New Chemicals Program Decision Framework for Hazard Identification of Skin Irritation and Corrosion

1 Purpose

This document describes the decision framework for use within EPA's New Chemicals Program (NCP) for identification of skin irritation or corrosion hazards for new chemical substances based on reproducible, human-relevant data.

2 Introduction

Multiple test methods are available to assess the skin irritation or corrosion potential of new chemical substances. The in vivo skin irritation test in rabbits, however, lacks reproducibility (Weil & Scala, 1971), particularly in the mild to moderate range of irritancy (Rooney et al., 2021). This inconsistency calls into question the utility of in vivo rabbit data for assessing risks to human health. Recently, Raabe et al (2025) evaluated skin irritation/corrosion New Approach Methodologies (NAMs)[1] for their reproducibility and relevance to mechanisms of human skin irritation/corrosion. These NAMs were found to perform as well as or better than the in vivo rabbit skin irritation test (Raabe et al., 2025), consistent with the Strategic Plan to Promote the Development and Implementation of Alternative Test Methods Within the TSCA Program (US EPA, 2018; hereafter referred to as the Strategic Plan). Specifically, the Strategic Plan, developed by EPA in 2018 in response to TSCA Section 4 (h)(2)(A), requires that EPA must ". . . *develop a strategic plan to promote the development and implementation of alternative test methods and strategies to reduce, refine, or replace vertebrate animal testing and provide information of equivalent or better scientific quality and relevance for assessing risks of injury to health or the environment . . .".* In alignment with the Strategic Plan, scientific relevance, reliability, and confidence of many of these NAMs have been evaluated and accepted for their accuracy, specificity, and sensitivity compared to in vivo rabbit skin irritation/corrosion data and the corresponding skin irritation/corrosion categories outlined in the Globally Harmonized System of Classification and Labeling of Chemicals (GHS) (United Nations, 2023).[2] This skin irritation and corrosion framework fulfills, in part, the long term goal of the Strategic Plan to integrate and prioritize NAMs in the TSCA decision process for risk evaluation to reduce and eventually replace the use of vertebrate animal testing.

Human cell-based tissue model NAMs, described in this framework, are given priority to assess the potential for skin irritation/corrosion hazard of a new chemical substance(s). If data from these models are unavailable or there are gaps, additional data from other skin

[1] Discussion of NAMs within this framework refers to in vitro, in chemico, and ex vivo methods.

[2] Although EPA's New Chemicals Program does not classify or label chemicals according to GHS categories, this document adopts GHS principles to guide the evaluation of skin irritation/corrosion hazards of new chemicals under Section 5 of TSCA.

irritation/corrosion NAMs that are reproducible and biologically relevant are then prioritized. In vivo data will be used only in the absence of sufficient NAMs data.

The NCP uses three categories to characterize skin irritation/corrosion potential for new chemical substances: (1) corrosive, (2) irritating, and (3) nonirritating. Currently, there is no NAM that can be used as a standalone test to distinguish across the three categories.[3] Rather, available NAMs provide two types of binary outcomes, which can be integrated, as appropriate, to identify skin irritation and corrosion as a hazard (Figure 1). There are methods that identify nonirritants but cannot distinguish between irritation and corrosion. Other methods can identify corrosives but cannot distinguish between irritants and nonirritants.

In alignment with TSCA Section 4(h)(2)(C) and the Strategic Plan, EPA maintains the List of Alternative Test Methods and Strategies (or NAMs) that includes representative NAMs EPA may consider. The National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) also maintains a list of Alternative Methods Accepted by US Agencies. The NCP will evaluate data from studies not approved by the Organisation for Economic Cooperation and Development (OECD), EPA, or NICEATM on a case-by-case basis for quality, reliability, and relevance to mechanisms of human skin irritation/corrosion. Table 2 provides the applicability domain for some established test methods.

