Section 4
Health effects

Test No. 442E
In Vitro Skin Sensitisation

In Vitro Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome Pathway for Skin Sensitisation

2 July 2026

OECD Guidelines for the Testing of Chemicals

OECD
BETTER POLICIES FOR BETTER LIVES

This work was approved and declassified by the Chemicals and Biotechnology Committee on 15/06/2026.

Please cite this publication as:

OECD (2026), *Test No. 442E: In Vitro Skin Sensitisation: In Vitro Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome Pathway for Skin Sensitisation,* OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264264359-en.

© OECD 2026

Attribution 4.0 International (CC BY 4.0)

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (https://creativecommons.org/licenses/by/4.0/).

Attribution – you must cite the work.

Translations – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of original work should be considered valid.*

Adaptations – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

Third-party material – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several Partner countries and from interested international organisations attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (https://www.oecd.org/en/topics/chemical-safety-and-biosafety.html).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank, Basel, Rotterdam and Stockholm Conventions and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

© OECD (2026)

Adopted:
25 June 2024
Corrected: 2 July 2026
(Figure 1 flowchart in each method annex for borderline ranges)

OECD GUIDELINES FOR THE TESTING OF CHEMICALS

Test No. 442E: *In Vitro* Skin Sensitisation: *In Vitro* Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome Pathway for Skin Sensitisation

© OECD (2026)

Table of contents

About the OECD 2

1 GENERAL INTRODUCTION 5

ANNEX 1 - IN VITRO SKIN SENSITISATION: HUMAN CELL LINE ACTIVATION TEST (H-CLAT) 9

ANNEX 2 - IN VITRO SKIN SENSITISATION: U937 CELL LINE ACTIVATION TEST (U-SENS™) 26

ANNEX 3 - IN VITRO SKIN SENSITISATION: INTERLEUKIN-8 REPORTER GENE ASSAY (IL-8 LUC ASSAY) 42

ANNEX 4 - IN VITRO SKIN SENSITISATION: GENOMIC ALLERGEN RAPID DETECTION (GARD™) FOR ASSESSMENT OF SKIN SENSITISERS (GARD™skin) 67

© OECD (2026)

1 GENERAL INTRODUCTION

Activation of dendritic cells Key Event based Test Guideline

1. A skin sensitiser refers to a substance that will lead to an allergic response following 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. 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 (DC), typically assessed by expression of specific cell surface markers, genomic transcripts, chemokines and cytokines. The fourth key event is T-cell activation and proliferation, which is indirectly assessed in the murine Local Lymph Node Assay (LLNA) (3).
2. This Test Guideline (TG) describes *in vitro* assays that address mechanisms described under the Key Event on activation of dendritic cells of the AOP for skin sensitisation (2). The TG 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 described in this TG are:

- Human Cell Line Activation test (h-CLAT)
- U937 cell line activation Test (U-SENS™)
- Interleukin-8 Reporter Gene Assay (IL-8 Luc assay)
- Genomic Allergen Rapid Detection (GARD™) for assessment of skin sensitisers (GARD™skin)

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

3. The assessment of skin sensitisation has typically 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 (TG 406) (4), assess both the induction and elicitation phases of skin sensitisation. The murine tests, the LLNA (TG 429) (3) and its two non-radioactive modifications,

© OECD (2026)

LLNA: DA (TG 442 A) (5) and LLNA: BrdU-ELISA (TG 442 B) (6), all assess the induction response exclusively, and have also 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.

4. Mechanistically-based *in chemico* and *in vitro* test methods addressing the first key event (OECD TG 442C (7)), and second key event (OECD TG 442D (8)) of the skin sensitisation AOP have been adopted for contributing to the evaluation of the skin sensitisation hazard potential of chemicals.

5. Skin sensitisers have been reported to induce the expression of cell membrane markers such as CD40, CD54, CD80, CD83, and CD86 in addition to induction of proinflammatory cytokines, such as IL-1β and TNF-α, and several chemokines including IL-8 (CXCL8) and CCL3 (9) (10) (11) (12), associated with DC activation (2). Test methods described in this TG either quantify the change in the expression of cell the surface marker(s) CD54 and CD86, the cytokine IL-8, or a series of genes (genomic biomarker signature) that are associated with the process of activation of monocytes and DC following exposure to sensitisers.

6. However, as DC activation represents only one key event of the skin sensitisation AOP (2) (13), information generated with test methods measuring markers of DC activation alone may not be sufficient as stand-alone methods 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 *in silico* modelling and read-across from chemical analogues (13). Examples of 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 the data to derive predictions, have been published (13) and are implemented in an OECD TG on defined approaches for skin sensitisation (14). If the intention is to use the data from these methods in a defined approach for skin sensitisation then refer to Test Guideline 497 for application of borderline criteria or other data interpretation procedures (14). For details on the application of the borderline criteria, refer to the Annex 1, Figure 1.

7. The test methods included in this Test Guideline might differ with regard to the procedures, the applicability domain and the limitations, but can each be used to address countries' requirements for test results on the Key Event on activation of dendritic cells of the AOP for skin sensitisation while benefiting from the OECD Mutual Acceptance of Data. In the context of Defined Approaches (DAs), methods are not automatically interchangeable; it is specified in TG 497 which methods/combination of methods should be used.

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

© OECD (2026)

substances and/or mixtures. Limited information is currently available on the applicability of the test methods to multi-constituent substances/mixtures (15) (16). The test methods are nevertheless 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.

LITERATURE

1. United Nations UN (2015), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Sixth revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [https://www.unece.org/trans/danger/publi/ghs/ghs_rev06/06files_e.html].
2. OECD (2012), *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, Part 1*, OECD Series on Testing and Assessment, No. 168, OECD Publishing, https://doi.org/10.1787/9789264221444-en.
3. OECD (2010), *Test No. 429: Skin Sensitisation: Local Lymph Node Assay*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264071100-en.
4. OECD (2022), *Test No. 406: Skin Sensitisation*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264070660-en.
5. OECD (2010), *Test No. 442A: Skin Sensitization: Local Lymph Node Assay: DA*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264090972-en.
6. OECD (2025), Test No. 442B: Skin Sensitization: Local Lymph Node Assay: BrdU-ELISA or –FCM, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264090996-en.
7. OECD (2026), *Test No. 442C: In Chemico Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264229709-en.
8. OECD (2026), *Test No. 442D: In Vitro Skin Sensitisation: Assays addressing the Adverse Outcome Pathway Key Event on Keratinocyte activation*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264229822-en .
9. Steinman RM. 1991. The dendritic cell system and its role in immunogenicity. Annu Rev Immunol 9:271-96.
10. Caux C, Vanbervliet B, Massacrier C, Azuma M, Okumura K, Lanier LL, and Banchereau J. 1994. B70/B7-2 is identical to CD86 and is the major functional ligand for CD28 expressed on human dendritic cells. J Exp Med 180:1841-7.
11. Aiba S, Terunuma A, Manome H, and Tagami H. 1997. Dendritic cells differently respond to haptens and irritants by their production of cytokines and expression of co-stimulatory molecules. Eur J Immunol 27:3031-8.

© OECD (2026)

12. Aiba S, Manome H, Nakagawa S, Mollah ZU, Mizuashi M, Ohtani T, Yoshino Y, and Tagami. H. 2003. p38 mitogen-activated protein kinase and extracellular signal-regulated kinases play distinct roles in the activation of dendritic cells by two representative haptens, NiCl2 and DNCB. J Invest Dermatol 120:390-8.
13. OECD (2016), *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,* OECD Series on Testing and Assessment, No. 256, OECD Publishing, Paris, https://doi.org/10.1787/9789264279285-en.
14. OECD. (2026), *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.
15. Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, NishiyamaN, Itagaki H. (2010), A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). Altern. Lab. Anim. 38, 275-284.
16. Piroird, C., Ovigne, J.M., Rousset, F., Martinozzi-Teissier, S., Gomes, C., Cotovio, J., Alépée, N. (2015). The Myeloid U937 Skin Sensitization Test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. Toxicol. In Vitro 29, 901-916.

© OECD (2026)

ANNEX 1 - IN VITRO SKIN SENSITISATION: HUMAN CELL LINE ACTIVATION TEST (H-CLAT)

INITIAL CONSIDERATIONS AND LIMITATIONS

1. The h-CLAT method quantifies changes in the expression of cell surface markers associated with the process of activation of monocytes and dendritic cells (DC) (i.e. CD86 and CD54), in the human monocytic leukaemia cell line THP-1, following exposure to sensitisers (1) (2). The measured expression levels of CD86 and CD54 cell surface markers are then used for supporting the discrimination between skin sensitisers and non-sensitisers.

2. The h-CLAT method has been evaluated in a European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)-coordinated validation study and subsequent independent peer review by the EURL ECVAM Scientific Advisory Committee (ESAC). Considering all available evidence and input from regulators and stakeholders, the h-CLAT was recommended by EURL ECVAM (3) to be used as part of an IATA to support the discrimination between sensitisers and non-sensitisers for the purpose of hazard classification and labelling. Examples of the use of h-CLAT data in combination with other information are reported in the literature (4) (5) (6) (7) (8) (9) (10) (11).

3. The h-CLAT method proved to be transferable to laboratories experienced in cell culture techniques and flow cytometry analysis. The level of reproducibility in predictions that can be expected from the test method is in the order of 80% within and between laboratories (3) (12). Results generated in the validation study (13) and other published studies (14) overall indicate that, compared with LLNA results, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers is 85% (N=142) with a sensitivity of 93% (94/101) and a specificity of 66% (27/41) (based on a re-analysis by EURL ECVAM (12) considering all existing data and not considering negative results for chemicals with a Log Kow greater than 3.5 as described in paragraph 4). False negative predictions with the h-CLAT are more likely to concern chemicals showing a low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals showing a high skin sensitisation potency (i.e. UN GHS subcategory 1A) (4) (13) (15). Taken together, this information indicates the usefulness of the h-CLAT method to contribute to the identification of skin sensitisation hazards. However, the accuracy values given here for the h-CLAT as a stand-alone test method are only indicative, since the test method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 6 and 8 in the General introduction. Furthermore, when evaluating non-animal methods for 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.

4. On the basis of the data currently available, the h-CLAT method was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, skin sensitisation potency (as determined in *in vivo* studies) and physicochemical properties (3) (14) (15). The h-CLAT

© OECD (2026)

method is applicable to test chemicals soluble or that form a stable dispersion (i.e. a colloid or suspension in which the test chemical does not settle or separate from the solvent/vehicle into different phases) in an appropriate solvent/vehicle (see paragraph 14). Test chemicals with a Log Kow greater than 3.5 tend to produce false negative results (14). Therefore negative results with test chemicals with a Log Kow greater than 3.5 should not be considered. However, positive results obtained with test chemicals with a Log Kow greater than 3.5 could still be used to support the identification of the test chemical as a skin sensitiser. Furthermore, because of the limited metabolic capability of the cell line used (16) and because of the experimental conditions, pro-haptens (i.e. substances requiring enzymatic activation for example via P450 enzymes) and pre-haptens (i.e. substances activated by oxidation) in particular with a slow oxidation rate may also provide negative results in the h-CLAT (15). Fluorescent test chemicals can be assessed with the h-CLAT (17), nevertheless, strong fluorescent test chemicals emitting at the same wavelength as fluorescein isothiocyanate (FITC) or as propidium iodide (PI), will interfere with the flow cytometric detection and thus cannot be correctly evaluated using FITC-conjugated antibodies or PI. In such a case, other fluorochrome-tagged antibodies or other cytotoxicity markers, respectively, can be used as long as it can be shown they provide similar results as the FITC-tagged antibodies (see paragraph 24) or PI (see paragraph 18) e.g. by testing the proficiency substances in Appendix II. In the light of the above, negative results should be interpreted in the context of the stated limitations and together with other information sources within the framework of IATA. In cases where there is evidence demonstrating the non-applicability of the h-CLAT method to other specific categories of test chemicals, it should not be used for those specific categories.

5. As described above, the h-CLAT method supports the discrimination between skin sensitisers from non-sensitisers. However, it may also potentially contribute to the assessment of sensitising potency (4) (5) (9) when used in defined approaches (14) or integrated approaches such as IATA.

6. Definitions are provided in Appendix I.

PRINCIPLE OF THE TEST

7. The h-CLAT method is an *in vitro* assay that quantifies changes of cell surface marker expression (i.e. CD86 and CD54) on a human monocytic leukemia cell line, THP-1 cells, following 24 hours exposure to the test chemical. These surface molecules are typical markers of monocytic THP-1 activation and may mimic DC activation, which plays a critical role in T-cell priming. The changes of surface marker expression are measured by flow cytometry following cell staining with fluorochrome-tagged antibodies. Cytotoxicity measurement is also conducted concurrently to assess whether upregulation of surface marker expression occurs at sub-cytotoxic concentrations. The relative fluorescence intensity of surface markers compared to solvent/vehicle control are calculated and used in the prediction model (see paragraphs 27-30), to support the discrimination between sensitisers and non-sensitisers

DEMONSTRATION OF PROFICIENCY

8. Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency, using the 10 Proficiency Substances listed in Appendix II. Moreover, test method users should maintain an historical database of data generated with the reactivity checks (see paragraph 11) and with the positive and solvent/vehicle controls (see paragraphs 20-22), and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

© OECD (2026)

PROCEDURE

9. This test method is based on the h-CLAT DataBase service on ALternative Methods to animal experimentation (DB-ALM) protocol no. 158 (18) which represents the protocol used for the EURL ECVAM-coordinated validation study. It is recommended that this protocol is used when implementing and using the h-CLAT method in the laboratory. The following is a description of the main components and procedures for the h-CLAT method, which comprises two steps: *dose finding assay* and *CD86/CD54 expression measurement*.

Preparation of cells

10. The human monocytic leukaemia cell line, THP-1, should be used for performing the h-CLAT method. It is recommended that cells (TIB-202™) are obtained from a well-qualified cell bank, such as the American Type Culture Collection.

11. THP-1 cells are cultured, at 37°C under 5% CO_2 and humidified atmosphere, in RPMI-1640 medium supplemented with 10% foetal bovine serum (FBS), 0.05 mM 2-mercaptoethanol, 100 units/mL penicillin and 100 µg/mL streptomycin. The use of penicillin and streptomycin in the culture medium can be avoided. However, in such a case users should verify that the absence of antibiotics in the culture medium has no impact on the results, for example by testing the proficiency substances listed in Appendix II. In any case, in order to minimise the risk of contamination, good cell culture practices should be followed independently of the presence or not of antibiotics in the cell culture medium. THP-1 cells are routinely seeded every 2-3 days at the density of 0.1 to 0.2 × 10^6 cells/mL. They should be maintained at densities from 0.1 to 1.0 × 10^6 cells/mL. Prior to using them for testing, the cells should be qualified by conducting a reactivity check. The reactivity check of the cells should be performed using the positive controls, 2,4-dinitrochlorobenzene (DNCB) (CAS n. 97-00-7, ≥ 99% purity) and nickel sulfate ($NiSO_4$) (CAS n. 10101-97-0, ≥ 99% purity) and the negative control, lactic acid (LA) (CAS n. 50-21-5, ≥ 85% purity), two weeks after thawing. Both DNCB and $NiSO_4$ should produce a positive response of both CD86 and CD54 cell surface markers, and LA should produce a negative response of both CD86 and CD54 cell surface markers. Only the cells which passed the reactivity check are to be used for the assay. Cells can be propagated up to two months after thawing. Passage number should not exceed 30. The reactivity check should be performed according to the procedures described in paragraphs 20-24.

12. For testing, THP-1 cells are seeded at a density of either 0.1 × 10^6 cells/mL or 0.2 × 10^6 cells/mL, and pre-cultured in culture flasks for 72 hours or for 48 hours, respectively. It is important that the cell density in the culture flask just after the pre-culture period be as consistent as possible in each experiment (by using one of the two pre-culture conditions described above), because the cell density in the culture flask just after pre-culture could affect the CD86/CD54 expression induced by allergens (19). On the day of testing, cells harvested from culture flask are resuspended with fresh culture medium at 2 × 10^6 cells/mL. Then, cells are distributed into a 24 well flat-bottom plate with 500 µL (1 × 10^6 cells/well) or a 96-well flat-bottom plate with 80 µL (1.6 × 10^5 cells/well).

Dose finding assay

13. A *dose finding assay* is performed to determine the CV75, being the test chemical concentration that results in 75% cell viability (CV) compared to the solvent/vehicle control. The CV75 value is used to determine the concentration of test chemicals for the *CD86/CD54 expression measurement* (see paragraphs 20-24).

© OECD (2026)

Preparation of test chemicals and control substances

14. The test chemicals and control substances are prepared on the day of testing. For the h-CLAT method, test chemicals are dissolved or stably dispersed (see also paragraph 4) in saline or medium as first solvent/vehicle options or dimethyl sulfoxide (DMSO, ≥ 99% purity) as a second solvent/vehicle option if the test chemical is not soluble or does not form a stable dispersion in the previous two solvents/vehicles, to final concentrations of 100 mg/mL (in saline or medium) or 500 mg/mL (in DMSO). Other solvents/vehicles than those described above may be used if sufficient scientific rationale is provided. Stability of the test chemical in the final solvent/vehicle should be taken into account.

15. Starting from the 100 mg/mL (in saline or medium) or 500 mg/mL (in DMSO) stock solutions of the test chemicals, the following dilution steps should be taken:

- For saline or medium as solvent/vehicle: Eight stock solutions (eight concentrations) are prepared, by two-fold serial dilutions using the corresponding solvent/vehicle. These stock solutions are then further diluted 50-fold into culture medium (working solutions). If the top final concentration in the plate of 1000 µg/mL is non-toxic, the maximum concentration should be re-determined by performing a new cytotoxicity test. The final concentration in the plate should not exceed 5000 µg/mL for test chemicals dissolved or stably dispersed in saline or medium.
- For DMSO as solvent/vehicle: Eight stock solutions (eight concentrations) are prepared, by two-fold serial dilutions using the corresponding solvent/vehicle. These stock solutions are then further diluted 250-fold into culture medium (working solutions).The final concentration in plate should not exceed 1000 µg/mL even if this concentration is non-toxic.

The working solutions are finally used for exposure by adding an equal volume of working solution to the volume of THP-1 cell suspension in the plate (see also paragraph 17) to achieve a further two-fold dilution (usually, the final range of concentrations in the plate is 7.81–1000 µg/mL).

16. The solvent/vehicle control used in the h-CLAT method is culture medium (for test chemicals solubilised or stably dispersed (see paragraph 4) either with medium or saline) or DMSO (for test chemicals solubilised or stably dispersed in DMSO) tested at a single final concentration in the plate of 0.2%. It undergoes the same dilution as described for the working solutions in paragraph 15.

Application of test chemicals and control substances

17. The culture medium or working solutions described in paragraphs 15 and 16 are mixed 1:1 (v/v) with the cell suspensions prepared in the 24-well or 96-well flat-bottom plate (see paragraph 12). The treated plates are then incubated for 24±0.5 hours at 37°C under 5% CO_2. Care should be taken to avoid evaporation of volatile test chemicals and cross-contamination between wells by test chemicals, e.g. by sealing the plate prior to the incubation with the test chemicals (20).

Propidium iodide (PI) staining

18. After 24±0.5 hours of exposure, cells are transferred into sample tubes and collected by centrifugation. The supernatants are discarded and the remaining cells are resuspended with 200 µL (in case of 96-well) or 600 µL (in case of 24-well) of a phosphate buffered saline containing 0.1% bovine serum albumin (staining buffer). 200 µL of cell suspension is transferred into 96-well round-bottom plate (in case of 96-well) or micro tube (in case of 24-well) and washed twice with 200 µL (in case of 96-well) or 600 µL (in case of 24-well) of staining buffer. Finally, cells are resuspended in staining buffer (e.g. 400 µL) and PI solution (e.g. 20 µL) is added (for example, final concentration of PI is 0.625 µg/mL). Other cytotoxicity markers, such as 7-Aminoactinomycin D (7-AAD), Trypan blue or others may be used if the

© OECD (2026)

alternative stains can be shown to provide similar results as PI, for example by testing the proficiency substances in Appendix II.

Cytotoxicity measurement by flow cytometry and estimation of CV75 value

19. The PI uptake is analysed using flow cytometry with the acquisition channel FL-3. A total of 10,000 living cells (PI negative) are acquired. The cell viability can be calculated using the following equation by the cytometer analysis program. When the cell viability is low, up to 30,000 cells including dead cells should be acquired. Alternatively, data can be acquired for one minute after the initiation of the analysis.

$$\textbf{Cell Viability} = \frac{\textbf{Number of living cells}}{\textbf{Total Number of acquired cells}} \times 100$$

The CV75 value (see paragraph 13), i.e. a concentration showing 75% of THP-1 cell survival (25% cytotoxicity), is calculated by log-linear interpolation using the following equation:

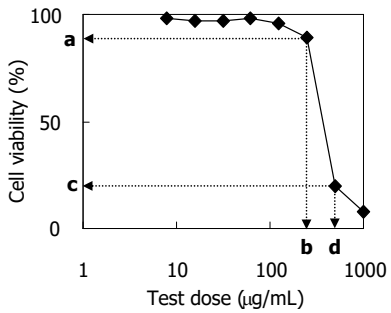
$$\textbf{Log CV75} = \frac{(75 - c) \times \text{Log}(b) - (75 - a) \times \text{Log}(d)}{a - c}$$

Where:

a is the minimum value of cell viability over 75%

c is the maximum value of cell viability below 75%

b and d are the concentrations showing the value of cell viability a and c respectively



Other approaches to derive the CV75 can be used as long as it is demonstrated that this has no impact on the results (e.g. by testing the proficiency substances).

© OECD (2026)

CD86/CD54 expression measurement

Preparation of the test chemicals and control substances

20. The appropriate solvent/vehicle (saline, medium or DMSO; see paragraph 14) is used to dissolve or stably disperse the test chemicals. The test chemicals are first diluted to the concentration corresponding to 100-fold (for saline or medium) or 500-fold (for DMSO) of the 1.2 × CV75 determined in the *dose finding assay* (see paragraph 19). If the CV75 cannot be determined (i.e. if sufficient cytotoxicity is not observed in the *dose finding assay*), the highest soluble or stably dispersed concentration of test chemical prepared with each solvent/vehicle should be used as starting concentration. Please note that the final concentration in the plate should not exceed 5000 µg/mL (in case of saline or medium) or 1000 µg/mL (in case of DMSO). Then, 1.2-fold serial dilutions are made using the corresponding solvent/vehicle to obtain the stock solutions (eight concentrations ranging from 100×1.2 × CV75 to 100×0.335 × CV75 (for saline or medium) or from 500×1.2 × CV75 to 500×0.335 × CV75 (for DMSO)) to be tested in the h-CLAT method (see DB-ALM protocol No. 158 for an example of dosing scheme). The stock solutions are then further diluted 50-fold (for saline or medium) or 250-fold (for DMSO) into the culture medium (working solutions). These working solutions are finally used for exposure with a further final two-fold dilution factor in the plate. If the results do not meet the acceptance criteria described in paragraphs 26 and 28 regarding cell viability, the *dose finding assay* may be repeated to determine a more precise CV75. Please note that only 24-well plates can be used for CD86/CD54 expression measurement.

21. The solvent/vehicle control is prepared as described in paragraph 16. The positive control used in the h-CLAT method is DNCB (see paragraph 11), for which stock solutions are prepared in DMSO and diluted as described for the stock solutions in paragraph 20. DNCB should be used as the positive control for *CD86/CD54 expression measurement* at a final single concentration in the plate (typically 4.0 µg/mL). To obtain a 4.0 µg/mL concentration of DNCB in the plate, a 2 mg/mL stock solution of DNCB in DMSO is prepared and further diluted 250-fold with culture medium to a 8 µg/mL working solution. Alternatively, the CV75 of DNCB, which is determined in each test facility, could be also used as the positive control concentration. Other suitable positive controls may be used if historical data are available to derive comparable run acceptance criteria. For positive controls, the final single concentration in the plate should not exceed 5000 µg/mL (in case of saline or medium) or 1000 µg/mL (in case of DMSO). The run acceptance criteria are the same as those described for the test chemical (see paragraph 26), except for the last acceptance criterion since the positive control is tested at a single concentration.

Application of test chemicals and control substances

22. For each test chemical and control substance, one experiment is needed to obtain a prediction. Each experiment consists of at least two independent runs for *CD86/CD54 expression measurement* (see paragraphs 26-28). Each independent run is performed on a different day or on the same day provided that for each run: a) independent fresh stock solutions and working solutions of the test chemical and antibody solutions are prepared and b) independently harvested cells are used (i.e. cells are collected from different culture flasks); however, cells may come from the same passage. Test chemicals and control substances prepared as working solutions (500 µL) are mixed with 500 µL of suspended cells ($1x10^6$ cells) at 1:1 ratio, and cells are incubated for 24±0.5 hours as described in paragraphs 20 and 21. In each run, a single replicate for each concentration of the test chemical and control substance is sufficient because a prediction is obtained from at least two independent runs.

© OECD (2026)

Cell staining and analysis

23. After 24±0.5 hours of exposure, cells are transferred from 24 well plate into sample tubes, collected by centrifugation and then washed twice with 1mL of staining buffer (if necessary, additional washing steps may be done). After washing, cells are blocked with 600 µL of blocking solution (staining buffer containing 0.01% (w/v) globulin (Cohn fraction II, III, Human: SIGMA, #G2388-10G)) and incubated at 4°C for 15 min. After blocking, cells are split in three aliquots of 180 µL into a 96-well round-bottom plate or micro tube.

24. After centrifugation, cells are stained with 50 µL of FITC-labelled anti-CD86, anti-CD54 or mouse IgG1 (isotype) antibodies at 4°C for 30 min. The antibodies described in the h-CLAT DB-ALM protocol no. 158 (18) should be used by diluting 3:25 (v/v, for CD86 (BD-PharMingen, #555657; Clone: Fun-1)) or 3:50 (v/v, for CD54 (DAKO, #F7143; Clone: 6.5B5) and IgG1 (DAKO, #X0927)) with staining buffer. These antibody dilution factors were defined by the test method developers as those providing the best signal-to-noise ratio. Based on the experience of the test method developers, the fluorescence intensity of the antibodies is usually consistent between different lots. However, users may consider titrating the antibodies in their own laboratory's conditions to define the best concentrations for use. Other fluorochrome-tagged anti-CD86 and/or anti-CD54 antibodies may be used if they can be shown to provide similar results as FITC-conjugated antibodies, for example by testing the proficiency substances in Appendix II. It should be noted that changing the clone or supplier of the antibodies as described in the h-CLAT DB-ALM protocol no. 158 (18) may affect the results. After washing twice or more with 150 µL of staining buffer, cells are resuspended in staining buffer (e.g. 400 µL), and the PI solution (e.g. 20 µL to obtain a final concentration of 0.625 µg/mL) or another cytotoxicity marker's solution (see paragraph 18) is added. The expression levels of CD86 and CD54, and cell viability are analysed using flow cytometry.

