Cutaneous and Ocular Toxicology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/icot20

Human relevance of in vivo and in vitro skin irritation tests for hazard classification of pesticides

Hans A. Raabe, Gertrude-Emilia Costin, David G. Allen, Anna Lowit, Marco Corvaro, Lindsay O'Dell, Julie Breeden-Alemi, Kathryn Page, Monique Perron, Tara Flint Silva, Walter Westerink, Elizabeth Baker & Kristie Sullivan

To cite this article: Hans A. Raabe, Gertrude-Emilia Costin, David G. Allen, Anna Lowit, Marco Corvaro, Lindsay O'Dell, Julie Breeden-Alemi, Kathryn Page, Monique Perron, Tara Flint Silva, Walter Westerink, Elizabeth Baker & Kristie Sullivan (24 Aug 2024): Human relevance of in vivo and in vitro skin irritation tests for hazard classification of pesticides, Cutaneous and Ocular Toxicology, DOI: 10.1080/15569527.2024.2387596

To link to this article: https://doi.org/10.1080/15569527.2024.2387596

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 24 Aug 2024.

Submit your article to this journal

Article views: 983

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=icot20

CUTANEOUS AND OCULAR TOXICOLOGY
https://doi.org/10.1080/15569527.2024.2387596

RESEARCH ARTICLE

OPEN ACCESS

Human relevance of in vivo and in vitro skin irritation tests for hazard classification of pesticides

Hans A. Raabe[a], Gertrude-Emilia Costin[a], David G. Allen[b#], Anna Lowit[c], Marco Corvaro[d], Lindsay O'Dell[c], Julie Breeden-Alemi[c], Kathryn Page[e], Monique Perron[c], Tara Flint Silva[c], Walter Westerink[f], Elizabeth Baker[g] and Kristie Sullivan[a]

[a]Institute for In Vitro Sciences, Gaithersburg, MD, USA; [b]Inotiv, Durham, NC, USA; [c]US Environmental Protection Agency, Washington, DC, USA; [d]Corteva Agriscience, Rome, Italy; [e]The Clorox Company, San Francisco, CA, USA; [f]Charles River Laboratories, Den Bosch, The Netherlands; [g]Physicians Committee for Responsible Medicine, Washington, DC, USA

ABSTRACT

Background: Test methods to inform hazard characterization and labeling of pesticides to protect human health are typically conducted using laboratory animals, and for skin irritation/corrosion the rabbit Draize test is currently required by many regulatory agencies. Although the Draize test is generally regarded to provide protective classifications for human health, new approach methodologies (NAMs) have been developed that offer more human relevant models that circumvent the uncertainty associated with species differences that exist between rabbits and humans. Despite wide applicability and use of these test methods across a broad range of chemicals, they have not been widely adopted for testing pesticides and pesticidal formulations. One of the barriers to adoption of these methods in this sector is low concordance with results from the Draize rabbit test, particularly for chemicals within the mild to moderate irritation spectrum.
Methods: This review compares and contrasts the extent to which available models used in skin irritation testing mimic the anatomy and physiology of human skin, and how each aligns with the known key events leading to chemically-induced adverse skin irritation and corrosion. Doing so fully characterizes the human relevance of each method.
Results: As alternatives to the rabbit Draize test, several protocols using *ex vivo*, *in chemico*, and *in vitro* skin models are available as internationally harmonized test guidelines. These methods rely on a variety of models of human skin, including excised rodent skin, synthetic biochemical models of barrier function, cell culture systems, and reconstructed human tissue models. We find these models exhibit biological and mechanistic relevance aligned with human skin irritation responses. Further, recent retrospective analyses have shown that the reproducibility of the Draize test is less than 50% for mild and moderate responses, with many of the replicate predictions spanning more than one category (*e.g.*, a moderate response reported in one study followed by a non-irritant response reported in another study).
Conclusions: Based on this comparative evaluation, we recommend top-down and bottom-up testing strategies that use the most human relevant *in vitro* test methods for skin irritation and corrosion classification of pesticides and pesticide formulations. To further discriminate among mild and non-irritant formulations, optimization of a cytokine release protocol and subsequent analyses of reference formulation test results is recommended.

ARTICLE HISTORY
Received 2 April 2024
Accepted 24 July 2024

KEYWORDS
skin irritation; skin corrosion; new approach methodologies; reconstructed human epidermis; pesticide; formulation; cytokine release; non-animal

Introduction

Pesticidal products are currently required to undergo a battery of acute mammalian assessments informally known as the acute 'six-pack', comprised of acute lethality *via* oral, dermal, and inhalation exposure routes, skin sensitisation, and irritation and corrosion testing of both the eye and skin. Tests evaluating both individual active ingredients and end-use formulations are used by regulatory agencies tasked with determining which hazard classifications, precautionary statements, first aid language, and/or personal protective equipment are necessary to protect consumers, professional handlers, and applicators. Historically, these are *in vivo* tests in mammals, but due to growing ethical and scientific concerns surrounding animal-based test methods, many alternatives for the acute 'six-pack' have been developed and evaluated. To validate alternatives for regulatory purposes, the expectation has been for alternative methods to show predictive performance as good as or better than the animal-based test. Whereas a comparison of predictive performance and human relevance would require a comparison of how well each method

CONTACT Kristie Sullivan ksullivan@iivs.org Institute for In Vitro Sciences, Gaithersburg, MD, USA.
[#]Current affiliation: International Collaboration on Cosmetic Safety

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

predicted human responses, sufficient human data are rarely available for comparison. As a surrogate, concordance with reference data obtained from testing of chemicals in the established animal tests has been used to evaluate the usefulness and limitations of the alternative assays.

