time of 4 hours). Depending on the RhE test method used and the exposure period assessed, the incubation temperature during exposure may vary between room temperature and 37ºC.

23. Concurrent negative and positive controls (PC) should be used in each run to demonstrate that viability (with negative controls), barrier function and resulting tissue sensitivity (with the PC) of the tissues are within a defined historical acceptance range. The suggested PC chemicals are glacial acetic acid or 8N KOH depending upon the RhE model used. It should be noted that 8N KOH is a direct MTT reducer that might require adapted controls as described in paragraphs 25 and 26. The suggested negative controls are 0.9% (w/v) NaCl or water.

***Cell Viability Measurements***

24. The MTT assay, which is a quantitative assay, should be used to measure cell viability under this Test Guideline (26). The tissue sample is placed in MTT solution of appropriate concentration (0.3 or 1 mg/mL) for 3 hours. The precipitated blue formazan product is then extracted from the tissue using a solvent (*e.g.* isopropanol, acidic isopropanol), and the concentration of formazan is measured by determining the OD at 570 nm using a filter band pass of maximum ± 30 nm, or by an HPLC/UPLC-spectrophotometry procedure (see paragraphs 30 and 31) (37).

25. Test chemicals may interfere with the MTT assay, either by direct reduction of the MTT into blue formazan, and/or by colour interference if the test chemical absorbs, naturally or due to treatment procedures, in the same OD range of formazan (570 ± 30 nm, mainly blue and purple chemicals). Additional controls should be used to detect and correct for a potential interference from these test chemicals such as the non-specific MTT reduction (NSMTT) control and the non-specific colour (NSC) control (see paragraphs 26 to 30). This is especially important when a specific test chemical is not completely removed from the tissue by rinsing or when it penetrates the *epidermis*, and is therefore present in the tissues when the MTT viability test is performed. Detailed description of how to correct direct MTT reduction and interferences by colouring agents is available in the SOPs for the test methods (33) (34) (35) (36).

26. To identify direct MTT reducers, each test chemical should be added to freshly prepared MTT medium (33) (34) (35) (36). If the MTT mixture containing the test chemical turns blue/purple, the test chemical is presumed to directly reduce the MTT, and further functional check on non-viable epidermis should be performed, independently of using the standard absorbance (OD) measurement or an HPLC/UPLC-spectrophotometry procedure. This additional functional check employs killed tissues that possess only residual metabolic activity but absorb the test chemical in similar amount as viable tissues. Each MTT reducing chemical is applied on at least two killed tissue replicates per exposure time, which undergo the whole skin corrosion test. The true tissue viability is then calculated as the percent tissue viability obtained with living tissues exposed to the MTT reducer minus the percent non-specific MTT reduction obtained with the killed tissues exposed to the same MTT reducer, calculated relative to the negative control run concurrently to the test being corrected (%NSMTT).

27. To identify potential interference by coloured test chemicals or test chemicals that become coloured when in contact with water or isopropanol and decide on the need for additional controls, spectral analysis of the test chemical in water (environment during exposure) and/or isopropanol (extracting solution) should be performed. If the test chemical in water and/or isopropanol absorbs light in the range of 570 ± 30 nm, further colorant controls should be performed or, alternatively, an HPLC/UPLC-spectrophotometry procedure should be used in which case these controls are not required (see paragraphs 30 and 31). When performing the standard absorbance (OD) measurement, each interfering coloured test chemical is applied on at least two viable tissue replicates per exposure time, which undergo the entire skin corrosion test but are incubated with medium instead of MTT solution during the MTT incubation step to generate a non-specific colour ($NSC_{living}$) control. The $NSC_{living}$ control needs to be performed concurrently per exposure time per coloured test chemical (in each run) due to the inherent biological variability of living tissues. The true tissue viability is then calculated as the percent tissue viability obtained with living tissues exposed to the interfering test chemical and incubated with MTT solution minus the percent non-specific colour obtained with living tissues exposed to the interfering test chemical and incubated with medium without MTT, run concurrently to the test being corrected (%$NSC_{living}$).

28. Test chemicals that are identified as producing both direct MTT reduction (see paragraph 26) and colour interference (see paragraph 27) will also require a third set of controls, apart from the NSMTT and $NSC_{living}$ controls described in the previous paragraphs, when performing the standard absorbance (OD)

measurement. This is usually the case with darkly coloured test chemicals interfering with the MTT assay (e.g., blue, purple, black) because their intrinsic colour impedes the assessment of their capacity to directly reduce MTT as described in paragraph 26. These test chemicals may bind to both living and killed tissues and therefore the NSMTT control may not only correct for potential direct MTT reduction by the test chemical, but also for colour interference arising from the binding of the test chemical to killed tissues. This could lead to a double correction for colour interference since the $NSC_{living}$ control already corrects for colour interference arising from the binding of the test chemical to living tissues. To avoid a possible double correction for colour interference, a third control for non-specific colour in killed tissues ($NSC_{killed}$) needs to be performed. In this additional control, the test chemical is applied on at least two killed tissue replicates per exposure time, which undergo the entire testing procedure but are incubated with medium instead of MTT solution during the MTT incubation step. A single $NSC_{killed}$ control is sufficient per test chemical regardless of the number of independent tests/runs performed, but should be performed concurrently to the NSMTT control and, where possible, with the same tissue batch. The true tissue viability is then calculated as the percent tissue viability obtained with living tissues exposed to the test chemical minus %NSMTT minus $\%NSC_{living}$ plus the percent non-specific colour obtained with killed tissues exposed to the interfering test chemical and incubated with medium without MTT, calculated relative to the negative control run concurrently to the test being corrected ($\%NSC_{killed}$).