DATA AND REPORTING

Data analyses

25. The expression of CD86 and CD54 is determined by flow cytometry with the acquisition channel FL-1. Based on the geometric mean fluorescence intensity (MFI), the relative fluorescence intensity (RFI) of CD86 and CD54 for positive control (ctrl) cells and chemical-treated cells are calculated according to the following equation:

$$\mathbf{RFI} = \frac{\textbf{MFI of chemical-treated cells} - \textbf{MFI of chemical-treated isotype control cells}}{\textbf{MFI of solvent/vehicle-treated ctrl cells} - \textbf{MFI of solvent/vehicle-treated isotype ctrl cells}} \times 100$$

The cell viability from the isotype control (ctrl) cells (which are stained with mouse IgG1 (isotype) antibodies) is also calculated according to the equation described in paragraph 19.

Acceptance criteria

26. The following acceptance criteria should be met to consider a run valid when using the h-CLAT method (22) (27).

© OECD (2026)

- The cell viabilities of medium and solvent/vehicle controls should be higher than 90%.
- In the solvent/vehicle control, RFI values of both CD86 and CD54 should not exceed the threshold for a positive criteria (CD86 RFI ≥ 150% and CD54 RFI ≥ 200%). RFI values of the solvent/vehicle control are calculated by using the formula described in paragraph 25 ("MFI of chemical" should be replaced with "MFI of solvent/vehicle", and "MFI of solvent/vehicle" should be replaced with "MFI of (medium) control").
- For both medium and solvent/vehicle controls, the MFI ratio of both CD86 and CD54 to isotype control should be > 105%.
- In the positive control (DNCB), RFI values of both CD86 and CD54 should cross the threshold for positivity (CD86 RFI ≥ 150 and CD54 RFI ≥ 200) and cell viability should be more than 50%.
- For the test chemical, the cell viability should be more than 50% in at least four tested concentrations in each run.

27. Negative results are acceptable only for test chemicals exhibiting a cell viability of less than 90% at the highest concentration tested (i.e. 1.2 × CV75 according to the serial dilution scheme described in paragraph 20). If the cell viability at 1.2 × CV75 is equal or above 90% the negative result should be discarded. In such a case it is recommended to try to refine the dose selection. Repeating the CV75 determination may also be considered. It should be noted that when 5000 µg/mL in saline (or medium or other solvents/vehicles), 1000 µg/mL in DMSO or the highest soluble concentration is used as the maximal test concentration of a test chemical, a negative result is acceptable even if the cell viability is above 90%.

Prediction model

28. Based on *CD86/CD54 expression measurement*, each test chemical is tested in at least two independent runs to derive a single prediction (POSITIVE, NEGATIVE or BORDERLINE).

An h-CLAT independent run prediction is considered POSITIVE if:

- The RFI of CD86 is > 184% in at least one tested concentration (with cell viability ≥ 50%); **and/or**
- The RFI of CD54 is > 255% in at least one tested concentration (with cell viability ≥ 50%).

An h-CLAT independent run prediction is considered NEGATIVE if the criteria outlined in paragraph 28 are met and:

- The maximum RFI of CD86 is ≤ 122%, **and**
- The maximum RFI of CD54 is ≤ 157%.

Furthermore, an h-CLAT prediction may also be considered NEGATIVE if any of the following conditions are met:

- The RFI of CD86 is > 122% and ≤ 184%, with cell viability < 50% at CD86 corresponding doses within this RFI range, **or**
- The RFI of CD86 is > 184%, with cell viability < 50% at CD86 corresponding doses, **and**
- The RFI of CD54 is > 157% and ≤ 255%, with cell viability < 50% at CD54 corresponding doses within this RFI range, **or**
- The RFI of CD54 is > 255%, with cell viability < 50% at CD54 corresponding doses.

Since confidence in predictions decreases as test results fall near a prediction threshold, statistically-derived borderline ranges have been established for evaluating the RFI for both the CD86 and CD54 markers. An h-CLAT independent trial prediction is considered BORDERLINE (BL) if neither of the criteria above for POSITIVE or NEGATIVE predictions are met, and both of the following conditions are met:

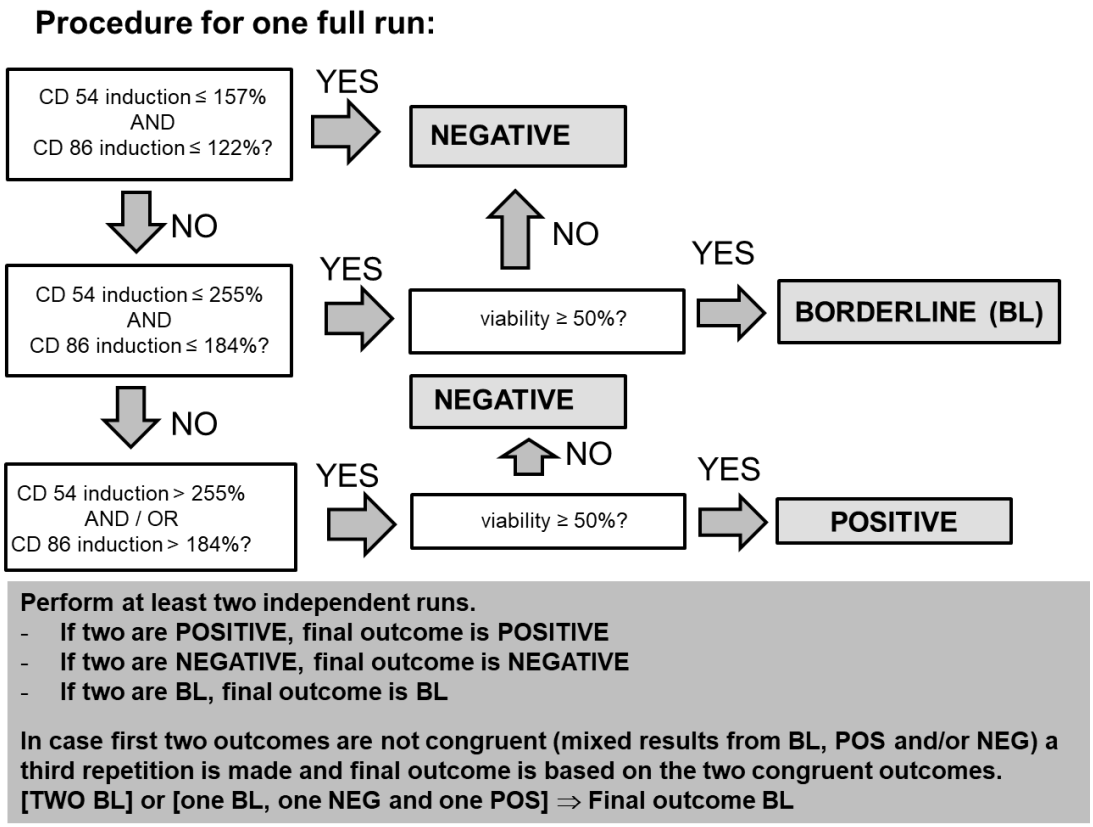
© OECD (2026)

- The maximum RFI of CD86 is > 122% and ≤ 184% (with cell viability ≥ 50% at the corresponding CD86 doses), **and**

The maximum RFI of CD54 is > 157% and ≤ 255% (with cell viability ≥ 50% at the corresponding CD54 doses).

29. Based on the above, if the first two runs are both positive for CD86 and/or are both positive for CD54, the h-CLAT prediction is considered POSITIVE and a third run does not need to be conducted. Similarly, if the first two runs are negative for both markers, the h-CLAT prediction is considered NEGATIVE (with due consideration of the provisions of paragraph 28) without the need for a third run. If neither of the two conditions above are met, and the results of two runs fall in the borderline range, the h-CLAT prediction is considered BORDERLINE. If however, the first two runs are not concordant for at least one of the markers (CD54 or CD86), a third run is needed and the final prediction will be based on the majority result of the three individual runs (i.e. 2 out of 3). In this respect, it should be noted that if two independent runs are conducted and one is only positive for CD86 and the other is only positive for CD54, a third run is required. If this third run is negative for both markers, the h-CLAT prediction is considered NEGATIVE. On the other hand, if the third run is positive for either marker or for both markers, the h-CLAT prediction is considered POSITIVE. Finally, if the outcome of three trials is one POSITIVE, one NEGATIVE and one BORDERLINE, the final outcome is BORDERLINE.

Figure 1. Flow-chart of the h-CLAT prediction model taking into account borderline ranges and multiple runs to conclude on borderline results.



© OECD (2026)

The original threshold for a positive classification is 150% induction of CD86 with a statistically derived borderline range around this threshold of 122 – 184% and 200% induction of CD54 with a statistically derived borderline range around this threshold of 157 – 255%.

30. For the test chemicals predicted as POSITIVE with the h-CLAT, optionally, two Effective Concentrations (EC) values, the EC150 for CD86 and EC200 for CD54, i.e. the concentration at which the test chemicals induced a RFI of 150% or 200%, may be determined. These EC values potentially could contribute to the assessment of sensitising potency (9) when used in defined approaches (14) or integrated approaches such as IATA (4) (5) (6) (7) (8). They can be calculated by the following equations:

$$\mathbf{EC150\ (for\ CD86) = B_{concentration} + [(150 - B_{RFI}) / (A_{RFI} - B_{RFI}) \times (A_{concentration} - B_{concentration})]}$$

where

$A_{concentration}$ is the lowest concentration in µg/mL with RFI > 150 (CD86) or 200 (CD54)

$B_{concentration}$ is the highest concentration in µg/mL with RFI < 150 (CD86) or 200 (CD54)

A_{RFI} is the RFI at the lowest concentration with RFI > 150 (CD86) or 200 (CD54)

B_{RFI} is the RFI at the highest concentration with RFI < 150 (CD86) or 200 (CD54)

For the purpose of more precisely deriving the EC150 and EC200 values, three independent runs for *CD86/CD54 expression measurement* may be required. The final EC150 and EC200 values are then determined as the median value of the ECs calculated from the three independent runs. When only two of three independent runs meet the criteria for positivity (see paragraph 27), the higher EC150 or EC200 of the two calculated values is adopted.

Test report

31. 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;
  - Physical appearance, Log Kow, water solubility, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - 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 solvent/vehicle for each test chemical.
- Multi-constituent substance, UVCB and mixture

© OECD (2026)

- Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
- Physical appearance, water solubility, DMSO solubility and additional relevant physicochemical properties, to the extent available;
- Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions 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 solvent/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, Log Kow, water solubility, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available and where applicable;
- Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
- 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.

  - Negative and solvent/vehicle control

- Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
- Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
- Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other control solvent/vehicle than those mentioned in the Test Guideline are used and to the extent available;
- Storage conditions and stability to the extent available;
- Justification for choice of solvent/vehicle for each test chemical.

Test method conditions

- Name and address of the sponsor, test facility and study director;
- Description of test method used;
- Cell line used, its storage conditions and source (e.g. the facility from which they were obtained);

© OECD (2026)

- Flow cytometry used (e.g. model), including instrument settings, globulin, antibodies and cytotoxicity marker used;
- A statement attesting to the proficiency of the laboratory in performing the test method (e.g. by testing of proficiency substances) or to demonstrate reproducible performance of the test method over time. The Testing Facility shall maintain documentation of the procedures and proficiency chemicals used to demonstration proficiency.

Test acceptance criteria

- Cell viability, MFI and RFI values obtained with the solvent/vehicle control in comparison to the acceptance ranges;
- Cell viability and RFI values obtained with the positive control in comparison to the acceptance ranges;
- Cell viability of all tested concentrations of the tested chemical.

Test procedure

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

Results

- Tabulation of the data, including CV75 (if applicable), individual geometric MFI, RFI, cell viability values, EC150/EC200 values (if applicable) obtained for the test chemical and for the positive control in each run, and an indication of the rating of the test chemical according to the prediction model;
- Description of any other relevant observations, if applicable.

Discussion of the results

- Discussion of the results obtained with the h-CLAT method, including any relevant factors (e.g., solubility limitations, or physico-chemical; characteristics) that may impact the prediction;
- Consideration of the test method results within the context of an IATA, if other relevant information is available.

Conclusions

© OECD (2026)

LITERATURE

1. Ashikaga T, Yoshida Y, Hirota M, Yoneyama K, Itagaki H, Sakaguchi H, Miyazawa M, Ito Y, Suzuki H, Toyoda H. (2006), Development of an in vitro skin sensitization test using human cell lines: The human Cell Line Activation Test (h-CLAT) I. Optimization of the h-CLAT protocol. Toxicol. In Vitro 20, 767–773.
2. Miyazawa M, Ito Y, Yoshida Y, Sakaguchi H, Suzuki H. (2007), Phenotypic alterations and cytokine production in THP-1 cells in response to allergens. Toxicol. In Vitro 21, 428-437.
3. EC EURL-ECVAM (2013), Recommendation on the human Cell Line Activation Test (h-CLAT) for skin sensitisation testing. Accessible at: https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations
4. Takenouchi O, Fukui S, Okamoto K, Kurotani S, Imai N, Fujishiro M, Kyotani D, Kato Y, Kasahara T, Fujita M, Toyoda A, Sekiya D, Watanabe S, Seto H, Hirota M, Ashikaga T, Miyazawa M. (2015), Test battery with the human cell line activation test, direct peptide reactivity assay and DEREK based on a 139 chemical data set for predicting skin sensitizing potential and potency of chemicals. J Appl Toxicol. 35, 1318-1332.
5. Hirota M, Fukui S, Okamoto K, Kurotani S, Imai N, Fujishiro M, Kyotani D, Kato Y, Kasahara T, Fujita M, Toyoda A, Sekiya D, Watanabe S, Seto H, Takenouchi O, Ashikaga T, Miyazawa M. (2015), Evaluation of combinations of in vitro sensitization test descriptors for the artificial neural network-based risk assessment model of skin sensitization. J Appl Toxicol. 35, 1333-1347.
6. Bauch C, Kolle SN, Ramirez T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B, Landsiedel R. (2012), Putting the parts together: combining in vitro methods to test for skin sensitizing potencials. Regul Toxicol Parmacol. 63, 489-504.
7. Van der Veen JW, Rorije E, Emter R, Natch A, van Loveren H, Ezendam J. (2014), Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. Regul Toxicol Pharmacol. 69, 371-379.
8. Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F, Natsch A, Emter R, Ashikaga T, Miyazawa M, Sakaguchi H. (2015), Assessing skin sensitization hazard in mice and men using non-animal test methods. Regul Toxicol Parmacol. 71, 337-351.
9. Jaworska JS, Natsch A, Ryan C, Strickland J, Ashikaga T, Miyazawa M. (2015), Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. Arch Toxicol. 89, 2355-2383.
10. Strickland J, Zang Q, Kleinstreuer N, Paris M, Lehmann DM, Choksi N, Matheson J, Jacobs A, Lowit A, Allen D, Casey W. (2016), Integrated decision strategies for skin sensitization hazard. J Appl Toxicol. DOI 10.1002/jat.3281.
11. Nukada Y, Ashikaga T, Miyazawa M, Hirota M, Sakaguchi H, Sasa H, Nishiyama N. (2012), Prediction of skin sensitization potency of chemicals by human Cell Line Activation Test (h-CLAT) and an attempt at classifying skin sensitization potency. Toxicol. In Vitro 26, 1150-60.
12. EC EURL ECVAM (2015), Re-analysis of the within and between laboratory reproducibility of the human Cell Line Activation Test (h-CLAT). Accessible at: https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-human-cell-line-activation-test-h-clat-for-skin-sensitisation-testing

© OECD (2026)

13. EC EURL ECVAM (2012), human Cell Line Activation Test (h-CLAT) Validation Study Report Accessible at: https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations

14. Takenouchi O, Miyazawa M, Saito K, Ashikaga T, Sakaguchi H. (2013), Predictive performance of the human Cell Line Activation Test (h-CLAT) for lipophilic with high octanol-water partition coefficients. J. Toxicol. Sci. 38, 599-609.

15. Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, NishiyamaN, Itagaki H. (2010), A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). Altern. Lab. Anim. 38, 275-284.

16. Fabian E., Vogel D., Blatz V., Ramirez T., Kolle S., Eltze T., van Ravenzwaay B., Oesch F., Landsiedel R. (2013), Xenobiotic metabolizin enzyme activities in cells used for testing skin sensitization in vitro. Arch Toxicol 87, 1683-1969.

17. Okamoto K, Kato Y, Kosaka N, Mizuno M, Inaba H, Sono S, Ashikaga T, Nakamura T, Okamoto Y, Sakaguchi H, Kishi M, Kuwahara H, Ohno Y. (2010), The Japanese ring study of a human Cell Line Activation Test (h-CLAT) for predicting skin sensitization potential (6th report): A study for evaluating oxidative hair dye sensitization potential using h-CLAT. AATEX 15, 81-88.

18. DB-ALM (INVITTOX) (2014), Protocol 158: human Cell Line Activation Test (h-CLAT), 23pp. Accessible at: http://ecvam-dbalm.jrc.ec.europa.eu/

19. Mizuno M, Yoshida M, Kodama T, Kosaka N, Okamato K, Sono S, Yamada T, Hasegawa S, Ashikaga T, Kuwahara H, Sakaguchi H, Sato J, Ota N, Okamoto Y, Ohno Y. (2008), Effects of pre-culture conditions on the human Cell Line Activation Test (h-CLAT) results; Results of the 4th Japanese inter-laboratory study. AATEX 13, 70-82.

20. Sono S, Mizuno M, Kosaka N, Okamoto K, Kato Y, Inaba H, , Nakamura T, Kishi M, Kuwahara H, Sakaguchi H, Okamoto Y, Ashikaga T, Ohno Y. (2010), The Japanese ring study of a human Cell Line Activation Test (h-CLAT) for predicting skin sensitization potential (7th report): Evaluation of volatile, poorly soluble fragrance materials. AATEX 15, 89-96.

21. OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, OECD Series on Testing and Assessment, No. 34, OECD Publishing, Paris, 96 pp, https://doi.org/10.1787/e1f1244b-en.

22. OECD (2012), *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins*, *Part 1,* OECD Series on Testing and Assessment, No. 168, OECD Publishing, https://doi.org/10.1787/9789264221444-en.

23. United Nations UN (2013), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Fifth revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html].

24. ECETOC (2003). Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).

25. Ashikaga T, Sakaguchi H, Okamoto K, Mizuno M, Sato J, Yamada T, Yoshida M, Ota N, Hasegawa S, Kodama T, Okamoto Y, Kuwahara H, Kosaka N, Sono S, Ohno Y. (2008), Assessment of the human Cell Line Activation Test (h-CLAT) for Skin Sensitization; Results of the First Japanese Inter-laboratory Study. AATEX 13, 27-35.

© OECD (2026)

APPENDIX I: 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 (21).

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 (22).

CV75: The estimated concentration showing 75% cell viability.

EC150: the concentrations showing the RFI values of 150 in CD86 expression

EC200: the concentrations showing the RFI values of 200 in CD54 expression

Flow cytometry: a cytometric technique in which cells suspended in a fluid flow one at a time through a focus of exciting light, which is scattered in patterns characteristic to the cells and their components; cells are frequently labeled with fluorescent markers so that light is first absorbed and then emitted at altered frequencies.

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 weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

Medium control: An untreated replicate containing all components of a test system. This sample is processed with test chemical-treated samples and other control samples to determine whether the solvent/vehicle interacts with the test system.

Mixture: 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).

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.

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.

Pre-haptens: chemicals which become sensitisers through abiotic transformation

Pro-haptens: chemicals requiring enzymatic activation to exert skin sensitisation potential

© OECD (2026)

Relative fluorescence intensity (RFI): Relative values of geometric mean fluorescence intensity (MFI) in chemical-treated cells compared to MFI in solvent/vehicle-treated cells.

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 (21).

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 (21).

Run: A run consists of one or more test chemicals tested concurrently with a solvent/vehicle control and with a positive control.

Sensitivity: The proportion of all positive/active 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 (21).

Staining buffer: A phosphate buffered saline containing 0.1% bovine serum albumin.

Solvent/vehicle control: An untreated sample containing all components of a test system except of the test chemical, but including the solvent/vehicle that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved or stably dispersed in the same solvent/vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent/vehicle interacts with the test system.

Specificity: The proportion of all negative/inactive 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 (21).

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 it composition.

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 (23).

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

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 (21).

© OECD (2026)

APPENDIX II: PROFICIENCY SUBSTANCES

Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency by correctly obtaining the expected h-CLAT prediction for the 10 substances recommended in Table 1 and by obtaining CV75, EC150 and EC200 values that fall within the respective reference range for at least 8 out of the 10 proficiency substances. Proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that the substances are commercially available, and that high-quality *in vivo* reference data as well as high quality *in vitro* data generated with the h-CLAT method are available. Also, published reference data are available for the h-CLAT method (3) (14).

Table 1. Recommended substances for demonstrating technical proficiency with the h-CLAT method

Proficiency substances	CASRN	Physical state	*In vivo* prediction[1]	CV75 Reference Range in µg/mL[2]	h-CLAT results for CD86 (EC150 Reference Range in µg/mL)[2]	h-CLAT results for CD54 (EC200 Reference Range in µg/mL)[2]
2,4-Dinitrochlorobenzene	97-00-7	Solid	Sensitiser (extreme)	2-12	Positive (0.5-10)	Positive (0.5-15)
4-Phenylenediamine	106-50-3	Solid	Sensitiser (strong)	5-95	Positive (<40)	Negative (>1.5)[3]
Nickel sulfate	10101-97-0	Solid	Sensitiser (moderate)	30-500	Positive (<100)	Positive (10-100)
2-Mercaptbenzothiazole	149-30-4	Solid	Sensitiser (moderate)	30-400	Negative (>10)[3]	Positive (10-140)
R(+)-Limonene	5989-27-5	Liquid	Sensitiser (weak)	>20	Negative (>5)[3]	Positive (<250)
Imidazolidinyl urea	39236-46-9	Solid	Sensitiser (weak)	25-100	Positive (20-90)	Positive (20-75)
Isopropanol	67-63-0	Liquid	Non-sensitiser	>5000	Negative (>5000)	Negative (>5000)
Glycerol	56-81-5	Liquid	Non-sensitiser	>5000	Negative (>5000)	Negative (>5000)
Lactic acid	50-21-5	Liquid	Non-sensitiser	1500-5000	Negative (>5000)	Negative (>5000)
4-Aminobenzoic acid	150-13-0	Solid	Non-sensitiser	>1000	Negative (>1000)	Negative (>1000)

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number

[1] The *in vivo* hazard and (potency) prediction is based on LLNA data (3) (14). The *in vivo* potency is derived using the criteria proposed by ECETOC (24).

[2] Based on historical observed values (13) (25).

[3] Historically, a majority of negative results have been obtained for this marker and therefore a negative result is mostly expected. The range provided was defined on the basis of the few historical positive results observed. In case a positive result is obtained, the EC value should be within the reported reference range.

© OECD (2026)

ANNEX 2 - IN VITRO SKIN SENSITISATION: U937 CELL LINE ACTIVATION TEST (U-SENS™)

INITIAL CONSIDERATIONS AND LIMITATIONS

1. The U-SENS™ method quantifies the change in the expression of a cell surface marker associated with the process of activation of monocytes and dendritic cells (DC) (i.e. CD86), in the human histiocytic lymphoma cell line U937, following exposure to sensitisers (1). The measured expression levels of CD86 cell surface marker in the cell line U937 is then used for supporting the discrimination between skin sensitisers and non-sensitisers.

2. The U-SENS™ method has been evaluated in a validation study (2) coordinated by L'Oreal and subsequently independent peer reviewed by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) Scientific Advisory Committee (ESAC) (3). Considering all available evidence and input from regulators and stakeholders, the U-SENS™ was recommended by EURL ECVAM (4) to be used as part of an IATA to support the discrimination between sensitisers and non-sensitisers for the purpose of hazard classification and labelling. In its guidance document on the reporting of structured approaches to data integration and individual information sources used within IATA for skin sensitisation, the OECD currently discusses a number of case studies describing different testing strategies and prediction models. One of the different defined approaches is based on the U-SENS assay (5). Examples of the use of U-SENS™ data in combination with other information, including historical data and existing valid human data (6), are also reported elsewhere in the literature (4) (5) (7).

3. The U-SENS™ method proved to be transferable to laboratories experienced in cell culture techniques and flow cytometry analysis. The level of reproducibility in predictions that can be expected from the test method is in the order of 90% and 84% within and between laboratories, respectively (8). Results generated in the validation study (8) and other published studies (1) overall indicate that, compared with LLNA results, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers is 86% (N=166) with a sensitivity of 91% (118/129) and a specificity of 65% (24/37). Compared with human results, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers is 77% (N=101) with a sensitivity of 100% (58/58) and a specificity of 47% (20/43). False negative predictions compared to LLNA with the U-SENS™ are more likely to concern chemicals showing a low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals showing a high skin sensitisation potency (i.e. UN GHS subcategory 1A) (1) (8) (9). Taken together, this information indicates the usefulness of the U-SENS™ method to contribute to the identification of skin sensitisation hazards. However, the accuracy values given here for the U-SENS™ as a stand-alone test method are only indicative, since the test method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore, when evaluating non-animal methods for 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.

© OECD (2026)

4. On the basis of the data currently available, the U-SENS™ method was shown to be applicable to test chemicals (including cosmetics ingredients e.g. preservatives, surfactants, actives, dyes) covering a variety of organic functional groups, of physicochemical properties, skin sensitisation potency (as determined in in vivo studies) and the spectrum of reaction mechanisms known to be associated with skin sensitisation (i.e. Michael acceptor, Schiff base formation, acyl transfer agent, substitution nucleophilic bi-molecular [SN2], or nucleophilic aromatic substitution [SNAr]) (1) (8) (9) (10). The U-SENS™ method is applicable to test chemicals that are soluble or that form a stable dispersion (i.e. a colloid or suspension in which the test chemical does not settle or separate from the solvent/vehicle into different phases) in an appropriate solvent/vehicle (see paragraph 13). Chemicals in the dataset reported to be pre-haptens (i.e. substances activated by oxidation) or pro-haptens (i.e. substances requiring enzymatic activation for example via P450 enzymes) were correctly predicted by the U-SENS™ (1) (10). Membrane disrupting substances can lead to false positive results due to a non-specific increase of CD86 expression, as 3 out of 7 false positives relative to the *in vivo* reference classification were surfactants (1). As such positive results with surfactants should be considered with caution whereas negative results with surfactants could still be used to support the identification of the test chemical as a non-sensitiser. Fluorescent test chemicals can be assessed with the U-SENS™ (1), nevertheless, strong fluorescent test chemicals emitting at the same wavelength as fluorescein isothiocyanate (FITC) or as propidium iodide (PI), will interfere with the flow cytometric detection and thus cannot be correctly evaluated using FITC-conjugated antibodies (potential false negative) or PI (viability not measurable). In such a case, other fluorochrome-tagged antibodies (e.g. Allophycocyanin APC) or other cytotoxicity markers, respectively, can be used as long as it can be shown they provide similar results as the FITC-tagged antibodies or PI (see paragraph 18) e.g. by testing the proficiency substances in Appendix II. In the light of the above, positive results with surfactants and negative results with strong fluorescent test chemicals should be interpreted in the context of the stated limitations and together with other information sources within the framework of IATA. In cases where there is evidence demonstrating the non-applicability of the U-SENS™ method to other specific categories of test chemicals, it should not be used for those specific categories.

