OECD KEY EVENT BASED GUIDELINE FOR THE TESTING OF CHEMICALS

DRAFT TG 442D - In vitro skin sensitisation assays addressing the Adverse Outcome Pathway Key Event on Keratinocyte activation

Note: Appendices IA and IB are not described in this draft document which **only includes the General introduction of TG 442D** (pages 1-12) **and the draft new Appendix I C on the Epidermal Sensitisation Assay (EpiSensA)** (pages 13-30). The other test methods included in TG 442D are available at: **Link**

GENERAL INTRODUCTION

Keratinocyte activation Key Event based Test Guideline

1. A skin sensitiser refers to a substance that will lead to an allergic response following repeated skin contact as defined by the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS) (1). There is general agreement on the key biological events underlying skin sensitisation. The current knowledge of the chemical and biological mechanisms associated with skin sensitisation has been summarised as an Adverse Outcome Pathway (AOP) (2), starting with the molecular initiating event through intermediate events to the adverse effect, namely allergic contact dermatitis. This AOP focuses on chemicals that react with thiol (i.e. cysteine) and primary amines (i.e. lysine) such as organic chemicals. In this instance, the molecular initiating event (i.e. the first key event) is the covalent binding of electrophilic substances to nucleophilic centres in skin proteins. The second key event in this AOP

takes place in the keratinocytes and includes inflammatory responses as well as changes in gene expression associated with specific cell signalling pathways such as the antioxidant/electrophile response element (ARE)-dependent pathways. The third key event is the activation of dendritic cells, typically assessed by expression of specific cell surface markers, chemokines and cytokines. The fourth key event is T-cell proliferation.

2. This Test Guideline describes in vitro assays that address mechanisms described under the second Key Event of the AOP for skin sensitisation, namely keratinocyte activation (2). The Test Guideline comprises test methods to be used for supporting the discrimination between skin sensitisers and non-sensitisers in accordance with the UN GHS (1). The test methods currently described in this Test Guideline include two in vitro ARE-Nrf2 luciferase test methods and a test method based on gene expression quantification:

- The ARE-Nrf2 luciferase KeratinoSens™ test method (Appendix IA),
- The ARE-Nrf2 luciferase LuSens test method (Appendix IB), and
- The Epidermal Sensitisation Assay – EpiSensA (Appendix IC)

3. These three test methods have been considered scientifically valid. The KeratinoSens™ test method first underwent a validation study followed by an independent peer-review by EURL ECVAM Scientific Advisory Committee (ESAC) and positive recommendations by EURL ECVAM, and is considered the validated reference method (VRM) with regards to ARE-Nrf2 luciferase test methods (3) (4) (5) (6). The LuSens test method later underwent a Performance Standard-based validation study based on which it was also reviewed and received positive opinion by ESAC (7) (8) (9) (10). The EpiSensA underwent validation studies (11) followed by an independent peer review (12) conducted by the Japanese Center for the Validation of Alternative Methods (JaCVAM). It is considered a VRM with regards to test methods quantifying changes in the expression of marker genes associated with keratinocyte activation (*ATF3*, *IL-8*, *GCLM*, and *DNAJB4*), using Reverse Transcription-quantitative PCR in reconstructed human epidermis (RhE) models. Performance Standards (13) are available for this type of method too, to facilitate the validation and assessment of similar and modified RhE-based test methods.

4. The test methods included in this Test Guideline may differ in relation to the procedure used to generate the data and the readouts measured but can be used indiscriminately to address countries' requirements for test results on the keratinocytes activation Key Event of the AOP for skin sensitisation while benefiting from the Mutual Acceptance of Data.

Background and principles of the test methods included in the Key Event based Test Guidelines

5. The assessment of skin sensitisation has historically involved the use of laboratory animals. The classical methods that use guinea-pigs, the Guinea Pig Maximisation Test (GPMT) of Magnusson and Kligman and the Buehler Test (OECD TG 406) (14), assess both the induction and elicitation phases of skin sensitisation. The murine tests, the LLNA (OECD TG 429) (15) and its three non-radioactive modifications, LLNA: DA (OECD TG 442A) (16) as well as LLNA: BrdU-ELISA and BrdU-FCM (OECD TG 442B) (17), all assess the induction response exclusively, and have gained acceptance since they provide an advantage over the guinea pig tests in terms of animal welfare together with an objective measurement of the induction phase of skin sensitisation.

6. Mechanistically-based in chemico and in vitro test methods addressing the first three key events of the skin sensitisation AOP have been adopted for contributing to the evaluation of the skin sensitisation hazard potential of chemicals: the OECD TG 442C describes the Direct Peptide Reactivity Assay (18) addressing the first key event; the present Test Guideline assesses keratinocyte activation addressing the second key event and the OECD TG 442E addresses the activation of dendritic cells, the third key event of the skin sensitisation AOP (19). Finally, the fourth key event representing T-cell proliferation is indirectly assessed in the murine Local Lymph Node Assay (LLNA) (15). It should be noted that not all key events need to be assessed to determine a skin sensitisation categorization.

7. As keratinocyte activation represents only one key event of the skin sensitisation AOP (2) (20), information generated with test methods developed to address this specific key event may not be sufficient to conclude on the presence or absence of skin sensitisation potential of chemicals. Therefore data generated with the test methods described in this Test Guideline are proposed to support the discrimination between skin sensitisers (i.e. UN GHS Category 1) and non-sensitisers when used within Integrated Approaches to Testing and Assessment (IATA), together with other relevant complementary information, e.g. derived from in vitro assays addressing other key events of the skin sensitisation AOP as well as non-testing methods, including read-across from chemical analogues (20). Examples on the use of data generated with these methods within Defined Approaches, i.e. approaches standardised both in relation to the set of information sources used and in the procedure applied to derive predictions have been published (20) and are implemented in an OECD TG on defined approaches for skin sensitisation (21).

8. The test methods described in this Test Guideline cannot be used on their own, neither to sub-categorise skin sensitisers into subcategories 1A and 1B as defined by UN GHS (1), for

authorities implementing these two optional subcategories, nor to predict potency for safety assessment decisions. However, depending on the regulatory framework, positive results generated with these methods may be used on their own to classify a chemical into UN GHS category 1.

9. The term "test chemical" is used in this Test Guideline to refer to what is being tested[1] and is not related to the applicability of the test methods to the testing of mono-constituent substances, multi-constituent substances and/or mixtures. When testing in submerged cultures, it should be determined that the test chemical is dissolved in the exposure medium or at least forms a stable dispersion (e.g. by visual inspection of the test chemical dissolved/prepared at the maximal final test concentration in the exposure medium, showing that no undissolved residues remain and that no precipitate or phase separation forms if the solution is left to settle for several hours).