Significant progress has been made towards the development of new approaches to replace animal tests. Recent examples include defined approaches for predicting skin sensitisation [1–4] and the GHS Mixtures Equation for determining the acute systemic toxicity of pesticide formulations [5]. Additionally, animal tests can be waived in certain circumstances, *e.g.* for acute dermal toxicity evaluation [6–9]. For assessing the eye irritation potential of cleaning products with antimicrobial claims (also known as antimicrobial cleaning products, or AMCPs), US EPA's Office of Pesticide Programs (US EPA OPP) has formally adopted an alternative test method framework which utilised *in vitro* and *ex vivo* eye irritation test methods in an integrated testing strategy [10]. More recently, a review of the human relevance and reliability of *in vivo*, *in vitro,* and *ex vivo* eye irritation and corrosion test methods for pesticides was conducted and found that *in vitro* and *ex vivo* methods are reflective of human biology, capture the necessary endpoints needed to assess the ocular toxicity of pesticide formulations, and are less variable than *in vivo* methods [11].

This review provides an assessment of the human relevance and reliability of the currently available skin irritation test methods, both the Draize rabbit skin test, and alternatives that have been developed and in many cases, adopted as OECD health effects test guidelines. Herein we identify the key anatomical structures and functional characteristics of human skin, as well as the biochemical mechanisms of skin damage, and the extent to which they are addressed by each test method. Doing so objectively demonstrates their applicability to identifying and classifying the acute skin irritation and corrosion potential of pesticide formulations.

Structure and function of human skin

In order to understand the human relevance of available *in vivo*, *ex vivo*, and *in vitro* methods, it is critical to first provide a characterisation of the structure and function of human skin, and how they relate to the key events in chemically-induced skin irritation and corrosion. This can then be compared and contrasted to rabbit, rat, and mouse skin, as well as artificial biobarrier, cell, and tissue models used in irritation and corrosion testing.

Vertebrate skin is comprised of two major functional layers: the nonvascular epidermis and the connective dermis, which is anchored to the subcutaneous hypodermis. Skin provides protection from excessive water and electrolyte loss, as well as protection from mechanical injury and exposure to xenobiotics, foreign materials, and UV light. In mammals, these functions are performed to varying degrees both by skin and hair, which necessarily vary with body location as well as the phenotype of the individual, leading to substantial inter- and intra-species differences in skin morphology [12].

In humans, the full thickness of skin is comprised of both the epidermis and underlying dermis and is reported to be approximately 3 mm [13]. The epidermis is divided into four to five distinct layers, with a total thickness of ca. 50 μm [13] comprised of ca. 10 to 20 cell layers deep at varying levels of differentiation depending upon the location on the body [14]. *In vivo,* the layers in the epidermis are continually produced, originating from the innermost epidermal layer, the *stratum basale* (also referred to as the *s. germinativum*) which is comprised of proliferative basal epidermal keratinocytes (Figure 1). As keratinocytes in the *s. basale* layer divide and are displaced superficially, they begin terminal differentiation to establish a series of strata: *s. spinosum*, *s. granulosum*, *s. lucidum*, and the outermost *s. corneum*. The *s. corneum*, formed by dead, keratinised cells, is characterised as a cross-linked network of intercellular lipids, ceramides, and cholesterol which provide the primary barrier against chemical penetration into the skin to protect the underlying viable epidermal and dermal cells. The *s. corneum* in humans can vary in thickness depending upon the location of the body; Jung and Maibach [13] report a representative thickness of 17 μm. Barrier function is a major factor in the ability of skin to tolerate exposures to irritating chemicals more effectively than other external tissues such as ocular or mucosal epithelia. In effect, the degree of barrier function determines the level of exposure of the underlying viable skin cells to offending chemicals. Other cell types including melanocytes,

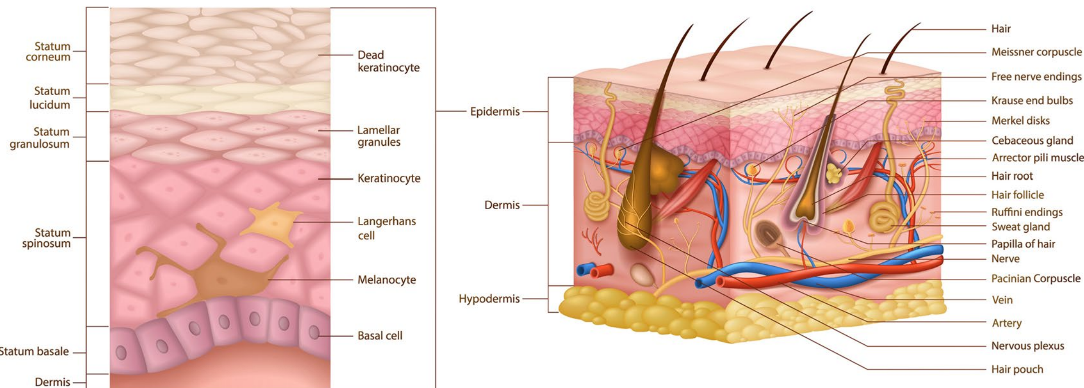


Figure 1. Structure of human skin. Diagrammatic cross section of human skin displaying epidermal layers, dermis, and appendageal skin structures. © Can Stock Photo Inc. / 10051252.

Langerhans cells, Merkel cells, and sensory nerves are resident in the epidermis in addition to keratinocytes; however, they do not contribute significantly to the skin's essential barrier function or biomass.