29. It is important to note that non-specific MTT reduction and non-specific colour interferences may increase the readoutsof the tissue extract above the linearity range of the spectrophotometer. On this basis, each laboratory should determine the linearity range of their spectrophotometer with MTT formazan (CAS # 57360-69-7) from a commercial source before initiating the testing of test chemicals for regulatory purposes. In particular, the standard absorbance (OD) measurement using a spectrophotometer is appropriate to assess direct MTT-reducers and colour interfering test chemicals when the ODs of the tissue extracts obtained with the test chemical without any correction for direct MTT reduction and/or colour interference are within the linear range of the spectrophotometer or when the uncorrected percent viability obtained with the test chemical already defined it as a corrosive (see paragraphs 35 and 36). Nevertheless, results for test chemicals producing %NSMTT and/or $\%NSC_{living} \geq 50\%$ of the negative control should be taken with caution.

30. For coloured test chemicals which are not compatible with the standard absorbance (OD) measurement due to too strong interference with the MTT assay, the alternative HPLC/UPLC-spectrophotometry procedure to measure MTT formazan may be employed (see paragraph 31) (37). The HPLC/UPLC-spectrophotometry system allows for the separation of the MTT formazan from the test chemical before its quantification (37). For this reason, $NSC_{living}$ or $NSC_{killed}$ controls are never required when using HPLC/UPLC-spectrophotometry, independently of the chemical being tested. NSMTT controls should nevertheless be used if the test chemical is suspected to directly reduce MTT or has a colour that impedes the assessment of the capacity to directly reduce MTT (as described in paragraph 26). When using HPLC/UPLC-spectrophotometry to measure MTT formazan, the percent tissue viability is calculated as percent MTT formazan peak area obtained with living tissues exposed to the test chemical relative to the MTT formazan peak obtained with the concurrent negative control. For test chemicals able to directly reduce MTT, true tissue viability is calculated as the percent tissue viability obtained with living tissues exposed to the test chemical minus %NSMTT. Finally, it should be noted that direct MTT-reducers that may also be colour interfering, which are retained in the tissues after treatment and reduce MTT so strongly that they lead to ODs (using standard OD measurement) or peak areas (using UPLC/HPLC-spectrophotometry) of the tested tissue extracts that fall outside of the linearity range of the spectrophotometer cannot be assessed, although these are expected to occur in only very rare situations.

31. HPLC/UPLC-spectrophotometry may be used also with all types of test chemicals (coloured, non-coloured, MTT-reducers and non-MTT reducers) for measurement of MTT formazan (37). Due to the

diversity of HPLC/UPLC-spectrophotometry systems, qualification of the HPLC/UPLC-spectrophotometry system should be demonstrated before its use to quantify MTT formazan from tissue extracts by meeting the acceptance criteria for a set of standard qualification parameters based on those described in the U.S. Food and Drug Administration guidance for industry on bio-analytical method validation (37) (38). These key parameters and their acceptance criteria are shown in Annex 4. Once the acceptance criteria defined in Annex 4 have been met, the HPLC/UPLC-spectrophotometry system is considered qualified and ready to measure MTT formazan under the experimental conditions described in this Test Guideline.

***Acceptability Criteria***

32. For each test method using valid RhE models, tissues treated with the negative control should exhibit OD reflecting the quality of the tissues as described in table 2 and should not be below historically established boundaries. Tissues treated with the PC, *i.e.* glacial acetic acid or 8N KOH, should reflect the ability of the tissues to respond to a corrosive chemical under the conditions of the test method (see Annex 2). The variability between tissue replicates of test chemical and/or control substances should fall within the accepted limits for each valid RhE model requirements (see Annex 2) (*e.g.* the difference of viability between the two tissue replicates should not exceed 30%). If either the negative control or PC included in a run fall out of the accepted ranges, the run is considered as not qualified and should be repeated. If the variability of test chemicals falls outside of the defined range, its testing should be repeated.

***Interpretation of Results and Prediction Model***

33. The OD values obtained for each test chemical should be used to calculate percentage of viability relative to the negative control, which is set at 100%. In case HPLC/UPLC-spectrophotometry is used, the percent tissue viability is calculated as percent MTT formazan peak area obtained with living tissues exposed to the test chemical relative to the MTT formazan peak obtained with the concurrent negative control. The cut-off percentage cell viability values distinguishing corrosive from non-corrosive test chemical (or discriminating between different corrosive sub-categories) are defined below in paragraphs 35 and 36 for each of the test methods covered by this Test Guideline and should be used for interpreting the results.

34. A single testing run composed of at least two tissue replicates should be sufficient for a test chemical when the resulting classification is unequivocal. However, in cases of borderline results, such as non-concordant replicate measurements, a second run may be considered, as well as a third one in case of discordant results between the first two runs.

35. The prediction model for the EpiSkin™ skin corrosion test method (9) (33) (22), associated with the UN GHS (1) classification system, is shown in Table 4:

**Table 4:EpiSkin™ prediction model**

| Viability measured after exposure time points (t=3, 60 and 240 minutes) | Prediction to be considered |
|---|---|
| < 35% after 3 min exposure | **Corrosive:**<br>• Optional Sub-category 1A * |
| ≥ 35% after 3 min exposure **AND**<br>< 35% after 60 min exposure<br>**OR**<br>≥ 35% after 60 min exposure **AND**<br>< 35% after 240 min exposure | **Corrosive:**<br>• A combination of optional Sub-categories 1B-and-1C |
| ≥ 35% after 240 min exposure | Non-corrosive |

*) According to the data generated in view of assessing the usefulness of the RhE test methods for supporting sub-categorisation, it was shown that around 22% of the Sub-category 1A results of the EpiSkin™ test method may actually constitute Sub-category 1B or Sub-category 1C substances/mixtures (i.e. over classifications) (see Annex 3).

