OECD/OCDE 439
Adopted:
28 July 2015

OECD GUIDELINES FOR THE TESTING OF CHEMICALS

In Vitro Skin Irritation: Reconstructed Human *Epidermis* Test Method

INTRODUCTION

1. Skin irritation refers to the production of reversible damage to the skin following the application of a test chemical for up to 4 hours [as defined by the United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS)](1). This Test Guideline (TG) provides an *in vitro* procedure that may be used for the hazard identification of irritant chemicals (substances and mixtures) in accordance with UN GHS Category 2 (1) (2). In member countries or regions that do not adopt the optional UN GHS Category 3 (mild irritants), this Test Guideline can also be used to identify non-classified chemicals. Therefore, depending on the regulatory framework and the classification system in use, this Test Guideline may be used to determine the skin irritancy of chemicals either as a stand-alone replacement test for *in vivo* skin irritation testing or as a partial replacement test within a testing strategy (3).

2. The assessment of skin irritation has typically involved the use of laboratory animals [OECD TG 404; originally adopted in 1981 and revised in 1992, 2002 and 2015] (4). For the testing of corrosivity, three validated *in vitro* test methods have been adopted as OECD TGs 430, 431 and 435 (5) (6) (7). 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 test and non-test data for the assessment of skin irritation and skin corrosion potentials of chemicals and (ii) proposes an approach when further testing is needed (3).

3. This Test Guideline addresses the human health endpoint skin irritation. It is based on the *in vitro* test system of reconstructed human *epidermis* (RhE), which closely mimics the biochemical and physiological properties of the upper parts of the human skin, *i.e.* the *epidermis.* The RhE test system uses human derived non-transformed keratinocytes as cell source to reconstruct an epidermal model with representative histology and cytoarchitecture. Performance Standards (PS) are available to facilitate the validation and assessment of similar and modified RhE-based test methods, in accordance with the principles of Guidance Document No. 34 (8) (9). This Test Guideline was originally adopted in 2010, updated in 2013 to include additional test methods using the RhE models, and updated in 2015 to refer to the IATA guidance document and introduce the use of an alternative procedure to measure viability.

4. Pre-validation, optimisation and validation studies have been completed for four commercially available *in vitro* test methods (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27)

(28) based on the RhE test system (sensitivity 80%, specificity 70%, and accuracy 75%). These four test methods are included in this TG and are listed in Annex 2, which also provides information on the type of validation study used to validate the respective test methods. As noted in Annex 2, the Validated Reference Method (VRM) have been used to develop the present TG and the Performance Standards (8).

5. Mutual Acceptance of Data will only be guaranteed for test methods, validated according to the Performance Standards (8), if these test methods have been reviewed and adopted by OECD. The test methods included in this TG can be used indiscriminately to address countries' requirements for test results from *in vitro* test method for skin irritation, while benefiting from the Mutual Acceptance of Data.

6. Definitions of terms used in this document are provided in Annex 1.

INITIAL CONSIDERATIONS AND LIMITATIONS

7. A limitation of the Test Guideline, as demonstrated by the full prospective validation study assessing and characterising RhE test methods (16), is that it does not allow the classification of chemicals to the optional UN GHS Category 3 (mild irritants) (1). Thus, the regulatory framework in member countries will decide how this Test Guideline will be used. 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 (3). It is recognized that the use of human skin is subject to national and international ethical considerations and conditions.

8. This Test Guideline addresses the human health endpoint skin irritation. While this Test Guideline does not provide adequate information on skin corrosion, it should be noted that OECD TG 431 on skin corrosion is based on the same RhE test system, though using another protocol (6). This Test Guideline is based on RhE-models using human keratinocytes, which therefore represent *in vitro* the target organ of the species of interest. It moreover directly covers the initial step of the inflammatory cascade/mechanism of action (cell and tissue damage resulting in localised trauma) that occurs during irritation *in vivo*. A wide range of chemicals has been tested in the validation underlying this Test Guideline and the database of the validation study amounted to 58 chemicals in total (16) (18) (23). The Test Guideline is applicable to solids, liquids, semi-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 pre-treatment of the sample is required. Gases and aerosols have not been assessed yet in a validation study (29). 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. 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. 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 Eskes et al. 2012 (30)), the Test Guideline should not be used for that specific category of mixtures. Similar care should be taken in case specific chemical classes or physico-chemical properties are found not to be applicable to the current Test Guideline.