3 Decision Framework

To predict the skin irritation or corrosion potential of new chemical substances,[4] a decision framework based on reproducible and human-relevant data is presented below. Full details of how data are prioritized are provided in *Sections 3.1–3.3* and in Figure 1. In brief, data are prioritized in the following order:

1. Data from human cell-based tissue model NAMs that have been demonstrated to be reproducible and relevant to skin irritation/corrosion.
2. Data from other skin irritation/corrosion NAMs that have been demonstrated to be reproducible and provide information on the mechanisms of toxicity relevant to skin irritation/corrosion.
3. Data from in vivo test methods.

3.1 Scientific Quality Review and Data Selection

Data are reviewed for scientific quality and applicability for evaluating the human skin irritation or corrosion potential of the new chemical substance. Scientific quality review may include, but

[3] This framework is expected to be generally applicable to future availability of NAMs that can distinguish across the three categories.

[4] This skin irritation/corrosion hazard identification does not apply to nanomaterials.

is not limited to, protocol adherence to test guidelines, inclusion of acceptable controls, sufficient sample size and appropriately applied statistical approaches. In vitro assays should be specifically evaluated for compliance with applicability domains (Table 2). Data from studies not approved by OECD, EPA or NICEATM will undergo a thorough review for quality, reliability, and relevance to mechanisms of human skin irritation/corrosion on a case-by-case basis. Data that do not meet scientific quality criteria or were gathered from assays that are not appropriate for the test substance type will not be further evaluated.

3.1.1 Data are prioritized based on availability and substance tested, as follows:

3.1.1.1 Skin irritation/corrosion data on the new chemical substance are preferred.

3.1.1.2 In the absence of sufficient data on the new chemical substance, analogue data may be incorporated throughout the decision framework, as appropriate, to address any gaps in the dataset.[5]

3.1.1.3 If skin irritation/corrosion data of sufficient quality for the new chemical substance or appropriate analogues are not available, physicochemical properties or other information such as structural alerts, other relevant test data, or chemical category conclusions from the TSCA New Chemicals Program Chemical Categories document (US EPA, 2010) will be considered. The complete absence of relevant data or structural alert information for the new chemical substance may preclude a hazard determination.

Following scientific quality review and selection of data on the new chemical substance and analogue(s), as appropriate, move to *Section 3.2* to begin hazard assessment.

3.2 Hazard Identification Using NAMs Data

Evaluate data from the new chemical substance and/or appropriate analogue(s) for relevance to skin irritation/corrosion endpoint and to humans. If skin irritation/corrosion NAMs data (e.g., see Table 1) are available, human cell-based tissue model NAMs data are evaluated first (*3.2.1*), followed by incorporation of other NAMs data (*3.2.2)*. In the absence of sufficient data to clearly identify hazard using NAMs data, in vivo data are considered as outlined in *3.3*. Table 1 summarizes the skin irritation/corrosion hazard identification capability of some established NAMs with OECD test guidelines. For additional details regarding test principles and applicability domains, see Table 2.

[5] Identification of appropriate analogues may include a number of approaches and considerations. Further information regarding analogue selection is available in the Sustainable Futures/P2 Framework Manual.

Table 1: Hazard Identification Capability of Some Established Test Methods

Guideline	Test Method	Hazard Classification Based on Percent Tissue Viability[6]		
Human Cell-Based Tissue Model NAMs based on In Vitro Reconstructed Human Epidermis (RhE)				
OECD TG 439 (Irritation)	RhE model	Nonirritating	> 50%	
		Irritating/corrosive	≤ 50%	
OECD TG 431 (Corrosion)	EpiSkin™ prediction model	Corrosive[7]	< 35% after 3 min, OR ≥ 35% after 3 min exposure AND < 35% after 60 min, OR ≥ 35% after 60 min exposure AND < 35% after 240min	
		Irritating/nonirritating	≥ 35% after 240 min	
	EpiDerm™ SkinEthic™ RHE epiCS® LabCyte EPI-MODEL24 SCT	Corrosive[7]	< 50% after 3 min, OR ≥ 50% after 3 min AND < 15% after 60 min	
		Irritating/nonirritating	≥ 50% after 3 min AND ≥ 15% after 60 min	
Other NAMs				
OECD TG 435	Corrositex (in chemico)	Acid/Alkaline Reserve	High	Low
		Corrosive[7]	≤ 240 min	≤ 60 min
		Irritating/nonirritating	> 240 min	> 60 min

3.2.1 Human Cell-based Tissue Model NAMs Test Data

This framework begins with an evaluation of the available human cell-based tissue NAMs test data and their use in determining the three categories.