36. The prediction models for the EpiDerm™ SCT (10) (34), the SkinEthic™ RHE (17) (18) (35), and the epiCS® (16) (36) skin corrosion test methods, associated with the UN GHS (1) classification system, are shown in Table 5:

**Table 5:EpiDerm™ SCT, SkinEthic™ RHE and epiCS®**

| Viability measured after exposure time points (t=3 and 60 minutes) | Prediction to be considered |
|---|---|
| < 50% after 3 min exposure | **Corrosive:**<br>• Optional Sub-category 1A* |
| ≥ 50% after 3 min exposure **AND**<br>< 15% after 60 min exposure | **Corrosive:**<br>• A combination of optional Sub-categories 1B-and-1C |
| ≥ 50% after 3 min exposure **AND**<br>≥ 15% after 60 min exposure | Non-corrosive |

*) According to the data generated in view of assessing the usefulness of the RhE test methods for supporting sub-categorisation, it was shown that around 42% of the Sub-category 1A results of the EpiDerm™ test method, and around 46% of the Sub-category 1A results of the SkinEthic™ and the epiCS® test method may actually constitute Sub-category 1B or Sub-category 1C substances/mixtures (i.e. over-classifications) (see Annex 3).

## DATA AND REPORTING

### *Data*

37. For each test, data from individual tissue replicates (*e.g.* OD values and calculated percentage cell viability for each test chemical, including classification) should be reported in tabular form, including data from repeat experiments as appropriate. In addition, means and ranges of viability and CVs between

tissue replicates for each test should be reported. Observed interactions with MTT reagent by direct MTT reducers or coloured test chemicals should be reported for each tested chemical.

***Test report***

38. The test report should include the following information:

*Test Chemical and Control Substances:*

- Mono-constituent substance: chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc;
- Multi-constituent substance, UVCB and mixture: characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents;
- Physical appearance, water solubility, and any additional relevant physicochemical properties;
- Source, lot number if available;
- Treatment of the test chemical/control substance prior to testing, if applicable (*e.g.* warming, grinding);
- Stability of the test chemical, limit date for use, or date for re-analysis if known;
- Storage conditions.

–

*RhE model and protocol used and rationale for it (if applicable)*

*Test Conditions:*

- RhE model used (including batch number);
- Calibration information for measuring device (*e.g.* spectrophotometer), wavelength and band pass (if applicable) used for quantifying MTT formazan, and linearity range of measuring device;
- Description of the method used to quantify MTT formazan;
- Description of the qualification of the HPLC/UPLC-spectrophotometry system, if applicable;
- Complete supporting information for the specific RhE model used including its performance. This should include, but is not limited to:
  i) Viability;
  ii) Barrier function;
  iii) Morphology;
  iv) Reproducibility and predictive capacity;
  v) Quality controls (QC) of the model;
- Reference to historical data of the model. This should include, but is not limited to acceptability of the QC data with reference to historical batch data;
- Demonstration of proficiency in performing the test method before routine use by testing of the proficiency substances.

*Test Procedure:*

- Details of the test procedure used (including washing procedures used after exposure period);
- Doses of test chemical and control substances used;
- Duration of exposure period(s) and temperature(s) of exposure;
- Indication of controls used for direct MTT-reducers and/or colouring test chemicals, if applicable;

- Number of tissue replicates used per test chemical and controls (PC, negative control, and NSMTT, $NSC_{living}$ and $NSC_{killed}$, if applicable), per exposure time;
- Description of decision criteria/prediction model applied based on the RhE model used;
- Description of any modifications of the test procedure (including washing procedures).

*Run and Test Acceptance Criteria:*

- Positive and negative control mean values and acceptance ranges based on historical data;
- Acceptable variability between tissue replicates for positive and negative controls;
- Acceptable variability between tissue replicates for test chemical.

*Results:*

- Tabulation of data for individual test chemicals and controls, for each exposure period, each run and each replicate measurement including OD or MTT formazan peak area, percent tissue viability, mean percent tissue viability, differences between replicates, SDs and/or CVs if applicable;
- If applicable, results of controls used for direct MTT-reducers and/or colouring test chemicals including OD or MTT formazan peak area, %NSMTT, $\%NSC_{living}$, $\%NSC_{killed}$, differences between tissue replicates, SDs and/or CVs (if applicable), and final correct percent tissue viability;
- Results obtained with the test chemical(s) and control substances in relation to the defined run and test acceptance criteria;
- Description of other effects observed;
- The derived classification with reference to the prediction model/decision criteria used.

*Discussion of the results*

*Conclusions*

**LITERATURE**

(1) UN. (2013). United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Fifth Revised Edition, UN New York and Geneva. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html]

(2) OECD. (2015). Guideline for Testing of Chemicals. (No. 404.): Acute Dermal Irritation, Corrosion, Organisation for Economic Cooperation and Development, Paris.

(3) OECD. (2015). Guideline for the Testing of Chemicals (No. 430.): *In Vitro* Skin Corrosion: Transcutaneous Electrical Resistance (TER). Organisation for Economic Cooperation and Development, Paris.

(4) OECD. (2015). Guideline for the Testing of Chemicals (No. 435.): *In Vitro* Membrane Barrier Test Method. Organisation for Economic Cooperation and Development, Paris. (5) OECD. (2015). Guideline for the Testing of Chemicals (No. 439.): *In Vitro* Skin Irritation: Reconstructed Human Epidermis Test Method. Organisation for Economic Cooperation and Development, Paris.