5. As described above, the U-SENS™ method supports the discrimination between skin sensitisers from non-sensitisers. However, it may also potentially contribute to the assessment of sensitising potency when used in integrated approaches such as IATA. Nevertheless, further work, preferably based on human data, is required to determine how U-SENS™ results may possibly inform potency assessment.

6. Definitions are provided in Appendix I.

PRINCIPLE OF THE TEST

7. The U-SENS™ method is an *in vitro* assay that quantifies changes of CD86 cell surface marker expression on a human histiocytic lymphoma cell line, U937 cells, following 45±3 hours exposure to the test chemical. The CD86 surface marker is one typical marker of U937 activation. CD86 is known to be a co-stimulatory molecule that may mimic monocytic activation, which plays a critical role in T-cell priming. The changes of CD86 cell surface marker expression are measured by flow cytometry following cell staining typically with fluorescein isothiocyanate (FITC)-labelled antibodies. Cytotoxicity measurement is also conducted (e.g. by using PI) concurrently to assess whether upregulation of CD86 cell surface marker expression occurs at sub-cytotoxic concentrations. The stimulation index (S.I.) of CD86 cell surface marker compared to solvent/vehicle control is calculated and used in the prediction model (see paragraph 19), to support the discrimination between sensitisers and non-sensitisers.

© OECD (2026)

DEMONSTRATION OF PROFICIENCY

8. Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency, using the 10 Proficiency Substances listed in Appendix II in compliance with the Good *in vitro* Method Practices (11). Moreover, test method users should maintain a historical database of data generated with the reactivity checks (see paragraph 11) and with the positive and solvent/vehicle controls (see paragraphs 15-16), and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

PROCEDURE

9. This test method is based on the U-SENS™ DataBase service on ALternative Methods to animal experimentation (DB-ALM) protocol no. 183 (12). The Standard Operating Procedures (SOP) should be employed when implementing and using the U-SENS™ method in the laboratory. An automated system to run the U-SENS™ can be used if it can be shown to provide similar results, for example by testing the proficiency substances in Appendix II. The following is a description of the main components and procedures for the U-SENS™ method.

Preparation of cells

10. The human histiocytic lymphoma cell line, U937 (13) should be used for performing the U-SENS™ method. Cells (clone CRL1593.2) should be obtained from a well-qualified cell bank such as the American Type Culture Collection.

11. U937 cells are cultured, at 37°C under 5% CO_2 and humidified atmosphere, in RPMI-1640 medium supplemented with 10% foetal calf serum (FCS), 2 mM L-glutamine, 100 units/mL penicillin and 100 µg/mL streptomycin (complete medium). U937 cells are routinely passaged every 2-3 days at the density of 1.5 or 3×10^5 cells/mL, respectively. The cell density should not exceed 2×10^6 cells/mL and the cell viability measured by trypan blue exclusion should be ≥ 90% (not to be applied at the first passage after thawing). Prior to using them for testing, every batch of cells, FCS or antibodies should be qualified by conducting a reactivity check. The reactivity check of the cells should be performed using the positive control, picrylsulfonic acid (2,4,6-Trinitro-benzene-sulfonic acid: TNBS) (CASRN 2508-19-2, ≥ 95% purity) and the negative control lactic acid (LA) (CASRN 50-21-5, ≥ 85% purity), at least one week after thawing. For the reactivity check, six final concentrations should be tested for each of the 2 controls (TNBS: 1, 12.5, 25, 50, 75, 100µg/mL and LA: 1, 10, 20, 50, 100, 200µg/mL). TNBS solubilised in complete medium should produce a positive and concentration-related response of CD86 (e.g. when a positive concentration, CD86 S.I. ≥ 150, is followed by a concentration with an increasing CD86 S.I), and LA solubilised in complete medium should produce negative response of CD86 (see paragraph 21). Only the batch of cells which passed the reactivity check 2 times should be used for the assay. Cells can be propagated up to seven weeks after thawing. Passage number should not exceed 21. The reactivity check should be performed according to the procedures described in paragraphs 18-22.

12. For testing, U937 cells are seeded at a density of either 3×10^5 cells/mL or 6×10^5 cells/mL, and pre-cultured in culture flasks for 2 days or 1 day, respectively. Other pre-cultured conditions than those described above may be used if sufficient scientific rationale is provided and if it can be shown to provide similar results, for example by testing the proficiency substances in Appendix II. In the day of testing, cells harvested from culture flask are resuspended with fresh culture medium at 5×10^5 cells/mL. Then, cells are distributed into a 96-well flat-bottom plate with 100 µL (final cell density of 0.5×10^5 cells/well).

© OECD (2026)

Preparation of test chemicals and control substances

13. Assessment of solubility is conducted prior to testing. For this purpose, test chemicals are dissolved or stably dispersed at a concentration of 50 mg/mL in complete medium as first solvent option or dimethyl sulfoxide (DMSO, ≥ 99% purity) as a second solvent/vehicle option if the test chemical is not soluble in the complete medium solvent/vehicle. For the testing, the test chemical is dissolved to a final concentration of 0.4 mg/mL in complete medium if the chemical is soluble in this solvent/vehicle. If the chemical is soluble only in DMSO, the chemical is dissolved at a concentration of 50 mg/mL. Other solvents/vehicles than those described above may be used if sufficient scientific rationale is provided. Stability of the test chemical in the final solvent/vehicle should be taken into account.

14. The test chemicals and control substances are prepared on the day of testing. Because a dose finding assay is not conducted, for the first run, 6 final concentrations should be tested (1, 10, 20, 50, 100 and 200 µg/mL) into the corresponding solvent/vehicle either in complete medium or in 0.4% DMSO in medium. For the subsequent runs, starting from the 0.4 mg/mL in complete medium or 50 mg/mL in DMSO, solutions of the test chemicals, at least 4 working solutions (i.e. at least 4 concentrations), are prepared using the corresponding solvent/vehicle. The working solutions are finally used for treatment by adding an equal volume of U937 cell suspension (see paragraph 11 above) to the volume of working solution in the plate to achieve a further 2-fold dilution (12). The concentrations (at least 4 concentrations) for any further run are chosen based on the individual results of all previous runs (8). The usable final concentrations are 1, 2, 3, 4, 5, 7.5, 10, 12.5, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180 and 200 µg/mL. The maximum final concentration is 200 µg/mL. In the case of a CD86 positive value at 1 µg/mL is observed, then 0.1 µg/mL is evaluated in order to find the concentration of the test chemical that does not induce CD86 above the positive threshold. For each run, the EC150 (concentration at which a chemical reaches the CD86 positive threshold of 150%, see paragraph 19) is calculated if a CD86 positive concentration-response is observed. Where the test chemical induces a positive CD86 response not concentration related, the calculation of the EC150 might not be relevant as described in the U-SENS™ DB-ALM protocol no. 183 (12). For each run, CV70 (concentration at which a chemical reaches the cytotoxicity threshold of 70%, see paragraph 19) is calculated whenever possible (12). To investigate the concentration response effect of CD86 increase, any concentrations from the usable concentrations should be chosen evenly spread around the EC150 (or the highest CD86 negative non cytotoxic concentration) and the CV70 (or the highest concentration allowed i.e. 200 µg/mL). A minimum of 4 concentrations should be tested per run with at least 2 concentrations being common with the previous run(s), for comparison purposes.

15. The solvent/vehicle control used in the U-SENS™ method is complete medium (for test chemicals solubilised or stably dispersed) (see paragraph 4) or 0.4% DMSO in complete medium (for test chemicals solubilised or stably dispersed in DMSO).

16. The positive control used in the U-SENS™ method is TNBS (see paragraph 11), prepared in complete medium. TNBS should be used as the positive control for CD86 expression measurement at a final single concentration in plate (100 µg/mL) yielding > 70% of cell viability. To obtain a 100 µg/mL final concentration of TNBS in plate, starting from a 5% in water (i.e. 50 mg/mL) stock solution of TNBS is diluted 250-fold with complete medium to a 200 µg/mL working solution. Lactic acid (LA, CAS 50-21-5) should be used as the negative control at 200 µg/mL solubilised in complete medium (from a 0.4 mg/mL stock solution). In each plate of each run, three replicates of complete medium untreated control, solvent/vehicle control, negative and positive controls are prepared (14). Other suitable positive controls (e.g. abietic acid) may be used if historical data are available to derive comparable run acceptance criteria. The run acceptance criteria are the same as described for the test chemical (see paragraph 22).

© OECD (2026)

Application of test chemicals and control substances

17. The solvent/vehicle control or working solutions described in paragraphs 14-16 are mixed 1:1 (v/v) with the cell suspensions prepared in the 96-well flat-bottom plate (see paragraph 12). The treated plates are then incubated for 45±3 hours at 37°C under 5% CO_2. Prior to incubation, plates are sealed with semi permeable membrane, to avoid evaporation of volatile test chemicals and cross-contamination between cells treated with test chemicals (12).

Cell staining

18. After 45±3 hours of exposure, cells are transferred into V-shaped microtiter plate and collected by centrifugation. Solubility interference is defined as crystals or drops observed under the microscope at 45 ± 3 hours post treatment (before the cell staining). The supernatants are discarded and the remaining cells are washed once with 100 µL of an ice-cold phosphate buffered saline (PBS) containing 5 % foetal calf serum (staining buffer). After centrifugation, cells are re-suspended with 100 µL of staining buffer and stained with 5 µL (e.g. 0.25 µg) of FITC-labelled anti-CD86 or mouse IgG1 (isotype) antibodies at 4°C for 30 min protected from light. The antibodies described in the U-SENS™ DB-ALM protocol no. 183 (12) should be used (for CD86: BD-PharMingen #555657 Clone: Fun-1, or Caltag/Invitrogen # MHCD8601 Clone: BU63; and for IgG1: BD-PharMingen #555748, or Caltag/Invitrogen # GM4992). Based on the experience of the test method developers, the fluorescence intensity of the antibodies is usually consistent between different lots. Other clones or supplier of the antibodies which passed the reactivity check may be used for the assay (see paragraph 11). However, users may consider titrating the antibodies in their own laboratory's conditions to define the best concentration for use. Other detection system e.g. fluorochrome-tagged anti-CD86 antibodies may be used if they can be shown to provide similar results as FITC-conjugated antibodies, for example by testing the proficiency substances in Appendix II. After washing with 100 µL of staining buffer two times and once with 100 µL of an ice-cold PBS, cells are resuspended in ice-cold PBS (e.g. 125 µL for samples being analysed manually tube by tube, or 50 µL using an auto-sampler plate) and PI solution is added (final concentration of 3 µg/mL). Other cytotoxicity markers, such as 7-Aminoactinomycin D (7-AAD) or Trypan blue may be used if the alternative stains can be shown to provide similar results as PI, for example by testing the proficiency substances in Appendix II (18).

Flow cytometry analysis

19. Expression level of CD86 and cell viability are analysed using flow cytometry. Cells are displayed within a size (FSC) and granularity (SSC) dot plot set to log scale in order to clearly identify the population in a first gate R1 and eliminate the debris. A targeting total of 10,000 cells in gate R1 are acquired for each well. Cells from the same R1 gate are displayed within an IP detector SSC dot plot. Viable cells are delineated by placing a second gate R2 selecting the population of propidium iodide-negative cells (FL3 or FL4 channel). The cell viability can be calculated using the following equation by the cytometer analysis program. When the cell viability is low, up to 20,000 cells including dead cells could be acquired. Alternatively, data can be acquired for one minute after the initiation of the analysis.

$$\text{Cell Viability} = \frac{\text{Number of living cells}}{\text{Total number of acquired cells}} \times 100$$

Percentage of FITC-positive cells is then measured among these viable cells gated on R2 (within R1). Cell surface expression of CD86 is analysed in a FITC detector / SSC dot plot gated on viable cells (R2).

© OECD (2026)

For the complete medium / IgG1 wells, the analysis marker is set close to the main population so that the complete medium controls have IgG1 within the target zone of 0.6 to 0.9%.

Colour interference is defined as a shift of the FITC-labelled IgG1 dot-plot (FITC IgG1 Geo Mean S.I. ≥ 150%).

The stimulation index (S.I.) of CD86 for controls cells (untreated or in 0.4% DMSO) and chemical-treated cells are calculated according to the following equation:

$$\text{S.I.} = \frac{\%\text{ of CD86}^{+}\text{ treated cells - \% of IgG1}^{+}\text{ treated cells}}{\%\text{ of CD86}^{+}\text{ control cells - \% of IgG1}^{+}\text{ control cells}} \times 100$$

% of IgG1⁺ untreated control cells: referred to as percentage of FITC-positive IgG1 cells defined with the analysis marker (accepted range of ≥ 0.6% and < 1.5%, see paragraph 22) among the viable untreated cells.

% of IgG1⁺/CD86⁺ control/treated cells: referred to as percentage of FITC-positive IgG1/CD86 cells measured without moving the analysis marker among the viable control/treated cells.

DATA AND REPORTING

Data evaluation

20. The following parameters are calculated in the U-SENS™ test method: CV70 value, i.e. a concentration showing 70% of U937 cell survival (30% cytotoxicity) and the EC150 value, i.e. the concentration at which the test chemicals induced a CD86 stimulation index (S.I.) of 150%.

CV70 is calculated by log-linear interpolation using the following equation:

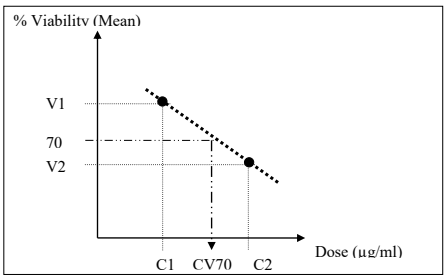
$$CV70 = C1 + [(V1 - 70) / (V1 – V2) * (C2 – C1)]$$

Where:

V1 is the minimum value of cell viability over 70%

V2 is the maximum value of cell viability below 70%

C1 and C2 are the concentrations showing the value of cell viability V1 and V2 respectively.



© OECD (2026)

Other approaches to derive the CV70 can be used as long as it is demonstrated that this has no impact on the results (e.g. by testing the proficiency substances).

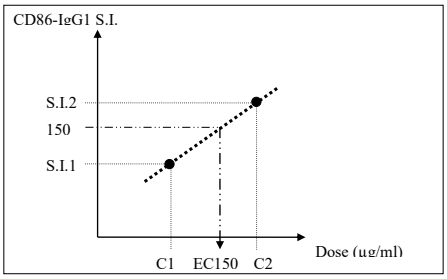
EC150 is calculated by log-linear interpolation using the following equation:

EC150 = C1 + [(150 – S.I.1) / (S.I.2 – S.I.1) * (C2 – C1)]

Where:

C1 is the highest concentration in µg/mL with a CD86 S.I. < 150% (S.I. 1)

C2 is the lowest concentration in µg/mL with a CD86 S.I. ≥ 150% (S.I. 2).



The EC150 and CV70 values are calculated

- for each run : the individual EC150 and CV70 values are used as tools to investigate the concentration response effect of CD86 increase (see paragraph 14),
- based on the average viabilities, the overall CV70 is determined (12) ,
- based on the average S.I. of CD86 values, the overall EC150 is determined for the test chemical predicted as POSITIVE with the U-SENS™ (see paragraph 21) (12).

Prediction model

21. The U-SENS prediction model is based on the measurement of CD86 induction (expressed as a Stimulation Index, SI) and cell viability. Each test chemical is tested in at least two independent runs to derive a single prediction (either POSITIVE, NEGATIVE, or BORDERLINE) (19).

A U-SENS independent run prediction is considered POSITIVE if:

- The cell viability is < 70% at all tested concentrations ≥ 10 µg/mL; or
- The SI of CD86 induction is > 176% at any tested concentration (with cell viability ≥ 70%).

A U-SENS independent run prediction is considered BORDERLINE (BL) if:

- The SI of CD86 induction is between 150% and 176% inclusive (150% ≤ SI CD86 ≤ 176%) with cell viability ≥ 70%; or

© OECD (2026)

- The SI of CD86 induction is between 128% and 150% (128% ≤ SI CD86 < 150%) at the highest non-cytotoxic concentration tested (with cell viability ≥ 70%).

A U-SENS independent run prediction is considered NEGATIVE if:

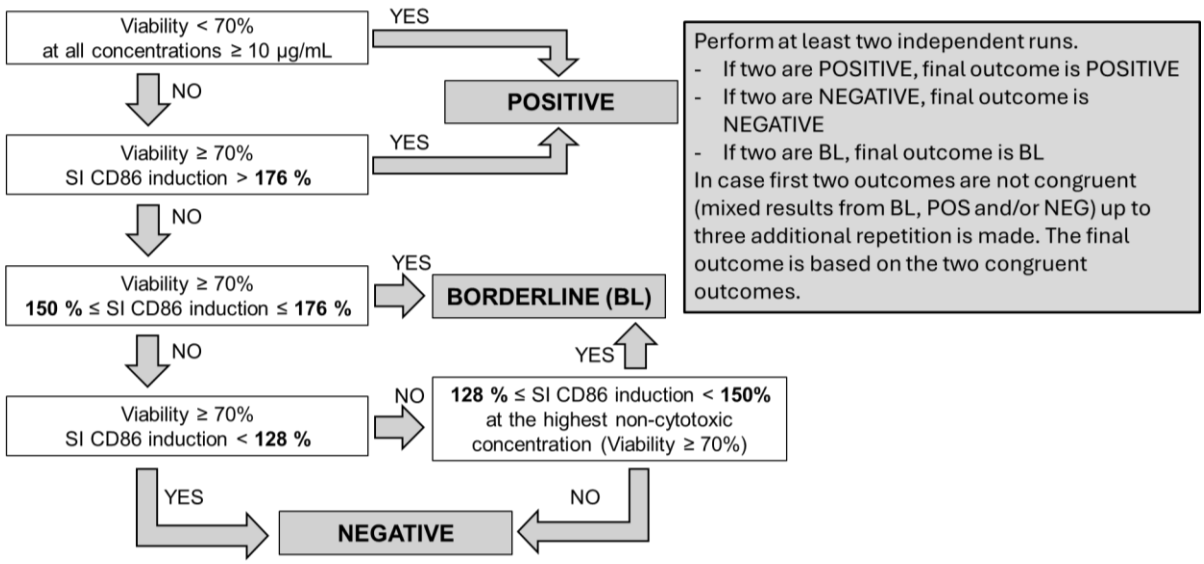
- The SI of CD86 induction is < 128% (with cell viability ≥ 70%) and the criteria for a Borderline result at the highest non-cytotoxic concentration are not met.

For the final outcome conclusion, at least two independent runs must be performed to reach a final conclusion:

- If the first two runs are POSITIVE, the final outcome is POSITIVE.
- If the first two runs are NEGATIVE, the final outcome is NEGATIVE.
- If the first two runs are BORDERLINE (BL), the final outcome is BORDERLINE.

In case the first two outcomes are not congruent (mixed results from BL, POS, and/or NEG), up to three additional repetitions are made. The final outcome is then based on the two congruent (majority) results (Figure 1).

Figure 1. Flow-chart of the U-SENS™ prediction model taking into account borderline ranges and multiple runs to conclude on borderline results.



Solubility or colour interference may require expert judgement to determine the final prediction.

The original threshold for a positive classification is 150% induction of the CD86 with a statistically derived borderline range around this threshold of 128% - 176% (19).

Acceptance criteria

22. The following acceptance criteria should be met when using the U-SENS™ method (12).

© OECD (2026)

- At the end of the 45±3 hours exposure period, the mean viability of the triplicate untreated U937 cells had to be > 90% and no drift in CD86 expression is observed. The basal expression (%IgG1 - %CD86) of untreated U937 cells had to be ≤ 25%.
- When DMSO is used as a solvent, the validity of the DMSO vehicle control is assessed by calculating a DMSO S.I. compared to untreated cells, and the mean viability of the triplicate cells had to be > 90%. The DMSO vehicle control is valid if the mean value of its triplicate CD86 S.I. was smaller than 250% of the mean of the triplicate CD86 S.I. of untreated U937 cells.
- The runs are considered valid if at least two out of three IgG1 values of untreated U937 cells fell within the range of ≥ 0.6% and < 1.5%.
- The concurrent tested negative control (lactic acid) is considered valid if at least two out of the three replicates were negative (CD86 S.I. < 150%) and non-cytotoxic (cell viability ≥ 70%).
- The positive control (TNBS) was considered as valid if at least two out of the three replicates were positive (CD86 S.I. ≥ 150%) and non-cytotoxic (cell viability ≥ 70%).

Test report

23. 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;
  - Physical appearance, complete medium solubility, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - 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 solvent/vehicle for each test chemical.
- Multi-constituent substance, UVCB and mixture:
  - Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
  - Physical appearance, complete medium solubility, DMSO solubility and additional relevant physicochemical properties, to the extent available;
  - Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions 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 solvent/vehicle for each test chemical.

© OECD (2026)

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, DMSO solubility, molecular weight, and additional relevant physicochemical properties, to the extent available and where applicable;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - 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.
- Negative and solvent/vehicle control
  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc.;
  - Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other control solvent/vehicle than those mentioned in the Test Guideline are used and to the extent available;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent/vehicle for each test chemical.

Test method Conditions

- Name and address of the sponsor, test facility and study director;
- Description of test method used;
- Cell line used, its storage conditions and source (e.g. the facility from which they were obtained);
- Flow cytometry used (e.g. model), including instrument settings, antibodies and cytotoxicity marker used;
- A statement attesting to the proficiency of the laboratory in performing the test method (e.g., by testing of proficiency substances) or to demonstrate reproducible performance of the test method over time. The Testing Facility shall maintain documentation of the procedures and proficiency chemicals used to demonstrate proficiency.

Test Acceptance Criteria

- Cell viability and CD86 S.I values obtained with the solvent/vehicle control in comparison to the acceptance ranges;
- Cell viability and S.I. values obtained with the positive control in comparison to the acceptance ranges;

© OECD (2026)

- Cell viability of all tested concentrations of the tested chemical.

Test procedure

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

Results

- Tabulation of the data, including CV70 (if applicable), S.I., cell viability values, EC150 values (if applicable) obtained for the test chemical and for the positive control in each run, and an indication of the rating of the test chemical according to the prediction model;
- Description of any other relevant observations, if applicable.

Discussion of the Results

- Discussion of the results obtained with the U-SENS™ method;
- Consideration of the test method results within the context of an IATA, if other relevant information is available.

Conclusions

© OECD (2026)

LITERATURE

1. Piroird, C., Ovigne, J.M., Rousset, F., Martinozzi-Teissier, S., Gomes, C., Cotovio, J., Alépée, N. (2015). The Myeloid U937 Skin Sensitization Test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. *Toxicol. In Vitro* 29, 901-916.
2. EURL ECVAM (2017). The U-SENS™ test method Validation Study Report. Accessible at: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations
3. EC EURL ECVAM (2016). ESAC Opinion No. 2016-03 on the L'Oréal-coordinated study on the transferability and reliability of the U-SENS™ test method for skin sensitisation testing. EUR 28178 EN; doi 10.2787/815737. Available at: [http://publications.jrc.ec.europa.eu/repository/handle/JRC103705].
4. EC EURL ECVAM (2017). EURL ECVAM Recommendation on the use of non-animal approaches for skin sensitisation testing. EUR 28553 EN; doi 10.2760/588955. Available at: [https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/eurl-ecvam-recommendation-use-non-animal-approaches-skin-sensitisation-testing].
5. Steiling, W. (2016). Safety Evaluation of Cosmetic Ingredients Regarding their Skin Sensitization Potential. doi:10.3390/cosmetics3020014.*Cosmetics* 3, 14.
6. OECD (2016), *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*, OECD Series on Testing and Assessment, No. 256, OECD Publishing, Paris, https://doi.org/10.1787/9789264279285-en.
7. Urbisch, D., Mehling, A., Guth, K., Ramirez, T., Honarvar, N., Kolle, S., Landsiedel, R., Jaworska, J., Kern, P.S., Gerberick, F., Natsch, A., Emter, R., Ashikaga, T., Miyazawa, M., Sakaguchi, H. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul. Toxicol. Pharmacol.* 71, 337-351.
8. Alépée, N., Piroird, C., Aujoulat, M., Dreyfuss, S., Hoffmann, S., Hohenstein, A., Meloni, M., Nardelli, L., Gerbeix, C., Cotovio, J. (2015). Prospective multicentre study of the U-SENS test method for skin sensitization testing. *Toxicol In Vitro* 30, 373-382.
9. Reisinger, K., Hoffmann, S., Alépée, N., Ashikaga, T., Barroso, J., Elcombe, C., Gellatly, N., Galbiati, V., Gibbs, S., Groux, H., Hibatallah, J., Keller, D., Kern, P., Klaric, M., Kolle, S., Kuehnl, J., Lambrechts, N., Lindstedt, M., Millet, M., Martinozzi-Teissier, S., Natsch, A., Petersohn, D., Pike, I., Sakaguchi, H., Schepky, A., Tailhardat, M., Templier, M., van Vliet, E., Maxwell, G. (2014). Systematic evaluation of non-animal test methods for skin sensitisation safety assessment. *Toxicol. In Vitro* 29, 259-270.
10. Fabian, E., Vogel, D., Blatz, V., Ramirez, T., Kolle, S., Eltze, T., van Ravenzwaay, B., Oesch, F., Landsiedel, R. (2013). Xenobiotic metabolizin enzyme activities in cells used for testing skin sensitization *in vitro*. *Arch. Toxicol.* 87, 1683-1696.
11. OECD (2025), *Guidance Document on Good In Vitro Method Practices (GIVIMP), Second Edition*, OECD Series on Testing and Assessment, No. 421, OECD Publishing, Paris, https://doi.org/10.1787/5ba6777b-en.
12. DB-ALM (2016). Protocol no. 183: Myeloid U937 Skin Sensitization Test (U-SENS™), 33pp. Accessible at: [http://ecvam-dbalm.jrc.ec.europa.eu/].
13. Sundström, C., Nilsson, K. (1976). Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int. J. Cancer* 17, 565-577.
14. OECD (2005), *Guidance Document on the Validation and International Acceptance of New or*

© OECD (2026)

Updated Test Methods for Hazard Assessment, OECD Series on Testing and Assessment, No. 34, OECD Publishing, Paris, 96 pp, https://doi.org/10.1787/e1f1244b-en.