10. Limited information is currently available on the applicability of the test methods to multi-constituent substances/mixtures (22) (23) (24). Although not evaluated in the validation studies, the test methods may nevertheless be technically applicable to the testing of multi-constituent substances and mixtures. When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically. Moreover, when testing multi-constituent substances or mixtures, consideration should be given to possible interference of cytotoxic constituents with the observed responses (e.g. the presence of a high content of non-sensitising cytotoxic constituents may mask the response of weakly sensitising components or sensitising components present at low concentration). It might, depending on the particular case, be scientifically justified to test either single main constituents forming the major fraction or several fractions of the mixture to conclude on the sensitisation potential of the complex mixture.

[1] In June 2013, the Joint Meeting agreed that where possible, a more consistent use of the term "test chemical" describing what is being tested should be applied in new and updated Test Guidelines.

Literature

1) United Nations (UN) (2017). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Seventh revised edition, New York and Geneva, United Nations Publications. Available at: [https://www.unece.org/trans/danger/publi/ghs/ghs_rev07/07files_e0.html]

2) OECD (2012). Series on Testing and Assessment No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Organisation for Economic Cooperation and Development, Paris. Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En

3) Emter R., Ellis G., Natsch A.(2010). Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. Toxicology and Applied Pharmacology 245, 281-290.

4) Natsch A., Bauch C., Foertsch L., Gerberick F., Normann K., Hilberer A., Inglis H., Landsiedel R., Onken S., Reuter H., Schepky A., Emter R. (2011). The intra- and inter-laboratory reproducibility and predictivity of the KeratinoSens assay to predict skin sensitizers in vitro: results of a ring-study in five laboratories. Toxicol. In Vitro 25, 733-744.

5) Natsch A., Ryan C.A., Foertsch L., Emter R., Jaworska J., Gerberick G.F., Kern P. (2013). A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. Journal of Applied Toxicology 33, 1337-1352.

6) EURL-ECVAM (2014). Recommendation on the KeratinoSens™ assay for skin sensitisation testing, 42 pp. Available at: [http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/recommendation-keratinosens-skin-sensitisation].

7) Ramirez T., Mehling A., Kolle S.N., Wruck C.J., Teubner W., Eltze T., Aumann A., Urbisch D., van Ravenzwaay B., Landsiedel R. (2014). LuSens: a keratinocyte based ARE reporter gene assay for use in integrated testing strategies for skin sensitization hazard identification. Toxicol In Vitro 28, 1482-1497.

8) Ramirez T., Stein N., Aumann A., Remus T., Edwards A., Norman K.G., Ryan C., Bader J.E., Fehr M., Burleson F., Foertsch L., Wang X., Gerberick F., Beilstein P., Hoffmann S., Mehling A., van Ravenzwaay B., Landsiedel R. (2016). Intra- and inter-laboratory reproducibility and accuracy of the LuSens assay: A reporter gene-cell line to detect keratinocyte activation by skin sensitizers. Toxicol In Vitro 32, 278-286.

9) ESAC (2016). ESAC opinion on the BASF-coordinated Performance Standards-based validation of the LuSens test method for skin sensitisation testing. Available at:

[http://publications.jrc.ec.europa.eu/repository/bitstream/JRC103706/esac_opinion_2016-04_lusens_final.pdf].

10) OECD (2015). Guidance Document No 213. Performance Standards for assessment of proposed similar or modified in vitro skin sensitisation ARE-NrF2 luciferase test methods. Series on Testing and Assessment. Available at: http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm

11) OECD (2023). Series on Testing & Assessment No. 384: Epidermal Sensitisation Assay (EpiSensA) Validation Study Report; Organisation for Economic Cooperation and Development, Paris.

12) OECD (2023). Series on Testing & Assessment No. 383: Epidermal Sensitisation Assay (EpiSensA) Peer review Report; Organisation for Economic Cooperation and Development, Paris.

13) OECD (). Guidance Document No . Performance Standards (in preparation)

14) OECD (1992). OECD Guidelines for the Testing of Chemicals No. 406. Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

15) OECD (2010). OECD Guidelines for Chemical Testing No. 429. Skin sensitization: Local Lymph Node assay. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

16) OECD (2010). OECD Guidelines for Chemical Testing No. 442A.Skin sensitization: Local Lymph Node assay: DA. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

17) OECD (2010). OECD Guidelines for Chemical Testing No. 442B. Skin sensitization: Local Lymph Node assay: BrdU-ELISA. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

18) OECD (2015). OECD Guideline for the Testing of Chemicals No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA). Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

19) OECD (2017). OECD Guideline for the Testing of Chemicals No. 442E: In Vitro Skin Sensitisation assays addressing the AOP Key Event on Activation of Dendritic Cells. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

20) OECD (2016). Series on Testing & Assessment No. 256: Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated

Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Annex 1 and Annex 2. ENV/JM/HA(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [https://community.oecd.org/community/iatass].

21) OECD (2023), *Guideline No. 497: Defined Approaches on Skin Sensitisation*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/b92879a4-en.

22) Andres E., Sa-Rocha V.M., Barrichello C., Haupt T., Ellis G., Natsch A. (2013). The sensitivity of the KeratinoSens™ assay to evaluate plant extracts: A pilot study. Toxicology In Vitro 27, 1220-1225.

23) Kolle, S.N., Mehling A., Teubner W., van Ravenzwaay B., Landsiedel R. (2013). Alternative method in practice: Post-validation experience of the skin sensitization in vitro test strategy. Toxicology Letters 221, S204.

24) Settivari RS, Gehen SC, Amado RA, Visconti NR, Boverhof DR, Carney EW (2015). Application of the KeratinoSens™ assay for assessing the skin sensitization potential of agrochemical active ingredients and formulations. Regul Toxicol Pharmacol; 72(2):350-60.

Annex: 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 (3).

AOP (Adverse Outcome Pathway): sequence of events from the chemical structure of a target chemical or group of similar chemicals through the molecular initiating event to an in vivo outcome of interest (2).

ARE: Antioxidant response element (also called EpRE, electrophile response element), is a response element found in the upstream promoter region of many cytoprotective and phase II genes. When activated by Nfr2, it mediates the transcriptional induction of these genes.

CV: Cell viability

Coefficient of variation: a measure of variability that is calculated for a group of replicate data by dividing the standard deviation by the mean. It can be multiplied by 100 for expression as a percentage.

CV75: The estimated concentration resulting in 75% cell viability.

EC1.5: Interpolated concentration resulting in a 1.5 fold luciferase induction.

Fold luciferase activity induction: Represents the ratio of luminescence of treated cells (minus blank) over the luminescence of the cells exposed to the concurrent solvent/vehicle control (minus blank).