3.2.1.1 If the data are from test methods that can identify the three skin irritation/corrosion categories, then integrate the test(s) with binary outcomes, as needed, to classify the new chemical substance into the appropriate skin irritation/corrosion category, as indicated in Figure 1. **STOP HERE.**

3.2.1.2 If the test(s) are only able to provide a binary identification of nonirritating vs. irritating/corrosive, and:

3.2.1.2.1 The data predicts nonirritancy, then select the nonirritating category. **STOP HERE.**

3.2.1.2.2 The data predicts irritancy/corrosivity and there are additional data to consider, see *Section 3.2.1.4*.

3.2.1.2.3 The data predicts irritancy/corrosivity and there are no other data to consider, then the corrosive category is selected. **STOP HERE.**

3.2.1.3 If the test(s) are only able to provide a binary identification of corrosive vs. irritating/nonirritating, and:

3.2.1.3.1 The data predicts corrosivity, then select the corrosive category.

[6] Published OECD Test Guidelines provide additional details on data interpretation.

[7] Test outcomes in the guideline that correspond to GHS classification as corrosive.

STOP HERE.

3.2.1.3.2 The data predicts noncorrosivity and there are additional data to consider, see *Section 3.2.1.4*.

3.2.1.3.3 The data predicts noncorrosivity and there are no other data to consider, then the irritation category is selected. **STOP HERE.**

3.2.1.4 If human cell-based tissue model NAMs data are unavailable, there are data gaps, or conflicting data, move to the next highest-prioritized data:

3.2.1.4.1 See *Section 3.3.2* for other NAMs test data.

3.2.1.4.2 In the absence of other NAMs test data, see *Section 3.3* for use of in vivo data.

3.2.2 Other NAMs Test Data

If sufficient human cell-based tissue NAMs test data are not available to determine the skin irritation/corrosion category, other NAMs test data (e.g., OECD TG 435) are considered.

3.2.2.1 If the data are from test methods that can identify the three skin irritation/corrosion categories, then integrate the test(s) with binary outcomes, as needed, to classify the new chemical substance into the appropriate skin irritation/corrosion category, as indicated in Figure 1. **STOP HERE.**

3.2.2.2 If the test(s) are only able to provide a binary identification of nonirritating vs. irritating/corrosive, and:

3.2.2.2.1 The data predicts nonirritancy, then select the nonirritating category. **STOP HERE.**

3.2.2.2.2 The data predicts irritancy/corrosivity and there are additional data to consider, see *Section 3.3.2.4.*

3.2.2.2.3 The data predicts irritancy/corrosivity and there are no additional data to consider, then the corrosive category is selected. **STOP HERE.**

3.2.2.3 If the test(s) are only able to provide a binary identification of corrosive vs. irritating/nonirritating, and:

3.2.2.3.1 The data predicts corrosivity, then select the corrosive category. **STOP HERE.**

3.2.2.3.2 The data predicts noncorrosivity and there are additional data to consider, see *Section 3.3.2.4*.

3.2.2.3.3 The data predicts noncorrosivity and there are no additional data to consider, then the irritation category is selected. **STOP HERE.**

3.2.2.4 If there are data gaps or conflicting data, move to the next highest-priority data:

3.2.2.4 See *Section 3.3* for use of in vivo data.

3.2.2.4 In the absence of in vivo data, see *Section 3.3.2*.

3.3 Hazard Identification Using In Vivo Data

In vivo data are incorporated in the absence of NAMs data or when the NAMs data do not provide a clear severity category (e.g., nonirritant vs. irritant).

3.3.1 If in vivo test data are available, and:

3.3.1.1 There are also NAMs data available, use the most severe category from NAMs data above. In limited cases, considering supporting in vivo data may be appropriate. For example, if NAMs data rules out corrosivity only, evaluating in vivo data to distinguish whether a substance is irritating or nonirritating may be appropriate. **STOP HERE**.