(6) OECD. (2014). Guidance Document on Integrated Approaches to Testing and Assessment of Skin Irritation/Corrosion. Environment, Health and Safety Publications, Series on Testing and Assessment, (No. 203.) Organisation for Economic Cooperation and Development, Paris.

(7) Botham P.A., Chamberlain M., Barratt M.D., Curren R.D., Esdaile D.J., Gardner J.R., Gordon V.C., Hildebrand B., Lewis R.W., Liebsch M., Logemann P., Osborne R., Ponec M., Regnier J.F., Steiling W., Walker A.P., and Balls M. (1995). A Prevalidation Study on *In Vitro* Skin Corrosivity Testing. The report and Recommendations of ECVAM Workshop 6.*ATLA* 23, 219-255.

(8) Barratt M.D., Brantom P.G., Fentem J.H., Gerner I., Walker A.P., and Worth A.P. (1998). The ECVAM International Validation Study on *In Vitro* Tests for Skin Corrosivity. 1. Selection and distribution of the Test Chemicals. *Toxicol.In Vitro* 12, 471-482.

(9) Fentem J.H., Archer G.E.B., Balls M., Botham P.A., Curren R.D., Earl L.K., Esdaile D.J., Holzhutter H.-G., and Liebsch M. (1998). The ECVAM International Validation Study on *In Vitro* Tests for Skin Corrosivity. 2. Results and Evaluation by the Management Team. *Toxicol.in Vitro* 12, 483-524.

(10) Liebsch M., Traue D., Barrabas C., Spielmann H., Uphill, P., Wilkins S., Wiemann C., Kaufmann T., Remmele M. and Holzhütter H. G. (2000). The ECVAM Prevalidation Study on the Use of EpiDerm for Skin Corrosivity Testing, *ATLA* 28, pp. 371-401.

(11) Balls M., Blaauboer B.J., Fentem J.H., Bruner L., Combes R.D., Ekwall B., Fielder R.J., Guillouzo A., Lewis R.W., Lovell D.P., Reinhardt C.A., Repetto G., Sladowski D., Spielmann H. et Zucco F. (1995). Practical Aspects of the Validation of Toxicity Test Procedures. The Report and Recommendations of ECVAM Workshops, *ATLA* 23, 129-147.(12) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). (1997). Validation and Regulatory Acceptance of Toxicological Test Methods. NIH Publication No. 97-3981. National Institute of Environmental Health Sciences, Research

Triangle Park, NC, USA. Available at: [http://www.iccvam.niehs.nih.gov/docs/guidelines/validate.pdf].

(13) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). (2002). ICCVAM evaluation of EpiDerm™ (EPI-200), EPISKIN™ (SM), and the Rat Skin Transcutaneous Electrical Resistance (TER) Assay: *In Vitro* Test Methods for Assessing Dermal Corrosivity Potential of Chemicals. NIH Publication No. 02-4502. National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.Available at:[http://www.iccvam.niehs.nih.gov/methods/epiddocs/epis_brd.pdf]

(14) EC-ECVAM. (1998). Statement on the Scientific Validity of the EpiSkin™ Test (an *In Vitro* Test for Skin Corrosivity), Issued by the ECVAM Scientific Advisory Committee (ESAC10), 3 April 1998. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

(15) EC-ECVAM. (2000). Statement on the Application of the EpiDerm™ Human Skin Model for Skin Corrosivity Testing, Issued by the ECVAM Scientific Advisory Committee (ESAC14), 21 March 2000. Available at: [http://ecvam.jrc.ec.europa.eu].

(16) Hoffmann J., Heisler E., Karpinski S., Losse J., Thomas D., Siefken W., Ahr H.J., Vohr H.W. and Fuchs H.W. (2005). Epidermal-Skin-Test 1000 (EST-1000)-A New Reconstructed Epidermis for *In Vitro* Skin Corrosivity Testing. *Toxicol.In Vitro* 19, 925-929.

(17) Kandárová H., Liebsch M., Spielmann,H., Genschow E., Schmidt E., Traue D., Guest R., Whittingham A., Warren N, Gamer A.O., Remmele M., Kaufmann T., Wittmer E., De Wever B., and Rosdy M. (2006). Assessment of the Human Epidermis Model SkinEthic RHE for *In Vitro* Skin Corrosion Testing of Chemicals According to New OECD TG 431. *Toxicol.In Vitro* 20, 547–559.

(18) Tornier C., Roquet M. and Fraissinette A.B. (2010). Adaptation of the Validated SkinEthic™ Reconstructed Human Epidermis (RHE) Skin Corrosion Test Method to 0.5 $cm^2$ Tissue Sample. *Toxicol. In Vitro* 24, 1379-1385.

(19) EC-ECVAM. (2006). Statement on the Application of the SkinEthic™ Human Skin Model for Skin Corrosivity Testing, Issued by the ECVAM Scientific Advisory Committee (ESAC25), 17 November 2006. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

(20) EC-ECVAM. (2009). ESAC Statement on the Scientific Validity of an *In-Vitro* Test Method for Skin Corrosivity Testing: the EST-1000, Issued by the ECVAM Scientific Advisory Committee (ESAC30), 12 June 2009. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

(21) OECD. (2013). Summary Document on the Statistical Performance of Methods in OECD Test Guideline 431 for Sub-categorisation. Environment, Health, and Safety Publications, Series on Testing and Assessment (No. 190.). Organisation for Economic Cooperation and Development, Paris.