Just beneath the mammalian epidermis lie the two zones of the dermis: the superficial papillary dermis immediately beneath the epidermis, and the deeper reticular dermis. The papillary dermis consists of a dense population of fibroblasts, featuring numerous ridges (papillae) and valleys that create an irregular border between the epidermis and dermis [15]. The primary functions of the papillary layer are to provide nutrients to the epidermis and to regulate temperature, both of which are accomplished *via* the fine capillary network within the papillary layer. The reticular layer is the thickest layer of skin and is comprised of dense collagen fibres supported by sparse dermal fibroblasts. The dermis also contains the blood vessels, lymphatics, nerves, and nerve endings that sustain the organ.

Appendageal (or adnexal) skin structures arise from the reticular dermis; in humans these include hair, sweat glands, and sebaceous glands, the first two of which are lined by epidermal cells and pass through the epidermis *via* follicles and ducts, respectively (Figure 1). Hair follicle density in human skin is notably less than in most other mammals, and is reported to be 0.2 to 0.3 follicles per mm^2 on the arms and legs [16]. Appendageal structures are relevant to chemical-induced skin irritation in that they provide an alternate avenue for chemicals to bypass the defensive barrier of the *s. corneum*. While the relative contributions of follicular and intracellular skin penetration pathways vary among species, follicular penetration has been shown to be a potentially significant pathway of absorption for topical applications of certain chemicals [17–19]. Notably, in furred mammals, densely packed hairs extend above the epidermal surface to minimise dermal exposure to xenobiotics. In humans, however, thicker epidermal and dermal layers exist in the absence of fur to provide increased barrier function to the exposed epidermis.

Key events involved in skin irritation and corrosion

Skin exposure to neat chemicals, mixtures, and formulations can lead to a wide range of adverse responses, from mild inflammation reactions such as transient erythema and edoema, to severe reactions resulting in necrosis and subsequent scar formation [20]. Chemicals fall onto a continuum of cytotoxic and tissue lytic potencies, thus defining whether they are likely mild or moderate skin irritants or corrosives [21]. Chemicals with corrosive potential are both highly cytotoxic and can rapidly penetrate into the dermis [21], while irritant chemicals may be less cytotoxic and/or less likely to permeate beyond the epidermis. Skin corrosion and irritation are manifestations covering a spectrum of tissue responses with varying degrees of severity, persistence and depth of injury. Consequently, highly irritant/marginally corrosive chemicals may manifest as either corrosive or highly irritated tissue responses under varying circumstances and exposure conditions.

The redness, swelling, and blistering that define irritant contact dermatitis, clinically known as erythema and edoema, result from the signalling cascade initiated by the penetration of the chemical through the *s. corneum* and into the epidermal cells. Irritation and corrosion are delineated from allergy by exclusion, encompassing any localised inflammatory response *not* due to an immune response, while irritation is distinguished from corrosion based on the reversibility of the tissue damage. In clinical testing, irritant contact dermatitis is diagnosed by a lack of T-cell involvement in the aetiology, rather than the involvement of a specified process or pathway [22]. Consequently, there are numerous biochemical pathways that can contribute to irritation [23,24]. Despite this complexity, it is useful to organise relevant mechanistic key

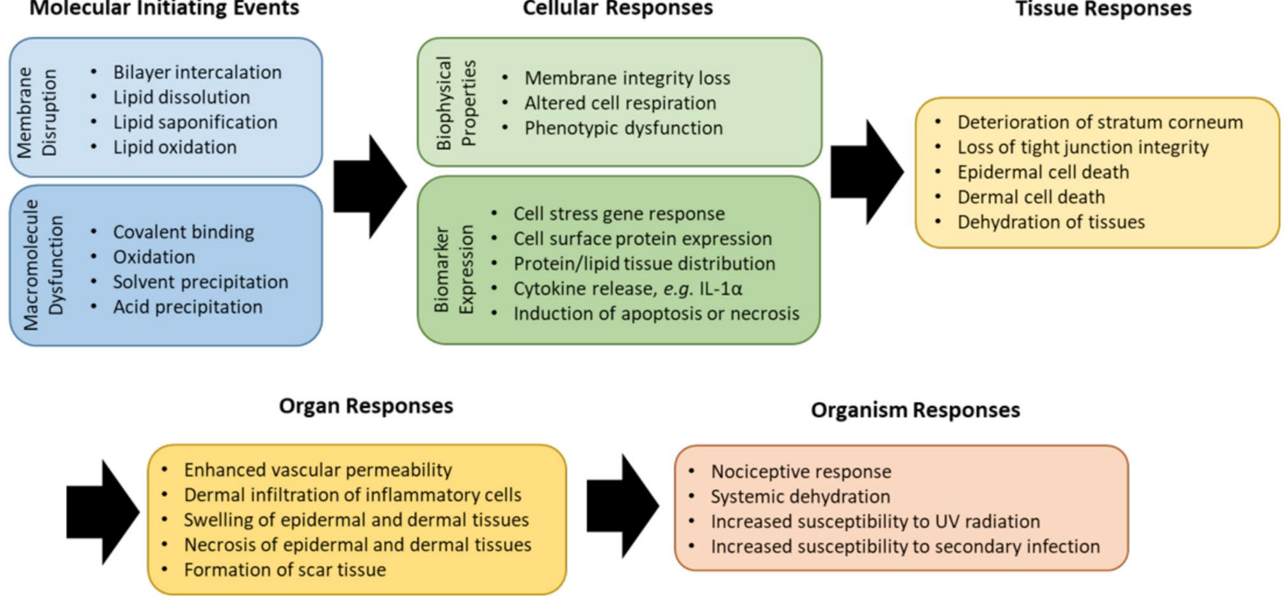


Figure 2. Generalised mechanisms of skin irritation and corrosion in mammals. Key events upstream of erythema and edoema that can be measured quantitatively include cytotoxicity, tissue dehydration, and cytokine release.

events into the framework of a generalised adverse outcome pathway (Figure 2), to help understand how available test methods might inform on the likelihood of an irritation or corrosion response in humans.