IC30: Concentration effecting a reduction of cellular viability by 30%.

IC50: Concentration effecting a reduction of cellular viability by 50%.

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

IATA (Integrated Approach to Testing and Assessment): A structured approach used for hazard identification (potential), hazard characterisation (potency) and/or safety assessment

(potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weighs all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

Imax: Maximal induction factor of luciferase activity compared to the solvent (negative) control measured at any test chemical concentration.

Keap1: Kelch-like ECH-associated protein 1, is a sensor protein that can regulate the Nrf2 activity. Under un-induced conditions the Keap1 sensor protein targets the Nrf2 transcription factor for ubiquitinylation and proteolytic degradation in the proteasome. Covalent modification of the reactive cysteine residues of Keap 1 by small molecules can lead to dissociation of Nrf2 from Keap1 (4) (5) (6).

Mixture: A mixture or a solution composed of two or more substances in which they do not react (1).

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

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.

Negative control: A sample containing all components of a test system and treated with a substance known not to induce a positive response in the test system. This sample is processed with test chemical-treated samples and other control samples.

Nrf2: nuclear factor (erythroid-derived 2)-like 2, is a transcription factor involved in the antioxidant response pathway. When Nrf2 is not ubiquitinylated, it builds up in the cytoplasm and translocates into the nucleus, where it combines to the ARE in the upstream promoter region of many cytoprotective genes, initiating their transcription (4) (5) (6).

Performance standards: Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are (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 (3).

Positive control: A sample 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.

Proficiency chemicals (substances): A subset of the Reference Chemicals included in the Performance Standards that can be used by laboratories to demonstrate technical competence with a standardised test method. Selection criteria for these substances typically include that they represent the range of responses, are commercially available, and have high quality reference data available.

Reference chemicals (substances): A set of chemicals to be used to demonstrate the ability of a new test method to meet the acceptability criteria demonstrated by the validated reference test method(s). These chemicals should be representative of the classes of chemicals for which the test method is expected to be used, and should represent the full range of responses that may be expected from the chemicals for which it may be used, from strong, to weak, to negative.

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 (3).

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability (3).

Reproducibility: The agreement among results obtained from testing the same substance using the same test protocol (see reliability) (3).

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 (3).

Solvent/vehicle control: A replicate containing all components of a test system except of the test chemical, but including the solvent that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved in the same solvent.

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 (3).

Substance: 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 (1).

Test chemical: The term "test chemical" is used to refer to what is being tested.

United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS): A system proposing the classification of chemicals (substances and mixtures) according to standardised 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).

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

Validated Reference Method (VRM): the first method(s) endorsed as scientific valid and used as a reference for performance-based validation studies.

Valid test method: A test method considered to have sufficient relevance and reliability for a specific purpose and which is based on scientifically sound principles. A test method is never valid in an absolute sense, but only in relation to a defined purpose (3).

Xeno-free: which does not contain any element that is not from the same species as the cells used, in this case, human.

Literature

1) United Nations (UN) (2017). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Seventh revised edition, New York and Geneva, United Nations Publications. Available at: [https://www.unece.org/trans/danger/publi/ghs/ghs_rev07/07files_e0.html]

2) OECD (2012). Series on Testing and Assessment No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Organisation for Economic Cooperation and Development, Paris. Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En

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

4) Natsch A. (2010). The Nrf2-Keap1-ARE Toxicity Pathway as a Cellular Sensor for Skin Sensitizers-Functional Relevance and Hypothesis on Innate Reactions to Skin Sensitizers. Toxicological Sciences 113, 284-292.

5) Dinkova-Kostova A.T., Holtzclaw W.D., Kensler T.W. (2005). The role of Keap1 in cellular protective responses. Chem. Res. Toxicol. 18, 1779-1791.

6) Kansanen E., Kuosmanen S.M., Leinonen H., Levonen A.L. (2013). The Keap1-Nrf2 pathway: Mechanisms of activation and dysregulation in cancer. Redox Biol. 1, 45-49.

Appendix IC: In Vitro Skin Sensitisation: Epidermal Sensitisation Assay (EpiSensA)

INITIAL CONSIDERATIONS AND LIMITATIONS

1. The EpiSensA method is proposed to address the second key event of the skin sensitisation Adverse Outcome Pathway (AOP) (1), namely keratinocyte activation, by quantifying changes in the expression of marker genes associated with keratinocyte activation in reconstructed human epidermis (RhE) models following exposure to sensitisers (2) (3). Two important keratinocyte responses occur in the skin sensitisation AOP: the inflammatory responses and the induction of cytoprotective gene pathways (2). Expression of the activating transcription factor 3 (*ATF3*) and interleukin-8 (*IL-8*) genes reflects the inflammatory response of keratinocytes, whereas expression of the glutamate-cysteine ligase modifier subunit (*GCLM*) and DnaJ (Hsp40) homolog subfamily B (*DNAJB4*) genes reflects the induction of cytoprotective gene pathways (3). In the EpiSensA, relative changes in marker gene (i.e. *ATF3*, *IL-8*, *GCLM*, and *DNAJB4*) expression are quantified using Reverse Transcription-quantitative PCR (RT-qPCR), and these data are then used for supporting the discrimination between skin sensitisers and non-sensitisers.

2. The EpiSensA test method underwent validation studies (4) followed by an independent peer review conducted by the Japanese Center for the Validation of Alternative Methods (JaCVAM) in cooperation with the International Collaboration on Alternative Test Methods (ICATM) (5). The EpiSensA test method was considered scientifically valid to be used as part of an Integrated Approaches to Testing and Assessment (IATA) to support the discrimination between skin sensitisers and non-sensitisers for the purpose of hazard identification.

3. The RhE model used in the EpiSensA test method is the LabCyte EPI-MODEL24. The EpiSensA test method is considered a Validated Reference Method (VRM) with regards to test methods quantifying changes in the expression of marker genes associated with keratinocyte activation (*ATF3*, *IL-8*, *GCLM*, and *DNAJB4),* using Reverse Transcription-quantitative PCR in reconstructed human epidermis (RhE) models. PS (6) are available to facilitate the validation of similar or modified EpiSensA test method and allow for timely amendment of this Test Guideline for their inclusion. A similar or modified EpiSensA test

method will only be added to the test guideline after review and agreement that all criteria described in the PS are met. Other RhE models can be used after a validation study based on the PS is conducted.