10. 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 cell viability measurements and need the use of adapted controls for corrections (see paragraphs 28-34).

11. A single testing run composed of three replicate tissues should be sufficient for a test chemical when the classification is unequivocal. However, in cases of borderline results, such as non-concordant replicate measurements and/or mean percent viability equal to 50 ± 5%, a second run should be considered, as well as a third one in case of discordant results between the first two runs.

PRINCIPLE OF THE TEST

12. 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 multilayered, highly differentiated model of the human *epidermis*. It consists of organized basal, spinous and granular layers, and a multilayered *stratum corneum* containing intercellular lamellar lipid layers representing main lipid classes analogous to those found *in vivo*.

13. Chemical-induced skin irritation, manifested mainly by erythema and oedema, is the result of a cascade of events beginning with penetration of the chemicals through the *stratum corneum* where they may damage the underlying layers of keratinocytes and other skin cells. The damaged cells may either release inflammatory mediators or induce an inflammatory cascade which also acts on the cells in the *dermis*, particularly the stromal and endothelial cells of the blood vessels. It is the dilation and increased permeability of the endothelial cells that produce the observed erythema and oedema (29). Notably, the RhE-based test methods, in the absence of any vascularisation in the *in vitro* test system, measure the initiating events in the cascade, e.g. cell / tissue damage (16) (17), using cell viability as readout.

14. Cell viability in RhE models is measured by enzymatic conversion of the vital dye MTT [3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue; CAS number 298-93-1], into a blue formazan salt that is quantitatively measured after extraction from tissues (31). Irritant chemicals are identified by their ability to decrease cell viability below defined threshold levels (*i.e.* ≤ 50%, for UN GHS Category 2). Depending on the regulatory framework and applicability of the Test Guideline, test chemicals that produce cell viabilities above the defined threshold level, may be considered non-irritants (*i.e.* > 50%, No Category).

DEMONSTRATION OF PROFICIENCY

15. Prior to routine use of any of the four validated test methods that adhere to this Test Guideline (Annex 2), laboratories should demonstrate technical proficiency, using the ten Proficiency Substances listed in Table 1. In situations where, for instance, a listed substance is unavailable, 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 (8)) provided that the same selection criteria as described in Table 1 are applied. Using an alternative proficiency substance should be justified.

16. As part of the proficiency testing, it is recommended that users verify the barrier properties of the tissues after receipt as specified by the RhE model producer. 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 acquired and 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 at regular intervals.

Table 1: Proficiency Substances[1]

Substance	CAS NR	*In vivo* score[2]	Physical state	UN GHS Category
NON-CLASSIFIED SUBSTANCES (UN GHS No Category)				
naphthalene acetic acid	86-87-3	0	Solid	No Cat.
isopropanol	67-63-0	0.3	Liquid	No Cat.
methyl stearate	112-61-8	1	Solid	No Cat.
heptyl butyrate	5870-93-9	1.7	Liquid	No Cat. (*Optional Cat. 3*)[3]
hexyl salicylate	6259-76-3	2	Liquid	No Cat. (*Optional Cat. 3*)[3]
CLASSIFIED SUBSTANCES (UN GHS Category 2)				
cyclamen aldehyde	103-95-7	2.3	Liquid	Cat. 2
1-bromohexane	111-25-1	2.7	Liquid	Cat. 2
potassium hydroxide (5% aq.)	1310-58-3	3	Liquid	Cat. 2
1-methyl-3-phenyl-1-piperazine	5271-27-2	3.3	Solid	Cat. 2
heptanal	111-71-7	3.4	Liquid	Cat. 2

[1] The Proficiency Substances are a subset of the substances used in the validation study and the selection is based on the following criteria; (i), the chemicals substances are commercially available; (ii), they are representative of the full range of Draize irritancy scores (from non-irritant to strong irritant); (iii), they have a well-defined chemical structure; (iv), they are representative of the chemical functionality used in the validation process; (v) they provided reproducible *in vitro* results across multiple testing and multiple laboratories; (vi) they were correctly predicted *in vitro*, and (vii) they are not associated with an extremely toxic profile (e.g. carcinogenic or toxic to the reproductive system) and they are not associated with prohibitive disposal costs.

[2] *In vivo* score in accordance with the OECD Test Guideline 404 (4).

[3] Under this Test Guideline, the UN GHS optional Category 3 (mild irritants) (1) is considered as No Category.