(22) Alépée N., Grandidier M.H., and Cotovio J. (2014). Sub-Categorisation of Skin Corrosive Chemicals by the EpiSkin™ Reconstructed Human Epidermis Skin Corrosion Test Method According to UN GHS: Revision of OECD Test Guideline 431. *Toxicol. In Vitro* 28, 131-145.

(23) OECD. (2015). Performance Standards for the Assessment of Proposed Similar or Modified In Vitro Reconstructed Human Epidermis (RHE) Test Methods For Skin Corrosion in Relation to OECD TG 431. Environmental Health and Safety Publications, Series on Testing and Assessment (No. 219).

Organisation for Economic Cooperation and Development, Paris

(24) OECD. (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. .Environment, Health and Safety Publications, Series on Testing and Assessment (No. 34.), Organisation for Economic Cooperation and Development, Paris.

(25) Eskes C. et al. (2012). Regulatory Assessment of *In Vitro* Skin Corrosion and Irritation Data Within the European Framework: Workshop Recommendations. *Regul.Toxicol.Pharmacol.* 62, 393-403.

(26) Mosmann T. (1983). Rapid Colorimetric Assay for Cellular Growth and Survival: Application to Proliferation and Cytotoxicity Assays. *J. Immunol. Methods* 65, 55-63.

(27) Tinois E., et al. (1994). The Episkin Model: Successful Reconstruction of Human *Epidermis In Vitro*. In: *In Vitro* Skin Toxicology. Rougier A.,. Goldberg A.M and Maibach H.I. (Eds): 133-140.

(28) Cannon C. L., Neal P.J., Southee J.A., Kubilus J. and Klausner M. (1994), New Epidermal Model for Dermal Irritancy Testing. *Toxicol.in Vitro* 8, 889 - 891.

(29) Ponec M., Boelsma E, Weerheim A, Mulder A, Bouwstra J and Mommaas M. *(*2000). Lipid and Ultrastructural Characterization of Reconstructed Skin Models. Inter. J. Pharmaceu. 203, 211 - 225.

(30) Tinois E., Tillier, J., Gaucherand, M., Dumas, H., Tardy, M. and Thivolet J. (1991). *In Vitro* and Post – Transplantation Differentiation of Human Keratinocytes Grown on the Human Type IV Collagen Film of a Bilayered Dermal Substitute. *Exp. Cell Res.* 193: 310-319.

(31) Parenteau N.L., Bilbo P, Nolte CJ, Mason VS and Rosenberg M. (1992). The Organotypic Culture of Human Skin Keratinocytes and Fibroblasts to Achieve Form and Function. *Cytotech*. 9, 163-171.

(32) Wilkins L.M., Watson SR, Prosky SJ, Meunier SF and Parenteau N.L. (1994). Development of a Bilayered Living Skin Construct for Clinical Applications. *Biotech.Bioeng*.43/8, 747-756.

(33) EpiSkin™. (December 2011).SOP, *INVITTOX* Protocol (No. 118.). EpiSkin™ Skin Corrosivity Test. Available at: [http://www.ecvam.jrc.ec.europa.eu.hm;].

(34) EpiDerm™ SOP. (February 2012). Version MK-24-007-0024Protocol for: *In Vitro* EpiDerm™ Skin Corrosion Test (EPI-200-SCT), for Use with MatTek Corporation's Reconstructed Human Epidermal Model EpiDerm. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

(35) SkinEthic™ RHE SOP, *INVITTOX* Protocol (January 2012). SkinEthic™ Skin Corrosivity Test. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

(36) EpiCS® SOP, Version 4.1 (January 2012). *In Vitro* Skin Corrosion: Human Skin Model Test Epidermal Skin Test 1000 (epiCS® ) CellSystems. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

(37) Alépée N., Barroso J., De Smedt A., De Wever B., Hibatallah J., Klaric M., Mewes K.R., Millet M., Pfannenbecker U., Tailhardat M., Templier M., and McNamee P. Use of HPLC/UPLC-spectrophotometry for Detection of MTT Formazan in In Vitro Reconstructed Human Tissue (RhT)-based Test Methods Employing the MTT Assay to Expand their Applicability to Strongly Coloured Test Chemicals. Manuscript in Preparation.

(38) US FDA. (2001). Guidance for Industry: Bioanalytical Method Validation. U.S. Department of Health and Human Services, Food and Drug Administration. (May 2001). Available at: [http://www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf].

## 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 (24).

**Cell viability:** Parameter measuring total activity of a cell population *e.g.* as ability of cellular mitochondrial dehydrogenases to reduce the vital dye MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue), which depending on the endpoint measured and the test design used, correlates with the total number and/or vitality of living cells.

**Chemical:** means a substance or a mixture.

**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 (24).

**$ET_{50}$:** Can be estimated by determination of the exposure time required to reduce cell viability by 50% upon application of the benchmark chemical at a specified, fixed concentration, see also $IC_{50}$.

**GHS (Globally Harmonized System of Classification and Labelling of Chemicals):** 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).

**HPLC**: High Performance Liquid Chromatography.

**IATA:** Integrated Approach on Testing and Assessment.

**$IC_{50}$:** Can be estimated by determination of the concentration at which a benchmark chemical reduces the viability of the tissues by 50% ($IC_{50}$) after a fixed exposure time, see also $ET_{50}$.

**Infinite dose:** Amount of test chemical applied to the *epidermis* exceeding the amount required to completely and uniformly cover the *epidermis* surface.

**Mixture:** means a mixture or solution composed of two or more substances in which they do not react.

**Mono-constituent substance**: A substance, defined by its quantitative composition, in which one main constituent is present to at least 80% (w/w).