4. The EpiSensA test method was shown to be transferable to laboratories experienced in cell culture techniques (4) (5) (8) (9) (10). The within-laboratory reproducibility of EpiSensA at two participating laboratories was 93.3% (14/15) and one was 86.7% (13/15). The calculated between-laboratory reproducibility at three participating laboratories for analyses of 27 test chemicals was 88.9% (9). Results generated in the validation study (4) and a published study (10) indicated that the accuracy of the EpiSensA in discriminating sensitisers (i.e. UN GHS Cat. 1) from non-sensitisers is 83.3% (120/144) with a sensitivity of 88.8% (95/107) and a specificity of 67.6% (25/37) when compared to LLNA results. The balanced accuracy is 78.2%. Compared to human results, the accuracy of the EpiSensA is 80.7% (71/88) with a sensitivity of 98.1% (53/54) and a specificity of 52.9% (18/34). The balanced accuracy is 75.5%. False negative EpiSensA predictions as compared with the LLNA are more likely for chemicals exhibiting low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals exhibiting high skin sensitisation potency (i.e. UN GHS subcategory 1A) (9). However, the accuracy values presented here for EpiSensA as a stand-alone test method are only suggestive, as results obtained with the test method should be considered in combination with information from other sources in the context of a Defined Approach or an IATA and in accordance with the provisions of paragraphs 7 and 8 of the General Introduction of this Test Guideline (11). Furthermore, when using *in vivo* data as a reference to evaluate non-animal methods for assessing skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in humans.

5. Currently available data indicate that the EpiSensA test method is applicable for the testing of chemicals representing a variety of organic functional groups, reaction mechanisms, skin sensitisation potencies (as determined through *in vivo* studies), and physico-chemical properties (3) (4) (5) (9) (10). EpiSensA is applicable for testing soluble chemicals or chemicals that form a stable dispersion in an appropriate vehicle (see paragraph 19 on *Vehicle selection and assessment of test chemical solubility*). In addition, as RhE represents a three-dimensional model involving an air-liquid interface, test chemicals are applied directly to the surface of the RhE model. Therefore, lipophilic vehicles can be used, and the test method can be used to test lipophilic chemicals (e.g. $\log K_{ow} > 3.5$) (3) (4) (5) (9) (10). Furthermore, the metabolic capacity of the RhE model is similar to that of human skin (12) (13). Therefore, pro-haptens (i.e. chemicals requiring enzymatic activation to exert skin sensitisation

potential) and pre-haptens (i.e. chemicals that become sensitisers via abiotic transformation) can be detected by the EpiSensA (3) (4) (5) (9) (10). Due to the limitations of the test exposure period, pro-haptens or pre-haptens that require more than 6 hours to be sufficiently metabolised or oxidised may not be detected by the EpiSensA.

6. Some surfactants can lead to false positive results due to non-specific expression of the *ATF3* and *IL-8* genes (4). In addition, testing chemicals at high concentrations (e.g. 100% or 50% (w/v)) using distilled water as the vehicle can induce high osmotic stress conditions that lead to non-specific expression of *ATF3* (4). Therefore, positive results in such cases should be interpreted with caution. However, negative results can still be used to support the classification of a test chemical as a non-sensitiser. When the solubility is assessed, if the chemical is not soluble or does not form a stable dispersion at 0.0122% (see paragraph 19 on *vehicle selection and assessment of test chemical solubility*), the chemical is not applicable for testing using EpiSensA. However, other vehicles can be used if sufficient scientific rationale can be provided. Test chemicals that significantly affect the expression of the endogenous control gene glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) at concentrations at which cell viability remains ≥ 80% (see *Acceptance criteria* paragraph) may not be applicable to the test method, as marker gene expression cannot be accurately measured by relative quantification using RT-qPCR. However, other endogenous control genes may be used if sufficient scientific rationale can be provided. In addition, test chemicals that either affect RNA itself (e.g. by inducing RNA degradation) or directly interfere with the RNA isolation system may not be applicable to the test method. In cases where there is evidence demonstrating the non-applicability of the EpiSensA test method to specific category of test chemicals, it should not be used for those specific categories of chemicals.

7. In addition to supporting the discrimination between skin sensitisers (i.e. UN GHS Category 1) and non-sensitisers, the EpiSensA test method also provides information (e.g. concentration-response data) that could contribute to the assessment of sensitising potency when used in integrated approaches such as a Defined Approach or an IATA. However, further data (preferably based on human research) are required to determine how EpiSensA test results might contribute to potency assessment.

8. Definitions are provided in the annex of the general introduction.

PRINCIPLE OF THE TEST

9. The EpiSensA test method is an *in vitro* assay that quantifies changes in the expression of four marker genes associated with keratinocyte activation (i.e. *ATF3*, *GCLM*, *DNAJB4*, and *IL-8*) in a RhE model following a 6-hour exposure to the test chemical of interest. Relative changes in marker gene expression are quantified using RT-qPCR. Cytotoxicity is also assessed concurrently to determine whether upregulated expression of the marker genes occurs at sub-cytotoxic concentrations (cell viability ≥ 80%). The relative induction of marker genes is calculated in comparison to vehicle controls. Test chemicals are considered positive in the EpiSensA test method if the expression of at least one marker gene exceeds the respective cut-off value (*ATF3*, 15-fold; *GCLM*, 2-fold; *DNAJB4*, 2-fold; *IL-8*, 4-fold) with cell viability remaining ≥ 80%. For this purpose, the mean maximum fold-induction (Imax) value is determined using data from concentrations at which mean cell viability remains ≥ 80%.

10. Prior to routine use of the EpiSensA test method, laboratories should demonstrate technical proficiency, using the ten proficiency substances listed in Annex 1 of this Appendix.

PROCEDURE

11. The *EpiSensA standard operating procedure* is available in the Tracking System for Alternative methods towards Regulatory acceptance (TSAR) (14) and should be employed when implementing and using the test method in the laboratory. The following paragraphs describe the main components and procedures of the EpiSensA test method, which comprises two steps: concentration-finding study and main study (Gene expression analysis).

General Test System Characterisation

12. Non-transformed human keratinocytes should be used to reconstruct the epithelium (15). Multiple layers of viable epithelial cells (*basal layer, stratum spinosum, stratum granulosum*) should be present under a functional *stratum corneum*. *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. the surfactant sodium lauryl sulphate (SLS) is used to test barrier function). 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

Barrier function

13. The RhE model developer/supplier should ensure that each batch of the RhE model meets defined quality control criteria for barrier function. The barrier function should be demonstrated and assessed by determination of the concentration at which a benchmark chemical (e.g. SLS) reduces the viability of the tissues by 50% (IC50) after a fixed exposure time, or by determination of the exposure time required to reduce cell viability by 50% (ET50) upon application of the benchmark chemical at a specified, fixed concentration.

Morphology

14. Histological examination of the RhE model should be provided by the RhE model developer/supplier demonstrating human epidermis-like structure (including multilayered stratum corneum as described in paragraph 13).