15. United Nations UN (2025). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Eleventh revised edition. New York & Geneva: United Nations Publications. ISBN: 9789211067699. Available at: [https://doi.org/10.18356/f8fbb7cb-en].
16. OECD (2012), *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins*, *Part 1,* OECD Series on Testing and Assessment, No. 168, OECD Publishing, https://doi.org/10.1787/9789264221444-en.
17. ECETOC (2003). Technical Report No. 87: Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals, Brussels. Available at: [https://ftp.cdc.gov/pub/Documents/OEL/06.%20Dotson/References/ECETOC_2003-TR87.pdf].
18. Ade, N., Teluob, S., Viricel, A., Piroird, C., Alépée, N. (2022). Amending the U-SENS™ skin sensitization test method for interfering auto-fluorescent chemicals. *Toxicol. In Vitro* 81, 105353. https://doi.org/10.1016/j.tiv.2022.105353
19. Nardelli, L., Tourneix, F., Carron, L., van Vliet, E., Alépée, N. (2025). Determination of U-SENS™ borderline range thresholds to address uncertainty for skin sensitization assessment in a regulatory context. *ALTEX*, 42(2) : 625-635. doi: 10.14573/altex.2411282.

© OECD (2026)

APPENDIX I: 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 (14).

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 (15).

CD86 Concentration response: There is concentration-dependency (or concentration response) when a positive concentration (CD86 S.I. ≥ 150) is followed by a concentration with an increasing CD86 S.I.

CV70: The estimated concentration showing 70% cell viability.

Drift: A drift is defined by i) the corrected %CD86$^+$ value of the untreated control replicate 3 is less than 50% of the mean of the corrected %CD86$^+$ value of untreated control replicates 1and 2; and ii) the corrected %CD86$^+$ value of the negative control replicate 3 is less than 50% of mean of the corrected %CD86$^+$ value of negative control replicates 1and 2.

EC150: the estimated concentrations showing the 150% S.I. of CD86 expression.

Flow cytometry: a cytometric technique in which cells suspended in a fluid flow one at a time through a focus of exciting light, which is scattered in patterns characteristic to the cells and their components; cells are frequently labeled with fluorescent markers so that light is first absorbed and then emitted at altered frequencies.

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 weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

Mixture: 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).

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.

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.

Pre-haptens: chemicals which become sensitisers through abiotic transformation, e.g. through oxidation.

Pro-haptens: chemicals requiring enzymatic activation to exert skin sensitisation potential.

© OECD (2026)

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 (14).

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 (14).

Run: A run consists of one or more test chemicals tested concurrently with a solvent/vehicle control and with a positive control.

Sensitivity: The proportion of all positive/active 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 (14).

S.I.: Stimulation Index. Relative values of geometric mean fluorescence intensity in chemical-treated cells compared to solvent-treated cells.

Solvent/vehicle control: An untreated sample containing all components of a test system except of the test chemical, but including the solvent/vehicle that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved or stably dispersed in the same solvent/vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent/vehicle interacts with the test system.

Specificity: The proportion of all negative/inactive 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 (14).

Staining buffer: A phosphate buffered saline containing 5% foetal calf serum.

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 it composition.

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 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 (16).

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

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 (14).

© OECD (2026)

APPENDIX II: PROFICIENCY SUBSTANCES

Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency by correctly obtaining the expected U-SENS™ prediction for the 10 substances recommended in Table 1 and by obtaining CV70 and EC150 values that fall within the respective reference range for at least 8 out of the 10 proficiency substances. Proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that the substances are commercially available, and that high-quality *in vivo* reference data as well as high quality *in vitro* data generated with the U-SENS™ method are available. Also, published reference data are available for the U-SENS™ method (1) (8).

Table 1. Recommended substances for demonstrating technical proficiency with the U-SENS™ method

Proficiency substances	CASRN	Physical state	*In vivo* prediction[1]	U-SENS™ Solvent/ Vehicle	U-SENS™ CV70 Reference Range in µg/mL[2]	U-SENS™ EC150 Reference Range in µg/mL[2]
4-Phenylenediamine	106-50-3	Solid	Sensitiser (strong)	Complete medium[3]	<30	Positive (≤10)
Picryl sulfonic acid	2508-19-2	Liquid	Sensitizer (strong)	Complete medium	>50	Positive (≤50)
Diethyl maleate	141-05-9	Liquid	Sensitiser (moderate)	DMSO	10-100	Positive (≤20)
Resorcinol	108-46-3	Solid	Sensitiser (moderate)	Complete medium	>100	Positive (≤50)
Cinnamic alcohol	104-54-1	Solid	Sensitiser (weak)	DMSO	>100	Positive (10-100)
4-Allylanisole	140-67-0	Liquid	Sensitiser (weak)	DMSO	>100	Positive (<200)
Saccharin	81-07-2	Solid	Non-sensitiser	DMSO	>200	Negative (>200)
Glycerol	56-81-5	Liquid	Non-sensitiser	Complete medium	>200	Negative (>200)
Lactic acid	50-21-5	Liquid	Non-sensitiser	Complete medium	>200	Negative (>200)
Salicylic acid	69-72-7	Solid	Non-sensitiser	DMSO	>200	Negative (>200)

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number

[1] The *in vivo* hazard and (potency) prediction is based on LLNA data (1) (8). The *in vivo* potency is derived using the criteria proposed by ECETOC (17).

[2] Based on historical observed values (1) (8).

[3] Complete medium: RPMI-1640 medium supplemented with 10% foetal calf serum, 2 mM L-glutamine, 100 units/mL penicillin and 100 µg/mL streptomycin (8).

© OECD (2026)

ANNEX 3 - IN VITRO SKIN SENSITISATION: INTERLEUKIN-8 REPORTER GENE ASSAY (IL-8 LUC ASSAY)

INITIAL CONSIDERATIONS AND LIMITATIONS

1. In contrast to assays analysing the expression of cell surface markers, the IL8-Luc assay quantifies changes in IL-8 expression, a cytokine associated with the activation of dendritic cells (DC) (1) (2). In the THP-1-derived IL-8 reporter cell line (THP-G8, established from the human acute monocytic leukemia cell line THP-1), IL-8 expression is measured following exposure to sensitisers (2) (3) (4). The expression of luciferase is then used to aid discrimination between skin sensitisers and non-sensitisers. In 2023, the test method was updated with a revised prediction model enhancing the applicability of the method to some poorly soluble chemicals, based on the evaluation of cytotoxicity using an inherent viability marker (5).

2. The IL-8 Luc method has been evaluated in a validation study (5) conducted by the Japanese Centre for the Validation of Alternatives Methods (JaCVAM), the Ministry of Economy, Trade and Industry (METI), and the Japanese Society for Alternatives to Animal Experiments (JSAAE) and subsequently subjected to independent peer review (6) under the auspices of JaCVAM and the Ministry of Health, Labour and Welfare (MHLW) with the support of the International Cooperation on Alternative Test Methods (ICATM). Considering all available evidence and input from regulators and stakeholders, the IL-8 Luc assay is considered useful as part of Integrated Approaches to Testing and Assessment (IATA) to discriminate sensitisers from non-sensitisers for the purpose of hazard classification and labelling. Examples of the use of IL-8 Luc assay data in combination with other information are reported in the literature (3) (7).

3. The IL-8 Luc assay proved to be transferable to laboratories experienced in cell culture and luciferase measurement. Within and between laboratory reproducibilities were 87.7% and 87.5%, respectively (5). Data generated in the validation study (5) and other published papers (3) (4) (8) presented the performance of the IL-8 Luc assay compared with the LLNA. In these studies, the IL-8 Luc assay judged 130 out of 143 chemicals as positive or negative and judged 13 chemicals as inconclusive. The accuracy of the IL-8 Luc assay in distinguishing skin sensitisers (UN GHS Cat. 1) from non-sensitisers (UN GHS No Cat.) is 83.6% (109/130) with a sensitivity of 92.0% (92/100) and a specificity of 56.7% (17/30). Excluding substances outside the applicability domain, such as surfactants, anhydrides and chemicals interfering with luciferase, described below (paragraph 6), the performance of the IL-8 Luc assay was 93.9% (92/98) for sensitivity, 68.0% (17/25) for specificity, and 88.6% (109/123) for accuracy. Using human data cited in Urbisch et al. (7), the IL-8 Luc assay judged 84 out of 90 chemicals as positive or negative and judged 6 chemicals as inconclusive. The performance was 89.7% (52/58) for sensitivity, 50.0% (13/26) for specificity, and 77.4% (65/84) for accuracy. Excluding substances outside the applicability domain, the IL-8 Luc assay judged 78 out of 84 chemicals as positive or negative and judged 6 chemicals as inconclusive and the accuracy is 82.1% (64/78) with a sensitivity of 89.7% (52/58) and a specificity of 60.0% (12/20). The difference in the performance of the IL-8 Luc assays for chemicals with $LogK_{ow}$ <3.5 and those with $LogK_{ow}$ ≥3.5 was examined and it was demonstrated that $LogK_{ow}$ ≥3.5 did not reduce the sensitivity of the IL-8 Luc assay (8).

4. False negative predictions with the IL-8 Luc assay are more likely to occur with chemicals

© OECD (2026)

showing low/moderate skin sensitisation potency (UN GHS subcategory 1B) than those with high potency (UN GHS subcategory 1A) (3) (4) (8). Together, the information supports a role for the IL-8 Luc assay in the identification of skin sensitisation hazards. The accuracy given for the IL-8 Luc assay as a standalone test method is only for guidance, as the method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore, when evaluating non-animal methods for skin sensitisation, it should be remembered that the LLNA and other animal tests may not fully reflect the situation in humans.

5. On the basis of the data currently available, the IL-8 Luc assay was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, skin sensitisation potency (as determined in in vivo studies) and physicochemical properties (3) (4) (5). The IL-8 Luc assay is also technically applicable to the testing of multi-constituent substances and mixtures. However, mixtures that do not dissolve completely at 20 mg/ml of X-VIVOTM 15 and are found to be non-sensitizers may contain hidden sensitizers due to the toxicity of other chemicals present. This should be taken into account along with the general concerns about in vitro sensitisation tests when examining mixtures described in paragraph 9 of the General Introduction.

6. A high false negative rate for anhydrides was seen in the validation study. Furthermore, because of the limited metabolic capability of the cell line (9) and the experimental conditions, pro-haptens (substances requiring metabolic activation) and pre-haptens (substances activated by air oxidation) might give negative results in the assay. However, although negative results for suspected pre/pro-haptens should be interpreted with caution, the IL-8 Luc assay correctly judged 11 out of 11 pre-haptens, 6/6 pro-haptens, and 5/8 pre/pro-haptens in the IL-8 Luc assay data set (5) (7). Based on the recent comprehensive review on three non-animal methods (the DPRA, the KeratinoSens™ and the h-CLAT) to detect pre and pro-haptens (10), and based on the fact that THP-G8 cells used in the IL-8 Luc assay is a cell line derived from THP-1 that is used in the h-CLAT, the IL-8 Luc assay may also contribute to increase the sensitivity of non-animal methods to detect pre and pro-haptens in the combination of other methods. Surfactants tested so far gave (false) positive results irrespective of their type (e.g. cationic, anionic or non-ionic). Finally, chemicals that interfere with luciferase can confound its activity/measurement, causing apparent inhibition or increased luminescence (11). For example, phytoestrogen concentrations higher than 1μM were reported to interfere with luminescence signals in other luciferase-based reporter gene assays due to over-activation of the luciferase reporter gene. Consequently, luciferase expression obtained at high concentrations of phytoestrogens or compounds suspected of producing phytoestrogen-like activation of the luciferase reporter gene needs to be examined carefully (12). Based on the above, surfactants, anhydrides and chemicals interfering with luciferase are outside the applicability domain of this assay. In cases where there is evidence demonstrating the non-applicability of the IL-8 Luc assay to other specific categories of test chemicals, the method should not be used for those specific categories.

7. As described above, the IL-8 Luc assay supports discrimination of skin sensitisers from non-sensitisers. Further work, preferably based on human data, is required to determine whether IL-8 Luc results can contribute to potency assessment when considered in combination with other information sources.

8. Definitions and abbreviations are provided in Table 1 and Appendix I.

PRINCIPLE OF THE TEST

9. The IL-8 Luc assay makes use of a human monocytic leukemia cell line THP-1 that was obtained from the American Type Culture Collection (Manassas, VA, USA). Using this cell line, the Dept. of Dermatology, Tohoku University School of Medicine, established a THP-1-derived IL-8 reporter cell line, THP-G8, that harbours the Stable Luciferase Orange (SLO) and Stable Luciferase Red (SLR) luciferase genes under the control of the IL-8 and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoters, respectively (2). This allows quantitative measurement of luciferase gene induction by

© OECD (2026)

detecting luminescence from well-established light producing luciferase substrates as an indicator of the activity of the IL-8 and GAPDH in cells following exposure to sensitising chemicals. The dual-colour assay system comprises an orange-emitting luciferase (SLO from *Rhagophthalmus ohbai*; λmax = 580 nm) (13) for the gene expression of the IL-8 promoter as well as a red-emitting luciferase (SLR from *Phrixothrix hirtus*; λmax = 630 nm) (14) for the gene expression of the internal control promoter, GAPDH. The two luciferases emit different colours upon reacting with firefly D-luciferin and their luminescence is measured simultaneously in a one-step reaction by dividing the emission from the assay mixture using an optical filter (15) (Appendix II). In addition, GAPDH mRNA is ubiquitously expressed at moderately abundant levels. It is frequently used as an endogenous control for quantitative real time polymerase chain reaction in some experimental systems, because its expression is constant at different times and after experimental manipulation (16) (17) (18). The inhibition of GAPLA (Inh-GAPLA) has proven to be a good marker of cell viability, with a strong correlation to propidium iodide (PI)-exclusion cells, a marker commonly used to determine cell viability by flow cytometry. Inh-GAPLA below 0.8 indicates cytotoxicity of the test chemical, which in turn suggests that the chemical dissolved in the culture medium. Therefore, the assay can be used to verify exposure with poorly soluble chemicals and to reduce the number of inconclusive results (8).

10. THP-G8 cells are treated for 16 hours with the test chemical, after which SLO luciferase activity (SLO-LA) reflecting IL-8 promoter activity and SLR luciferase activity (SLR-LA) reflecting GAPDH promoter activity are measured. To make the abbreviations easy to understand, SLO-LA and SLR-LA are designated as IL8LA and GAPLA, respectively. Table 3 gives a description of the terms associated with luciferase activity in the IL-8 Luc assay. The measured values are used to calculate the normalised IL8LA (nIL8LA), which is the ratio of IL8LA to GAPLA; the induction of nIL8LA (Ind-IL8LA), which is the ratio of the arithmetic means of quadruple-measured values of the nIL8LA of THP-G8 cells treated with a test chemical and the values of the nIL8LA of untreated THP-G8 cells; and Inh-GAPLA, which is the ratio of the arithmetic means of quadruple-measured values of the GAPLA of THP- G8 cells treated with a test chemical and the values of the GAPLA of untreated THP-G8 cells, and used as an indicator for cytotoxicity.

Table 1. Description of terms associated with the luciferase activity in the IL-8 Luc assay

Abbreviations	Definition
GAPLA	SLR luciferase activity reflecting GAPDH promoter activity
IL8LA	SLO luciferase activity reflecting IL-8 promoter activity
nIL8LA	IL8LA / GAPLA
Ind-IL8LA	nIL8LA of THP-G8 cells treated with chemicals / nIL8LA of untreated cells
Inh-GAPLA	GAPLA of THP-G8 treated with chemicals / GAPLA of untreated cells
CV05	The lowest concentration of the chemical at which Inh-GAPLA becomes <0.05.
PI reduction	% of PI-excluding cells among THP-G8 cells treated with chemicals/ % of PI-excluding cells among THP-G8 cells without chemical treatment.

DEMONSTRATION OF PROFICIENCY

11. Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency, using the 9 Proficiency Substances listed in Appendix III in compliance with the Good in vitro Method Practices (19). Moreover, test method users

© OECD (2026)

should maintain a historical database of data generated with the reactivity checks and with the positive and solvent/vehicle controls, and use these data to confirm the reproducibility of the test method in their laboratory is maintained over time.

PROCEDURE

12. The Standard Operating Procedure (SOP) for the IL-8 Luc assay is available and should be employed when performing the test (5). Laboratories willing to perform the test can obtain the recombinant THP-G8 cell line from Tottori Bioscience Promotion Organization, Tottori, Japan, upon signing a Material Transfer Agreement (MTA) in line with the conditions of the OECD template. The following paragraphs provide a description of the main components and procedures of the assay.

Preparation of cells

13. The THP-G8 cell line should be used for performing the IL-8 Luc assay (see paragraphs 9 and 12). On receipt, cells are propagated (2-4 passages) and stored frozen as a homogeneous stock. Cells from this stock can be propagated up to a maximum of 12 passages or a maximum of 6 weeks. The medium used for propagation is the RPMI-1640 culture medium containing 10% foetal bovine serum (FBS), antibiotic/antimycotic solution (100U/mL of penicillin G, 100µg/mL of streptomycin and 0.25µg/mL of amphotericin B in 0.85% saline) (e.g. GIBCO Cat#15240-062), 0.15µg/mL Puromycin (e.g. CAS:58-58-2) and 300µg/mL G418 (e.g. CAS:108321-42-2).

14. Prior to use for testing, the cells should be qualified by conducting a reactivity check. This check should be performed 1-2 weeks or 2-4 passages after thawing, using the positive control, 4-nitrobenzyl bromide (4-NBB) (CAS:100-11-8, ≥ 99% purity) and the negative control, lactic acid (CAS:50-21-5, ≥85% purity). 4-NBB should produce a positive response to Ind-IL8LA (≥1.4), while lactic acid should produce a negative response to Ind-IL8LA (<1.4). If this condition cannot be met, it is recommended to use a new frozen stock vial and recheck the concentration of 4-NBB. Only cells that pass the reactivity check are used for the assay. The check should be performed according to the procedures described in paragraphs 21-23.

15. For testing, THP-G8 cells are seeded at a density of 2 to 5 × 10^5 cells/mL, and pre-cultured in culture flasks for 48 to 96 hours. On the day of the test, cells harvested from the culture flask are washed with RPMI-1640 containing 10% FBS without any antibiotics, and then, resuspended with RPMI-1640 containing 10% FBS without any antibiotics at 1 × 10^6 cells/mL. Then, cells are distributed into a 96-well flat-bottom black plate (e.g. Costar Cat#3603) with 50µL (5 × 10^4 cells/well).

Preparation of the test chemical and control substances

16. The test chemical and control substances are prepared on the day of testing. For the IL-8 Luc assay, test chemicals are dissolved in X-VIVO™ 15, a commercially available serum-free medium (Lonza[2], 04-418Q), to the final concentration of 20 mg/mL. X-VIVO™ 15 is added to 20 mg of test chemical (regardless of the chemical's solubility) in a microcentrifuge tube and brought to a volume of 1mL and then vortexed vigorously and shaken on a rotor at a maximum speed of 8 rpm for 30 min at an ambient temperature of about 20°C. Furthermore, if solid chemicals are still insoluble, the tube is sonicated until the chemical is dissolved completely or stably dispersed. For test chemicals soluble in X-VIVO™ 15, the solution is diluted by a factor of 5 with X-VIVO™ 15 and used as an X-VIVO™ 15 stock solution of the test chemical (4 mg/mL). For test chemicals not fully soluble at 20 mg/mL in X-VIVO™ 15, the mixture is rotated again for at least 30 min, then centrifuged at 15,000 rpm (≈20,000g) for 5 min; the

[2] https://bioscience.lonza.com/lonza_bs/IT/en/Culture-Media-and-Reagents/p/000000000000217685/X-VIVO-15-Serum-free-Hematopoietic-Cell-Medium

© OECD (2026)

resulting supernatant is used as an X-VIVO™ 15 stock solution of the test chemical. A scientific rationale should be provided for the use of other solvents, such as DMSO, water, or the culture medium. The detailed procedure for dissolving chemicals is shown in Appendix V. The X-VIVO™ 15 solutions described in paragraphs 17-22 are mixed 1:1 (v/v) with the cell suspensions prepared in a 96-well flat-bottom black plate (see paragraph 15).

17. The first test run is aimed to determine the cytotoxic concentration and to examine the skin sensitising potential of chemicals. Using X-VIVO™ 15, serial dilutions of the X-VIVO™ 15 stock solutions of the test chemicals are made at a dilution factor of two (see Appendix V) using a 96-well assay block (e.g. Costar Cat#EW-01729-03). Next, 50 µL/well of diluted solution is added to 50 µL of the cell suspension in a 96-well flat-bottom black plate. Thus, for test chemicals that are soluble in X-VIVO™ 15, the final concentrations of the test chemicals range from 0.002 to 2 mg/mL (Appendix V). For test chemicals that are not soluble in X-VIVO™ 15 at 20 mg/mL, only dilution factors that range from 2 to 210, are determined, although the actual final concentrations of the test chemicals remain uncertain and are dependent on the saturated concentration of the test chemicals in the X-VIVO™ 15 stock solution.

18. In subsequent test runs (i.e. the second, third, and fourth replicates), the X-VIVO™ 15 stock solution is made at the concentration 4 times higher than the concentration of cell viability 05 (CV05; the lowest concentration at which the Inh-GAPLA becomes <0.05) in the first experiment. If Inh-GAPLA does not decrease below 0.05 at the highest concentration in the first run, the X-VIVO™ 15 stock solution is made at the first run highest concentration. The concentration of CV05 is calculated by dividing the concentration of the stock solution in the first run by dilution factor for CV05 (X) (dilution factor CV05 (X); the dilution factor required to dilute stock solution to CV05) (see Appendix V). For test substances not soluble in X-VIVO at 20 mg/ml, CV05 is determined by the concentration of the stock solution x 1/X. For run 2 to 4, a second stock solution is prepared as 4 x CV50 (Appendix V).

19. Serial dilutions of the X-VIVO™ 15 second stock solutions are made at a dilution factor of 1.5 using a 96-well assay block. Next, 50 µL/well of diluted solution is added to 50 µL of the cell suspension in the wells of a 96-well flat-bottom black plate. Each concentration of each test chemical should be tested in 4 wells. The samples are then mixed on a plate shaker and incubated for 16 hours at 37°C and 5% CO_2, after which the luciferase activity is measured as described below.

20. The solvent control is the mixture of 50 µL/well of X-VIVOTM 15 and 50 µL/well of cell suspension in RPMI-1640 containing 10% FBS.

21. The recommended positive control is 4-NBB. 20 mg of 4-NBB is prepared in a 1.5-mL microfuge tube, to which X-VIVOTM 15 is added up to 1 mL. Since 4-NBB is not fully soluble at 20 mg/mL in X-VIVO 15 (see paragraph 16), the tube is vortexed vigorously and shaken on a rotor at a maximum speed of 8 rpm for at least 30 min. After centrifugation at 20,000g for 5 min, the supernatant is diluted by a factor of 4 with X-VIVOTM 15, and 500 µL of the diluted supernatant is transferred to a well in a 96-well assay block. The diluted supernatant is further diluted with X-VIVO™ 15 at factors of 2 and 4, and 50 µL of the solution is added to 50 µL of THP-G8 cell suspension in the wells of a 96-well flat-bottom black plate (Appendix VI). Each concentration of the positive control should be tested in 4 wells. The plate is agitated on a plateshaker, and incubated in a CO_2 incubator for 16 hours (37°C, 5% CO_2), after which the luciferase activity is measured as described in paragraph 28.

22. The recommended negative control is lactic acid. 20 mg of lactic acid prepared in a 1.5-mL microfuge tube, to which X-VIVO™ 15 is added up to 1 mL (20 mg/ mL). Twenty mg/mL of lactic acid solution is diluted by a factor of 5 with X-VIVO™ 15 (4 mg/mL); 500 µL of this 4 mg/mL lactic acid solution is transferred to a well of a 96-well assay block. This solution is diluted by a factor of 2 with X-VIVO™ 15 and then diluted again by a factor of 2 to produce 2 mg/mL and 1 mg/mL solutions. 50 µL of these 3 solutions and vehicle control (X-VIVO™ 15) are added to 50 µl of THP-G8 cell suspension in the wells of a 96-well flat-bottom black plate. Each concentration of the negative control is tested in 4 wells. The plate is agitated on a plate shaker and incubated in a CO_2 incubator for 16 hours (37°C, 5% CO_2), after which

© OECD (2026)

the luciferase activity is measured as described in paragraph 28.

23. Other suitable positive or negative controls may be used if historical data are available to derive comparable run acceptance criteria.

24. Care should be taken to avoid evaporation of volatile test chemicals and cross-contamination between wells by test chemicals, e.g. by sealing the plate prior to the incubation with the test chemicals.

25. The test chemicals and solvent control require 2 to 4 runs to derive a positive or negative prediction (see Table 2). Each run is performed on a different day with fresh X-VIVO™ 15 stock solution of test chemicals and independently harvested cells. Cells may come from the same passage.

Luciferase activity measurements

26. Luminescence is measured using a 96-well microplate luminometer equipped with optical filters, e.g. Phelios (ATTO, Tokyo, Japan), Tristan 941 (Berthold, Bad Wildbad, Germany) and the ARVO series (PerkinElmer, Waltham, MA, USA). The luminometer must be calibrated for each test to ensure reproducibility. Recombinant orange and red emitting luciferases are available for this calibration.

27. 100µL of pre-warmed Tripluc® Luciferase assay reagent (Tripluc) is transferred to each well of the plate containing the cell suspension treated with or without chemical. The plate is shaken for 10 min at an ambient temperature of about 20°C. The plate is placed in the luminometer to measure the luciferase activity. Bioluminescence is measured for 3 sec each in the absence (F0) and presence (F1) of the optical filter. Justification should be provided for the use of alternative settings, e.g. depending on the model of luminometer used.