3.3.1.2 The in vivo data are the only data available, use the most severe category from in vivo data to provide hazard identification. **STOP HERE**.

3.3.2 If sufficient in vivo data are not available to address the data gap(s) or there are conflicting data, use the most severe category from NAMs data above. **STOP HERE**.

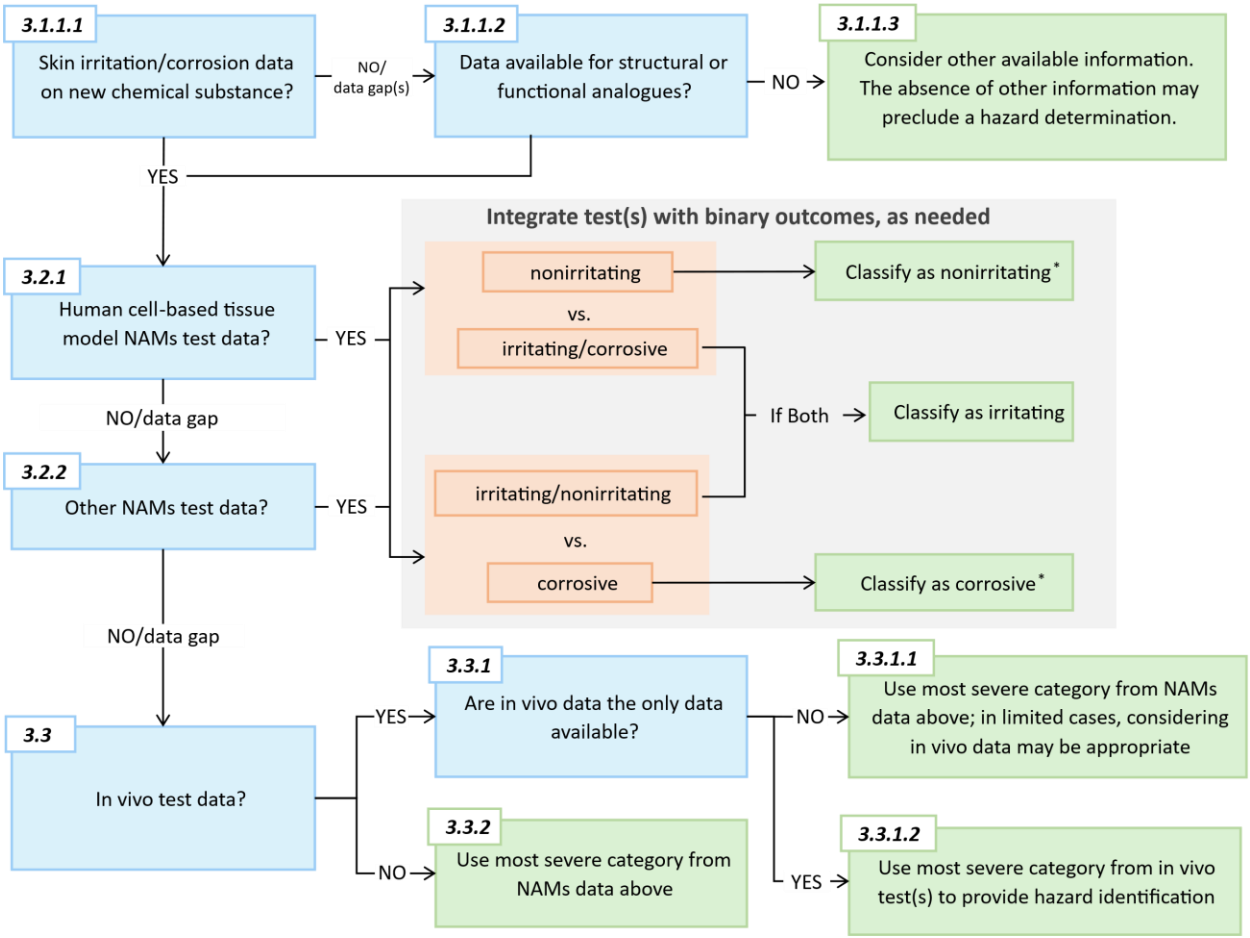

Figure 1. Decision Framework to Identify Skin Irritation/Corrosion Based on Data Availability. Flowchart demonstrating the basic parameters and prioritization order of data within the decision framework for identification of skin irritation/corrosion potential for of new chemical substances. Blue boxes indicate major decisions based on prioritization of data type within the framework, orange boxes represent NAMs by hazard identification capability, and green boxes indicate outcomes. *In the case of conflicting data (e.g., results of nonirritating in one test and

corrosive in another test method), consider additional test data and other information. Further details and references are provided within the text.