Reproducibility

15. The RhE model developer/supplier should maintain a database of the QC release test results of the viability and barrier function tests to monitor reproducibility over time. It is recommended that the EpiSensA test method user maintain a database of the EpiSensA positive and vehicle (i.e. negative) control results to monitor reproducibility of test method execution over time.

Quality control (QC)

16. The RhE model should only be used if the RhE model developer/supplier demonstrates that each batch of the RhE model used meets defined production release criteria, among which those for barrier function (paragraph 13) and morphology (paragraph 14) 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 IC50 or ET50 should be established by the RhE model developer/supplier (15). Only results produced with qualified tissues can be accepted for reliable prediction. The acceptability range for the test method included in Appendix I C is given in Table 1.

Table 1. QC batch release criteria of the RhE models included in Appendix I C

RhE model	Lower acceptance limit	Upper acceptance limit
LabCyte EPI-MODEL24 (18 hours treatment with SLS) (13)	IC_{50}=1.4 mg/mL	IC_{50}=4.0 mg/mL

Preparation of RhE model

17. The EpiSensA test method should be conducted utilising an RhE model. The LabCyte EPI-MODEL24 kit (#401124), which can be obtained from Japan Tissue Engineering Co., Ltd. (J-TEC), is currently the sole model that can be used in the EpiSensA test method. Other RhE models can be used after a validation study based on the PS (6) is conducted.

18. RhE models are cultured at 37°C with 5% CO_2 in a humidified atmosphere in the assay medium included with the LabCyte EPI-MODEL24 kit (#401124).

Vehicle selection and assessment of test chemical solubility

19. Assessment of solubility is conducted prior to testing. The solubility of each chemical is evaluated and confirmed visually. For this purpose, test chemicals are dissolved or stably dispersed at a concentration of 50% in acetone: olive oil; 4:1 v/v (AOO) as a first vehicle option, distilled water (DW) as a second vehicle option, or 50 v/v % ethanol in DW (50% EtOH) as a third vehicle option. For example, 0.1g of test chemical is measured, and 0.1mL of AOO is added. If the test chemical is not soluble or does not stably disperse (i.e. a colloid or suspension in which the test chemical does not settle or separate from the vehicle into different phases within 10 minutes of preparation at room temperature) at a concentration of 50% in any of the vehicles, the highest soluble concentration should be determined by 2-fold serial dilutions beginning with 50% down to 0.0122%. If the test chemical is not soluble or does not form a stable dispersion at 0.0122%, the chemical is not applicable for testing using EpiSensA. The appropriate vehicle is defined as the vehicle that dissolves the test chemical or forms a stable dispersion at the highest concentration tested. It should be verified whether the highest concentration determined can be prepared at weight per volume in a volumetric flask. If the highest soluble or stably dispersed concentration is determined to be 0.0488%, 0.0244%, or 0.0122%, the subsequent concentration-finding study (paragraphs 20-26) can be skipped, and main study should be performed (see paragraph 27). In cases in which a vehicle other than AOO, DW, or 50% EtOH is used, appropriate scientific rationale for use of that

vehicle should be provided.

Concentration-finding study

20. A concentration-finding assay is performed to determine the concentrations of test chemical to be used for the main study (see *Main study (Gene expression analysis)* paragraphs). In the main study, test chemical concentrations that show ≥ 80% mean cell viability should be used. Therefore, the lowest test chemical concentration that induces a < 80% cell viability is determined in concentration-finding study.

Preparation of test chemicals and control substances for the concentration-finding study

21. Test chemicals are prepared on the day of testing and dissolved or stably dispersed in an appropriate vehicle at the highest concentration determined as specified in paragraph 19. Starting from the highest concentration, 4-fold serial dilutions are prepared to 0.0122 or 0.0244% (w/v) in the corresponding vehicle. Thus, depending on the starting concentration, a varying number of dilutions are prepared and tested as exemplified in Table 2. The corresponding vehicles utilised for the preparation of the test chemicals are used as the vehicle controls. Both non-treated control and killed control are used for calculation of cell viability. Non-treated control is used to define 100% cell viability, and killed control is used to define 0% cell viability (see paragraph 25). Triton X-100 is used as the control substance for killed control in the EpiSensA test method. Triton X-100 should be prepared as a 10% (w/v) solution in DW.

Application of test chemicals and control substances for the concentration-finding study

22. For each test chemical, one run is needed to determine the concentration to be used in the main study (Gene expression analysis). One RhE tissue for each test chemical concentration and non-treated control and two tissue units for the killed control are used for the cell viability assay (Table 2). Test chemicals prepared as a working solution (5 µL) and Triton X-100 solution (10 µL) are applied to the centre of each epidermis surface using a positive-displacement pipette and tips. The treated tissue units are then incubated for 6 hours at 37°C with 5% CO_2 in a humidified atmosphere.

Table 2. Example of plate layout for the concentration-finding study using 24-well plate

1. Non-treated	2. Killed Ctrl.	3. Killed Ctrl.	13. Test chemical B 0.012% w/v	14. Test chemical B 0.049% w/v	15. Test chemical B 0.20% w/v
4. AOO	5. DW	6. 50% EtOH	16. Test chemical B 0.78% w/v	17. Test chemical B 3.13% w/v	
7. Test chemical A 0.024% w/v	8. Test chemical A 0.098% w/v	9. Test chemical A 0.39% w/v			
10. Test chemical A 1.56% w/v	11. Test chemical A 6.25% w/v	12. Test chemical A 25% w/v			

AOO: acetone: olive oil; 4:1 v/v

DW: distilled water

50% EtOH: 50 v/v % ethanol in DW

Cytotoxicity assessment

23. Cell viability is measured by a lactate dehydrogenase (LDH) assay utilising formazan as the dye. LDH is a stable cytoplasmic enzyme present in all cell types, and it is released into the cell culture medium as a result of damage to the plasma membrane. The LDH assay measures the amount of formazan dye produced by released LDH. The criteria for interference of test chemical in LDH assay (i.e. inhibition of LDH reaction) is described in the TSAR (14).

24. After a 6-hour exposure, 50 µL of the medium for each sample is placed into the wells of a 96-well plate, and an equal volume (i.e. 50 µL) of substrate solution containing lactate and tetrazolium salt is added to each well. The plate is incubated for 30 minutes at room temperature with protection from light, and the reaction is stopped by adding 25 µL/well of 1 mol/L hydrochloric acid (HCl). The absorbance of each well is then measured at 490 or 492 nm along with the reference wavelength (≥ 600 nm) using a 96-well plate absorbance reader. Δabs. is calculated by subtracting the absorbance at reference wavelength from the

absorbance at 490 or 492 nm. The absorbance should be measured immediately (no longer than 1 hour) after the addition of HCl.