28. Parameters for each concentration are calculated from the measured values, e.g. IL8LA, GAPLA, nIL8LA, Ind-IL8LA, Inh-GAPLA, the mean ±SD of IL8LA, the mean ±SD of GAPLA, the mean ±SD of nIL8LA, the mean ±SD of Ind-IL8LA, the mean ±SD of Inh-GAPLA, and the 95% confidence interval of Ind-IL8LA. Definitions of the parameters used in this paragraph are provided in Appendices I and IV, respectively.

29. Prior to measurement, colour discrimination in multi-colour reporter assays is generally achieved using detectors (luminometer and plate reader) equipped with optical filters, such as sharp-cut (long-pass or short-pass) filters or band-pass filters. The transmission coefficients of the filters for each bioluminescence signal colour should be calibrated prior to testing, per Appendix II.

DATA AND REPORTING

Data evaluation

30. The criteria for a positive/negative decision in each run in the IL-8 Luc assay are the following:

- an IL-8 Luc assay prediction is judged positive if a test chemical has a Ind-IL8LA > 1.4 and the lower limit of the 95% confidence interval of Ind-IL8LA < 1.0
- an IL-8 Luc assay prediction is judged negative if a test chemical has a Ind-IL8LA < 1.4 and/or the lower limit of the 95% confidence interval of Ind-IL8LA < 1.0

Prediction model

31. Test chemicals that provide two positive results from among the 1st, 2nd, 3rd. and 4th runs are identified as positives whereas those that give three negative results from among the 1st, 2nd, 3rd and 4th

© OECD (2026)

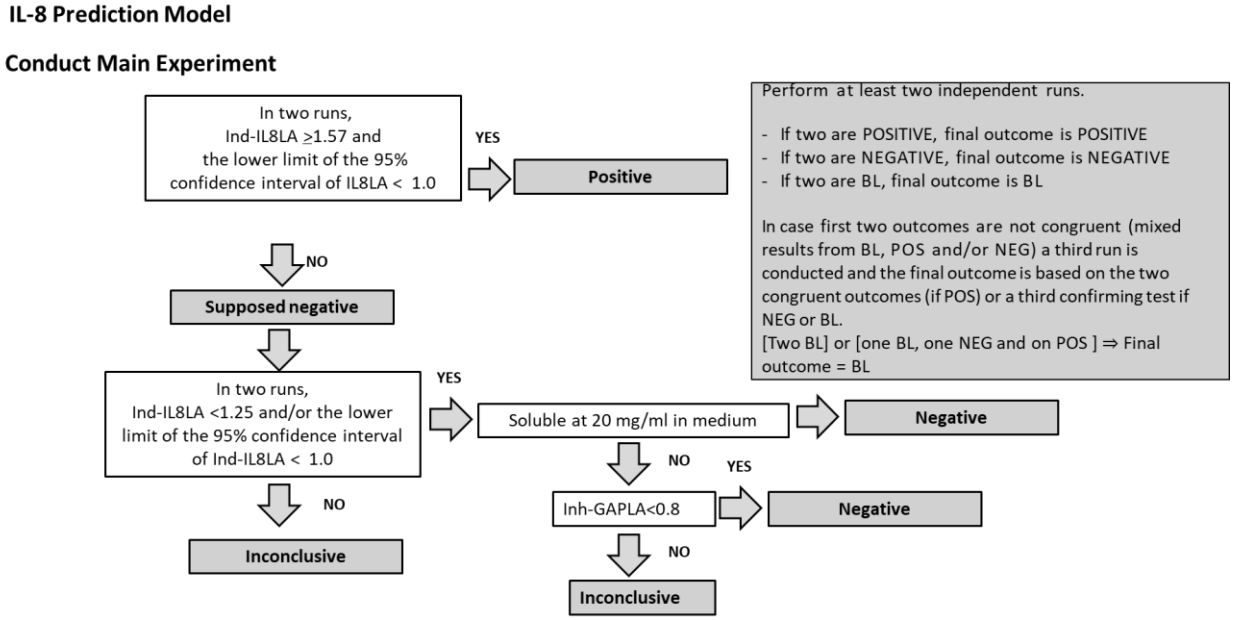
runs are identified as supposed negative (Table 2). Among supposed negative chemicals, if chemicals are dissolved at 20 mg/ml in X-VIVO™ 15, they are judged as negative. If chemicals are not dissolved at 20 mg/ml in X-VIVO™ 15, or there is cause for uncertainty regarding the extent of dissolution or miscibility, chemicals that give less than 0.8 of Inh-GAPLA are judged as negative, while those that give 0.8 or higher of Inh-GAPLA are judged as inconclusive (Figure 1).

Table 2. Criteria for identifying positive and supposed negative

1st run	2nd run	3rd run	4th run	Judge
Positive	Positive	-	-	Positive
Positive	Negative	Positive	-	Positive
Positive	Negative	Negative	Positive	Positive
Positive	Negative	Negative	Negative	Supposed negative
Negative	Positive	Positive	-	Positive
Negative	Positive	Negative	Positive	Positive
Negative	Positive	Negative	Negative	Supposed negative
Negative	Negative	Positive	Positive	Positive
Negative	Negative	Positive	Negative	Supposed negative
Negative	Negative	Negative	-	Supposed negative

© OECD (2026)

Figure 1. Prediction model for IL-8 Luc Assay taking into account borderline ranges and multiple runs to conclude on borderline results



The original threshold for a positive classification is 1.4 -fold luciferase induction with a statistically derived borderline range around this threshold of 1.25 – 1.57.

Acceptance criteria

32. The following acceptance criteria should be met when using the IL-8 Luc assay:

- Ind-IL8LA should be more than 5.0 at least in one concentration of the positive control, 4-NBB, in each run.
- Ind-IL8LA should be less than 1.4 at any concentration of the negative control, lactic acid, in each run.
- Data from plates where GAPLA in control wells containing cells and Tripluc but no chemicals is less than five times GAPLA in wells containing only test medium and Tripluc should be rejected.
- Data from plates with Inh-GAPLA of less than 0.05 for all concentrations of test chemical should be rejected. In this case, the initial test should be repeated so that the highest final concentration of the repeated test is the lowest final concentration of the previous test.

TEST REPORT

33. The test report should include the following information:

Test chemicals

- Mono-constituent substance:

  - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI

© OECD (2026)

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 as appropriate and practically feasible, etc;
- Treatment prior to testing, if applicable (e.g. warming, grinding);
- Solubility in X-VIVO™ 15. For chemicals that are insoluble in X-VIVO™ 15, whether precipitation or flotation are observed after centrifugation;
- Concentration(s) tested;
- Storage conditions and stability to the extent available;
- Justification for choice of solvent/vehicle for each test chemical if X-VIVO™ 15 has not been used.

- Multi-constituent substance, UVCB and mixture:

  - Characterisation as far as possible by e.g. chemical identity (see above), purity, 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 case of mixtures/polymers of known compositions or other information relevant for the conduct of the study;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Solubility in X-VIVO™ 15. For chemicals that are insoluble in X-VIVO™ 15, whether precipitation or flotation are observed after centrifugation;
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available.
  - Justification for choice of solvent/vehicle for each test chemical, if X-VIVO™ 15 has not been used.

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 and where applicable;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
  - Treatment prior to testing, if applicable (e.g. warming, grinding);
  - Concentration(s) tested;
  - Storage conditions and stability to the extent available;

© OECD (2026)

- Reference to historical positive control results demonstrating suitable acceptance criteria, if applicable.

- Negative control:

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

Test method conditions

- Name and address of the sponsor, test facility and study director;
- Description of test method used;
- Cell line used, its storage conditions, and source (e.g. the facility from which it was obtained);
- Lot number and origin of FBC, supplier name, lot number of 96-well flat-bottom black plate, and lot number of Tripluc reagent;
- Passage number and cell density used for testing;
- Cell counting method used for seeding prior to testing and measures taken to ensure homogeneous cell number distribution;
- Luminometer used (e.g. model), including instrument settings, luciferase substrate used, and demonstration of appropriate luminescence measurements based on the control test described in Appendix II;
- A statement attesting to the proficiency of the laboratory in performing the test method (e.g., by testing of proficiency substances) or to demonstrate reproducible performance of the test method over time. The Testing Facility shall maintain documentation of the procedures and proficiency chemicals used to demonstrate proficiency.

Test procedure

- Number of replicates and runs performed;
- Test chemical concentrations, application procedure and exposure time (if different from those recommended);
- Description of evaluation and decision criteria used;
- Description of study acceptance criteria used;
- Description of any modifications of the test procedure.

© OECD (2026)

Results

- Measurements of IL8LA and GAPLA;
- Calculations for nIL8LA, Ind-IL8LA, and Inh-GAPLA;
- The 95% confidence interval of Ind-IL8LA;
- A graph depicting dose-response curves for induction of luciferase activity and viability;
- Description of any other relevant observations, if applicable.

Discussion of the results

- Discussion of the results obtained with the IL-8 Luc assay;
- Consideration of the assay results in the context of an IATA, if other relevant information is available.

Conclusion

© OECD (2026)

LITERATURE

1. Toebak, M.J., Pohlmann, P.R., Sampat-Sardjoepersad, S.C., von Blomberg, B.M., Bruynzeel, D.P., Scheper, R.J., Rustemeyer, T., Gibbs, S. (2006), CXCL8 secretion by dendritic cells predicts contact allergens from irritants. *Toxicol In Vitro* 20, 117-124, 10.1016/j.tiv.2005.06.039
2. Takahashi, T., Kimura, Y., Saito, R., et al. (2011), An in vitro test to screen skin sensitizers using a stable THP-1-derived IL-8 reporter cell line, THP-G8. *Toxicological sciences : an official journal of the Society of Toxicology* 124: 359-369, 10.1093/toxsci/kfr237
3. Kimura, Y., Fujimura, C., Ito, Y., et al. (2015), Optimization of the IL-8 Luc assay as an in vitro test for skin sensitization. *Toxicol In Vitro* 29: 1816-1830, 10.1016/j.tiv.2015.07.006
4. Kimura, Y., Watanabe, M., Suzuki, N., et al. (2018), The performance of an in vitro skin sensitisation test, IL-8 Luc assay (OECD442E), and the integrated approach with direct peptide reactive assay (DPRA). *J Toxicol Sci* 43: 741-749, 10.2131/jts.43.741
5. OECD (2023), *Second edition of the Validation report for the international validation study on the IL-8 Luc assay as a test evaluating the skin sensitizing potential of chemicals conducted by the IL-8 Luc Assay*, Series on Testing and Assessment, No. 267. OECD Environment, Health and Safety, Paris. Available at: https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/publications-on-testing-and-assessment-of-chemicals.html .
6. OECD (2017), *Report of the Peer Review Panel for the IL-8 Luciferase (IL-8 Luc) Assay for in vitro skin sensitisation*. Series on Testing and Assessment, No. 268. OECD Environment, Health and Safety, Available at: https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/publications-on-testing-and-assessment-of-chemicals.html .
7. Urbisch, D., Mehling, A., Guth, K., et al. (2015), Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul Toxicol Pharmacol* 71: 337-351, 10.1016/j.yrtph.2014.12.008
8. Kimura, Y., Fujimura, C., Aiba, S. (2021), The modified IL-8 Luc assay, an in vitro skin sensitisation test, can significantly improve the false-negative judgment of lipophilic sensitizers with logKow values > 3.5. *Arch Toxicol* 95: 749-758, 10.1007/s00204-020-02934-9
9. Ashikaga, T., Sakaguchi, H., Sono, S., et al. (2010), A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Alternatives to laboratory animals : ATLA* 38: 275-284
10. Patlewicz, G., Casati, S., Basketter, D.A., et al. (2016), Can currently available non-animal methods detect pre and pro-haptens relevant for skin sensitization? *Regul Toxicol Pharmacol* 82: 147-155, 10.1016/j.yrtph.2016.08.007
11. Thorne, N., Inglese, J., Auld, D.S. (2010), Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology. *Chem Biol* 17: 646-657, 10.1016/j.chembiol.2010.05.012
12. OECD (2021*), Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists,* OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264265295-en.
13. Viviani, V., Uchida, A., Suenaga, N., et al. (2001), Thr226 is a key residue for bioluminescence spectra determination in beetle luciferases. Biochem Biophys Res Commun 280: 1286-1291, 10.1006/bbrc.2001.4254
14. Viviani, V.R., Bechara, E.J., Ohmiya, Y. (1999), Cloning, sequence analysis, and expression of active Phrixothrix railroad-worms luciferases: relationship between bioluminescence spectra and primary structures. Biochemistry 38: 8271-8279, 10.1021/bi9900830
15. Nakajima, Y., Kimura, T., Sugata, K., et al. (2005), Multicolor luciferase assay system: one-step monitoring of multiple gene expressions with a single substrate. *Biotechniques* 38: 891-894, 05386ST03 [pii]
16. Edwards, D.R., Denhardt, D.T. (1985), A study of mitochondrial and nuclear transcription with cloned

© OECD (2026)

cDNA probes. Changes in the relative abundance of mitochondrial transcripts after stimulation of quiescent mouse fibroblasts. *Exp Cell Res* 157: 127-143,

17. Mori, R., Wang, Q., Danenberg, K.D., et al. (2008), Both beta-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer. *Prostate* 68: 1555-1560, 10.1002/pros.20815
18. Winer, J., Jung, C.K., Shackel, I., et al. (1999), Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes in vitro. Anal Biochem 270: 41-49, 10.1006/abio.1999.4085
19. OECD (2025), *Guidance Document on Good In Vitro Method Practices (GIVIMP), Second Edition*, OECD Series on Testing and Assessment, No. 421, OECD Publishing, Paris, https://doi.org/10.1787/5ba6777b-en.
20. United Nations UN (2025). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Eleventh revised edition. New York & Geneva: United Nations Publications. ISBN: 9789211067699. Available at: [https://doi.org/10.18356/f8fbb7cb-en].

© OECD (2026)

APPENDIX I: 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.

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.

CV05: Cell viability 05. Minimum concentration at which chemicals show less than 0.05 of Inh-GAPLA.

FInSLO-LA: Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to Ind-IL8LA. See Ind-IL8LA for definition.

GAPLA: Luciferase Activity of Stable Luciferase Red (SLR) (λmax = 630 nm), regulated by GAPDH promoter and demonstrates cell viability and viable cell number.

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 weights all relevant data to inform regulatory decision regarding potential hazard and/or risk and/or the need for further targeted and therefore minimal testing.

II-SLR-LA: Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to Inh-GAPLA. See Inh-GAPLA for definition

IL-8 (Interleukin-8): A cytokine derived from endothelial cells, fibroblasts, keratinocytes, macrophages, and monocytes that causes chemotaxis of neutrophils and T-cell lymphocytes.

IL8LA: Luciferase Activity of Stable Luciferase Orange (SLO) (λmax = 580 nm), regulated by IL-8 promoter.

Ind-IL8LA: Fold induction of IL8LA. It is obtained by dividing the nIL8LA of THP-G8 cells treated with chemicals by that of non-stimulated THP-G8 cells and represents the induction of IL-8 promoter activity by chemicals.

Inh-GAPLA: Inhibition of GAPLA. It is obtained by dividing GAPLA of THP-G8 treated with chemicals with GAPLA of non-treated THP-G8 and represents cytotoxicity of chemicals.

Minimum induction threshold (MIT): the lowest concentration at which a chemical satisfies the positive criteria

Mixture: 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).

Multi-constituent substance: A substance, defined by its quantitative composition, in which more than one of the main constituents 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

© OECD (2026)

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.

nIL8LA: The SLO luciferase activity reflecting IL-8 promoter activity (IL8LA) normalised by the SLR luciferase activity reflecting GAPDH promoter activity (GALPA). It represents IL-8 promoter activity after considering cell viability or cell number.

nSLO-LA: Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to nIL8LA. See nIL8LA for definition

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.

Pre-haptens: Chemicals which become sensitisers through abiotic transformation.

Pro-haptens: Chemicals requiring enzymatic activation to exert skin sensitisation potential.

Propidium-iodide (PI) exclusion cells: The fluorescent marker propidium-iodide (PI) is not taken up by living cells. Therefore, in fluorescence staining and flow cytometry, cells that are not stained with PI are considered live cells and cells that are stained are considered dead cells.

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.

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.

Run: A run consists of one or more test chemicals tested concurrently with a solvent/vehicle control and with a positive control.

Sensitivity: The proportion of all positive/active 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.

SLO-LA: Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to IL8LA. See IL8LA for definition.

SLR-LA: Abbreviation used in the validation report and in previous publications regarding the IL-8 Luc assay to refer to GAPLA. See GAPLA for definition.

Solvent/vehicle control: An untreated sample containing all components of a test system except of the test chemical, but including the solvent/vehicle that is used. It is used to establish the baseline response for the samples treated with the test chemical dissolved or stably dispersed in the same solvent/vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent/vehicle interacts with the test system.

Specificity: The proportion of all negative/inactive 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.

© OECD (2026)

Substance: Chemical elements and their compounds in the natural state or obtained by any production process, inducing 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 it composition.

Surfactant: Also called surface-active agent, this is a substance, such as a detergent, that can reduce the surface tension of a liquid and thus allow it to foam or penetrate solids; it is also known as a wetting agent. (TG437)

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

THP-G8: An IL-8 reporter cell line used in IL-8 Luc assay. The human macrophage-like cell line THP-1 was transfected the SLO and SLR luciferase genes under the control of the IL-8 and GAPDH promoters, respectively.

United Nations Globally Harmonized System of Classification and Labeling 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.

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

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.

© OECD (2026)

APPENDIX II: PRINCIPLE OF MEASUREMENT OF LUCIFERASE ACTIVITY AND DETERMINATION OF THE TRANSMISSION COEFFICIENTS OF OPTICAL FILTER FOR SLO AND SLR

MultiReporter Assay System -Tripluc- can be used with a microplate-type luminometer with a multi-colour detection system, which can equip an optical filter (e.g. Phelios AB-2350 (ATTO), ARVO (PerkinElmer), Tristar LB941 (Berthold)). The optical filter used in measurement is 600–620 nm long or short pass filter, or 600–700 nm band pass filter.

(1) Measurement of two-colour luciferases with an optical filter.

This is an example using Phelios AB-2350 (ATTO). This luminometer is equipped with a 600 nm long pass filter (R60 HOYA Co.), 600 nm LP, Filter 1) for splitting SLO (λmax = 580 nm) and SLR (λmax = 630 nm) luminescence.

To determine transmission coefficients of the 600 nm LP, first, using purified SLO and SLR luciferase enzymes, measure i) the intensity of SLO and SLR bioluminescence intensity without filter (F0), ii) the SLO and SLR bioluminescence intensity that passed through 600 nm LP (Filter 1), and iii) calculate the transmission coefficients of 600 nm LP for SLO and SLR listed below.

Transmission Coefficients

Transmission coefficients		Abbreviation	Definition
SLO	Filter 1 Transmission coefficients	κO_{R60}	The filter's transmission coefficient for the SLO
SLR	Filter 1 Transmission coefficients	κR_{R60}	The filter's transmission coefficient for the SLR

When the intensity of SLO and SLR in test sample are defined as O and R, respectively, i) the intensity of light without filter (all optical) F0 and ii) the intensity of light that transmits through 600 nm LP (Filter 1) F1 are described as below.

$$F0=O+R$$

$$F1=\kappa O_{R60} \times O + \kappa R_{R60} \times R$$

These formulas can be rephrased as follows:

$$\begin{pmatrix} F0 \\ F1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \kappa O_{R60} & \kappa R_{R60} \end{pmatrix} \begin{pmatrix} O \\ R \end{pmatrix}$$

Then using calculated transmittance factors (κO_{R60} and κR_{R60}) and measured F0 and F1, you can calculate O and R-value as follows:

$$\begin{pmatrix} O \\ R \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \kappa O_{R60} & \kappa R_{R60} \end{pmatrix}^{-1} \begin{pmatrix} F0 \\ F1 \end{pmatrix}$$

© OECD (2026)

Materials and methods for determining transmission coefficients

(1) Reagents

· Single purified luciferase enzymes:

Lyophilised purified SLO enzyme

Lyophilised purified SLR enzyme

(which for the validation work were obtained from Tottori Bioscience Promotion Organization, Tottori, Japan with THP-G8 cell line)

· Assay reagent:

Tripluc® Luciferase assay reagent (for example from TOYOBO Cat#MRA-301)

· Medium: for luciferase assay (30 ml, stored at 2 – 8°C)

Reagents Concentration and Required Amount

Reagent	Conc.	Final conc. in medium	Required amount
RPMI-1640	-	-	27 ml
FBS	-	10 %	3 ml

(2) Preparation of enzyme solution

Dissolve lyophilised purified luciferase enzyme in tube by adding 200 µl of 10 ~ 100 mM Tris/HCl or Hepes/HCl (pH 7.5 ~ 8.0) supplemented with 10% (w/v) glycerol, divide the enzyme solution into 10 µl aliquots in 1.5 ml disposable tubes and store them in a freezer at -80°C. The frozen enzyme solution can be used for up to 6 months. When used, add 1 ml of medium for luciferase assay (RPMI-1640 with 10% FBS) to each tube containing the enzyme solutions (diluted enzyme solution) and keep them on ice to prevent deactivation.

(3) Bioluminescence measurement

Thaw Tripluc® Luciferase assay reagent (Tripluc) and keep it at room temperature either in a water bath or at ambient air temperature. Power on the luminometer 30 min before starting the measurement to allow the photomultiplier to stabilise. Transfer 100 µl of the diluted enzyme solution to a black 96 well plate (flat bottom) (the SLO reference sample to #B1, #B2, #B3, the SLR reference sample to #D1, #D2, #D3). Then, transfer 100 µl of pre-warmed Tripluc to each well of the plate containing the diluted enzyme solution using a pipetman. Shake the plate for 10 min at room temperature (about 25°C) using a plate shaker. Remove bubbles from the solutions in wells if they appear. Place the plate in the luminometer to measure the luciferase activity. Bioluminescence is measured for 3 sec each in the absence (F0) and presence (F1) of the optical filter.

Transmission coefficient of the optical filter was calculated as follows:

Transmission coefficient (SLO (κO_{R60}))= (#B1 of F1+ #B2 of F1+ #B3 of F1) / (#B1 of F0+ #B2 of F0+ #B3 of F0)

Transmission coefficient (SLR (κR_{R60}))= (#D1 of F1+ #D2 of F1+ #D3 of F1) / (#D1 of F0+ #D2 of F0+ #D3 of F0)

© OECD (2026)

Calculated transmittance factors are used for all the measurements executed using the same luminometer.

Quality control of equipment

The procedures described in the IL-8 Luc protocol should be used (5)

© OECD (2026)

APPENDIX III: PROFICIENCY SUBSTANCES

Prior to routine use of the test method described in this Annex to Test Guideline 442E, laboratories should demonstrate technical proficiency by obtaining the expected IL-8 Luc assay prediction for the 9 substances recommended in Table 1 and by obtaining values that fall within the respective reference range for at least 8 out of the 9 proficiency substances (selected to represent the range of responses for skin sensitisation hazards). Other selection criteria were that the substances are commercially available, and that high-quality *in vivo* reference data as well as high quality *in vitro* data generated with the IL-8 Luc assay are available. Also, published reference data are available for the IL-8 Luc assay (6) (1).

Table 1. Recommended substances for demonstrating technical proficiency with the IL-8 Luc assay

Proficiency substances	CAS no.	State	Solubility in X-VIVO15 at 20 mg/mL	*In vivo* prediction[1]	IL-8 Luc prediction[2]	Reference range (µg/mL)[3]	
						CV05[4]	IL-8 Luc MIT[5]
2,4-Dinitrochlorobenzene	97-00-7	Solid	Insoluble[6]	Sensitiser (Extreme)	Positive	2.3-3.9	0.5-2.3
Formaldehyde	50-00-0	Liquid	Soluble	Sensitiser (Strong)	Positive	9-30	4-9
2-Mercaptobenzothiazole	149-30-4	Solid	Insoluble[6]	Sensitiser (Moderate)	Positive	250-290	60-250
Ethylenediamine	107-15-3	Liquid	Soluble	Sensitiser (Moderate)	Positive	500-700	0.1-0.4
Ethyleneglycol dimethacrylate	97-90-5	Liquid	Insoluble[6]	Sensitiser (Weak)	Positive	>2000	0.04-0.1
Citral	5392-40-5	Liquid	Insoluble[6]	Sensitiser (Weak)	Positive	12-30	4-12
Streptomycin sulphate	3810-74-0	Solid	Soluble	Non-sensitiser	Negative	>2000	>2000
Glycerol	56-81-5	Liquid	Soluble	Non-sensitiser	Negative	>2000	>2000
Isopropanol	67-63-0	Liquid	Soluble	Non-sensitiser	Negative	>2000	>2000

Abbreviations: CAS no. = Chemical Abstracts Service Registry Number

[1] The *in vivo* potency is derived using the criteria proposed by ECETOC (20).

[2] Based on historical observed values (2) (9).

[3] CV05 and IL-8 Luc MIT were calculated using water solubility given by EPI Suite™.

[4] CV05: the minimum concentration at which chemicals show less than 0.05 of Inh-GAPLA.[5] MIT: the lowest concentrations at which a chemical satisfies the positive criteria.

[6] Insoluble or not fully soluble

© OECD (2026)

APPENDIX IV: INDEXES AND JUDGMENT CRITERIA

nIL8LA (nSLO-LA)

The j-th repetition (j = 1-4) of the i-th concentration (i = 0-11) is measured for IL8LA (SLO-LA) and GAPLA (SLR-LA) respectively. The normalised IL8LA, referred to as nIL8LA (nSLO-LA), and is defined as:

$nIL8LA_{ij} = IL8LA_{ij} / GAPLA_{ij}$.

This is the basic unit of measurement in this assay.

Ind-IL8LA (FInSLO-LA)

The fold increase of the averaged nIL8LA (nSLO-LA) for the repetition on the i-th concentration compared with it at the 0 concentration, Ind-IL8LA, is the primary measure of this assay. This ratio is written by the following formula:

$$\text{Ind-IL8LA}_i = \left\{(1/4)\times\sum_j nIL8LA_{ij}\right\} / \left\{(1/4)\times\sum_j nIL8LA_{0j}\right\}.$$

The lead laboratory has proposed that a value of 1.4 corresponds to a positive result for the tested chemical. This value is based on the investigation of the historical data of the lead laboratory. Data management team then used this value through all the phases of validation study. The primary outcome, Ind-IL8LA, is the ratio of 2 arithmetic means as shown in equation.