At the molecular level, initiating events causing membrane disruption and macromolecule dysfunction can occur through direct binding, inhibition, oxidation, precipitation, dissolution, or saponification of intracellular or extracellular components. Lipids, organic solvents, and surfactants can intercalate membrane bilayers and dissolve lipids that make up the extracellular matrix of the *s. corneum*. Oxidants and alkylators can affect lipids and proteins, inducing structural changes that lead to loss of macromolecular function. Bases saponify lipids, while acids can induce protein precipitation. While these mechanisms represent numerous possible initiating events, they inform directly on the types of materials that pose a risk to epithelial cells – electrophiles and oxidants, strong acids and bases, lipophilic solvents, *et cetera*.

Upon damaging or breaching the barrier function of the *s. corneum*, chemicals can injure the keratinocytes in the epidermis [22,25], and in more severe cases penetrate deeper affecting the stromal fibroblasts and accessory cells originating in the dermis. As denatured proteins and disrupted membranes fail to protect the underlying cells, these effects can increase the depth of penetration through epidermal and dermal tissues leading to enhanced damage. At the cellular level, the same molecular mechanisms that disrupt the barrier also disrupt cellular structures and organelles, leading to cell death. Following sufficient damage to cellular processes, local inflammation of dermal tissue is facilitated by the keratinocyte release of pre-formed membrane bound cytokine Interleukin-1 alpha (IL-1α) [26], as well as by local activation of small molecule, protein-based and lipid-based autacoids, reactive oxygen species, and other interleukins, and glycolipids [20,27]. IL-1α is a unique member of the IL-1 family, behaving as a dual-function cytokine and transcription factor that is constitutively present in healthy cells [28,29]. Its precursor can rapidly shuttle between cytosol and the nucleus, activating inflammatory cascades in the cytosol, while nuclear localisation inhibits inflammation [30]. Extracellularly, when released from cells under necrotic or hypoxic stress, IL-1α binds to the IL-1R1 receptor, activating neighbouring cells to produce a protective response to cell stress [31]. In response to these stress signals from dying epidermal cells, nearby cells rapidly induce transcription of secondary inflammatory mediators such as IL-6, IL-8, and Tumour Necrosis Factor alpha (TNF-α) [32,33] in neighbouring keratinocytes.

While damaged epidermal cells are the initiators of inflammatory cascades, the ultimate targets of these mediators are the endothelial and stromal layers of local blood vessels in the dermal layer. Circulating macrophages and neutrophils are recruited locally in response to IL-8, while IL-1α and TNF-α induce expression of intercellular adhesion molecules (ICAM-1) on endothelial cells and fibroblasts to allow adhesion and migration of the neutrophils to the site of chemical injury. These signalling cascades result in increased permeability of endothelia, allowing for dilation of capillaries and fluid accumulation leading to erythema and edoema, which are the observable primary apical events in skin irritation.

Skin corrosion is differentiated from skin irritation based on the lack of reversibility of the tissue damage. In corrosion, the depth of the cellular damage and necrosis substantially involves cellular and architectural components of the dermis including dermal fibroblasts and the collagen matrix. Whereas the epidermis can rapidly recover by epidermal keratinocyte sheet migration from surrounding unaffected epidermal tissues to cover the site of the wound, chemical damage into the dermis can induce fibroblast collagen deposition and scar formation. Furthermore, significant damage into the dermis may cause permanent local loss of accessory structures including hair follicles, sweat and sebaceous glands, and in extreme cases induce destruction of the capillary beds. While eschar is typically observed at the site of exposure to corrosives, it is not unusual for edoema and erythema to be observed at the periphery to the corrosive damage, as one would expect a gradient of decreasing concentration of the offending chemical distally from the site of exposure with progressively less tissue damage.

Structure and function of skin models used for irritation and corrosion testing

Having characterised the structure and function of human skin and identified both the upstream key events and the downstream apical manifestations associated with chemical-induced skin irritation and corrosion, in this section the similarities and differences between human skin and that of the *in vivo*, *ex vivo*, *in chemico* and *in vitro* models used for skin irritation and corrosion testing are described. Additionally, the mechanistic basis for each of the models as they are used to inform on skin irritation and corrosion are presented, as well as the key events and endpoints utilised.

The current test methods addressing skin corrosion and irritation in regulatory guidelines as well as those in development are listed in Tables 1–3. The test methods are divided into three tables to facilitate comparison of methods that have similar advantages and limitations. The regulatory use (where applicable) is described alongside assay and protocol characteristics influencing human relevance for comparative reference. For most of the alternative test methods, sensitivity, specificity and accuracy data were determined relative to the available rabbit Draize data, and consequently these performance statistics do not directly convey their abilities to predict human skin irritation potential. Whereas most of the alternative test methods listed have undergone varying levels of evaluation or validation for their ability to replace the Draize test, not all test methods are used for regulatory submissions. This review seeks to present the available methods to inform the endpoint of skin irritation and corrosion, whether or not they have undergone a specific validation process or are included in OECD test guidelines, as such methods may still be useful towards our efforts.