PROCEDURE

17. The following is a description of the components and procedures of a RhE test method for skin irritation assessment (See also Annex 3 for parameters related to each test method). Standard Operating Procedures (SOPs) for the four test methods complying with this TG are available (32) (33) (34) (35).

RHE TEST METHOD COMPONENTS

General conditions

18. 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*. *Stratum corneum* should be multilayered 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. 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

19. The assay used for quantifying viability is the MTT-assay (31). The viable cells of the RhE tissue construct can 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 optical density (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 (36). 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 (in the Skin Irritation Test Method conditions) are established by the RhE model developer/supplier. Acceptability ranges 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 the negative control are stable in culture (provide similar viability measurements) for the duration of the test exposure period.

Table 2: Acceptability ranges for negative control OD values of the test methods included in this TG

	Lower acceptance limit	Upper acceptance limit
EpiSkin™ (SM)	≥ 0.6	≤ 1.5
EpiDerm™ SIT (EPI-200)	≥ 0.8	≤ 2.8
SkinEthic™ RHE	≥ 0.8	≤ 3.0
LabCyte EPI-MODEL24 SIT	≥ 0.7	≤ 2.5

Barrier function

20. The *stratum corneum* and its lipid composition should be sufficient to resist the rapid penetration of cytotoxic benchmark chemicals, *e.g.* SDS or Triton X-100, as estimated by IC_{50} or ET_{50} (Table 3).

Morphology

21. Histological examination of the RhE model should be provided demonstrating human *epidermis*-like structure (including multilayered *stratum corneum*).

Reproducibility

22. The results of the positive and negative controls of the test method should demonstrate reproducibility over time.

Quality control (QC)

23. 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 19), *barrier function* (paragraph 20) and *morphology* (paragraph 21) are the most relevant. These data should be provided to the test method users, so that they are able to include this information in the test report. An acceptability range (upper and lower limit) for the IC_{50} or the ET_{50} should be established by the RhE model developer/supplier. Only results produced with qualified tissues can be accepted for reliable prediction of irritation classification. The acceptability ranges for the four test methods included in this TG are given in Table 3.

Table 3: QC batch release criteria of the test methods included in this TG

	Lower acceptance limit	Upper acceptance limit
EpiSkin™ (SM) (18 hours treatment with SDS) (32)	IC_{50} = 1.0 mg/ml	IC_{50} = 3.0 mg/ml
EpiDerm™ SIT (EPI-200) (1% Triton X-100) (33)	ET_{50} = 4.0 hr	ET_{50} = 8.7 hr
SkinEthic™ RHE (1% Triton X-100) (34)	ET_{50} = 4.0 hr	ET_{50} = 10.0 hr
LabCyte EPI-MODEL24 SIT (18 hours treatment with SDS) (35)	IC_{50} = 1.4 mg/ml	IC_{50} = 4.0 mg/ml

Application of the Test Chemical and Control Substances

24. At least three replicates should be used for each test chemical and for the controls in each run. 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.* ranging from 26 to 83 $\mu L/cm^2$ or mg/cm^2 (see Annex 3). For solid chemicals, the *epidermis* surface should be moistened with deionised or distilled water before application, to improve contact between the test chemical and the *epidermis* surface. Whenever possible, solids should be tested as a fine powder. A nylon mesh may be used as a spreading aid in some cases (see Annex 3). At the end of the exposure period, the test chemical should be carefully washed from the *epidermis* surface with aqueous buffer, or 0.9% NaCl. Depending on the RhE test methods used, the exposure period ranges between 15 and 60 minutes, and the incubation temperature between 20 and 37°C. These exposure periods and temperatures are optimized for each individual RhE test method and represent the different intrinsic properties of the test methods (*e.g.* barrier function) (see Annex 3).

25. Concurrent negative control (NC) and positive control (PC) should be used in each run to demonstrate that viability (using the NC), barrier function and resulting tissue sensitivity (using the PC) of the tissues are within a defined historical acceptance range. The suggested PC is 5% aqueous SDS. The suggested NCs is either water or phosphate buffered saline (PBS).

Cell Viability Measurements

26. According to the test procedure, it is essential that the viability measurement is not performed immediately after exposure to the test chemical, but after a sufficiently long post-treatment incubation period of the rinsed tissue in fresh medium. This period allows both for recovery from weak cytotoxic

effects and for appearance of clear cytotoxic effects. A 42 hours post-treatment incubation period was found optimal during test optimisation of two of the RhE-based test methods underlying this TG (11) (12) (13) (14) (15).