**MTT**: 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; Thiazolyl blue tetrazolium bromide.

**Multi-constituent substance:** A substance, defined by its quantitative composition, in which more than one main constituent is present in a concentration ≥ 10% (w/w) and < 80% (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

**NC**: Non corrosive.

**$NSC_{killed}$ control**: Non-Specific Colour control in killed tissues.

**$NSC_{living}$control** : Non-Specific Colour control in living tissues.

**NSMTT**: Non-Specific MTT reduction.

**OD:** Optical Density

**PC**: Positive Control, a replicate containing all components of a test system and treated with a substance 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.

**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 similar levels of reliability and accuracy, 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 (24).

**Relevance:** Description of relationship of the test method to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test method correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (24).

**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 (24).

**Run:** A run consists of one or more test chemicals tested concurrently with a negative control and with a PC.

**Sensitivity:** The proportion of all positive/active chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (24).

**Skin corrosion *in vivo:*** The production of irreversible damage of the skin; namely, visible necrosis through the *epidermis* and into the dermis, following the application of a test chemical for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of observation at 14 days, by discoloration due to blanching of the skin, complete areas of alopecia, and scars. Histopathology should be considered to evaluate questionable lesions.

**Specificity:** The proportion of all negative/inactive chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (24).

**Substance:** means 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.

**Test chemical:** means what is being tested.

**UPLC:** Ultra-High Performance Liquid Chromatography.

**UVCB:** substances of unknown or variable composition, complex reaction products or biological materials.

## ANNEX 2

### MAIN TEST METHOD COMPONENTS OF THE RhE TEST METHODS VALIDATED FOR SKIN CORROSION TESTING

| Test Method Components | EpiSkin™ | EpiDerm™ SCT | SkinEthic™ RHE | epiCS® |
|---|---|---|---|---|
| **Model surface** | 0.38 $cm^2$ | 0.63 $cm^2$ | 0.5 $cm^2$ | 0.6 $cm^2$ |
| **Number of tissue replicates** | At least 2 per exposure time | 2-3 per exposure time | At least 2 per exposure time | At least 2 per exposure time |
| **Treatment doses and application** | Liquids and viscous: 50 µL ± 3 µL (131.6 µL/cm2)<br><br>Solids: 20 ± 2 mg (52.6 mg/cm2) + 100 µL ± 5µL NaCl solution (9 g/L)<br><br>Waxy/sticky: 50 ± 2 mg (131.6 mg/cm2) with a nylon mesh | Liquids: 50 µL (79.4 µL/$cm^2$) with or without a nylon mesh<br>*Pre-test compatibility of test chemical with nylon mesh*<br><br>Semisolids: 50 µL (79.4 µL/$cm^2$)<br><br>Solids: 25 µL $H_2O$ (or more if necessary) + 25 mg (39.7 mg/$cm^2$)<br><br>Waxes: flat "disc like" piece of ca. 8 mm diameter placed atop the tissue wetted with 15 µL $H_2O$. | Liquids and viscous: 40 µL ± 3µl (80 µL/$cm^2$) using nylon mesh<br>*Pre-test compatibility of test chemical with nylon mesh*<br><br>Solids: 20 µL ± 2µl $H_2O$ + 20 ± 3 mg (40 mg/$cm^2$)<br><br>Waxy/sticky: 20 ± 3 mg (40 mg/$cm^2$) using nylon mesh | Liquids: 50 µL (83.3 µL/$cm^2$) using nylon mesh<br>*Pre-test compatibility of test chemical with nylon mesh*<br><br>Semisolids: 50 µL (83.3 µL/$cm^2$)<br><br>Solids: 25 mg (41.7 mg/$cm^2$) + 25 µL $H_2O$ (or more if necessary)<br><br>Waxy: flat "cookie like" piece of ca. 8 mm diameter placed atop the tissue wetted with 15 µL $H_2O$ |
| **Pre-check for direct MTT reduction** | 50 µL (liquid) or 20 mg (solid) + 2 mL MTT<br>0.3 mg/mL solution for 180 ± 5 min at 37°C, 5% $CO_2$, 95% RH<br>→ if solution turns blue/purple, water-killed adapted controls should be performed | 50 µL (liquid) or 25 mg (solid) + 1 mL MTT<br>1 mg/mL solution for 60 min at 37°C, 5% $CO_2$, 95% RH<br>→ if solution turns blue/purple, freeze-killed adapted controls should be performed | 40 µL (liquid) or 20 mg (solid) + 1 mL MTT<br>1 mg/mL solution for 180± 15 min at 37°C, 5% $CO_2$, 95% RH<br>→ if solution turns blue/purple, freeze-killed adapted controls should be performed | 50 µL (liquid) or 25 mg (solid) + 1 mL MTT<br>1 mg/mL solution for 60 min at 37°C, 5% $CO_2$, 95% RH<br>→ if solution turns blue/purple, freeze-killed adapted controls should be performed |