4 OECD In Vitro, Ex Vivo and In Chemico Skin Irritation/Corrosion Test Guideline Summaries

Information has been extracted from the OECD's *Guidance Document (GD) 203 on an Integrated Approach on Testing on Assessment (IATA) for Skin Corrosion and Irritation* and additional details can be found in GD 203 (OECD, 2017) and the OECD TGs (OECD, 2004c, 2015b, 2019, 2021). Table 2 excludes test methods not published by OECD and the list is not intended to be an exhaustive list of skin irritation/corrosion NAMs applicable for TSCA decisions.

Table 2: Test Principles, Applicability Domains and Test Outcomes for Skin Irritation and Corrosion Test Methods

Method or Approach	Principle of the Test	Applicability Domain and Limitations [8]	Categorization and Outcome Type
OECD TG 439: *In Vitro* Skin Irritation: Reconstructed Human Epidermis Test Methods (OECD, 2021)	Test material(s) are applied to reconstructed human epidermis (RhE), a multilayered, highly differentiated model of the upper layers of human skin (i.e., epidermis). Cell viability is quantitatively measured as conversion of the vital dye MTT (3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide, Thiazolyl blue; CASRN 298-93-1) into blue formazan salt. The test is based on the premise that irritant materials penetrate the stratum corneum and cause damage to keratinocytes in underlying layers. Damaged keratinocytes release mediators that initiate an inflammatory reaction that causes dilation and increased permeability of the endothelial cells which produces the erythema and oedema observed in vivo. Irritant chemicals are identified using initiating events in the cascade (i.e., tissue damage) measured as decreased tissue viability below 50% of the negative controls.	• Applicable to monoconstituent substances and mixtures. • Applicable to soluble or insoluble solids, aqueous or nonaqueous liquids, semi-solids, and waxes. • Not applicable to gases or aerosols. Careful consideration of vapour pressure is advised. • Substances absorbing light in the same range as formazan dye and substances able to directly reduce the vital dye MTT to blue formazan salt may interfere with the tissue viability measurements and require the use of adapted controls for corrections. <u>Limitations</u>: • Cannot identify mild irritants (GHS category 3) • Cannot discriminate between irritants and corrosives.	*Outcome type:* Binary As a stand-alone test method, it can identify the following outcomes: • Nonirritating (GHS No Category) • Irritating/corrosive

[8] Published OECD Test Guidelines provide additional details in the "*Initial Considerations and Limitations*" Section.