27. The MTT assay is a standardised quantitative method which should be used to measure cell viability under this Test Guideline. It is compatible with use in a three-dimensional tissue construct. The tissue sample is placed in MTT solution of appropriate concentration (*e.g.* 0.3 - 1 mg/mL) for 3 hours. The MTT is converted into blue formazan by the viable cells. 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 using an HPLC/UPLC-spectrophotometry procedure (see paragraph 34) (36).

28. Optical properties of the test chemical or its chemical action on MTT (e.g. chemicals may prevent or reverse the colour generation as well as cause it) may interfere with the assay leading to a false estimate of viability. This may occur when a specific test chemical is not completely removed from the tissue by rinsing or when it penetrates the *epidermis*. If a test chemical acts directly on the MTT (*e.g.* MTT-reducer), is naturally coloured, or becomes coloured during tissue treatment, additional controls should be used to detect and correct for test chemical interference with the viability measurement technique (see paragraphs 29 and 33). Detailed description of how to correct direct MTT reduction and interferences by colouring agents is available in the SOPs for the four validated test methods included in this Test Guideline (32) (33) (34) (35).

29. To identify direct MTT reducers, each test chemical should be added to freshly prepared MTT solution. If the MTT mixture containing the test chemical turns blue/purple, the test chemical is presumed to directly reduce MTT and a further functional check on non-viable RhE tissues 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 a similar way as viable tissues. Each MTT reducing test chemical is applied on at least two killed tissue replicates which undergo the entire testing procedure to generate a non-specific MTT reduction (NSMTT) (32) (33) (34) (35). A single NSMTT control is sufficient per test chemical regardless of the number of independent tests/runs performed. 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).

30. 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 33 and 34). When performing the standard absorbance (OD) measurement,each interfering coloured test chemical is applied on at least two viable tissue replicates, which undergo the entire testing procedure 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 to the testing of the coloured test chemical and in case of multiple testing, an independent NSC_{living} control needs to be conducted with each test performed (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 ($\%NSC_{living}$).

31. Test chemicals that are identified as producing both direct MTT reduction (see paragraph 29) and colour interference (see paragraph 30) 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 29. 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, 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 ($\%NSC_{killed}$).

32. It is important to note that non-specific MTT reduction and non-specific colour interferences may increase the readouts of 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. 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 is already $\leq 50\%$. Nevertheless, results for test chemicals producing %NSMTT and/or $\%NSC_{living} \geq 50\%$ of the negative control should be taken with caution as this is the cut-off used to distinguish classified from not classified chemicals (see paragraph 36).

33. 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 34) (36). The HPLC/UPLC-spectrophotometry system allows for the separation of the MTT formazan from the test chemical before its quantification (36). 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 29). 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.

34. 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 (36). 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 (36) (37). 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

35. For each test method using valid RhE model batches (see paragraph 23), tissues treated with the negative control should exhibit OD reflecting the quality of the tissues that followed shipment, receipt steps and all protocol processes. Control OD values should not be below historically established boundaries. Similarly, tissues treated with the PC, *i.e.* 5% aqueous SDS, should reflect their ability to respond to an irritant chemical under the conditions of the test method (see Annex 3 and for further information SOPs of the four test methods included in this TG (32) (33) (34) (35)). Associated and appropriate measures of variability between tissue replicates, i.e., standard deviations (SD) should fall within the acceptance limits established for the test method used (see Annex 3).

Interpretation of Results and Prediction Model

36. The OD values obtained with each test chemical can be used to calculate the percentage of viability normalised to the negative control, which is set to 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 value of percentage cell viability distinguishing irritant from non-classified test chemicals and the statistical procedure(s) used to evaluate the results and identify irritant chemicals should be clearly defined, documented, and proven to be appropriate (see SOPs of the test methods for information). The cut-off values for the prediction of irritation are given below:

- The test chemical is identified as requiring classification and labelling according to UN GHS (Category 2 or Category 1) if the mean percent tissue viability after exposure and post-treatment incubation is less than or equal (≤) to 50%. Since the RhE test methods covered by this TG cannot resolve between UN GHS Categories 1 and 2, further information on skin corrosion will be required to decide on its final classification [see also the IATA Guidance Document (3)]. In case the test chemical is found to be non-corrosive (e.g., based on TG 430, 431 or 435), and shows tissue viability after exposure and post-treatment incubation is less than or equal (≤) to 50%, the test chemical is considered to be irritant to skin in accordance with UN GHS Category 2.