All of the models used for skin irritation/corrosion testing differ from native human skin to various degrees in their structural and functional characteristics. Histological images of mammalian skin models that are used in irritation and

Table 1. *In Vivo* and *Ex Vivo* Test Methods.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
OECD TG 404: Acute Dermal Irritation/Corrosion [57] / EPA OPPTS 870.2500	*Overview:* The test method was developed in 1944 [34]. The test involves clipping away dorsal fur approximately 24 hours before applying test material to the newly exposed skin. The skin is then covered with a gauze patch, which is held in place for up to 4 hours [35]. The test results are qualitatively assigned to skin irritation scores based on subjective examination of the nature and severity of lesions; reversibility of lesions is assessed for up to 14 days. *Classification:* The test method is used to assign one of four EPA categories (I, II, III, or IV), or one of three GHS categories (1, 2, 3, or No Category). The EPA and GHS criteria for designation are independent and are not harmonised. While the Draize test is conducted identically for both OECD [35] and US EPA submissions, the evaluation of the dermal irritation scores are ultimately interpreted in a different manner by the two regulatory systems [36]. For GHS and EPA, dermal corrosion in one animal can trigger the category 1 classification, however, in the GHS system the results are averaged over several animals. The Office of Pesticide Program (OPP) within US EPA previously provided two options for the calculation of the Primary Irritation Index (PII) for Dermal Irritation studies: 1) each animal's erythema and edoema scores for the 1, 24, 48 and 72 hours scoring intervals are added separately; then all six values are added together and divided by the (number of test sites x 4 scoring intervals); 2) the 1, 24, 48, and 72 hours erythema and edoema scores for all animals are added and divided by the (number of test sites x 4 scoring intervals)[37]. However, the current harmonised guideline does not recommend averages but rather a full reporting of individual rabbit scores and full description of observations during the entire timespan of the test for each animal to assist in interpreting the numeric scores recorded [38]. The GHS also provides two options for the calculation of the PII as follows: 1) for both erythema and edoema end points, all 24, 48 and 72 hours scores for each animal are added together and divided by the (number of test sites x 3 scoring intervals); 2) for both erythema and edoema, all 24, 48, and 72 hours scores of each animal individually are added and divided by the number of scoring intervals [39]. *Performance:* Qualitative scoring of erythema and edoema is subjective and not particularly reproducible. A recent analysis of over 900 substances tested *in vivo* more than once found that chemicals classified as moderate or mild irritants at least once (in both US EPA and GHS schemes) have approximately a 40-60% chance of that same classification being reported in a subsequent study. Substances were often re-classified to a non-adjacent category, including a considerable proportion (5-10%) of re-classification between corrosive and mild or non-irritant categories [40].	*Test system:* The test is conducted in live albino rabbits. *Mode of action:* The test method evaluates apical outcomes in rabbit skin, typically including erythema, edoema, and eschar formation, and any other notable lesions or changes. *Exposure time and washing:* Three test patches are applied to rabbit skin at a dose rate of 500 µL or 500 mg/6 cm^2 – generally under semi-occlusive conditions. Up to three rabbits may be used. The first patch is removed after 3 min. If no serious skin reaction is observed, the second patch is removed after 1 hour. If observations indicate that exposure can be continued humanely, the third patch is removed after 4 hours and the responses graded. If a corrosive effect is observed after either 3 min or 1 hour of exposure, the test is immediately terminated by removal of the remaining patches and euthanasia of the animal. The skin may be rinsed after 4 hours, and the exposure sites evaluated and scored after 60 minutes, 24, 48, and 72 hours, and up to 14 days to evaluate for reversibility of effects [35]. According to EPA TG OPPTS 870.2500 and OECD TG 404, the post-treatment evaluation period should be sufficient to fully evaluate the reversibility or irreversibility of the effects observed to delineate between corrosion and irritation. *Test material limitations:* The subjectivity of visual evaluation is especially prominent in the case of compounds that colour the skin [20]. Volatility may affect the amount of chemical in contact with skin, and the test is not considered applicable to the testing of gases and aerosols [41]. *Human relevance:* the endpoints of the test method include apical outcomes similar to those observed clinically in humans. The mode of action for corrosivity is expected to be highly relevant, but the pathways for skin irritation may differ between species. The model uses native non-human mammalian skin complete with adnexal structures similar to those observed in human skin, but the significantly thinner skin, and high density of hair follicles relative to humans makes the test method prone to notable over-prediction.

(*Continued*)

Table 1. Continued.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
OECD TG 430: *In Vitro* Skin Corrosion: Transcutaneous Electrical Resistance Test Method (TER) [42]	*Overview:* This test method utilises *ex vivo* rat skin discs to identify corrosives by their ability to produce a loss of normal *s. corneum* integrity and barrier function as measured by decreases in transcutaneous electrical resistance (TER) below a threshold. Since as many as fifteen skin discs can be obtained from one rat, the test method allows for a reduction and refinement in animal use for corrosion testing [42]. *Classification:* The test method can discriminate between corrosive (GHS Category 1) and non-corrosive substances but cannot further subcategorise corrosive substances. For non-corrosive substances, other test methods must be utilised to delineate the categories of irritation [42]. *Performance:* Based upon a validation study [43] and other published studies [44,45], an overall sensitivity of 94% (51/54) and specificity of 71% (48/68) when compared to the *in vivo* rabbit test is presented for a database of 122 substances.	*Test system:* The test method uses skin discs, harvested post-mortem from 28 to 30-day old rats. *Mode of action:* Intact skin discs exhibit high (ca. ≥10 kΩ) TER. The rat skin discs undergo breakdown and breach of the *s. corneum* barrier function following exposure to corrosive material. *Exposure time and washing:* Test substances are applied to the epidermal side of the skin discs and maintained at room temperature for 24 hours, after which the substances are rinsed with copious water, and prepared for post-treatment TER determinations. *Test material limitations:* Incomplete removal of test material can interfere with TER measurements and lead to readings > 20 kΩ [42]. Further, non-corrosive detergents, surfactants and other materials can lead to reduced TER by masking lipids and making the *ex vivo* skin discs more ion-permeable. When TER readings are near or below 5 kΩ but visual damage to the skin disc is not observed, the penetration of sulforhodamine B dye can be assessed to identify such false positives [42]. *Human relevance:* the mode of action for corrosivity is expected to be highly relevant, but the test method endpoint may not be highly relevant. The model uses native non-human mammalian skin complete with adnexal structures similar to those observed in human skin, but the significantly thinner skin, and high density of hair follicles relative to humans makes the test method prone to notable over-prediction.
Skin Integrity Function Test (SIFT) for Skin Irritation	*Overview:* The test method is based upon measuring changes in the barrier function in *ex vivo* mouse skin using the endpoints of trans-epidermal water loss (TEWL) and electrical resistance (ER). *Classification:* Chemicals tested in the SIFT assay are predicted skin irritants if the ratios of the pre- and post-application values for either TEWL or ER are greater than five-fold. Conversely, if the ratio of both parameters is less than five-fold then the chemical is deemed non-irritant [46,47]. *Performance:* The test method was initially evaluated by ECVAM in a multiphase validation. Twenty chemicals (9 irritants and 11 non-irritants) were tested in Phase I in the lead lab to evaluate protocol improvements made as a result of prevalidation activities. Based upon the results of Phase I, the Validation Management Team concluded that the predictive ability of the test method was not sufficient to proceed Phase II [48].	*Test system:* The test system uses *ex vivo* mouse skin. *Mode of action:* The test measures TEWL and changes in electrical resistance as the barrier function of the *s. corneum* is compromised [20]. *Exposure time and washing:* A validated protocol has not yet been developed for this test method. *Test material limitations:* Incomplete removal of test material can interfere with TER and TEWL measurements. *Human relevance:* neither the single mode of action for skin irritation, nor the test method endpoint are relevant. The model uses native non-human mammalian skin complete with adnexal structures similar to those observed in human skin, but the significantly thinner skin, and high density of hair follicles relative to humans makes the test method prone to notable over-prediction.