Method or Approach	Principle of the Test	Applicability Domain and Limitations [8]	Categorization and Outcome Type
OECD TG 431: *In Vitro* Skin Corrosion: Reconstructed Human Epidermis Test Method (OECD, 2019)	Test material(s) are applied to reconstructed human epidermis (RhE), a multilayered, highly differentiated model of the upper layers of human skin (i.e., epidermis). Cell viability is quantitatively measured as conversion of the vital dye MTT into blue formazan salt immediately after exposure. The test is based on the premise that corrosive materials penetrate the stratum corneum and are cytotoxic to the cells in underlying layers. Corrosive chemicals are identified based on their ability to decrease cell viability below defined thresholds.	• Applicable to monoconstituent substances and mixtures. • Applicable to soluble or insoluble solids, aqueous or nonaqueous liquids, semisolids, and waxes. • Not applicable to gases or aerosols. Careful consideration of vapour pressure is advised. • Substances absorbing light in the same range as formazan dye and substances able to directly reduce the vital dye MTT to blue formazan salt may interfere with the tissue viability measurements and need the use of adapted controls for corrections. <u>Limitations</u>: • Cannot discriminate between skin irritants (GHS Category 2) and chemicals not requiring classification (GHS No Category) • Can discriminate between corrosive GHS Sub-categories 1A versus 1B-and-1C versus No Category, but not between Sub-category 1B and 1C unless using other test methods within an IATA (OECD, 2017).	*Outcome type:* Binary with partial subcategorization for corrosive properties As a stand-alone test method, it can identify the following outcomes: • Corrosive (GHS Category 1A or Category 1B-and-1C) • Irritating/nonirritating
OECD TG 430: *In Vitro* Skin Corrosion: Transcutaneous Electrical Resistance Test Method (OECD, 2004c)	Test material is applied to rat skin discs in a two-compartment test system for up to 24 hours. Loss of normal stratum corneum (SC) integrity barrier function, caused by corrosive effects on the skin and erosion of the SC, is measured as a reduction in the transcutaneous electrical resistance (TER) below a threshold level.	• Applicable to monoconstituent substances and mixtures. • Applicable to soluble or insoluble solids, aqueous or nonaqueous liquids, semisolids, and waxes. • Not applicable to gases or aerosols. Careful consideration of vapour pressure is advised. <u>Limitations</u>: • Cannot distinguish among the skin corrosive subcategories (GHS Category 1A, 1B, 1C) • Cannot discriminate between skin irritants (GHS Category 2) and chemicals not requiring classification (GHS No Category) unless using other test methods within an IATA (OECD, 2017).	*Outcome type*: Binary As a standalone test method, it can identify the following outcomes: • Corrosive, if: 1) mean TER ≤ 5 kΩ AND skin discs are obviously damaged (e.g., perforated), OR 2) mean TER ≤ 5 kΩ, AND • skin discs show no obvious damage, AND • mean disc dye content: test chemical ≥ 10M HCl positive control • Irritating/nonirritating if: 1) mean TER > 5 kΩ, OR

Method or Approach	Principle of the Test	Applicability Domain and Limitations [8]	Categorization and Outcome Type
			2) mean TER ≤ 5 kΩ, AND • skin discs show no obvious damage, AND • mean disc dye content: test chemical < 10M HCl positive control
OECD TG 435: In Chemico Membrane Barrier Test Method for Skin Corrosion (OECD, 2015b)	This Corrositex test method uses a proteinaceous macromolecular aqueous gel and an artificial membrane as a model for in vivo membranes (i.e., skin). Test chemicals are applied to the surface of the membrane and corrosivity-induced membrane barrier damage is detected. The damage in the membrane barrier may be measured by various procedures, including the use of pH sensitive dyes or some other property of the indicator solution below the barrier.	• Applicable to monoconstituent substances and mixtures, to soluble or insoluble solids, aqueous or nonaqueous liquids, and emulsions. • Applicable to organic and inorganic acids, acid derivatives, and bases. • Not applicable to chemicals with a pH in the range of 4.5 to 8.5 • Not applicable to gases and aerosols. Careful consideration of vapour pressure is advised. <u>Limitations</u>: • Test materials not causing detectable changes in the chemical detection system may not be tested (Eskes et al., 2012) • Cannot discriminate between skin irritants (GHS Category 2) and chemicals not requiring classification (GHS No Category)	*Outcome type:* Binary with full subcategorization for corrosive properties As a stand-alone test method, it can identify the following outcomes: • Corrosive (GHS Category 1A, 1B, or 1C) • Irritating/nonirritating

References

Eskes, C., Detappe, V., Koëter, H., Kreysa, J., Liebsch, M., Zuang, V., Amcoff, P., Barroso, J., Cotovio, J., Guest, R., Hermann, M., Hoffmann, S., Masson, P., Alépée, N., Arce, L.A., Brüschweiler, B., Catone, T., Cihak, R., Clouzeau, J., D'Abrosca, F., Delveaux, C., Derouette, J.P., Engelking, O., Facchini, D., Fröhlicher, M., Hofmann, M., Hopf, N., Molinari, J., Oberli, A., Ott, M., Peter, R., Sá-Rocha, V.M., Schenk, D., Tomicic, C., Vanparys, P., Verdon, B., Wallenhorst, T., Winkler, G.C., Depallens, O. (2012). Regulatory assessment of in vitro skin corrosion and irritation data within the European framework: Workshop recommendations. *Regul Toxicol Pharmacol, 62*(2), 393-403. doi:10.1016/j.yrtph.2011.10.015