- Depending on the regulatory framework in member countries, the test chemical may be considered as non-irritant to skin in accordance with UN GHS No Category if the tissue viability after exposure and post-treatment incubation is more than (>) 50%.

DATA AND REPORTING

Data

37. For each run, data from individual replicate tissues (*e.g.* OD values and calculated percentage cell viability data for each test chemical, including classification) should be reported in tabular form, including data from repeat experiments as appropriate. In addition means ± SD for each run should be reported. Observed interactions with MTT reagent and 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 the choice, 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 predictivity;
  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); Does of test chemical and control substances used;
- Duration and temperature of exposure and post-exposure incubation period;
- 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);
- Description of decision criteria/prediction model applied based on the RhE model used;
- Description of any modifications to 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 chemical for each run and each replicate measurement including OD or MTT formazan peak area, percent tissue viability, mean percent tissue viability and SD;
- If applicable, results of controls used for direct MTT-reducers and/or colouring test chemicals including OD or MTT formazan peak area, %NSMTT, $\%NSC_{living}$, $\%NSC_{killed}$, SD, 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) United Nations (UN). (2013). Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Second Revised Edition, UN New York and Geneva, 2013. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html].

2) EURL-ECVAM. (2009). Statement on the "Performance Under UN GHS of Three *In Vitro* Assays for Skin Irritation Testing and the Adaptation of the Reference Chemicals and Defined Accuracy Values of the ECVAM Skin Irritation Performance Standards", Issued by the ECVAM Scientific Advisory Committee (ESAC30), 9 April 2009. Available at: [http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing]

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

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

5) OECD. (2015). OECD Guideline for the Testing of Chemicals (No. 430.): *In Vitro* Skin Corrosion: Transcutaneous Electrical Resistance (TER).

6) OECD. (2015). OECD Guideline for the Testing of Chemicals (No. 431.): *In Vitro* Skin Model. Organisation for Economic Cooperation and Development, Paris.

7) OECD. (2015). OECD Guideline for the Testing of Chemicals (No. 435.); *In Vitro* Membrane Barrier Test Method. Organisation for Economic Cooperation and Development, Paris.

8) OECD. (2015). Performance Standards for the Assessment of Proposed Similar or Modified *In Vitro* Reconstructed Human *Epidermis* (RhE) Test Methods for Skin Irritation in Relation to TG 439. Environment, health and Safety Publications, Series on Testing and Assessment (No. 220). Organisation for Economic Cooperation and Development, Paris.

9) 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.

10) Fentem, J.H., Briggs, D., Chesné, C., Elliot, G.R., Harbell, J.W., Heylings, J.R., Portes, P., Roguet, R., van de Sandt, J.J. M. and Botham, P. (2001). A Prevalidation Study on *In Vitro* Tests for Acute Skin Irritation, Results and Evaluation by the Management Team, *Toxicol. in Vitro* 15, 57-93.

11) Portes, P., Grandidier, M.-H., Cohen, C. and Roguet, R. (2002). Refinement of the EPISKIN Protocol for the Assessment of Acute Skin Irritation of Chemicals: Follow-Up to the ECVAM Prevalidation Study, *Toxicol. in Vitro* 16, 765–770.

12) Kandárová, H., Liebsch, M., Genschow, E., Gerner, I., Traue, D., Slawik, B. and Spielmann, H. (2004). Optimisation of the EpiDerm Test Protocol for the Upcoming ECVAM Validation Study on *In Vitro* Skin Irritation Tests, *ALTEX* 21, 107–114.

13) Kandárová, H., Liebsch, M., Gerner, I., Schmidt, E., Genschow, E., Traue, D. and Spielmann, H. (2005), The EpiDerm Test Protocol for the Upcoming ECVAM Validation Study on *In Vitro* Skin Irritation Tests – An Assessment of the Performance of the Optimised Test, *ATLA* 33, 351-367.

14) Cotovio, J., Grandidier, M.-H., Portes, P., Roguet, R. and Rubinsteen, G. (2005). The *In Vitro* Acute Skin Irritation of Chemicals: Optimisation of the EPISKIN Prediction Model Within the Framework of the ECVAM Validation Process, *ATLA* 33, 329-349.