95% confidence interval (95% CI)

The 95% confidence interval (95% CI) based on the ratio can be estimated to show the precision of this primary outcome measure. The lower limit of the 95% CI ≥ 1 indicates that the nIL8LA with the i-th concentration is significantly greater than that with solvent control. There are several ways to construct the 95% CI. We used the method known as Fieller's theorem in this study. This 95% confidence interval theorem is obtained from the following formula:

© OECD (2026)

$$\left[\frac{-B-\sqrt{B^2-4AC}}{2A}, \frac{-B+\sqrt{B^2-4AC}}{2A}\right],$$

where $A=\bar{x}_0^2-t_{0.975(\nu)}^2\times\frac{sd_0^2}{n_0}$, $B=-2\times\bar{x}\times\bar{y}$, $C=\bar{y}_i^2-t_{0.975(\nu)}^2\times\frac{sd_{y_i}^2}{n_{y_i}}$, and n_0 =4,

$\bar{x}_0=(1/n_0)\times\sum_j nIL8LA_{0j}$, $sd_0^2=\{1/(n_0-1)\}\times\sum_j(nIL8LA_{0j}-\bar{x}_0)^2$,

n_{y_i} =4, $\bar{y}_i=(1/n_{y_i})\times\sum_j(nIL8LA_{ij})$, $sd_{y_i}^2=\{1/(n_{y_j}-1)\}\times\sum_j(nIL8LA_{ij}-\bar{y}_i)^2$.

$t_{0.975(\nu)}$ is 97.5 percentile of the central t distribution with the ν of the degree of freedom, where

$$\nu=\left(\frac{sd_0^2}{n_0}+\frac{sd_{y_i}^2}{n_{y_i}}\right)^2/\left\{\left(\frac{sd_0^2}{n_0}\right)^2/(n_0-1)+\left(\frac{sd_{y_i}^2}{n_{y_i}}\right)^2/(n_{y_i}-1)\right\}.$$

Inh-GAPLA (II-SLR-LA)

The Inh-GAPLA is a ratio of the averaged GAPLA (SLR-LA) for the repetition of the i-th concentration compared with that with solvent control, and this is written by

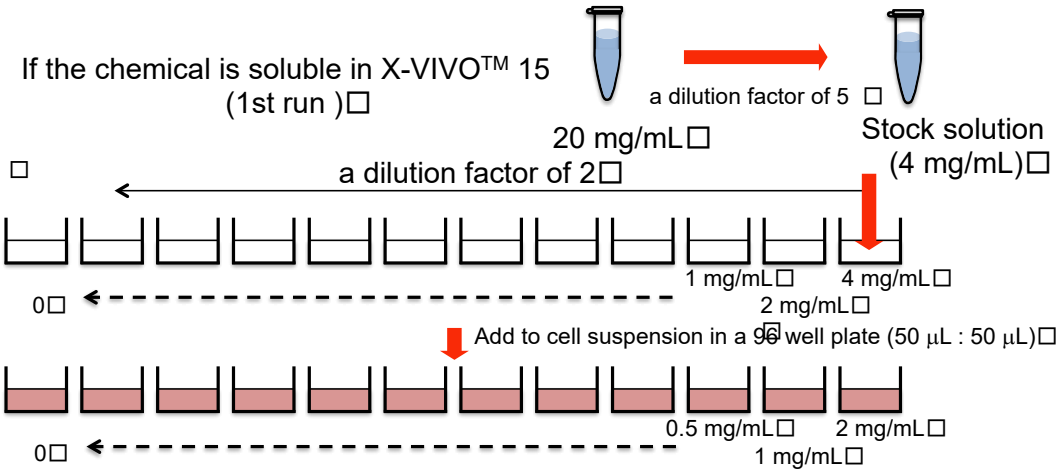
Inh-GAPLA $_i$ = $\{(1/4)\times\sum_j GAPLA_{ij}\}/\{(1/4)\times\sum_j GAPLA_{0j}\}$.

Since the GAPLA is the denominator of the nIL8LA, an extremely small value causes large variation in the nIL8LA. Therefore, Ind-IL8LA values with an extremely small value of Inh-GAPLA (less than 0.05) might be considered poor precision.

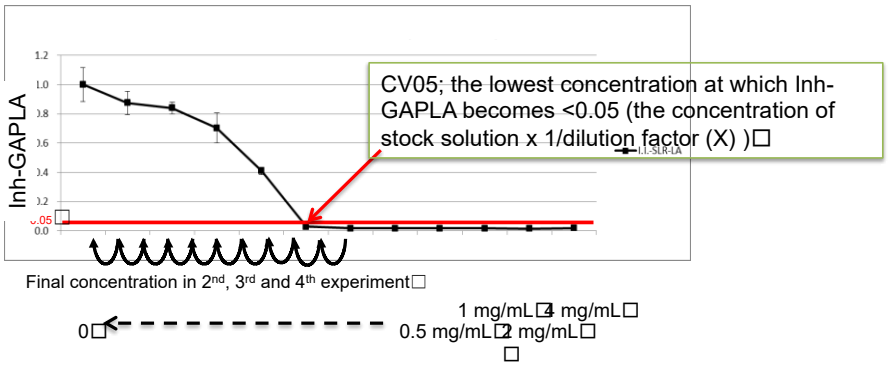
© OECD (2026)

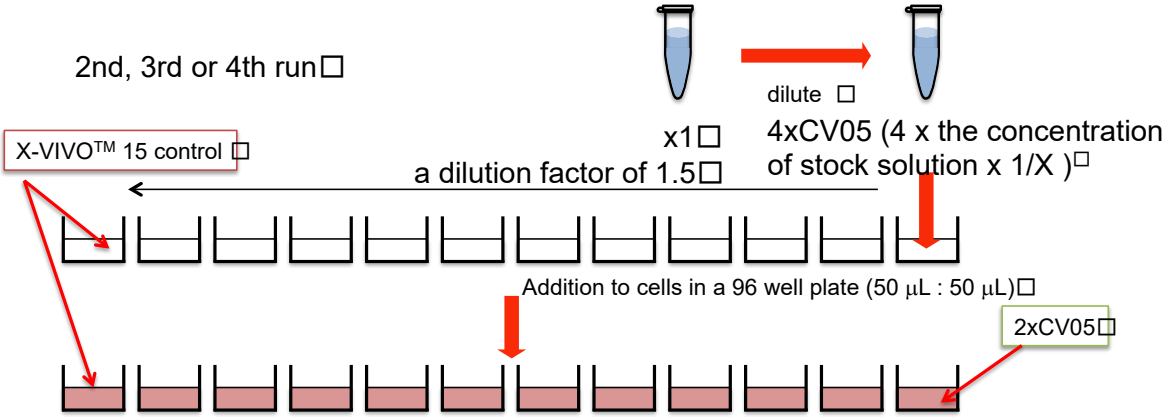
APPENDIX V: THE SCHEME OF THE METHODS TO DISSOLVE CHEMICALS FOR THE IL-8 LUC ASSAY.

a) For chemicals dissolved in X-VIVO™ 15 at 20 mg/mL



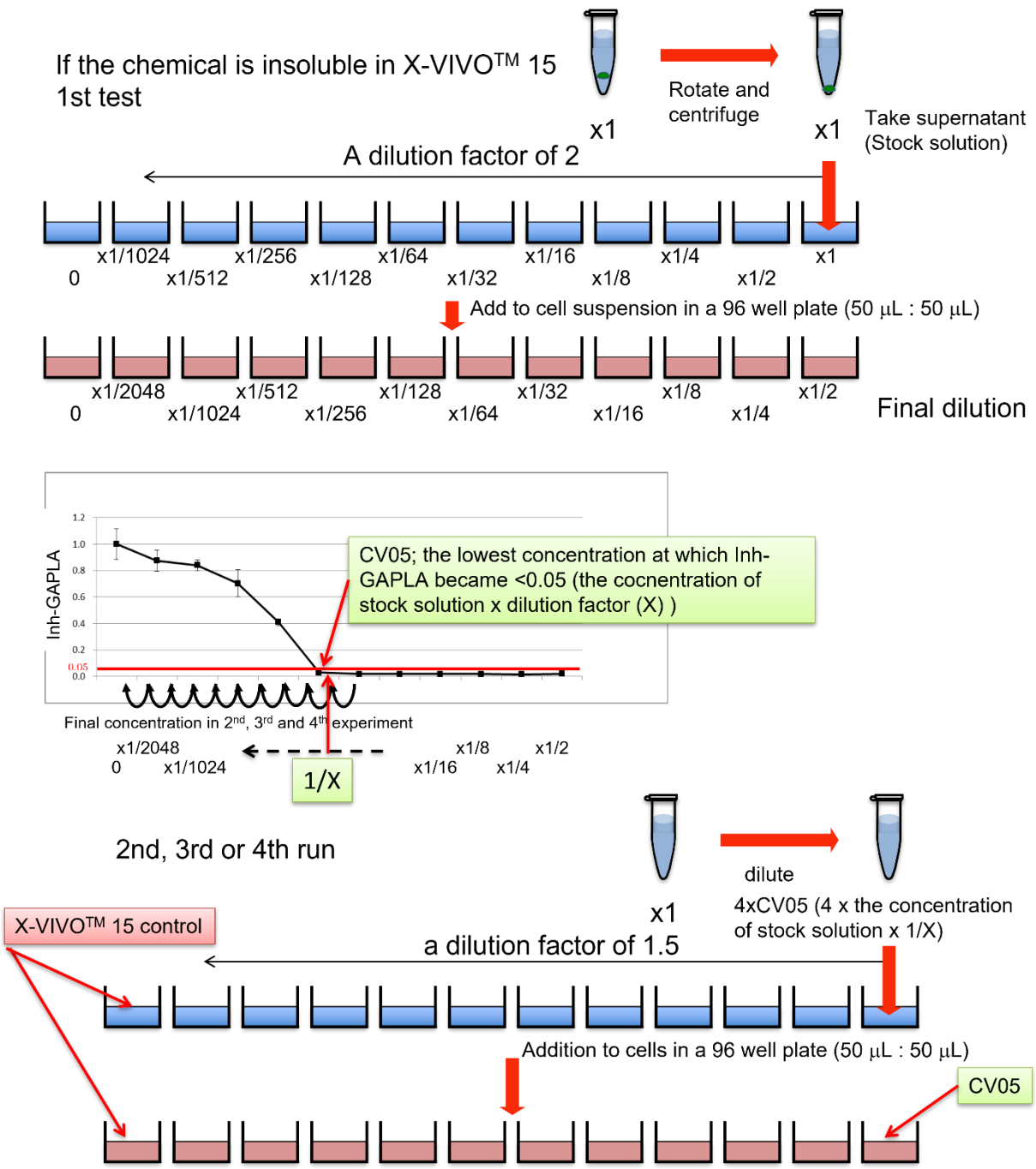
Determine the highest concentration of the following experiments





b) For chemicals insoluble in X-VIVO™ 15 at 20 mg/mL

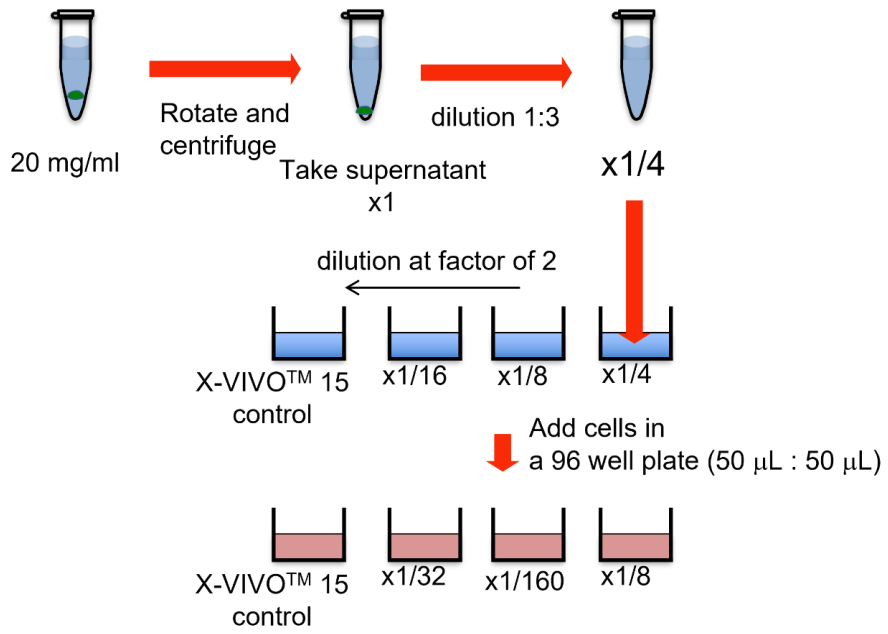
© OECD (2026)


If the chemical is insoluble in X-VIVO™ 15
1st test
x1
Rotate and centrifuge
x1
Take supernatant (Stock solution)
A dilution factor of 2
x1/1024
x1/256
x1/64
x1/16
x1/4
x1
0
x1/512
x1/128
x1/32
x1/8
x1/2
Add to cell suspension in a 96 well plate (50 µL : 50 µL)
x1/2048
x1/512
x1/128
x1/32
x1/8
x1/2
0
x1/1024
x1/256
x1/64
x1/16
x1/4
Final dilution
Inh-GAPLA
1.2
1.0
0.8
0.6
0.4
0.2
0.05
0.0
CV05; the lowest concentration at which Inh-GAPLA became <0.05 (the cocnentration of stock solution x dilution factor (X))
Final concentration in 2nd, 3rd and 4th experiment
x1/2048
0
x1/1024
1/X
x1/8
x1/2
x1/16
x1/4
2nd, 3rd or 4th run
x1
dilute
4xCV05 (4 x the concentration of stock solution x 1/X)
X-VIVO™ 15 control
a dilution factor of 1.5
Addition to cells in a 96 well plate (50 µL : 50 µL)
CV05

© OECD (2026)

APPENDIX VI: THE SCHEME OF THE METHOD TO DISSOLVE 4-NBB FOR THE POSITIVE CONTROL OF THE IL-8 LUC ASSAY.

The positive control : 4-NBB (not fully soluble in X-VIVO™ 15)



© OECD (2026)

ANNEX 4 - IN VITRO SKIN SENSITISATION: GENOMIC ALLERGEN RAPID DETECTION (GARD™) FOR ASSESSMENT OF SKIN SENSITISERS (GARD™skin)

INITIAL CONSIDERATIONS AND LIMITATIONS

1. The GARD™skin method provides binary hazard identification of skin sensitisers (i.e. UN GHS Category 1 versus non-sensitisers). The method evaluates the transcriptional patterns of an endpoint-specific genomic biomarker signature, referred to as the GARDskin Genomic Prediction Signature (GPS), in the SenzaCell™ cell line (1) (2), a subclone of the myeloid leukaemia cell line MUTZ-3 (3) (4) (5), exposed to test chemicals.

2. The GARDskin GPS (N genes = 196) was identified by genome-wide data-driven analysis of a discovery dataset based on the human surrogate DC-like SenzaCell cell line exposed to a panel of well-characterised skin sensitisers (UN GHS Category 1) (N=20) and non-sensitisers (N=20) (6). The GPS monitors mechanistic events associated with xenobiotic recognition, generation of immunological danger signals and DC activation, as described by KE3 of the OECD AOP. Of note, certain mechanistic events associated with the GARDskin GPS may also be associated with other KE:s, albeit monitored in a DC cell line. For further details on the origin and the biological functions of the GPS, please refer to the Supporting document to the Test Guideline for the GARDskin test method (7). The potential utilisation of the GARDskin GPS in a predictive assay was proposed (8), and the functionality was demonstrated in a GARDskin application based on the GeneChip® microarray platform (9). Following the evaluation of alternative technological platforms for targeted gene expression analysis (10), GARDskin was transferred to a NanoString nCounter® system (11) format, on which it was demonstrated to exhibit retained predictive performance, as well as improved resource effectiveness (12).

3. The GARDskin method has been evaluated in a validation study (13) (14) coordinated by SenzaGen AB and subsequently independently peer reviewed by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) Scientific Advisory Committee (ESAC) (15). Considering all available evidence, the GARDskin was recommended to be used as part of an IATA to support the discrimination between skin sensitisers and non-sensitisers for the purpose of hazard classification and labelling.

4. In addition to a review of data produced within the context of the ring trial, the ESAC also performed an in-depth review of the complete GARDskin analysis pipeline and its bioinformatical components, as hosted in the GARD Data Analysis Application (GDAA) and described in the Supporting document to the Test Guideline for the GARDskin test method (7). The ESAC was able to reproduce the prediction

© OECD (2026)

algorithm from the training dataset and verify and reproduce all steps from raw data to final classifications of test chemicals (15).

5. GARDskin was demonstrated to be transferable to naïve laboratories experienced in routine cell culture and molecular biology techniques, including flow cytometry and isolation of RNA (test chemicals N=28). The levels of within-laboratory reproducibility (WLR) for GARDskin obtained in the validating ring trial ranged between 78.6-89.2% when considering also the concordance of missing data points[3] and 82.1-88.9% when excluding missing data points. Similarly, an estimation of between-laboratory reproducibility (BLR) was calculated to 82.1% when considering also the concordance of missing data points and 92.0% when excluding missing data points (13) (14) (15).

6. Results generated in the validation study (13) (14) overall indicated that, when compared with an expert judgement-based classification reference data set, using weight of evidence incorporating human (16) and LLNA (17) data sources as summarised and presented by the ESAC (15), the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers was 91.7% (N=28) with a sensitivity of 92.4% (N=19) and a specificity of 90.1% (N=9). The balanced accuracy was 91.2%. Omitting test chemicals which overlap with the GARDskin training dataset, thereby only considering truly naïve test chemicals, the accuracy in distinguishing skin sensitisers (i.e. UN GHS Cat.1) from non-sensitisers was 95.4% (N=17) with a sensitivity of 96.6% (N=13) and a specificity of 91.7% (N=4). The balanced accuracy was 94.1%.

7. Following the submission of the GARDskin method, OECD published the guideline on Defined Approaches for Skin Sensitisation (18). With this publication, an extended dataset of chemicals with curated LLNA and human reference data became public (19). Taking this curated reference data into account, the predictive performance of GARDskin was calculated, using GARDskin data generated in the validation study (13) (14) or available in other published studies (20). Calculations were performed both when including and excluding test chemicals overlapping with chemicals used during method development, as summarised in Tables 1A and 1B, respectively. As these figures are based on imbalanced datasets, the measure of specificity should be regarded as uncertain.

Table 1A. Performance of GARDskin in comparison to LLNA or human reference data (19). Calculations include test chemicals overlapping with chemicals used during method development.

		LLNA			Human	
		NS (N=11)	S (N=64)		NS (N=9)	S (N=27)
GARDskin[1]	NS	9.89	8.16		4.94	3.68
	S	1.11	55.9		4.06	23.3
Accuracy		87.6%			78.5%	
Sensitivity		87.2%			86.4%	
Specificity		89.9%			54.9%	
Balanced accuracy		88.6%			70.7%	
N		75			36	

[1] Confusion matrices are based on weighted calculations of GARDskin results, as implemented by the ESAC (15).

[3] Chemicals that failed to generate a valid result due to failed acceptance criteria, as described in section procedures.

© OECD (2026)

Table 1B. Performance of GARDskin in comparison to LLNA or human reference data (19). Calculations exclude test chemicals overlapping with chemicals used during method development.

		LLNA			Human	
		NS (N=5)	S (N=47)		NS (N=6)	S (N=18)
GARDskin[1]	NS	4.33	4.92		2.50	2.89
	S	0.667	42.1		3.50	15.1
Accuracy		89.3%			73.4%	
Sensitivity		89.5%			84.0%	
Specificity		86.7%			41.7%	
Balanced accuracy		88.1%			62.8%	
N		52			24	

[1] Confusion matrices are based on weighted calculations of GARDskin results, as implemented by the ESAC (15).

8. Taken together, this information indicates the usefulness of the GARDskin method to contribute to the identification of skin sensitisation hazards. However, the test method should be considered in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction.

9. Known limitations of the method are mainly associated with solubility issues and compatibility with vehicles as well as the aqueous cell system. In addition, autoflourescent test chemicals may interfere with flow cytometry-based cytotoxicity assessments. The known limitations are listed in Appendix II, together with potential circumventions. The GARDskin method has been shown to be applicable to test chemicals covering a wide variety of organic functional groups, reaction mechanisms, skin sensitisation potencies and physicochemical properties (12) (14) (20). On the basis of currently available data, there are no specific classes and/or types of chemistries excluded from the applicability domain. Similar conclusions were drawn from an independent expert/expert systems review of reported data, which was also included in the data package submitted for ESAC peer-review (21). Of particular note, method applicability and predictive performance are maintained in certain chemical space subsets that are otherwise considered inherently difficult to accurately assess. This includes e.g. lipophilic compounds (logP > 3.5) (20), indirectly acting haptens (20) and metal compounds (22).

10. The GARDskin method was validated for assessment of mono-constituent chemical substances. Although not evaluated in the validation studies, the test method is nevertheless technically applicable to the testing of multi-constituent substances and mixtures (20) (23). Definitions are provided in Appendix I.

DEMONSTRATION OF LABORATORY PROFICIENCY

11. Prior to routine use of GARDskin, laboratories should demonstrate technical proficiency in conducting the test method. Proficiency is demonstrated by testing of a specified set of proficiency chemicals with known sensitising properties, as listed in Appendix III. This testing will also confirm the responsiveness of the test system. Testing of the proficiency chemicals should be carried out in full adherence to the herein described procedure, and the results should be consistent with the listed classifications in Appendix III. A historical database of data generated with the proficiency chemicals shall be maintained at the test facility to confirm the reproducibility of the test method over time.

© OECD (2026)

PRINCIPLE OF THE TEST METHOD

12. The GARDskin method utilises the SenzaCell cell line, a subclone of the myeloid leukaemia cell line MUTZ-3, as an *in vitro* surrogate model of DC. Following test chemical exposure, at test chemical-specific exposure concentrations for 24 h, the quantifiable readout of the assay is the gene expression levels of the GARDskin GPS, obtained from measurements of isolated total RNA from exposed cell cultures, and assessed by the NanoString nCounter® system.

13. The high-dimensional data is analysed using the GDAA, hosting a Support Vector Machine (SVM) prediction algorithm (24), appropriately trained and frozen during assay development (12). Based on obtained gene expression levels in cell cultures exposed to test chemicals, the output from the GARDskin prediction algorithm predicts each test chemical as being a skin sensitiser (UN GHS Category 1) or a non-sensitiser.

CLOUD-BASED SOFTWARE

14. The GARDskin data analysis pipeline is based on a cloud-based and version-controlled software referred to as the GDAA, which facilitates the entire data analysis-workflow, from raw-data preprocessing to final classification of test chemicals. The GDAA is designed to ensure data integrity in accordance with published guidance (25) (26).

15. Test facilities should periodically and/or before use (based on a risk assessment), check all functions of the cloud-based GDAA software (25) (26). A historic reference test dataset should therefore be uploaded to the cloud-based system and processed/analysed by the software. The test chemicals of this reference test dataset are specified by the test facility, but each test chemical should generate exact and reproducible test results over time, in terms of generated decision values and Message-Digest algorithm 5 (MD5) checksums (27). For further details and explanations of decision values and MD5 checksums, see section *Data Analysis and Reporting* of this TG. The results from this periodic and before use testing of the software shall be documented, tracking the stability of the (computerised) system over time.

PROCEDURE

16. The *GARDskin Assay Protocol* is publicly available in the Tracking System for Alternative methods towards Regulatory acceptance (TSAR) (28). The protocol should be used when implementing the GARDskin method in the laboratory. The following paragraphs provide a description of the main components and procedures of the GARDskin test method and a graphical outline of the consecutive steps of the GARDskin method is presented in Figure 1.

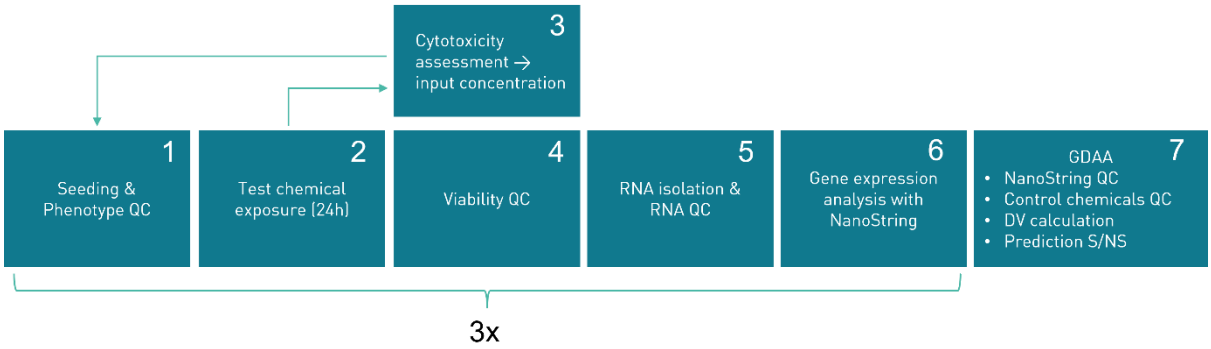
17. In the GARDskin method, chemical exposures are performed in two subsequent types of experiments. In a first step, cytotoxicity assessment experiments are performed to identify a suitable and test chemical-specific exposure concentration derived from the cytotoxic properties of the test chemical, referred to as the GARD input concentration. In a second step, main stimulation experiments are performed using the previously defined GARD input concentration in order to harvest RNA for downstream analysis.

18. Three independent and biologically replicate main stimulations shall be performed. Within the context of this description of the GARDskin procedure, independent and biologically replicate experiments are defined as identical experiments being performed using i) separate cell cultures (i.e. *cell batches*), and ii) separate and independent preparations of test chemicals and controls.

© OECD (2026)

19. The endpoint measurement of GARDskin, i.e. the quantification of the GARDskin GPS mRNA transcripts, is performed using the NanoString nCounter analysis system, using a CodeSet comprising probes corresponding to the genes of the GARDskin GPS. Generated raw data of gene expression levels are analysed using the GARD Data Analysis Application (GDAA), and each test chemical is classified by the GARDskin prediction model as either a sensitiser or a non-sensitiser.

Figure 1. Graphical outline of the consecutive steps of the GARDskin procedure.



Note: A cytotoxicity assessment experiment consists of the sequential combination of element 1-4, whereas a main stimulation experiment consists of the sequential combination of element 1, 2, 4-6. Following the completion of three independent main stimulation experiments, all data analysis is performed using the GDAA, as outlined in element 7. QC: Quality Control. DV: Decision Value.

Cells

20. The human myeloid leukemia-derived cell line, SenzaCell, should be used in the GARDskin method. The SenzaCell cell line is made available from SenzaGen AB[4], following appropriate licensing of the GARD technology. The SenzaCell cell line is provided on dry ice and should be stored at <-136°C in accordance with the Guidance Document on Good In Vitro Method Practices (GIVIMP) (29). The SenzaCell cell line should be expanded and frozen in liquid nitrogen at a concentration of $7x10^6$ cells / mL in cell medium supplemented with 10% v/v DMSO (molecular biology grade, ≥99%).