corrosion test methods are shown for comparison to human skin in Figure 3.

Structure and function of Rabbit and rodent skin

Non-human mammalian species have been used historically for determining the skin irritation and corrosion potential of chemicals, and the *in vivo* rabbit has been used extensively following the methods first described by Draize [34] to cover the full spectrum of irritation potentials from identifying non-irritants to discriminating among corrosive subcategories. Test methods utilising *ex vivo* rodent skin have been developed, namely, the Transcutaneous Electrical Resistance test method (TER) for skin corrosion testing, and the Skin Integrity Function Test (SIFT) for skin irritation, using *ex vivo* rat and mouse skin, respectively. These methods offer improvements over the Draize rabbit test in both animal welfare and quantitative measurement.

Rabbit and rodent skin differ from human skin due to major species differences in the structure and function of the epidermis (Figure 3A). In rabbits, the epidermis develops from the germination layer (*s. basale, or s. germinativum*), but it differs from human epidermis in that it is limited to three strata: *s. basale, s. spinosum, and s. corneum*. Further, the rabbit *s. corneum*, as well as the full epidermis and dermis, are observed to be thinner than in humans, with the most substantial differences observed in the *s. corneum* and dermis. The thickness of New Zealand White rabbit skin tissue was reported to

Table 2. *In Chemico* Membrane Test Methods.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
OECD TG 435: *In Vitro* Membrane Barrier Test Method for Skin Corrosion (Corrositex®) [49]	*Overview:* The test method uses an *in chemico* membrane barrier that can be used to identify corrosive chemicals and allows the sub-categorisation of corrosive test chemicals according to the GHS and US Department of Transportation (DOT) classification systems. Prior to routine use of this method, laboratories must test twelve proficiency substances to demonstrate technical proficiency. *Classification:* Based upon the breakthrough time, the test method can classify into all three GHS corrosion subcategories, as well as identify non-corrosive substances. Those highly corrosive substances that penetrate the bio-barrier in ≤ 3 minutes are predicted to be GHS Category 1 A corrosive, while substances that do not penetrate the bio-barrier within the established observation time of up to 4 hours are predicted non-corrosive. However, for non-corrosive substances the test method does not provide further information on the skin irritation potential; other test methods must be utilised to discriminate the irritation potential of non-corrosive substances [49]. *Performance:* Validation studies have been completed for the Corrositex® test method [50] showing an overall accuracy to predict skin corrosivity of 79% (128/163), a sensitivity of 85% (76/89), and a specificity of 70% (52/74) for a database of 163 substances and mixtures [51]. Corrositex® meets the performance standards outlined by the National Toxicology Program Interagency Centre for the Evaluation of Alternative Toxicological Methods [52].	*Test system:* The test system comprises two components: a synthetic macromolecular bio-barrier and a chemical detection system (CDS). The bio-barrier membrane model is composed of lipids, proteins, glycoproteins, and low molecular weight substances that model the *s. corneum* barrier of skin. *Mode of action:* The test method addresses the rate of bio-barrier membrane penetration as a result of dissolution, denaturation, or other mechanisms associated with skin barrier protein and lipid denaturation. The rates for breaching the bio-barrier are correlated to the corrosion subcategories. *Exposure time and washing:* A test chemical is applied to the bio-barrier membrane and the test system is monitored for up to 4 hours for the first indications of test substance breakthrough into the CDS. No rinsing is performed. *Test material limitations:* The assay principle is limited to chemicals that can disrupt the bio-barrier. Chemicals tested must be outside the pH range 4.5 – 8.5. Multiple colorimetric indicators are compatible to overcome spectral interference with coloured test materials. *Human relevance:* the mode of action for corrosivity is expected to be highly relevant. The model is not based upon any cellular or biological structure and is not human relevant.
In Vitro Macromolecular DERMAL Irritection® Test Method	*Overview:* The Dermal Irritection® assay utilises an *in chemico* membrane substrate comprised of a covalently crosslinked mixture of keratin and collagen in the presence of an indicator dye as a model of skin barrier function. *Classification:* The extent of indicator dye release and protein denaturation may be quantitated by measuring the changes in optical density of the reagent solution at 450 nm. The changes in the optical density of the reagent solution are evaluated to determine a skin irritancy score designed to relate to the potential skin irritancy of the test chemical [53]. *Performance:* This test method has not undergone formal validation for regulatory application and is not currently approved for international regulatory classification and labelling purposes.	*Test system:* The test system comprises a covalently crosslinked protein membrane of keratin and collagen in the presence of an indicator dye. *Mode of action:* When applied to the synthetic membrane, irritant chemicals disrupt the ordered structure of keratin and collagen and thereby release the bound indicator dye into the indicator reagent solution under the membrane. Additionally, irritant chemicals also induce conformational changes in the globular proteins found in the reagent solution. *Exposure time and washing:* A validated protocol has not yet been developed for this test method. *Test material limitations:* Although this test method is not yet approved for use to identify skin irritants, test material limitations will likely be similar to those for the Ocular Irritection® assay, namely some coloured chemicals, protein precipitating chemicals, concentrated surfactants, and highly volatile chemicals [54]. *Human relevance:* the single mode of action for skin irritation is not expected to be especially relevant. The model is not based upon any cellular or biological structure and is not human relevant.