15) Zuang, V., Balls, M., Botham, P.A., Coquette, A., Corsini, E., Curren, R.D., Elliot, G.R., Fentem, J.H., Heylings, J.R., Liebsch, M., Medina, J., Roguet, R., van De Sandt, J.J.M., Wiemann, C. and Worth, A. (2002). Follow-Up to the ECVAM Prevalidation Study on *In Vitro* Tests for Acute Skin Irritation, The European Centre for the Validation of Alternative Methods Skin Irritation Task Force report 2, *ATLA* 30, 109-129.

16) Spielmann, H., Hoffmann, S., Liebsch, M., Botham, P., Fentem, J., Eskes, C., Roguet, R., Cotovio, J., Cole, T., Worth, A., Heylings, J., Jones, P., Robles, C., Kandárová, H., Gamer, A., Remmele, M., Curren, R., Raabe, H., Cockshott, A., Gerner, I. and Zuang, V. (2007). The ECVAM International Validation Study on *In Vitro* Tests for Acute Skin Irritation: Report on the Validity of the EPISKIN and EpiDerm Assays and on the Skin Integrity Function Test, *ATLA* 35, 559-601.

17) Hoffmann S. (2006). ECVAM Skin Irritation Validation Study Phase II: Analysis of the Primary Endpoint MTT and the Secondary Endpoint IL1-α. Available at: [http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing]

18) Eskes C., Cole, T., Hoffmann, S., Worth, A., Cockshott, A., Gerner, I. and Zuang, V. (2007). The ECVAM International Validation Study on *In Vitro* Tests for Acute Skin Irritation: Selection of Test Chemicals, *ATLA* 35, 603-619.

19) Cotovio, J., Grandidier, M.-H., Lelièvre, D., Roguet, R., Tinois-Tessonneaud, E. and Leclaire, J. (2007). *In Vitro* Acute Skin Irritancy of Chemicals Using the Validated EPISKIN Model in a Tiered Strategy - Results and Performances with 184 Cosmetic Ingredients, *AATEX*, 14, 351-358.

20) EURL-ECVAM. (2007). Statement on the Validity of *In Vitro* Tests for Skin Irritation, Issued by the ECVAM Scientific Advisory Committee (ESAC26), 27 April 2007. Available at: [http://www.ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing].

21) EURL-ECVAM. (2007). Performance Standards for Applying Human Skin Models to *In Vitro* Skin Irritation Testing. Available at: [http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing] *N.B. These are the original PS used for the validation of two test methods. These PS should not be used any longer as an updated version (8) is now available.*

22) EURL-ECVAM. (2008). Statement on the Scientific Validity of *In Vitro* Tests for Skin Irritation Testing, Issued by the ECVAM Scientific Advisory Committee (ESAC29), 5 November 2008. Available at: [http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing].

23) OECD. (2010). Explanatory Background Document to the OECD Draft Test Guideline on *In Vitro* Skin Irritation Testing. Environment, Health and Safety Publications. Series on Testing and Assessment, (No. 137.), Organisation for Economic Cooperation and Development, Paris.

24) Katoh, M., Hamajima, F., Ogasawara, T. and Hata K. (2009). Assessment of Human Epidermal Model LabCyte EPI-MODEL for *In Vitro* Skin Irritation Testing According to European Centre for the Validation of Alternative Methods (ECVAM)-Validated Protocol, *J Toxicol* Sci, 34, 327-334

25) Katoh, M. and Hata K. (2011). Refinement of LabCyte EPI-MODEL24 Skin Irritation Test Method for Adaptation to the Requirements of OECD Test Guideline 439, *AATEX*, 16, 111-122

26) OECD. (2011). Validation Report for the Skin Irritation Test Method Using LabCyte EPI-MODEL24. Environment, Health and Safety Publications, Series on Testing and Assessment (No. 159.), Organisation for Economic Cooperation and Development, Paris.

27) OECD. (2011). Peer Review Report of Validation of the Skin Irritation Test Using LabCyte EPI-MODEL24. Environment, Health and Safety Publications, Series on Testing and Assessment (No. 155.), Organisation for Economic Cooperation and Development, Paris.