25. Cell viability can be calculated using the following equation:

$$\textbf{Cell viability}\ (\%) = \mathbf{100} - \frac{\boldsymbol{\Delta}\textbf{abs. of test chemical treatment} - \boldsymbol{\Delta}\textbf{abs. of non-treated control}}{\textbf{mean } \boldsymbol{\Delta}\textbf{abs. of killed control} - \boldsymbol{\Delta}\textbf{abs. of non-treated control}} \times \mathbf{100}$$

26. If the LDH assay is not applicable to a test chemical of interest, another cytotoxicity assay (e.g. MTT assay or ATP assay) may be used. In the MTT assay, the activation of the metabolism in a mitochondria in a cell 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. ATP assay is a homogeneous method, in which tissues are lysed and the number of viable cells is determined based on quantitation of the ATP present in tissues. Therefore, total RNA cannot be isolated from the tissues used for the MTT assay or ATP assay. For this reason, tissues used for the cytotoxicity assessment are required besides tissues used for the gene expression analysis when the MTT assay or ATP assay is used. Other methods (e.g. XTT assay) can be used if sufficient scientific rationale is provided based on the PS (16).

Main study (Gene expression analysis)

Preparation of test chemicals and control substances for the main study

27. An appropriate vehicle (AOO, DW, or 50% EtOH; see paragraph 19) should be used to dissolve or stably disperse the test chemical. The lowest concentration that resulted in < 80% cell viability in the *concentration-finding study* should serve as the highest concentration (i.e. the starting concentration) in the *main study* for each chemical and is used in the negative judgement (see paragraph 39). If the cell viability was ≥ 80% at any of the tested concentrations in the *concentration-finding study*, the highest soluble or stably dispersed concentration of the test chemical should be used as the starting concentration. Based on the starting concentration, 2-fold serial dilutions are prepared using the corresponding vehicle to obtain working solutions (at least 3 concentrations are used, including the lowest concentration that resulted in < 80% cell viability in the *concentration-finding study* or the highest soluble or stably dispersed concentration). If the highest soluble or stably dispersed concentration determined in the *vehicle selection and assessment of test chemical solubility*

(paragraph 19) is 0.0488%, only 3 concentrations (0.0488, 0.0244, and 0.0122% w/v) are used. If the concentration determined in the solubility check is 0.0244% or 0.0122%, only 2 concentrations (0.0244, and 0.0122% w/v) or only 1 concentration (0.0122% w/v) are used, respectively. Likewise, if the lowest test chemical concentration that induces a < 80% cell viability is 0.0244% or 0.0122% in the *concentration-finding study*, only 2 concentrations (0.0244, and 0.0122% w/v) or only 1 concentration (0.0122% w/v) are used, respectively. The vehicle control is prepared as described in paragraph 21. Clotrimazole (CAS no. 23593-75-1, ≥ 98% purity) and 4-nitrobenzyl bromide (4-NBB) (CAS no. 100-11-8, ≥ 98% purity) are used as positive controls in the EpiSensA test method, and 0.78% (w/v) clotrimazole and 0.10% (w/v) 4-NBB solutions are prepared in AOO (working solution). To calculate cell viability, non-treated and killed controls are prepared as described in paragraph 21.

Application of test chemicals and control substances for the main study

28. For each test chemical, one run is required to obtain a prediction. Three tissue units for each test chemical concentration, positive control substance and vehicle controls, two tissue units for the killed control, and one tissue unit for the non-treated control are used for the gene expression analysis. Other than the number of tissue units, application is conducted in the same condition as described in paragraph 22 (Table 3).

Table 3. Example of plate layout for the main study

1. Non-treated	2. Killed Ctrl.	3. Killed Ctrl.	13. Test chemical 1.56% w/v	14. Test chemical A 1.56% w/v	15. Test chemical A 1.56% w/v
4. AOO	5. AOO	6. AOO	16. Test chemical A 3.13% w/v	17. Test chemical A 3.13% w/v	18. Test chemical A 3.13% w/v
7. Clotrimazole 0.78% w/v	8. Clotrimazole 0.78% w/v	9. Clotrimazole 0.78% w/v	19. Test chemical A 6.25% w/v	20. Test chemical A 6.25% w/v	21. Test chemical A 6.25% w/v
10. 4-NBB 0.10% w/v	11. 4-NBB 0.10% w/v	12. 4-NBB 0.10% w/v			

AOO: acetone: olive oil; 4:1 v/v

4-NBB: 4-nitrobenzyl bromide

Cytotoxicity assessment

29. After a 6-hour exposure to the test chemical, cell viability is determined as described in paragraphs 23-26.

RNA isolation

30. For gene expression analysis, the tissue surface is washed three times with phosphate-buffered saline, and the tissue is collected and lysed using one of two lysis methods that were used during test method development and validation (TRIzol reagent and a vortex mixer, or a shredder column and centrifuge).

31. Total RNA, including mRNA, is isolated from lysed RhE tissue samples using a commercially available kit and reagents (e.g. RNeasy Mini kit which was used during test method development and validation).

32. The RNA concentration is quantified, and the RNA quality is analysed from each sample

using an RNA analysis equipment, e.g. NanoDrop™ (Thermo Fisher Scientific), with following the protocols provided by the instrument supplier. More than 500 ng of RNA is required for complementary DNA (cDNA) synthesis. RNA concentration and quality should correspond to the recommendations described by the supplier of the reagents which are used in subsequent RT-qPCR (e.g. ≥ 100 ng/µl RNA concentration and A260/A280 in range 1.8-2.0).

RT-qPCR

33. cDNA is synthesised using the commercially available reagents (e.g. Superscript III First-Strand Synthesis System which was used during test method development and validation).

34. After cDNA synthesis, the expression levels of marker genes (i.e. *ATF3*, *GCLM*, *DNAJB4*, and *IL-8*) and the endogenous control gene (i.e. *GAPDH*) are analysed using RT-qPCR. The method described in the *EpiSensA standard operating procedure* (14) should be used (i.e. TaqMan Gene Expression Assay and TaqMan Universal PCR Master Mix). Another reagent for gene expression analysis can be used if appropriate scientific rationale for use of that reagent is provided. Other methods that quantify changes in the gene expression can be used if sufficient scientific rationale is provided based on the PS (6).