be 1.21 ± 0.04 mm, while the epidermis and dermis were reported to be approximately 0.03 and 1.18 mm, respectively [74]. Similarly, the full thickness of rat skin is reported to be ca. 2.1 mm, and the epidermis and *s. corneum* are reported to be 0.032 and 0.018 mm, respectively, while for mouse skin the full thickness is reported to be 0.8 mm, and the epidermis and *s. corneum* are reported to be 0.013 and 0.005 mm, respectively [13,75]. Whereas these metrics are representative of adult skin derived from the ventral or dorsal tissues and may vary considerably with location on the body and age of the animals, the test method protocols for the rat TER and mouse SIFT specify the age of the animals and body location for collection. In contrast, these layers are two- to three-fold thicker in human skin in order to provide sufficient protection from xenobiotics [76].

Further, methods utilising mammalian skin necessitate the removal of fur by shaving or shearing to allow for application of test material to the skin surface, exposing not only the epidermis but also inadvertently exposing hair follicles to test material application. Hair follicle density differs significantly between rabbit, rodent, and human skin: in humans, hair follicle density is reported to be 0.2 to 0.3 follicles per mm^2 on the arms and legs [16], while in rabbits and rodents, the dermis contains approximately 20-fold higher density of hair follicles, with 4 to 5 follicles per mm^2 [76]. Similarly, the trans-epidermal permeation rates measured for key reference chemicals differ by well over 10-fold between rodent and human skin [77] with the following general rank order of permeability determined for select chemicals: permeability in rabbit > rat > human [78,79]. In summary, all of the