28) Kojima, H., Ando, Y., Idehara, K., Katoh, M., Kosaka, T., Miyaoka, E., Shinoda, S., Suzuki, T., Yamaguchi, Y., Yoshimura, I., Yuasa, A., Watanabe, Y. and Omori, T. (2012). Validation Study of the In Vitro Skin Irritation Test with the LabCyte EPI-MODEL24, *Altern Lab Anim*, 40, 33-50.

29) Welss, T., Basketter, D.A. and Schröder, K.R. (2004). *In Vitro* Skin Irritation: Fact and Future. State of the Art Review of Mechanisms and Models, *Toxicol. In Vitro* 18, 231-243.

30) 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).

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

32) EpiSkin™. (February 2009). SOP, Version 1.8ECVAM Skin Irritation Validation Study: Validation of the EpiSkin™ Test Method 15 min - 42 hours for the Prediction of acute Skin Irritation of Chemicals. Available at: [http://www.ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing.html].

33) EpiDerm™. (Revised March 2009). SOP, Version 7.0, Protocol for: *In Vitro* EpiDerm™ Skin Irritation Test (EPI-200-SIT), for Use with MatTek Corporation's Reconstructed Human Epidermal Model EpiDerm (EPI-200). Available at: [http://www.ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing.html].

34) SkinEthic™ RHE SOP, Version 2.0 (February 2009), SkinEthic Skin Irritation Test-42bis Test Method for the Prediction of Acute Skin Irritation of Chemicals: 42 Minutes Application + 42 Hours Post-Incubation. Available at: [http://www.ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing.html].

35) LabCyte. (June 2011). EPI-MODEL24 SIT SOP, Version 8.3, Skin Irritation Test Using the Reconstructed Human Model "LabCyte EPI-MODEL24" Available at:

36) 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.

37) 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].

38) Harvell, J.D., Lamminstausta, K., and Maibach, H.I. (1995). Irritant Contact Dermatitis, in: Practical Contact Dermatitis, pp 7-18, (Ed. Guin J. D.). Mc Graw-Hill, New York.

39) EURL-ECVAM. (2009). Performance Standards for *In Vitro* Skin Irritation Test Methods Based on Reconstructed Human Epidermis (RhE). Available at: [http://www.ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing] *N.B. This is the current version of the ECVAM PS, updated in 2009 in view of the implementation of UN GHS. These PS should not be used any longer as an updated version (8) is now available related to the present TG.*

40) EURL-ECVAM. (2009). ESAC Statement on the Performance Standards (PS) for *In Vitro* Skin Irritation Testing Using Reconstructed Human Epidermis, Issued by the ECVAM Scientific Advisory Committee (ESAC31), 8 July 2009. Available at: [http://www.ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing].

41) EC. (2001). Commission Directive 2001/59/EC of 6 August 2001 Adapting to Technical Progress for the 28th Time Council Directive 67/548/EEC on the Approximation of Laws, Regulations and Administrative Provisions Relating to the Classification, Packaging and Labelling of Dangerous Substances, Official Journal of the European Union L225, 1-333.

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 (9).

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 (9).

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

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 a 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.

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

NSC: Non-Specific Colour in living tissues.

NSMTT: Non-Specific MTT reduction.

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

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.
Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (9).

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 (9).

Replacement test: A test which is designed to substitute for a test that is in routine use and accepted for hazard identification and/or risk assessment, and which has been determined to provide equivalent or improved protection of human or animal health or the environment, as applicable, compared to the accepted test, for all possible testing situations and chemicals (9).

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 test chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (9).

Skin irritation *in vivo*: The production of reversible damage to the skin following the application of a test chemical for up to 4 hours. Skin irritation is a locally arising reaction of the affected skin tissue and appears shortly after stimulation (38). It is caused by a local inflammatory reaction involving the innate (non-specific) immune system of the skin tissue. Its main characteristic is its reversible process involving inflammatory reactions and most of the clinical characteristic signs of irritation (erythema, oedema, itching and pain) related to an inflammatory process.

Specificity: The proportion of all negative/inactive test chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (9).