DATA AND REPORTING

Data evaluation

35. Relative gene expression is analysed using RT-qPCR. Based on the threshold cycle (Ct) value, the ΔCt and ΔΔCt values as well as fold-induction are calculated according to the following equation:

$$\textbf{ΔCt value of marker gene} = \textbf{Ct value of marker gene} - \textbf{Ct value of GAPDH}$$

$$\textbf{ΔΔCt value of marker gene} = \textbf{ΔCt value of marker gene} - \textbf{ΔCt value of markergene (vehicle control)}$$

$$\textbf{fold induction} = 2^{-\textbf{ΔΔCt value of marker gene}}$$

Cell viability is also calculated according to the equation provided in paragraph 25.

Acceptance criteria

36. The following acceptance criteria should be met for a run to be considered valid:

- The cell viability of at least two tissue units of the vehicle control should be ≥ 95%. If the cell viability of only one vehicle control is < 95%, the Ct values obtained from the remaining two tissue units should be used.

- The mean cell viability of both positive controls (i.e. 0.78% [w/v] clotrimazole and 0.10% [w/v] 4NBB) should be ≥ 80%.

- In the 0.78% (w/v) clotrimazole positive control, the mean fold-induction values for *ATF3* and *IL-8* should exceed the cut-off value (i.e. the *ATF3* fold-induction value should be > 15, and the *IL-8* fold-induction value should be > 4).

- In the 0.10% (w/v) 4NBB positive control, the mean fold-induction values for *GCLM* and *DNAJB4* should exceed the cut-off value (i.e. the *GCLM* fold-induction value should be > 2, and the *DNAJB4* fold-induction value should be > 2).

37. The following acceptance criteria should be met in order to consider a tested concentration's result valid:

- The result of at least one tested concentration should shows ≥ 80% mean cell viability. If the mean cell viability is < 80% for a given tested concentration, the result for that tested concentration should be excluded for a positive prediction but might be used for negative prediction (see para 39).

- When the mean *GAPDH* Ct value for a given test chemical concentration is within ±1 of the mean *GAPDH* Ct value of the corresponding vehicle control, the result obtained at that concentration is acceptable.

Prediction model

38. Each test chemical is evaluated in one run to derive a prediction (positive or negative). An EpiSensA prediction is considered positive if at least one of the following conditions is met:

- The Imax for *ATF3* is > 15 for at least one tested concentration.
- The Imax for *GCLM* is > 2 for at least one tested concentration.
- The Imax for *DNAJB4* is > 2 for at least one tested concentration.
- The Imax for *IL-8* is > 4 for at least one tested concentration.

39. The EpiSensA prediction is considered negative if:
    - The mean fold-induction value of the marker genes does not exceed the respective cut-off values for any of the four genes, at any of the tested concentrations and
    - At least one mean cell viability at the tested concentrations is < 80%

40. Cases can occur in which the mean fold-induction value of all four marker genes does not exceed the respective cut-off values at the tested concentration but the mean cell viability at all tested concentrations is ≥ 80%. In such cases, an additional main study should be performed using 2-fold serial dilutions beginning with the concentration greater than the highest concentration used in the first main study. However, if the test chemical does not produce a mean cell viability of < 80% at either the highest soluble or stably dispersed concentration (for solid substances) or 100% (for liquids), the test chemical result is judged as negative.

41. If all mean cell viabilities are < 80% at the tested concentrations greater than or equal to 0.0122% (w/v), the prediction is considered inconclusive.

42. Other cases can occur in which the fold-induction value of a marker gene exceeds the cut-off value only at the lowest concentration showing < 80% mean cell viability. In such a case, the test chemical should be retested using a narrower concentration-response analysis and lower dilution factor (e.g. √2 [=1.41]-fold dilution) in order to determine whether induction has occurred at a cytotoxic level (80 to 95% mean cell viability).

Test report

43. The test report should include the following information:

Test chemical

- Mono-constituent substance
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers, such as batch/lot number and expiration date;
  - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
  - Purity, chemical identity of impurities, etc., as appropriate and practically feasible;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability, to the extent available;
  - Justification for choice of vehicle for each test chemical.

- Multi-constituent substance, UVCB and mixture:
  - Characterisation as far as possible, for example, by chemical identity (see above), purity, and quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
  - Physical appearance, water solubility, and additional relevant physicochemical properties, to the extent available;
  - Molecular weight or apparent molecular weight in cases of mixtures/polymers of known composition, or other information relevant for the conduct of the study;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability, to the extent available;
  - Justification for choice of vehicle for each test chemical.

Controls

- Positive control
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;

- Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
- Purity, chemical identity of impurities, etc., as appropriate and practically feasible;
- Treatment prior to testing, if applicable (e.g. warming, grinding);
- Concentration(s) tested;
- Storage conditions and stability, to the extent available;
- Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.

- Vehicle control
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), and/or other identifiers;
  - Purity, chemical identity of impurities, etc., as appropriate and practically feasible;
  - Physical appearance, molecular weight, and additional relevant physicochemical properties in cases in which vehicles controls other than those mentioned in this Appendix are used and to the extent available;
  - Storage conditions and stability, to the extent available;
  - Justification for choice of vehicle for each test chemical.

Testing conditions

- Name and address of the sponsor, test facility, and study director;
- Description of test method used;
- RhE model used (including batch number);
- 96-well plate absorbance reader equipped for reading at 490 (or 492) nm and ≥ 600 nm;
- RNA extraction method used;
- Spectral photometer for measurement of RNA concentration;
- Thermal cycler and RT-qPCR system used (e.g. model), including instrument settings, primers, and reverse transcription (RT) and PCR reagents;
- 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.
- A statement of the proficiency of the laboratory in performing the test method (e.g. by testing of proficiency substances) or to demonstrate reproducibility of the test method over time.

Testing procedure

- Test chemical concentrations, application procedure, and exposure time used (if different than that recommended);
- Description of evaluation and decision criteria used;
- Description of study acceptance criteria used;
- Description of any modifications of the test procedure.

Results

- Tabulation of data, including individual Ct, ΔCt, ΔΔCt, fold-induction, and cell viability values obtained for the test chemical and for the positive control, and an indication of the rating of the test chemical according to the prediction model;
- A graph depicting concentration-response curves for induction of gene expression and cell viability;
- Description of any other relevant observations, if applicable.

Discussion of the results

- Discussion of the results obtained with the EpiSensA test method;

Conclusions

LITERATURE

(1) OECD (2012)10/PART1. The adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins. Part 1: Scientific evidence. Series on Testing and Assessment. ENV/JM. Mono, 168.

(2) Saito, K., Nukada, Y., Takenouchi, O., Miyazawa, M., Sakaguchi, H., Nishiyama, N. (2013). Development of a new in vitro skin sensitization assay (Epidermal Sensitization Assay; EpiSensA) using reconstructed human epidermis. Toxicology in Vitro, 8, 2213-24, doi: 10.1016/j.tiv.2013.08.007.