| Test Method Components | EpiSkin™ | EpiDerm™ SCT | SkinEthic™ RHE | epiCS® |
|---|---|---|---|---|
| **Pre-check for colour interference** | 10 µL (liquid) or 10 mg (solid) + 90 µL $H_2O$ mixed for 15 min at RT<br>→ if solution becomes coloured, living adapted controls should be performed | 50 µL (liquid) or 25 mg (solid) + 300 µL $H_2O$<br>for 60 min at 37°C, 5% $CO_2$, 95% RH<br>→ if solution becomes coloured, living adapted controls should be performed | 40 µL (liquid) or 20mg (solid) + 300 µL $H_2O$ mixed for 60 min at RT<br>→ if test chemical is coloured, living adapted controls should be performed | 50 µL (liquid) or 25 mg (solid) + 300 µL $H_2O$<br>for 60 min at 37°C, 5% $CO_2$, 95% RH<br>→ if solution becomes coloured, living adapted controls should be performed |
| **Exposure time and temperature** | 3 min, 60 min (± 5 min)<br>and 240 min (± 10 min)<br>In ventilated cabinet<br>Room Temperature (RT, 18-28°C) | 3 min at RT, and 60 min<br>at 37°C, 5% $CO_2$, 95% RH | 3 min at RT, and 60 min<br>at 37°C, 5% $CO_2$, 95% RH | 3 min at RT, and 60 min<br>at 37°C, 5% $CO_2$, 95% RH |
| **Rinsing** | 25 mL 1x PBS (2 mL/throwing) | 20 times with a constant soft stream of 1x PBS | 20 times with a constant soft stream of 1x PBS | 20 times with a constant soft stream of 1x PBS |
| **Negative control** | 50 µL NaCl solution (9 g/L)<br>Tested with every exposure time | 50 µL $H_2O$<br>Tested with every exposure time | 40 µL $H_2O$<br>Tested with every exposure time | 50 µL $H_2O$<br>Tested with every exposure time |
| **Positive control** | 50 µL Glacial acetic acid<br>Tested only for 4 hours | 50 µL 8N KOH<br>Tested with every exposure time | 40 µL 8N KOH<br>Tested only for 1 hour | 50 µL 8N KOH<br>Tested with every exposure time |
| **MTT solution** | 2 mL 0.3 mg/mL | 300 µL 1 mg/mL | 300 µL 1 mg/mL | 300 µL 1 mg/mL |

| Test Method Components | EpiSkin™ | EpiDerm™ SCT | SkinEthic™ RHE | epiCS® |
|---|---|---|---|---|
| **MTT incubation time and temperature** | 180 min (± 15 min) at 37°C, 5% $CO_2$, 95% RH | 180 min at 37°C, 5% $CO_2$, 95% RH | 180 min (± 15 min) at 37°C, 5% $CO_2$, 95% RH | 180 min at 37°C, 5% $CO_2$, 95% RH |
| **Extraction solvent** | 500 µL acidified isopropanol<br><br>(0.04 N HCl in isopropanol)<br>(isolated tissue fully immersed) | 2 mL isopropanol<br>(extraction from top and bottom of insert) | 1.5 mL isopropanol<br>(extraction from top and bottom of insert) | 2 mL isopropanol<br>(extraction from top and bottom of insert) |
| **Extraction time and temperature** | Overnight at RT, protected from light | Overnight without shaking at RT or for 120 min with shaking (~120 rpm) at RT | Overnight without shaking at RT or for 120 min with shaking (~120 rpm) at RT | Overnight without shaking at RT or for 120 min with shaking (~120 rpm) at RT |
| **OD reading** | 570 nm (545 - 595 nm)<br>without reference filter | 570 nm (or 540 nm)<br>without reference filter | 570 nm (540 - 600 nm)<br>without reference filter | 540 - 570 nm<br>without reference filter |
| **Tissue Quality Control** | 18 hours treatment with SDS<br>1.0 mg/mL ≤ $IC_{50}$ ≤ 3.0 mg/mL | Treatment with 1% Triton X-100<br>4.08 hours ≤ $ET_{50}$ ≤ 8.7 hours | Treatment with 1% Triton X-100<br>4.0 hours ≤ $ET_{50}$ ≤ 10.0 hours | Treatment with 1% Triton X-100<br>2.0 hours ≤ $ET_{50}$ ≤ 7.0 hours |

| Test Method Components | EpiSkin™ | EpiDerm™ SCT | SkinEthic™ RHE | epiCS® |
|---|---|---|---|---|
| **Acceptability Criteria** | 1. Mean OD of the tissue replicates treated with the negative control (NaCl) should be ≥ 0.6 and ≤ 1.5 for every exposure time<br>2. Mean viability of the tissue replicates exposed for 4 hours with the positive control (glacial acetic acid), expressed as % of the negative control, should be ≤ 20%<br>3. In the range 20-100% viability and for ODs≥ 0.3, difference of viability between the two tissue replicates should not exceed 30%. | 1. Mean OD of the tissue replicates treated with the negative control ($H_2O$) should be ≥ 0.8 and ≤ 2.8 for every exposure time<br>2. Mean viability of the tissue replicates exposed for 1 hour with the positive control (8N KOH), expressed as % of the negative control, should be < 15%<br>3. In the range 20 - 100% viability, the Coefficient of Variation (CV) between tissue replicates should be ≤ 30% | 1. Mean OD of the tissue replicates treated with the negative control ($H_2O$) should be ≥ 0.8 and ≤ 3.0 for every exposure time<br>2. Mean viability of the tissue replicates exposed for 1 hour (and 4 hours, if applicable) with the positive control (8N KOH), expressed as % of the negative control, should be < 15%<br>3. In the range 20-100% viability, and for ODs ≥ 0.3, difference of viability between the two tissue replicates should not exceed 30% | 1. Mean OD of the tissue replicates treated with the negative control ($H_2O$) should be ≥ 0.8 and ≤ 2.8 for every exposure time<br>2. Mean viability of the tissue replicates exposed for 1 hour with the positive control (8N KOH), expressed as % of the negative control, should be < 20%<br>3. In the range 20-100% viability, and for ODs ≥ 0.3, difference of viability between the two tissue replicates should not exceed 30% |

## ANNEX 3

### PERFORMANCE OF TEST METHODS FOR SUB-CATEGORISATION

The table below provides the performances of the four test methods calculated based on a set of 80 chemicals tested by the four test developers. Calculations were performed by the OECD Secretariat, reviewed and agreed by an expert subgroup (21).