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

TEST METHODS INCLUDED IN THIS TG

Nr.	Test method name	Validation study type	References
1	**EpiSkin™**	Full prospective validation study (2003-2007). The test method components of this method were used to define the essential test method components of the original and updated ECVAM PS (39) (40) (21)*. Moreover, the method's data relating to identification of non-classified vs classified substances formed the main basis for defining the specificity and sensitivity values of the original PS*.	(2) (10) (11) (14) (15) (16) (17) (18) (19) (20) (21) (23) (32) (39) (40)
2	**EpiDerm™ SIT (EPI-200)**	EpiDerm™ (*original*): Initially the test method underwent full prospective validation together with Nr. 1. from 2003-2007. The test method components of this method were used to define the essential test methods components of the original and updated ECVAM PS (39) (40) (21)*. **EpiDerm™ SIT (EPI-200):** A modification of the original EpiDerm™ was validated using the original ECVAM PS (21) in 2008*	(2) ((10) (12) (13) (15) (16) (17) (18) (20) (21) (23) (33) (39) (40) (2) (21) (22) (23) (33)
3	**SkinEthic™ RHE**	Validation study based on the original ECVAM Performance Standards (21) in 2008*.	(2) (21) (22) (23) (31)
4	**LabCyte EPI-MODEL24 SIT**	Validation study (2011-2012) based on the Performance Standards (PS) of OECD TG 439 (8) which are based on the updated ECVAM PS* (39) (40).	(24) (25) (26) (27) (28) (35) (39) (40) and PS of this TG (8)*

*) The original ECVAM Performance Standards (PS) (21) were developed in 2007 upon completion of the prospective validation study (16) which had assessed the performance of test methods Nr 1 and 2 in reference to the classification system as described in the 28th amendment to the EU Dangerous Substances Directive (41). In 2008 the UN GHS was introduced (1), effectively shifting the cut-off value for distinguishing non-classified from classified substances from an *in vivo* score of 2.0 to 2.3. To adapt to this changed regulatory requirement, the accuracy values and reference chemical list of the ECVAM PS were updated in 2009 (2) (39) (40). As the original PS, also the updated PS were largely based data from methods Nr. 1 and 2 (16), but additionally used data on reference chemicals from method Nr. 3. In 2010, the updated ECVAM PS were used for stipulating the PS related to this TG (8). For the purpose of this Test Guideline, EpiSkin™ is considered the VRM, due to the fact that it was used to develop all the criteria of the PS.. Detailed information on the validation studies, a compilation of the data generated as well as background to the necessary adaptations of the PS as a consequence of the UN GHS implementation can be found in the ECVAM/BfR explanatory background document to this OECD TG (23).

SIT: Skin Irritation Test
RHE: Reconstructed Human Epidermis

ANNEX 3

PROTOCOL PARAMETERS SPECIFIC TO EACH OF THE TEST METHODS INCLUDED IN THIS TG

The RhE methods do show very similar protocols and notably all use a post-incubation period of 42 hours (32) (33) (34) (35). Variations concern mainly three parameters relating to the different barrier functions of the test methods and listed here: A) pre-incubation time and volume, B) Application of test chemicals and C) Post-incubation volume.

	EpiSkin™ (SM)	**EpiDerm™ SIT (EPI-200)**	**SkinEthic RHE™**	**LabCyte EPI-MODEL24 SIT**
A) Pre-incubation				
Incubation time	18- 24 hours	18-24 hours	< 2 hours	15-30 hours
Medium volume	2mL	0.9mL	0.3 or 1mL	0.5mL
B) Test chemical application				
For liquids	10µL (26µL/cm^2)	30µL (47µL/cm^2)	16µL (32µL/cm^2)	25µL (83µL/cm^2)
For solids	10mg (26mg/cm^2) + DW (5µL)	25mg (39mg/cm^2) + DPBS (25µL)	16mg (32mg/cm^2) + DW (10µL)	25mg (83mg/cm^2) + DW (25µL)
Use of nylon mesh	Not used	If necessary	Applied	Not used
Total application time	15 minutes	60 minutes	42 minutes	15 minutes
Application temperature	RT	a) at RT for 25 minutes b) at 37ºC for 35 minutes	RT	RT
C) Post-incubation volume				
Medium volume	2 mL	0.9mL x 2	2 mL	1 mL
D) Maximum acceptable variability				
Standard deviation between tissue replicates	SD≤18	SD≤18	SD≤18	SD≤18

RT: Room temperature
DW: distilled water
DPBS: Dulbecco's Phosphate Buffer Saline

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 (36) (37)	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 Quality Controls: %Dev ≤ 15% and CV ≤ 15%
Reproducibility (inter-day)	Day 1: 1 calibration curve and Quality Controls in isopropanol (n=3) Day 2: 1 calibration curve and Quality Controls in isopropanol (n=3) 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.