(3) Saito, K., Takenouchi, O., Nukada, Y., Miyazawa, M., Sakaguchi, H. (2017). An in vitro skin sensitization assay termed EpiSensA for broad sets of chemicals including lipophilic chemicals and pre/pro-haptens. Toxicology in Vitro: An International Journal Published in Association with BIBRA, 40, 11-25. doi: 10.1016/j.tiv.2016.12.005.

(4) OECD (2023). Series on Testing & Assessment No. 384: Epidermal Sensitisation Assay (EpiSensA) Validation Study Report; Organisation for Economic Cooperation and Development, Paris. Available at: https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm

(5) OECD (2023). Series on Testing & Assessment No. 383: Epidermal Sensitisation Assay (EpiSensA) Peer review Report; Organisation for Economic Cooperation and Development, Paris. Available at https://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm

(6) OECD (). Guidance Document No . Performance Standards (in preparation)

(7) OECD (2005), OECD Series on Testing and Assessment No. 34. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment.

(8) Mizumachi, H., Sakuma, M., Ikezumi, M., Saito, K., Takeyoshi, M., Imai, N., Okutomi, H., Umetsu, A., Motohashi, H., Watanabe, M., Miyazawa, M. (2018). Transferability and within- and between-laboratory reproducibilities of EpiSensA for predicting skin sensitization potential in vitro: A ring study in three laboratories. Journal of Applied Toxicology: JAT, 38(9), 1233-1243. doi: 10.1002/jat.3634.

(9) Mizumachi, H., Watanabe, M., Ikezumi, M., Kajiwara, M., Yasuda, M., Mizuno, M., Imai, N., Sakuma, M., Shibata, M., Watanabe, S., Motoyama, J., Basketter, D., Eskes, C., Hoffmann, S., Lehman, D., Ashikaga, T., Sozu, T., Takeyoshi, M., Suzuki, S., Miyazawa, M., Kojima, H. (2023), The inter-laboratory validation study of EpiSensA for predicting skin sensitization potential. Journal of Applied Toxicology: doi: 10.1002/jat.4559.

(10) Mizumachi, H., Suzuki, S., Sakuma, M., Natsui, M., Imai, N., Miyazawa, M. (2023), Reconstructed human epidermis-based Testing Strategy (RTS) of skin sensitization hazard and

potency classification using EpiSensA and *in silico* data. Journal of Applied Toxicology: doi: 10.1002/jat.4551

(11) OECD (2023), *Guideline No. 497: Defined Approaches on Skin Sensitisation*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/b92879a4-en

(12) Tokudome, Y., Katayanagi, M., Hashimoto, F. (2015). Esterase activity and intracellular localization in reconstructed human epidermal cultured skin models. Annals of Dermatology, 27(3), 269-274. doi: 10.5021/ad.2015.27.3.269.

(13) Oesch, F., Fabian, E., Landsiedel, R. (2018). Xenobiotica-metabolizing enzymes in the skin of rat, mouse, pig, guinea pig, man, and in human skin models. Archives of Toxicology, 92(8), 2411-2456. doi: 10.1007/s00204-018-2232-x.

(14) EURL ECVAM. (2023). EpiSensA standard operating procedure. Available at: https://tsar.jrc.ec.europa.eu/test-method/tm2018-01-0.

(15) OECD (2021), *Test No. 439: In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method*, OECD Publishing, Paris. https://doi.org/10.1787/9789264242845-en

ANNEX 1 - PROFICIENCY SUBSTANCES

In Vitro Skin Sensitisation: Epidermal Sensitisation Assay (EpiSensA)

Prior to routine use of the test method described in this Annex to Test Guideline 442D, laboratories should demonstrate technical proficiency by testing at least three of the four fixed concentrations of the 10 proficiency substances recommended in Table 1. Then, the results obtained for each marker gene should be consistent with those specified in Table 1 for 8 out of the 10 proficiency substances. These proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that they are commercially available, that curated *in vivo* reference data and high quality EpiSensA data are available, and that they were used during pre-validation ring study or the JaCVAM-coordinated validation study.

Table 1: Recommended substances for demonstrating technical proficiency with the EpiSensA method

No.	Proficiency substances	CAS No.	Physical state	*in vivo* prediction[1]	Vehicle	Test concentration (w/v%)	EpiSensA results for each marker gene[2]			
							ATF3	*GCLM*	*DNAJB4*	*IL-8*
1	2,4-Dinitrochlorobenzene	97-00-7	Solid	Sensitiser (GHS Cat. 1A)	AOO	0.39, 0.20, 0.10, 0.05	p	p	p	p/n
2	p-Phenylenediamine	106-50-3	Solid	Sensitiser (GHS Cat. 1A)	AOO	1.56, 0.78, 0.39, 0.20	p/n	p	p	p/n
3	Methyl heptine carbonate	111-12-6	Liquid	Sensitiser (GHS Cat. 1A)	AOO	3.13, 1.56, 0.78, 0.39	p	p	p	p
4	Metol	55-55-0	Solid	Sensitiser (GHS Cat. 1A)	DW	3.13, 1.56, 0.78, 0.39	p	p	p	p/n
5	Abietic acid	514-10-3	Solid	Sensitiser (GHS Cat. 1B)	AOO	12.5, 6.25, 3.13, 1.56	p	p	p	p
6	Farnesol	4602-84-0	Liquid	Sensitiser (GHS Cat. 1B)	AOO	12.5, 6.25, 3.13, 1.56	p	p/n	p	p
7	Amyl cinnamic aldehyde	122-40-7	Liquid	Sensitiser (GHS Cat. 1B)	AOO	100, 50, 25, 12.5	p	n	p	p
8	Cetrimide	57-09-0	Solid	Non-sensitiser (Not classified)	50%EtOH	1.56, 0.78, 0.39, 0.20	n	n	n	n
9	Lactic acid[3]	50-21-5	Liquid	Non-sensitiser (Not classified)	DW	6.25, 3.13, 1.56, 0.78	n	n	n	n
10	Hexane	110-54-3	Liquid	Non-sensitiser (Not classified)	AOO	100, 50, 25	n	n	n	n

[1]: The *in vivo* hazard and potency prediction is based on LLNA data (TG497, SD Annex3) (Urbisch, 2015). The *in vivo* potency is derived using the criteria based on UN GHS Sub-categorisation.

[2]: "p" indicates that the fold-induction of marker gene exceeds the cut-off value with ≥ 80% viability. "n" indicates that the fold-induction of marker gene doesn't exceed the cut-off value with ≥ 80% viability. "p/n" means both "p" and "n" are acceptable.

[3]: MTT assay should be performed instead of LDH assay.