EpiSkin™, EpiDerm™ , SkinEthic™ and epiCS® test methods are able to sub-categorize (i.e. 1A versus 1B-and-1C versus NC) but differences are observed between EpiSkin™ and the three other test methods, EpiDerm™, SkinEthic™ and epiCS®, for sub-categorization. Results from EpiSkin™ can be directly used based on the outcoming results, whereas results from EpiDerm™, SkinEthic™ and epiCS®, should take into account high over-classification rates from those three test methods for 1B-and-1C sub-category. Therefore, for EpiDerm™, SkinEthic™ and epiCS®, chemicals that are classified as 1B-and-1C can be considered as 1B-and-1C, and chemicals for which cell viability at 3 minutes is below 50% should be considered as 1, that is to say that either under the prediction principle they could be claimed as 1A or they should undergo further testing to be possibly confirmed as 1B-and-1C. The regulatory framework in member countries will decide how this Test Guideline will be used.

Performances, Overclassification rates, Underclassification rates, and Accuracy (Predictive capacity) of the four test methods based on a set of 80 chemicals all tested over 2 or 3 runs in each test method:

**STATISTICS ON PREDICTIONS OBTAINED ON THE ENTIRE SET OF CHEMICALS**

**(n= 80 chemicals tested over 2 or 3 independent runs, i.e. 159* or 240 classifications)**

*one chemical was tested once because of no availability

| | **EpiSkin™** | **EpiDerm™** | **SkinEthic™** | **epiCS®** |
|---|---|---|---|---|
| **Overclassifications:** | | | | |
| 1B-and-1C overclassified 1A | 21.50% | 41.94% | 46.24% | 45.90% |
| NC overclassified 1B-and-1C | 20.72% | 23.42% | 24.32% | 28.38% |
| NC overclassified 1A | 0.00% | 2.70% | 2.70% | 0.00% |
| overclassified Corr. | 20.72% | 26.13% | 27.03% | 28.38% |
| Global overclassification rate (all categories) | 17.92% | 28.33% | 30.42% | 30.82% |
| **Underclassifications:** | | | | |
| 1A underclassified 1B-and-1C | 16.67% | 8.33% | 13.89% | 8.33% |
| 1A underclassified NC | 0.00% | 0.00% | 0.00% | 0.00% |
| 1B-and-1C underclassified NC | 2.15% | 0.00% | 7.53% | 6.56% |
| Global underclassification rate (all categories) | 3.33% | 2.47% | 5.00% | 3.77% |
| **Correct Classifications:** | | | | |
| 1A correctly classified | 83.33% | 91.67% | 86.11% | 91.67% |
| 1B-and-/1C correctly classified | 76.34% | 58.06% | 46.24% | 47.54% |
| NC correctly classified | 79.28% | 73.87% | 72.97% | 71.62% |
| **Accuracy** (Predictive capacity) | 78.75% | 70.42% | 64.58% | 65.41% |

NC: Non-corrosive

**ANNEX 4**

Key parameters and acceptance criteria for qualification of an HPLC/UPLC-spectrophotometry system for measurement of MTT formazan extracted from RhE tissues

| Parameter | Protocol Derived from FDA Guidance (37)(38) | Acceptance Criteria |
|---|---|---|
| Selectivity | Analysis of isopropanol, living blank (isopropanol extract from living RhE tissues without any treatment), dead blank (isopropanol extract from killed RhE tissues without any treatment) | $Area_{interference} \leq 20\%$ of $Area_{LLOQ}$[1] |
| Precision | Quality Controls (i.e., MTT formazan at 1.6 µg/mL, 16 µg/mL and 160 µg/mL ) in isopropanol (n=5) | CV ≤ 15% or ≤ 20% for the LLOQ |
| Accuracy | Quality Controls in isopropanol (n=5) | %Dev ≤ 15% or ≤ 20% for LLOQ |
| Matrix Effect | Quality Controls in living blank (n=5) | 85% ≤ Matrix Effect % ≤ 115% |
| Carryover | Analysis of isopropanol after an ULOQ[2] standard | $Area_{interference} \leq 20\%$ of $Area_{LLOQ}$ |
| Reproducibility (intra-day) | 3 independent calibration curves (based on 6 consecutive 1/3 dilutions of MTT formazan in isopropanol starting at ULOQ, i.e., 200 µg/mL); Quality Controls in isopropanol (n=5) | Calibration Curves: %Dev ≤ 15% or ≤ 20% for LLOQ<br><br>Quality Controls: %Dev ≤ 15% and CV ≤ 15% |
| Reproducibility (inter-day) | Day 1: 1 calibration curve and Quality Controls in isopropanol (n=3)<br>Day 2: 1 calibration curve and Quality Controls in isopropanol (n=3)<br>Day 3: 1 calibration curve and Quality Controls in isopropanol (n=3) | |
| Short Term Stability of MTT Formazan in RhE Tissue Extract | Quality Controls in living blank (n=3) analysed the day of the preparation and after 24 hours of storage at room temperature | %Dev ≤ 15% |
| Long Term Stability of MTT Formazan in RhE Tissue Extract, if required | Quality Controls in living blank (n=3) analysed the day of the preparation and after several days of storage at a specified temperature (e.g., 4ºC, -20ºC, -80ºC) | %Dev ≤ 15% |

[1]LLOQ: Lower Limit of Quantification, defined to cover 1-2% tissue viability, i.e., 0.8 µg/mL.
[2]ULOQ: Upper Limit of Quantification, defined to be at least two times higher than the highest expected MTT formazan concentration in isopropanol extracts from negative controls i.e., 200 µg/mL.