21. Cell work should be performed under sterile conditions, free of antibiotics. Centrifugations with the SenzaCell cell line should be performed at 300-315xg, 5 min, 2-8°C. Incubation of the SenzaCell cell line should be performed at 37°C under 5% CO_2 and humidified atmosphere. The SenzaCell cell line should be grown in cell culture flasks for maintenance and expansion, or cell culture plates for chemical exposure. The SenzaCell cell line should be grown in MEM/Alpha medium (with L-Glutamine, with Ribo- and Deoxyribonucleosides) supplemented with 20% (v/v) Fetal Bovine Serum (FBS) and 40 ng/mL GM-CSF (Premium grade, purity >97%, endotoxin level <0.1 EU/µg cytokine, and activity of $\geq 5x10^6$ IU/mg). Cell cultures should be counted and split to a concentration of $0.2x10^6$ cells / mL every 3-4 days for maximum 16 cell passages after thawing. The cells should be seeded for test chemical exposure directly following a cell split, i.e. test chemical exposure experiments should be scheduled to coincide with routine cell culture maintenance. Cells are seeded in flat-bottomed 12-well or 24-well plates, with a final total well volume of

[4] SenzaGen AB
Medicon Village
SE-223 81 Lund, Sweden
info@senzagen.com

© OECD (2026)

4 ml and 2 ml, respectively. Other types and plate sizes may be used if equivalent and reproducible results can be demonstrated. Cytotoxicity assessment experiments should be performed at passage number 4 to 16 and main stimulation experiments at passage number 6 to 12.

Procedures for Phenotypic Quality Control and acceptance criteria

22. The same day as performing a chemical exposure experiment, the phenotype of the untreated cells should be evaluated. This is done to ensure that cells are maintained in an inactivated state and to detect phenotypic drift.

23. All washing steps in the flow cytometer analysis should be performed in wash buffer, i.e. PBS with 0.5-1% (w/w) BSA (Cohn fraction V), 0.2 µm filter sterilisation is required. Cells should be stained with labeled monoclonal antibodies towards human antigens CD1a, CD14, CD34, CD54, CD80, CD86 and HLA-DR, as well as with relevant polyclonal isotype controls. Recommended antibodies include the following fluorescein isothiocyanate (FITC)-labelled antibodies: anti-CD86 (BD Biosciences, #555657), anti-HLA-DR (BD Biosciences, #347400), anti-CD34 (BD Biosciences, #555821), anti-CD1a (Agilent Dako, #F714101-2) and mouse polyclonal anti-IgG1-FITC (BD Biosciences, #555748). Furthermore, recommended antibodies include the following phycoerythrin (PE)-labelled antibodies: anti-CD54 (BD Biosciences, #555511), anti-CD14 (Agilent Dako, #R086401-2), anti-CD80 (BD Biosciences, #340294), and mouse polyclonal anti-IgG1-PE (BD Biosciences, #555749). In addition, Propidium Iodide, 50 µg/mL (PI) (BD Biosciences, #556463) is used for cell viability analysis. However, equivalent antibodies and viability markers may be used, provided their functional similarities can be demonstrated and documented. Note that each new lot of antibodies requires titration using the SenzaCell cell line to determine antibody concentration giving saturation. The staining of the antibodies may preferably be done by pairwise staining, using one FITC-labelled and one PE-labelled antibody per sample. In each staining sample, ~ 0.2 $x10^6$ cells are washed twice before staining. After incubation, the stained cells are washed again and resuspended in wash buffer.

24. The samples are analysed with a flow cytometer (with capability to detect PE and FITC, as applicable based on choice of antibodies) and a minimum of 10 000 events should be recorded. Gating analysis can be performed on the flow cytometer software or other related analysis software, according to instructions by the provider. For details on gating procedures and quantification of cell surface expression of phenotypic biomarkers, please refer to the GARDskin assay protocol (28).

25. Generated results shall meet the acceptance criteria listed in Table 2. If any biomarker is out of the specified ranges, the cell batch should not be used for the purpose of chemical exposure experiments and the properties of the used antibodies may need to be verified in separate assessments.

© OECD (2026)

Table 2. Acceptance criteria of the viability and phenotypic quality control

Parameter	Acceptance criteria (%)[1]
Phenotypic Biomarker	
CD86	10-40
CD54	+ (>90)
HLA-DR	+ (>60)
CD80	<10
CD34	+ (35-70)
CD14	+ (5-50)
CD1a	+ (10-60)
Viability stain	
PI negative cells (absolute viability)	≥84.5

[1] "+" indicates the presence of positive cells (>0 %). An entirely positive cell population is not required. Numbers given in parentheses are expected ranges based on historical data of the developing laboratory but is not part of the acceptance criteria. As the SenzaCell cell line is known to be heterogenous, variations are expected.

Acceptance criteria of GARD Controls

26. With each GARDskin assessment, a set of controls should be analysed. The unstimulated control (i.e., cell culture medium) and the negative control (i.e., test chemical solvent) should be analysed in each cytotoxicity assessment experiment. The unstimulated control, the negative control and the positive control (i.e., *p*-Phenylenediamine, PPD, CAS# 106-50-3) should be analysed in each of the three replicate main stimulation experiments.

27. The unstimulated control is used for determination of absolute cell viability of cell batches, calculations of the relative cell viability in cytotoxicity assessment experiments and the main stimulation experiments and for normalisation purposes in the Data analysis workflow, as further described in the Supporting document to the Test Guideline for the GARDskin test method (7) and in section *Analysis of Data* below.

28. The negative control, should have a Relative viability ≥95.5 % in the cytotoxicity assessment experiment(s) and main stimulation experiments and be classified as a non-sensitiser by the GARDskin prediction model, as defined in paragraph 76, to verify that cells have not become activated in any steps of the method's experimental procedures.

29. The positive control (PPD) should have a Relative viability 84.5 - 95.4 % in the main stimulation experiments and be classified as a sensitiser by the GARDskin prediction model, as defined in paragraph 76, to demonstrate that the cells used during an experiment are responsive and can become activated upon exposure of a sensitiser.

Preparation of the test chemicals and control substances

30. The test chemical and control substances should be stored according to instructions from the sponsor or supplier to ensure stability. Preparation of the test chemical and control substance(s) should be performed on the day of cellular exposure experiments.

31. Test chemicals should be dissolved in a compatible solvent as appropriate stocks of target in-well concentration. Compatible solvents used during method validation are listed in Table 3, together with maximum target in-well concentrations for which a non-detectable impact on genome-wide gene

© OECD (2026)

expression levels have been confirmed. Corresponding in-well concentrations should be used for the negative control. The positive control is preferably dissolved in DMSO at 1000x of the target in-well concentration. Other solvents than those listed in Table 3 or direct solution in cell media may be used if method compatibility can be demonstrated and sufficient scientific rationale is provided. The solvent should not cause cell cytotoxicity and should be classified as a non-sensitiser at the proposed in-well concentration.

Table 3. List of GARDskin compatible solvents used during validation of the method.

Solvent	CASRN	Maximum in-well concentration (%)
DMSO ≥99%	67-68-5	0.1
Water[1]	-	0.1

[1] Cell culture grade.

32. The maximum target in-well concentration of any test chemical is 500 µM. For a test chemical which has no defined molecular weight, a maximum in-well concentration of 100 µg/mL is defined as default, as derived from empirical studies (23) unless a rationale for otherwise preferred concentrations can be given.

33. Solubility of the test chemical in both the selected solvent and all downstream dilutions in cell media should be ensured by a visual inspection of the solution. If required, extensive vortexing and heat (37°C) can be applied to achieve complete dissolution, as long as it can be ensured that the stability of the test chemical is not compromised by doing so. If the test chemical is not soluble to the maximum in-well concentration of 500 µM, the solvent that generates the highest test chemical in-well concentration should be used.

34. Stock solutions with an appropriate concentration should be prepared in the selected solvent, considering dilution effects and target in-well concentrations of both test chemicals and solvents. Typically, considering the solvents listed in Table 3, such a stock concentration may be prepared at 1000x target in-well concentrations. In examples below, such a stock solution is referred to as Stock A. A Stock A of a test chemical may preferably be further diluted in medium (in examples below referred to as Stock B) before adding the test chemical to the cell culture.

35. If the Stock A is poorly soluble in Stock B (typically identified as a precipitation in medium), the highest soluble concentration in Stock B is used.

36. Note that if scientifically justified and motivated by practical benefits (e.g. increased observed solubility of certain test chemicals), the dilution scheme involving Stock A and B described above may be omitted, as long as compliance with the herein described in-well cell concentration and the maximum in-well concentrations of both test chemicals and used solvents is maintained. In such instances, a direct dilution from Stock A into the well may constitute an acceptable alternative. Similarly, if the use of an alternative solvent with different limitations in regard to maximum in-well concentration is scientifically justified and proven compatible, the concentrations of both Stock A and B may differ.

Cytotoxicity assessment experiment

37. The goal of a cytotoxicity assessment experiment is to define a test chemical-specific exposure concentration, referred to as a GARD input concentration, to be used in downstream main stimulations.

© OECD (2026)

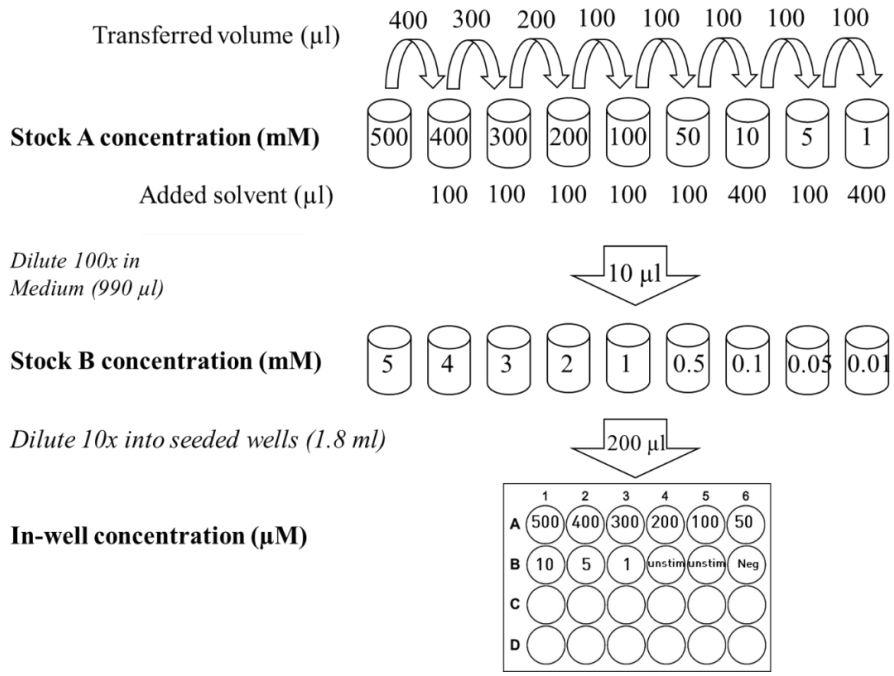
The GARD input concentration is defined based on solubility and cytotoxic properties of the test chemical, both of which are investigated in the herein described procedure.

38. For a schematic example of a typical cytotoxicity assessment experiment, see Figure 2. A serial dilution of the test chemical is performed in the selected solvent, from the default maximum in-well concentration of 500 µM (or the otherwise highest soluble concentration below 500 µM), to get a range of Stock A concentrations. Mixing and vortexing between each dilution step is recommended. From Stock A, a range of Stock B concentrations should be prepared by adding appropriate volume of Stock A to medium. Extensive vortexing and heat (37°C) can be applied as required to optimise dissolution. In addition, a Stock B concentration of the utilised solvent (negative control) in medium should be prepared to achieve the corresponding in-well concentration of the solvent.

39. Cells are seeded for chemical exposure directly following a cell split, at an appropriate cell concentration with regards to the dilution that occurs upon addition of stock solution(s) of test chemical and/or controls. The final in-well cell concentration, after addition of test chemical, should be 0.2×10^6 cells/mL.

© OECD (2026)

Figure 2. A schematic example of chemical preparations and seeding of a typical cytotoxicity assessment experiment in a 24-well plate, illustrating the serial dilution of Stock A, the conversion of Stock A:s into Stock B:s by dilution in medium and a typical plate layout following cell seeding and addition of test chemical. Note the inclusion of unstimulated and negative controls in each experiment.



40. The plate(s) with cell cultures exposed to test chemical(s) and controls are covered with plastic plate-lid(s) and incubated for 24 h at 37°C under 5% CO_2 and humidified atmosphere.

41. Following 24 h of incubation, the relative viability of exposed cell cultures (as compared to the absolute viability of unexposed cell cultures) are to be investigated. If a PI assay based on flow cytometry is used, the following procedure is recommended.

42. Each well suspension is split into duplicate flow cytometry samples. Staining and washing steps for the flow cytometry analysis are performed in wash buffer, as described in paragraph 23. Wash the cells twice and stain each sample with wash buffer and PI, 50:1, as described in paragraph 23. The samples are incubated in the dark at 2-8°C for ~15 min. The cells are washed once with ~1 mL and resuspended in an appropriate volume of wash buffer.

43. Duplicate sets of unstimulated cell cultures are required, in order to obtain technical duplicate flow cytometry samples of i) unstained unstimulated control samples and ii) stained unstimulated control samples. The unstained unstimulated control samples are used to set gates during analysis, while the stained unstimulated control samples are used for calculations of relative viability, as described below. The inclusion of duplicate unstimulated cell cultures is illustrated in Figure 2.

44. Prepared samples are analysed with a flow cytometer as described in paragraph 24.

45. Analysis of PI-stained samples should be done without any exclusion of dead cells and debris.

© OECD (2026)

46. The unstimulated unstained sample is used to set a gate for PI-positive and -negative cells, by outlining the contours of the cell population in a PE/FITC scatter plot. The PI-positive and -negative gates are then applied in the analysis of all PI-stained samples in a PE/FITC scatter plot. The percentage of PI-negative cells are recorded for each sample, representing an estimation of absolute viability. The relative viability of each sample is calculated according to Equation 1. For each test chemical concentration of the dilution range and for each control, calculate the mean value of the duplicate samples.

$$Rv = \frac{V_S}{V_C} \cdot 100 \quad (1)$$

where

Rv is the relative viability of the sample in %.
VS is the absolute viability of the sample in %.
VC is the mean absolute viability of the two PI-stained unstimulated control samples in %.

47. The controls must pass the following criteria; unstimulated control: mean absolute viability ≥84.5% and negative control: mean relative viability ≥95.5%. (Note that these criteria are also included as part of the viability control acceptance criteria following main stimulation experiments, as further detailed in section *Main stimulations* and Table 4).

48. The GARD input concentration used for main stimulations of a test chemical should be selected as follows:

i) A test chemical that induces cytotoxicity should be used at the concentration that induces 84.5%-95.4% mean relative viability. This concentration ensures bioavailability of the test chemical, while not impairing immunological responses. If multiple concentrations fulfill the acceptance criterion, the concentration that yields the Relative viability closest to 90% is chosen as the GARD input concentration. If the Relative viability decreases from ≥95.5% to <84.5% between two data points within the dilution range, repeated cytotoxicity assessment experiment(s) with additional concentrations within the critical concentration range is needed. Interpolation between data points is not recommended, as linearity cannot be assumed.

ii) A test chemical that is not cytotoxic (Relative viability ≥95.5%) should be used at a concentration of 500 µM, or at the highest soluble concentration.

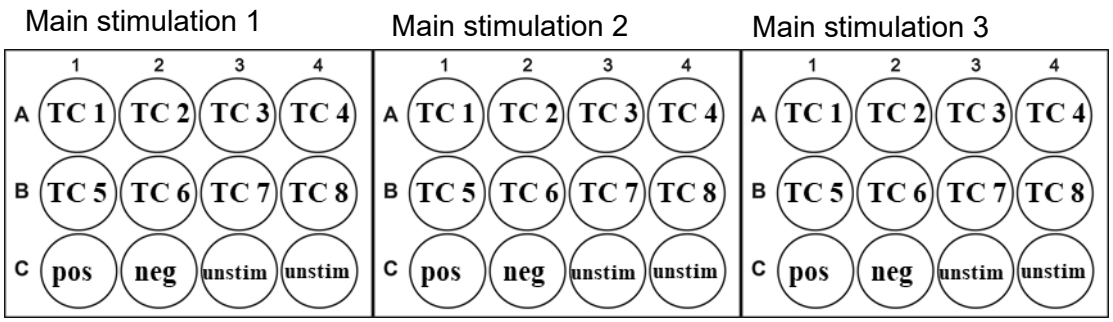
iii) A test chemical that has solubility issues in Stock A or Stock B and is not cytotoxic should be re-evaluated to control if any other solubility method, including e.g. application of heat or change of vehicle, can be used to increase the in-well concentration to get closer to the maximum in-well concentration of 500 µM.

Main stimulations

49. Once the input concentration for a test chemical is established, main stimulations are repeated in three valid independent experiments with independent preparations of the test chemical and controls (unstimulated control, negative control and positive control) and independent cell cultures originating from separate batches of cells to achieve three valid biological replicate samples. The three main stimulations can either be run in parallel or sequentially, but always with independent stock solutions of both test chemical and controls. If several test chemicals are analysed in the same experiment, the same set of controls should be used, independently of number of plates, provided that all test chemicals are dissolved in the same vehicle. If different vehicles are used for different test chemicals within the same experiment, additional negative controls are required corresponding to each vehicle utilised in the experiment. In Figure 3, a schematic example of three main stimulation experiments with eight test chemicals and three controls are visualised, including one extra well with unstimulated controls. In this example, it is assumed that all test chemicals are dissolved in the same vehicle, therefore, one negative control is included.

© OECD (2026)

Figure 3. A schematic example of eight test chemicals and controls stimulated in the three replicate main stimulations using 12-well plates. TC; Test chemical.



50. The seeding procedures of main stimulations are typically identical with those described for cytotoxicity assessment experiments, as described in paragraphs 38-39, with the exception that only one concentration is investigated for each test chemical. A brief summary of a typical procedure is provided below.

51. Appropriate volume of Stock A of the test chemical is prepared in appropriate solvent as established in the preparation of the test chemical. Appropriate measures, e.g. vortexing and heat (37°C) should be applied if necessary, to achieve complete dissolution.

52. The Stock B concentration is prepared by adding appropriate volume of Stock A to cell medium (depending on the maximum target in-well concentration of solvent). Appropriate measures, e.g. vortexing and heat (37°C) may be applied if necessary, to achieve complete dissolution.

53. In addition, the positive and negative controls should be prepared to achieve appropriate in-well concentrations.

54. Cells are seeded for chemical exposure directly following a cell split, at an appropriate cell concentration with regards to the dilution that occurs upon addition of stock solution(s) of test chemical and/or controls. The final in-well cell concentration, after addition of test chemical, should be 0.2×10^6 cells/mL.

55. Note that if justified and motivated, the same option to omit the Stock B dilution step as described in section *Cytotoxicity assessment experiment* above, applies to main stimulation experiments as well. Alternative test chemical dilution schemes, not based on the herein described A and B stock solutions, are acceptable provided the target in-well test chemical concentration, cell concentration and maximum solvent concentration are met.

56. The plate(s) are covered with plastic plate-lid(s) and incubated for 24 h at 37°C under 5% CO_2 and humidified atmosphere.

57. After 24 h of incubation, the cell culture is mixed by carefully pipetting up and down and each cell culture from separate wells is divided into RNase-free micro tubes and duplicate flow cytometry samples.

58. Samples in micro tubes will be used for RNA isolation. For this purpose, cell pellets are lysed using an appropriate and fit-for-purpose reagent, e.g. TRIzol reagent (Ambion, #15596018), according to instructions provided by the supplier. Cell lysate samples may be stored at ≤-70°C up to a year.

59. For each test chemical and control, several cell lysate samples may be generated from each of the three main stimulations. However, only one cell lysate sample from each of the three main stimulations

© OECD (2026)

is required for RNA isolation and further analysed using the NanoString nCounter system. Any remaining cell lysate replicates may be stored (≤-70°C) as backup samples due to the possibility of having insufficient RNA concentration or RNA quality in the primary cell lysate sample.

60. For the flow cytometry samples, the same washing, staining and analysis procedures as described in paragraphs 42-46 for the cytotoxicity assessment experiment should be followed.

Acceptance criteria of the Viability Quality Control

61. The PI-stained samples are used as Quality Control of the viability to ensure that the test chemical and controls show a Relative or Absolute viability within the Quality Control criteria described in Table 4. If a test chemical fails the described acceptance criteria, it should not be used for downstream analysis. If any control sample fails the described acceptance criteria, all samples from the main stimulation experiment from which they originate are to be disregarded and not used for downstream analysis.

Table 4. Viability control acceptance criteria

Test chemical or control	Acceptance criteria[1]
Unstimulated control	Absolute viability of ≥84.5%
Negative control	Relative viability of ≥95.5%
Positive control	Relative viability 84.5% - 95.4%
Test chemical with expected cytotoxicity	Relative viability 84.5% - 95.4%
Test chemical assayed at 500 µM or highest soluble concentration	Relative viability ≥84.5%

[1] Listed acceptance criteria for unstimulated and negative controls apply to both cytotoxicity assessment experiments and main stimulation experiments, while criteria for the positive control and test chemical are only applicable in main stimulation experiments.

RNA isolation

62. Total RNA, including mRNA, is isolated from the lysed cell samples using commercially available kit and reagents, e.g. Direct-zol RNA MiniPrep (Zymo Research, # R2052) was used during test method development and validation.

63. Quantify the RNA concentration and analyze the RNA quality from each sample using an RNA analysis equipment, e.g. with an Agilent Bioanalyzer 2100, or an equivalent instrument (i.e. an instrument measuring RNA quality and RNA concentration in the range ~5-500 ng/µL). Follow protocols provided by the instrument supplier. RNA concentration and quality should correspond to NanoString recommendations. During test method development and validation, a sample with an RNA Integrity Number (RIN) of 8.0 and above, as derived from the Agilent Bioanalyzer 2100, was considered a sample of high quality. Corresponding or otherwise equivalent RNA quality metrics may be used to assure high quality RNA.

Endpoint measurement; gene expression analysis using the NanoString nCounter® system

64. The endpoint measurement of the GARDskin assay is the mRNA quantification of the endpoint-specific GPS, using the NanoString nCounter system. The NanoString nCounter protocol starts with manual processing including a hybridisation step using a thermal cycler, i.e. the nCounter XT CodeSet gene expression assay. A custom made CodeSet (i.e. sets of oligonucleotide probes representing the

© OECD (2026)

genes of the GARDskin GPS, the individual genes of which are presented in the Supporting document to the Test Guideline for the GARDskin test method (7), is provided by NanoString under a license agreement with SenzaGen AB. Manufacturer's instructions for the nCounter XT CodeSet gene expression assay should be followed.

65. The nCounter XT CodeSet gene expression assay is followed by automated sample processing, immobilising the probe/target on the nCounter Cartridge, and digital data acquisition, counting the color codes on the probe/targets immobilised on the cartridge, using the nCounter® instrument. Corresponding instructions for the nCounter instrument should be followed. The highest possible resolution and sensitivity mode in the nCounter instrument should be used.

66. For each RNA sample analysed in the NanoString nCounter system, a NanoString raw data file, a Reporter Code Count (RCC)-file with tabulated counts of each target molecule, is generated.

DATA ANALYSIS AND REPORTING

GARD Data Analysis Application

67. Following generation of the RCC-files, all downstream data preprocessing, normalisation and analysis are performed in the GDAA software as summarised below. For an in-depth review of all such steps, refer to the Supporting document to the Test Guideline for the GARDskin test method (7). Both the GARDskin analysis pipeline and the GDAA were extensively evaluated by the ESAC during method review, who concluded that the software was fit-for-purpose as well as user-friendly.

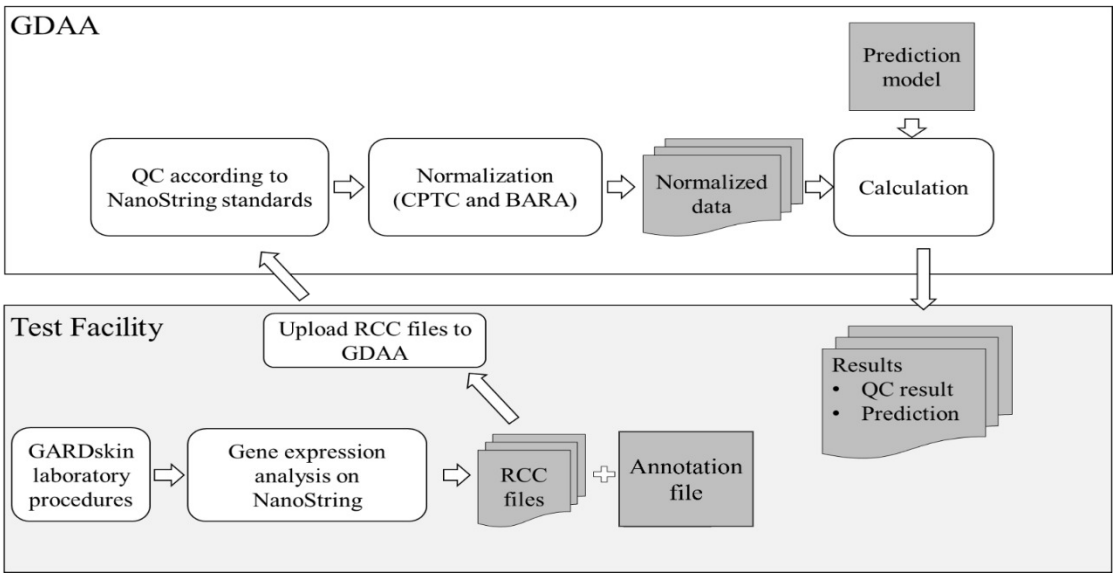
68. GDAA is a cloud-based application (Shinyapps on Amazon Web Services) requiring an internet connected computer with an installed web browser, e.g. Google Chrome, Mozilla Firefox or Microsoft Edge. Access to the GDAA requires a service level agreement and valid login credentials, both acquired from SenzaGen AB (www.senzagen.com)[5].