OECD. (2004a). *Test Guideline 430, In Vitro Skin Corrosion: Transcutaneous Electrical Resistance Test Method (TER)*. Retrieved from Paris, France: https://www.oecd-ilibrary.org/environment/test-no-430-in-vitro-skin-corrosion-transcutaneous-electrical-resistance-test-ter_9789264071124-en

OECD. (2004b). *Test Guideline 431, In Vitro Skin Corrosion: Reconstructed Human Epidermis (RhE) Test Method*. Retrieved from Paris, France: https://www.oecd-ilibrary.org/environment/test-no-431-in-vitro-skin-corrosion-human-skin-model-test_9789264071148-en

OECD. (2004c). *Test No. 430: In Vitro Skin Corrosion: Transcutaneous Electrical Resistance Test (TER)*.

OECD. (2006). *Test Guideline 435, In Vitro Membrane Barrier Test Method for Skin Corrosion*. Retrieved from Paris, France: https://www.oecd-ilibrary.org/environment/test-no-435-in-vitro-membrane-barrier-test-method-for-skin-corrosion_9789264067318-en

OECD. (2013). *Test Guideline 439, In Vitro Skin Irritation: Reconstructed Human Epidermis Test Methods*. Retrieved from Paris, France: https://www.oecd-ilibrary.org/environment/test-no-439-in-vitro-skin-irritation-reconstructed-human-epidermis-test-method_9789264203884-en

OECD. (2015a). *Test No. 404: Acute Dermal Irritation/Corrosion*.

OECD. (2015b). *Test No. 435: In Vitro Membrane Barrier Test Method for Skin Corrosion*.

OECD. (2017). Guidance Document 203 on an Integrated Approach on Testing and Assessment (IATA) for Skin Corrosion and Irritation. Retrieved from https://www.oecd.org/publications/guidance-document-on-an-integrated-approach-on-testing-and-assessment-iata-for-skin-corrosion-and-irritation-9789264274693-en.htm

OECD. (2019). *Test No. 431: In Vitro Skin corrosion: reconstructed human epidermis (RHE) test method*.

OECD. (2021). *Test No. 439: In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method*.

Raabe, H. A., Costin, G. E., Allen, D. G., Lowit, A., Corvaro, M., O'Dell, L., Breeden-Alemi, J., Page, K., Perron, M., Flint Silva, T., Westerink, W., Baker, E., & Sullivan, K. (2025). Human relevance of in vivo and in vitro skin irritation tests for hazard classification of pesticides. *Cutaneous and ocular toxicology*, *44*(1), 1–21. https://doi.org/10.1080/15569527.2024.2387596

Rooney, J. P., Choksi, N. Y., Ceger, P., Daniel, A. B., Truax, J., Allen, D., & Kleinstreuer, N. (2021). Analysis of variability in the rabbit skin irritation assay. *Regul Toxicol Pharmacol, 122*, 104920. doi:10.1016/j.yrtph.2021.104920

United Nations. (2023). Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Tenth revised edition. Retrieved from https://unece.org/transport/dangerous-goods/ghs-rev10-2023

US EPA. (2010). *TSCA New Chemicals Program (NCP) Chemical Categories*. Retrieved from https://www.epa.gov/sites/default/files/2014-10/documents/ncp_chemical_categories_august_2010_version_0.pdf

US EPA. (2018). Strategic Plan to Promote the Development and Implementation of Alternative Test Methods Within the TSCA Program. June 22, 2018. U.S. Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention, Washington, DC. Retrieved from: https://www.epa.gov/sites/default/files/2018-06/documents/epa_alt_strat_plan_6-20-18_clean_final.pdf

Weil, C. S., & Scala, R. A. (1971). Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. *Toxicol Appl Pharmacol, 19*(2), 276-360. doi:10.1016/0041-008x(71)90112-8