69. GDAA performs all data analysis required for generating predictions using the GARDskin method. The functionality of GDAA includes the reading of RCC-files, checking the NanoString nCounter quality control of each uploaded file, normalising the read file's gene expression values by stepwise application of a counts-per-total counts (CPTC) (12) algorithm followed by Batch Adjustment by Reference Alignment (BARA) (30). Lastly, individual samples are evaluated with the GARDskin prediction algorithm, allowing for the final classification of the test chemical by the GARDskin prediction model. For a schematic of the processes performed by the GDAA and how they relate to other steps of the procedure, see Figure 4.

[5] SenzaGen AB
Medicon Village
SE-223 81 Lund, Sweden
info@senzagen.com

© OECD (2026)

Figure 4. A schematic of the GDAA workflow, including input and output files. RCC; Reporter Code Count (raw gene expression data), CPTC; Counts-per-total-counts (RNA content normalisation), BARA; Batch Adjustment by Reference Alignment (batch adjustment normalisation)



70. The analysis with the GDAA requires upload of two different types of files: the RCC files (containing raw data of gene expression levels) and an Annotation file (containing sample information used to map each control and test chemical to specific RCC files). Note that the RCC files for each test chemical must be analysed together with the RCC-files of the unstimulated, positive and negative controls from the same main stimulation experiments, in order to enable both the BARA normalisation process (as further described in the Supporting document to the Test Guideline for the GARDskin test method (7), as well as the classification of negative and positive controls, in order to evaluate if acceptance criteria are met (as defined in section *Summary of acceptance criteria*).

71. After uploading the files, each RCC file is automatically quality checked in GDAA. The quality criteria listed in Table 5 are adapted from the recommended (default) acceptance criteria of the instrument supplier. Details of each quality metric listed in Table 5 are provided in the Supporting document to the Test Guideline for the GARDskin test method (7). Samples that fail any of the below described quality control criteria are not used for further analysis in the GARD data analysis and GDAA automatically rejects samples that fail the NanoString nCounter® Quality Control acceptance criteria.

© OECD (2026)

Table 5. Summary of the NanoString nCounter® quality control acceptance criteria.[1]

Quality Metric	Description	Acceptance criteria
Imaging Quality	Imaging quality describes the fraction of successfully processed sections of the NanoString cartridge.	> 0.75
Linearity	Linearity is expressed as an R2 value that is estimated using the positive spike-in controls.	> 0.95
Limit of Detection	The limit of detection is evaluated by comparing the acquired counts of the positive spike-in probe POS_E to the counts of the negative control probes. Probe POS_E is the positive spike-in probe with the lowest concentration that is expected to be observed above noise levels.	< POS_E
Binding Density	Binding density reports on the level of image saturation that was observed during cartridge processing. The value is dependent on the amount of sample that was loaded and the effectiveness of the NanoString hybridisation.	0.05 - 2.25

[1] For details, refer to the Supporting document to the Test Guideline for the GARDskin test method (7).

72. The last steps of a GARDskin analysis includes the application of a prediction algorithm, which in turn provides input to the GARDskin prediction model, as described in section *Prediction model* below.

73. In addition to facilitating the complete GARDskin analysis pipeline, the GDAA functionality includes verification of integrity of transferred data using algorithms for calculating MD5 checksums. The MD5 algorithm takes input data of arbitrary length and calculates a 128-bit fingerprint. It will always produce the same fingerprint for a specific input, and it is highly unlikely that two different data inputs would generate the same output values. These properties of the algorithm make it useful for verifying integrity of data. For example, the integrity of a transferred file can be ensured by comparing the 128-bit fingerprints calculated prior to the transfer with a fingerprint calculated following transfer. These MD5 checksums are evaluated as part of the periodic and before use testing of the computerised system, as described in section *Periodic testing of cloud-based software*.

Prediction model

74. The GARDskin prediction algorithm is a SVM hosted in the GDAA, appropriately trained and frozen during method development. The output of the prediction algorithm is referred to as a decision value (DV). Unique DVs are calculated for each replicate sample generated by test chemicals and controls, as described by equation 2.

$$DV = b + \sum_{i=1}^{n} w_i x_i \qquad (2)$$

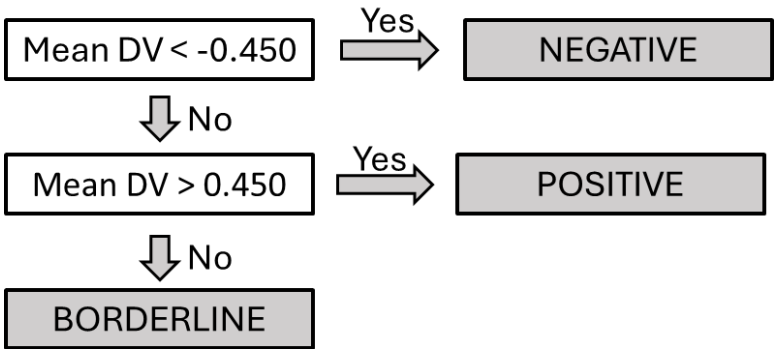
where n is the number of variables (genes, i.e., 196 for GARDskin), b is a constant (i.e., the SVM's intercept), w_i the weight for variable i, and x_i the normalised gene expression value for variable i. For an in-depth review of how the prediction model was defined, please refer to the Supporting document to the Test Guideline for the GARDskin test method (7).

75. The DVs of the three individual replicate samples are then used as input to the GARDskin prediction model. The GARDskin prediction model is defined as follows:

© OECD (2026)

76. Any test chemical with a calculated mean DV ≥ 0 is classified as a sensitiser (UN GHS category 1), whereas any test chemical assigned a mean DV < 0 is classified as a non-sensitiser. A schematic representation of the GARDskin prediction model is provided in Figure 5.

Figure 5. Schematic of the GARDskin prediction model, taking into account borderline ranges and multiple runs to conclude on borderline results.



The original threshold for a positive classification is a Mean DV > 0 with a statistically derived borderline range of -0.450 < Mean DV < 0.450.

Summary of Acceptance criteria

77. Below is a summary of the acceptance criteria that are specified for the GARDskin method.

i) All cell exposure experiments should have been performed with a batch of SenzaCell cells that passed the acceptance criteria of the Phenotypic Quality Control (Table 2). This applies to both cytotoxicity assessment and main stimulation experiments.

ii) All generated RNA samples should originate from cellular experiments, which have passed the acceptance criteria of the Viability Quality Control (Table 4). This applies to all test chemical and positive, negative and unstimulated control samples, from the three included (i.e., valid) main stimulation experiments. Similarly, cytotoxicity assessment experiment(s) from which a GARDskin input concentration is derived should fulfill all applicable Viability Quality Control criteria (Table 4).

iii) All generated RNA samples should pass the Acceptance criteria of the NanoString nCounter® Quality Control (Table 5). This applies to all test chemical and positive, negative and unstimulated control samples, from the three included (i.e., valid) main stimulation experiments.

iv) The final classification should be made using three valid biological replicates which have all passed acceptance criteria i-iii.

v) The positive control and negative control should be accurately classified as a sensitiser and a non-sensitiser, respectively, by the GARDskin prediction model.

Test Report

78. The following information should be reported. The results should be tabulated and include, when applicable, individual test results for each performed experiment as well as the overall results from all three experiments.

© OECD (2026)

General information

- Name and address of sponsor, test facility and study director.
- Reference and description of the test method used.

Demonstration of proficiency

- Statement that the test facility has demonstrated proficiency in the performance of the test method before routine use by testing of proficiency chemicals.

Demonstration of GDAA system stability over time

- Statement that periodic and/or before use testing of the GDAA have been performed using a historic dataset and have passed required criteria.

Test chemical and controls

- Source, batch/lot number, expiry date. Chemical identification, such as IUPAC name, CAS registry number, SMILES or InChI code, structural formula, and/or other identifiers like batch/lot number and expiry date.
- Physical appearance, solvent solubility as applicable, molecular weight, and additional physical chemical properties to the extent available.
- Statement on (in)solubility or stable dispersion in exposure media.
- Purity, chemical identity of impurities as appropriate and practically feasible.
- Procedure(s) used to dissolve test chemical(s).
- Storage conditions and stability to the extent available.
- Justification for choice of solvent/vehicle for each test chemical.
- Solvent (including source) used for each test chemical and control.

Test method conditions

- Cell line used, cell culture ID, its storage conditions and source.
- Cell media components (including source) used in the study.
- Flow cytometry equipment used.
- Antibodies and viability markers (including sources) used in the study.
- RNA isolation kit, RNA qualification kit and NanoString nCounter GARDskin CodeSets (including sources) used in the study.

Test acceptance criteria results

- Phenotypic Quality Control data from each experiment (percentage of positive cells for each phenotypic biomarker, as well as absolute viability of cells).
- Cell viability Quality Control data of the test chemical and negative, positive and unstimulated controls.
- NanoString nCounter Quality control data (imaging quality, linearity, limit of detection and binding density) of the test chemical as well as negative and positive controls.
- Classifications of negative and positive controls.

Cytotoxicity assessment results

- Test concentrations with justifications.
- Relative viability for each test chemical concentration
- Justification for selected GARD input concentration.

Main stimulation results

- Measured RNA-quality of test chemical and control RNA-samples
- Gene expression levels (content of RCC files) for test chemical and controls, in a format compliant with available guidance (31) (32).

© OECD (2026)

- GDAA version number used within the study.
- Statement on matching MD5 checksums of uploaded RCC-files and MD5 checksums from downloaded GDAA report.
- Decision Values (individual samples as well as mean) obtained for the test chemical and positive and negative controls.
- Classification of the test chemical.
- Description of any other relevant observations, if applicable.

Discussion of the results

Conclusion

© OECD (2026)

Literature

1. WO 2019/057977. Available at:
https://worldwide.espacenet.com/publicationDetails/biblio?DB=EPODOC&II=0&ND=3&adjacent=true&locale=en_EP&FT=D&date=20190328&CC=WO&NR=2019057977A1&KC=A1

2. Larsson K, Lindstedt M, Borrebaeck CA. Functional and transcriptional profiling of MUTZ-3, a myeloid cell line acting as a model for dendritic cells. Immunology. 2006;117(2):156-166. doi:10.1111/j.1365-2567.2005.02274.x

3. Hu, Z. B., Ma W., Zaborski, M., MacLeod, R., Quentmeier, H. and Drexler, H.G. (1996). Establishment and characterization of two novel cytokine-responsive acute myeloid and monocytic leukemia cell lines, MUTZ-2 and MUTZ-3. Leukemia, 10: 1025-40.

4. Quentmeier, H., Duschl, A., Hu, Z.B., Schnarr, B., Zaborski, M. and Drexler, H.G. (1996). MUTZ-3, a monocytic model cell line for interleukin-4 and lipopolysaccharide studies. Immunology, 89: 606-612.

5. Masterson, A.J., Sombroek, C.C., De Gruijl, T.D., Graus, Y.M., van der Vliet, H. J., Lougheed, S.M., van den Eertwegh, A.J., Pinedo, H.M. and Scheper, R.J. (2002). MUTZ-3, a human cell line model for the cytokine-induced differentiation of dendritic cells from CD34+ precursors. Blood, 100: 701-703.

6. Johansson, H., Lindstedt, M., Albrekt, A. S. and Borrebaeck, C. A. (2011). A genomic biomarker signature can predict skin sensitizers using a cell-based in vitro alternative to animal tests. BMC Genomics, 12: 399

7. OECD (2022), *Supporting document to the Genomic Allergen Rapid Detection Test Method for Skin Sensitisation (GARD™skin), described in Test Guideline 442E*, Series on Testing and Assessment, No. 357, OECD Environment, Health and Safety, Paris, available at: https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/publications-on-testing-and-assessment-of-chemicals.html.

8. Johansson, H., Albrekt, A. S., Borrebaeck, C. A. K. and Lindstedt, M. (2013). The GARD assay for assessment of chemical skin sensitizers. Toxicol In Vitro., 27 (3): 1163-9.

9. Johansson, H., Rydnert, F., Kuhnl, J., Schepky, A., Borrebaeck, C. and Lindstedt, M. (2014). Genomic allergen rapid detection in-house validation - a proof of concept. Toxicol Sci., 139 (2): 362-70.

10. Forreryd, A., Johansson, H., Albrekt, A. S. and Lindstedt, M. (2014). Evaluation of high throughput gene expression platforms using a genomic biomarker signature for prediction of skin sensitization. BMC Genomics, 15 (1): 379.

11. Geiss, G. K., Bumgarner, R. E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D. L., Fell, H. P., Ferree, S., George, R. D., Grogan, T., James, J. J., Maysuria, M., Mitton, J. D., Oliveri, P., Osborn, J. L., Peng, T., Ratcliffe, A. L., Webster, P. J., Davidson, E. H., Hood, L. and Dimitrov, K. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol., 26 (3): 317-25.

12. Forreryd, A., Zeller, K. S., Lindberg, T., Johansson, H. and Lindstedt, M. (2016). From genome-wide arrays to tailor-made biomarker readout - Progress towards routine analysis of skin sensitizing chemicals with GARD. Toxicol In Vitro., 37: 177-88.

13. EURL ECVAM. (2018). Genomic Allergen Rapid Detection (GARD) assay for skin sensitization. Report on the GARDskin validation study. Available at: https://tsar.jrc.ec.europa.eu/test-method/tm2011-09

14. Johansson, H., Gradin, R., Johansson, A., Adriaens, E., Edwards, A., Zuckerstatter, V., Jerre, A., Burleson, F., Gehrke, H. and Roggen, E. L. (2019). Validation of the GARDskin Assay for Assessment of Chemical Skin Sensitizers: Ring Trial Results of Predictive Performance and Reproducibility. Toxicol Sci., 170 (2): 374-381

15. ESAC opinion on the scientific validity of the GARDskin and GARDpotency test methods. Available at: https://publications.jrc.ec.europa.eu/repository/handle/JRC125963. DOI: 10.2760/626728

© OECD (2026)

16. Basketter, D.A., Alépée, N., Ashikaga, T., Barroso, J., Gilmour, N., Goebel, C., Hibatallah, J., Hoffmann, S., Kern, P., Martinozzi-Teissier, S., Maxwell, G., Reisinger, K., Sakaguchi, H., Schepky, A., Tailhardat, M. and Templier, M. (2014). Categorization of chemicals according to their relative human skin sensitizing potency. Dermatitis, 25 (1): 11-21. doi: 10.1097/DER.0000000000000003. PMID: 24407057..

17. Natsch, A., Ryan, C.A., Foertsch, L., Emter, R., Jaworska, J., Gerberick, F. and Kern, P. (2013). A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. J Appl Toxicol, 33 (11): 1337-52. doi: 10.1002/jat.2868. Epub 2013 Apr 9. PMID: 23576290.

18. OECD (2026), *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.

19. OECD (2021), *Annex 2 of the Supporting document to the Guideline (GL) on Defined Approaches (DAs) for Skin Sensitisation,* Series on Testing and Assessment, No. 336, OECD Environment, Health and Safety, Paris, available at https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/publications-on-testing-and-assessment-of-chemicals.html.

20. Forreryd A, Gradin R, Humfrey C, Sweet L, Johansson H. Exploration of the GARD™skin applicability domain: Indirectly acting haptens, hydrophobic substances and UVCBs. ALTEX. 2023;40(1):53-60. doi: 10.14573/altex.2201281. Epub 2022 Apr 21. PMID: 35502631.

21. Macmillan, D. 2018. GARD Applicability domain project. Supporting document for method peer-review, available online in the Tracking System for Alternative methods towards Regulatory acceptance (TSAR). Available at: https://tsar.jrc.ec.europa.eu/test-method/tm2011-09

22. Forreryd A, Gradin R, Larne O, Rajapakse N, Deag E, Johansson H. The GARD™skin assay: Investigation of the applicability domain for metals. ALTEX. 2023;40(3):425-438. doi: 10.14573/altex.2203021. Epub 2022 Nov 3. PMID: 36343115.

23. Corvaro Marco, Henriquez Joseph, Settivari Raja, Mattson Ulrika, Forreryd Andy, Gradin Robin, Johansson Henrik, Gehen Sean (2024) GARD™skin and GARD™potency: A proof-of-concept study investigating applicability domain for agrochemical formulations. Regulatory Toxicology and Pharmacology. 148. 105595. 10.1016/j.yrtph.2024.105595.

24. Cortes, C. and Vapnik, V. (1995). Support-Vector Networks. Machine Learning, 20 (3): 273-297.

25. OECD (2016), *Application of GLP Principles to Computerised Systems*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 17, OECD Publishing, Paris, https://doi.org/10.1787/e3c9cda5-en.

26. OECD (2021), *GLP Data Integrity*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 22, OECD Publishing, Paris, https://doi.org/10.1787/45779212-en.

27. Rivest, R. (1992). The MD5 Message-Digest Algorithm. IETF RCF1321. Available at: https://www.ietf.org/rfc/rfc1321.txt

28. EURL ECVAM. (2021). GARDskin Assay Protocol. Available at: https://tsar.jrc.ec.europa.eu/system/files/Published/GARDskin%20Assay%20Protocol%20TSAR.pdf .

29. OECD (2025), *Guidance Document on Good In Vitro Method Practices (GIVIMP), Second Edition*, OECD Series on Testing and Assessment, No. 421, OECD Publishing, Paris, https://doi.org/10.1787/5ba6777b-en.

30. Gradin, R., Lindstedt, M. and Johansson, H. (2019). Batch adjustment by reference alignment (BARA): Improved prediction performance in biological test sets with batch effects. PLoS One, 14 (2): e0212669.

31. OECD (2021), Omics technologies in chemical testing, available at: https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/omics-technologies-chemical-testing.html .

© OECD (2026)

32. Harrill, J.A., Viant, M.R., Yauk, C.L., Sachana, M., Gant, T.W., Auerbach, S.S., Beger, R.D., Bouhifd, M., O'Brien, J., Burgoon, L., Caiment, F., Carpi, D., Chen, T., Chorley, B.N., Colbourne, J., Corvi, R., Debrauwer, L., O'Donovan, C., Ebbels, T.M.D., Ekman, D.R., Faulhammer, F., Gribaldo, L., Hilton, G.M., Jones, S.P., Kende, A., Lawson, T.N., Leite, S.B., Leonards, P.E.G., Luijten, M., Martin, A., Moussa, L., Rudaz, S., Schmitz, O., Sobanski, T., Strauss, V., Vaccari, M., Vijay, V., Weber, R.J.M., Williams, A.J., Williams, A., Thomas, R.S. and Whelan, M. (2021) Progress towards an OECD reporting framework for transcriptomics and metabolomics in regulatory toxicology. Regul Toxicol Pharmacol. Oct; 125:105020. doi: 10.1016/j.yrtph.2021.105020.

© OECD (2026)

Appendix I - Definitions and abbreviations.

Abbreviations that are defined by the test method developers and/or specific for certain instrumentation utilised by the GARD methods are *defined in italic font*.

ARE Antioxidant Response Element

BARA Batch Adjustment by Reference Alignment

An algorithm for removal of batch-effects observed between datasets.

BLR Between-Laboratory Reproducibility

BSA Bovine Serum Albumin

CD Cluster of Differentiation

Cell Batch

Within the context of this TG, a unique cell batch is defined as:

- *cells originating from different frozen vials, or…*
- *cells originating from the same frozen vial, which have been cultivated separately. A division of cell cultures for the purpose of achieving separate cell batches should be done no sooner than passage 3 after thawing, and no later than at least 2 passages prior to exposure experiments.*

CPTC Counts-Per-Total-Counts

An algorithm for RNA content normalisation.

DB-ALM DataBase service on ALternative Methods to animal experimentation

DC Dendritic Cells

DMSO Dimethyl Sulfoxide

DV Decision Value

A quantifiable output from a Support Vector Machine.

ESAC ECVAM Scientific Advisory Committee

EURL ECVAM European Union Reference Laboratory for alternatives to animal testing

FBS Fetal Bovine Serum

FITC Fluorescein IsoThioCyanate

GARD Genomic Allergen Rapid Detection

A series of predictive assays for immunotoxicological endpoints, the main subject of this Test Guideline.

GARDskin GARD Test Method for Skin Sensitisation

© OECD (2026)

The specific subject of this TG. A method used for hazard assessment of skin sensitisers

GARDskin prediction algorithm

An algorithm that, based on raw gene expression data, provides DVs as output. The output of the prediction algorithm is in turn used as input in the prediction model. The GARDskin prediction algorithm is an SVM. See e.g., DV, GARDskin prediction model, SVM.

GARDskin prediction model

A heuristic that, based on triplicate DVs from a test chemical or control, provides a GARDskin classification of the same.

GDAA GARD Data Analysis Application

A cloud-based software for fit-for-purpose and automated data processing and analysis of all raw data generated with the GARD methods.

GHS Globally Harmonized System

GM-CSF Granulocyte Macrophage Colony Stimulating Factor

GPS Genomic Prediction Signature

A set of gene identities that collectively compose the set of predictors, the gene expression values of which are used as the input in the GARD prediction models, i.e., the endpoint-specific Support Vector Machine(s), each appropriately trained and frozen during test method development. Each GARD method (e.g. GARDskin) for different endpoints utilises a different GPS.

LLNA Local Lymph Node Assay

LOD Limit of Detection

A parameter of the NanoString instrumentation.

MD5 Message-Digest algorithm 5

A function that creates digital fingerprints of input data. Within the context of the GARDskin method, such fingerprints are used to verify integrity of data.

OECD Organisation for Economic Co-operation and Development

PBS Phosphate Buffered Saline

PE Phycoerythrin

PI Propidium Iodide

RCC Reporter Code Count

A filetype created by the NanoString instrumentation. Stores raw gene expression data.

© OECD (2026)

RIN RNA Integrity Number

An RNA quality parameter utilised by the Agilent instrumentation.

SVM Support Vector Machine

Supervised prediction models with associated learning algorithms that analyze data for classification and regression analysis

TG Test Guideline

WLR Within-Lab Reproducibility

© OECD (2026)

Appendix II - Known limitations of the GARDskin method.

Table AII.1. A summary of known limitations of the GARDskin method and possible adaptations.

Substance class / interference	Possible consequence of interference	Possible adaptations	Example substance
Test chemicals absorbing and/or autoflourescing light at the wavelengths of PI-detection.	May influence cytotoxicity assessment results and may lead to inappropriately defined GARD input concentrations.	May be circumvented by use of alternative reagents for assessment of cytotoxicity, if demonstrated to generate equivalent results to those of the herein proposed methods.	Citral (CAS #5392-40-5)
Substances with unknown precise molecular weight. A Test chemical is preferably defined by a known molecular weight, as appropriate GARD input concentrations are defined by molar concentrations.	May lead to inappropriate GARD input concentrations, which may in turn lead to misclassifications.	May be circumvented by -Use of weight-based concentrations (e.g. ppm) (1). A vast majority of skin sensitisers are detected <100 ppm (1). -Approximation of an apparent molecular weight of the complex mixture,	UVCBs, chemical emissions, products or formulations with variable or not fully known composition, natural extracts.
Test chemicals which cannot be dissolved in an appropriate solvent at a final in-well concentration of 500 µM and do not exhibit cytotoxic properties at investigated max concentration.	A sufficient exposure concentration for detection of weak skin sensitisers may not be guaranteed. May cause false negatives. However, available data suggest that a vast majority of sensitisers are detected <100 µM (1).	Test chemicals that do not induce cytotoxicity with a maximum soluble concentration below 500 µM may be further analysed according to downstream GARDskin procedures and positive results from such testing can be used to support the identification of the test chemical as a skin sensitiser.	n/a
Test chemicals Incompatible with vehicles.	Insolubility or reactive interference with Test chemical, which may in turn lead to inappropriate GARD input concentrations and possible misclassifications, or complete incompatibility with the method.	If a scientific rationale is available, alternative and otherwise compatible vehicles may be used (2) (3). Compatibility of such alternative vehicles should be confirmed by inclusion of the blank vehicle as a negative control, at identical exposure concentrations. If a Test chemical remains insoluble, see handling of not sufficiently dissolved Test chemicals above.	n/a
Test chemicals that hydrolyse rapidly in cell system.	A sufficient in-well concentration of Test chemicals may not be guaranteed. May cause false negatives		Hydrazine (CAS #2644-70-4)

Literature

1. Gradin, R., Forreryd, A., Mattson, U., Jerre, A. and Johansson, H. (2021). Quantitative Assessment of Sensitizing Potency using a dose-response adaptation of GARDskin. Sci. Rep, 176 (2): 423-432.
2. Forreryd A, Gradin R, Humfrey C, Sweet L, Johansson H. Exploration of the GARD™skin applicability domain: Indirectly acting haptens, hydrophobic substances and UVCBs. ALTEX. 2023;40(1):53-60. doi: 10.14573/altex.2201281. Epub 2022 Apr 21. PMID: 35502631.
3. Corvaro Marco, Henriquez Joseph, Settivari Raja, Mattson Ulrika, Forreryd Andy, Gradin Robin, Johansson Henrik, Gehen Sean (2024) GARD™skin and GARD™potency: A proof-of-concept study investigating applicability domain for agrochemical formulations. Regulatory Toxicology and Pharmacology. 148. 105595. 10.1016/j.yrtph.2024.105595..

© OECD (2026)

Appendix III - Proficiency Substances.

Table AIII.1. Substances for demonstrating technical proficiency with GARDskin.

Chemical ID	CASRN	Physical state	Reference classification[1]		Expected GARDskin values (reference range)		
			LLNA	Human	Classification	(DV)	(Input concentration, µM)
4-nitrobenzyl bromide	100-11-8	Solid	S (extreme)	*NA*	S	(0 – 10)	<25
Propyl gallate	121-79-9	Solid	S (strong)	*NA*	S	(2 – 13)	25 - 400
Isoeugenol	97-54-1	Liquid	S (moderate)	S	S	(2 – 13)	>100
3-(Dimethylamino)-1-propylamine	109-55-7	Liquid	S (moderate)	*NA*	S	(0 – 11)	>250
Eugenol	97-53-0	Solid	S (weak)	S	S	(0 – 10)	>100
Ethylene glycol dimethacrylate	97-90-5	Liquid	S (weak)	*NA*	S	(2 – 20)	>25
Glycerol	56-81-5	Liquid	NS	*NA*	NS	(<0)	≥250
Hexane	110-54-3	Liquid	NS	NS	NS	(<0)	≥125
1-Butanol	71-36-3	Liquid	NS	*NA*	NS	(<0)	≥250

[1] Extracted from Annex 2 of the supporting document to OECD TG 497 (1). S: Sensitiser. NS: Non-sensitiser NA: missing value.

Literature

1. OECD (2021). *Annex 2 of the Supporting document to the Guideline (GL) on Defined Approaches (DAs) for Skin Sensitisation*, Series on Testing and Assessment, No. 336, OECD Environment, Health and Safety, Paris, available at https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/publications-on-testing-and-assessment-of-chemicals.html.

© OECD (2026)