Table 3. *In Vitro* Cell Culture Test Methods.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
OECD TG 431: *In Vitro* Skin Corrosion: Reconstructed Human Epidermis (RhE) Test Method [55]	*Overview:* This test method quantitatively measures a substance's ability to induce cytotoxicity in three-dimensional reconstructed human epidermis (RhE) models, based on the premise that corrosive chemicals are able to rapidly penetrate the *s. corneum* and are highly cytotoxic to the cells in the underlying layers. Currently five commercially-available RhE tissues (EpiSkin™ Standard Model, EpiDerm™ SCT, SkinEthic™ RHE1, epiCS®, and LabCyte EPI-MODEL24) have been approved within the OECD Test Guidelines program for identifying corrosive (Category 1) and non-corrosive (No Category) substances, as well as subcategorization (between Category 1 A and 1B/1C) of corrosive substances within the GHS classification system [55]. *Classification:* Chemicals which do not reduce tissue viability below the test method-specific threshold can be identified as non-corrosive according to both the GHS and EPA classification systems. For non-corrosive substances, other test methods must be utilised to identify irritation potential. OECD TG 431 can be used in a bottom-up or top-down approach [41]. *Performance:* Validation studies of ca. 79 substances in the 5 tissue models demonstrated that the overall accuracy varied from 69.8% to 78.8%; the rate of correct non-corrosive predictions varied from 71.6% to 79.3%; the rate of correct GHS 1B/1C predictions varied from 60.7% to 76.3%; and the rate of correct 1 A predictions varied from 83.3% to 87.5%. Furthermore, no GHS 1 A substances were under-predicted as non-corrosive [55].	*Test system:* This test method uses a reconstructed human epidermis derived from human keratinocytes cultured at the air-liquid interface, complete with an intact *s. corneum*. *Mode of action:* The RhE tissue models essential events in skin corrosion (i.e. breach of the *s. corneum* barrier function and keratinocyte cell death), by measuring changes in cellular respiratory metabolism after a fixed exposure (via quantitative succinate dehydrogenase activity using MTT). *Exposure time and washing:* Exposure times vary for different RhE models, but are generally short exposures ranging from 3 to 60 minutes where discrimination between non-corrosives, and subcategories of corrosive substances is relevant (the EpiSkin™ test method also includes a 240 minute exposure time). After exposure, the RhE models are thoroughly rinsed prior to conducting the MTT viability assessment. No post-treatment expression incubations are conducted. *Test material limitations:* While coloured compounds and those that interact directly with the MTT molecule can interfere with results, these confounding effects can be measured and corrected using specific controls when necessary. Where indicated, high-performance liquid chromatography is used to separate interfering compounds before quantifying the absorption of the formazan product. OECD TG 431 indicates that physicochemical properties are otherwise not excluded from applicability [55]. *Human relevance:* The mode of action for corrosivity is expected to be highly relevant. The model uses human-sourced cells to reconstruct human skin architecture. Although the skin models are generally more permeable than native human skin, and no adnexal structures are present, the model is highly human relevant.
OECD TG 439: *In Vitro* Skin Irritation: Reconstructed Human Epidermis (RhE) Test Method [56]	*Overview:* This test method quantitatively measures a substance's ability to induce cytotoxicity in RhE tissue models, based on the premise that skin irritant chemicals are able to penetrate the *s. corneum* and induce cytotoxic effects in the cells within the epidermis. Also known as the Skin Irritation Test (SIT). Currently seven commercially-available RhE tissues (EpiSkin™, EpiDerm™ SIT EPI-200, SkinEthic™ RHE, epiCS®, LabCyte EPI-MODEL24 SIT, Skin+®, and KeraSkin™ SIT) have been approved within the OECD Test Guidelines program for discriminating between GHS Category 2 skin irritants and those not requiring classification within the GHS classification system and within the framework of an Integrated Approach to Testing and Assessment (IATA) [41, 56]. *Classification:* Chemicals which do not reduce tissue viability below the test method-specific threshold post-exposure are classified as No Category by GHS. For remaining substances, the test method does not provide further information on the skin corrosion potential; thus other test methods must be utilised to discriminate between GHS Category 2 skin irritants and GHS Category 1 corrosive substances. The test method cannot independently distinguish between GHS Category 2 and Category 3, but can be used to identify Category 2 irritants when used within an IATA framework. Furthermore, this method has not been formally evaluated for its ability to classify substances into EPA categories, although efforts are ongoing [57]. OECD TG 439 can be used in a bottom-up or top-down approach [41]. *Performance:* Each of the approved tissue-specific protocols have achieved the following minimum performance criteria: sensitivity 80%, specificity 70%, and accuracy 75% [56], and as such a wide range of chemicals and products have been tested in these assays. Of relevance to certain pesticides and mixtures, a study comparing *in vitro* and *in vivo* data for 65 agrochemical formulations revealed an overall accuracy of 54% (based on 65 agrochemical formulations), a sensitivity of 44% (based on 25 formulations) and a specificity of 60% (based on 40 formulations) as compared to the rabbit test [58]. Further analyses reveal that of the 14 formulations classified as Cat 2 by the rabbit test, but predicted as non-irritant by the *in vitro* SIT, 7 formulations resulted in <85% relative viability demonstrating induction of cytotoxic effects, albeit above the prediction threshold.	*Test system:* This test method uses a reconstructed human epidermis derived from human keratinocytes cultured at the air-liquid interface, complete with an intact *s. corneum*. *Mode of action:* The RhE tissues model essential events in skin irritation (i.e. breach of the *s. corneum* barrier function and keratinocyte cytotoxicity), by measuring changes in cellular respiratory metabolism after a fixed exposure (via quantitative succinate dehydrogenase activity using MTT). *Exposure time and washing:* Single fixed exposure times vary for the different RhE models but are generally shorter exposures ranging from 15 to 60 minutes where discrimination between moderate skin irritants and mild/non-irritants is relevant. After exposure, the RhE models are thoroughly rinsed, and incubated in a post-treatment expression incubation of 42 hours, prior to conducting the MTT viability assessment. *Test material limitations:* While coloured compounds and those that interact directly with the MTT molecule can interfere with results, these confounding effects can be measured and corrected using specific controls when necessary. Where indicated, high-performance liquid chromatography (HPLC) is used to separate interfering compounds before quantifying the absorption of the formazan product. OECD TG 439 indicates that physicochemical properties are otherwise not excluded from applicability [56]. *Human relevance*: The modes of action modelled for skin irritation are expected to be highly relevant. The model uses human-sourced cells to reconstruct human skin architecture. Although the skin models are generally more permeable than native human skin, and no adnexal structures are present, the model is highly human relevant. Furthermore, the combination of the initial chemical exposure time, and the subsequent post-treatment expression incubation time allows for multiple mechanisms of skin irritation to be modelled. This protocol lends itself well to adding downstream cell response endpoints such as induction of gene expression changes and protein / cytokine production.

(*Continued*)

Table 3. Continued.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
Time-to-Toxicity (ET_{50}) Methods using RhE models	*Overview:* The Time-to-Toxicity protocol quantitatively measures a substance's ability to induce cytotoxicity in three-dimensional reconstructed human epidermis (RhE) models, based on the premise that highly irritant or corrosive chemicals will more rapidly penetrate the *s. corneum* and induce cytotoxic effects in epidermal cells relative to less irritant chemicals. Whereas highly irritant chemicals will be tolerated by the RhE tissues for relatively short exposure times, mild or non-irritant chemicals will be tolerated over notably longer exposure times [21]. By testing typically 3 to 5 exposure times covering a range of exposure times up to 24 hours, an ET_{50} value (the exposure time found to reduce cell viability by 50%) may be determined. The time-to-toxicity approach using a reconstructed human corneal epithelium is utilised in the US EPA eye irritation classification of antimicrobial cleaning products [10]. This test has been broadly used by industry for rank ordering of irritation potential and general irritation categorisation of substances and candidate formulations, particularly when used in conjunction with well-characterised benchmark materials [59,60]. *Classification:* No test methods utilising a Time-to-Toxicity protocol have been approved for any regulatory applications. However, a specific protocol using the MatTek EpiDerm™ EPI-200 was prevalidated for identifying skin irritants using an ET_{50} threshold of 60 minutes to discriminate R38 from non-classified substances [47]. (The authors note that risk-phrase categories like R38 (irritating to skin) are no longer used in the EU and can be interpreted as UN GHS Category 2.) This method has not been evaluated for its ability to classify substances into EPA or GHS categories. *Performance:* Since the pre-validation results cited above showed an overall accuracy of 58%, with an over-prediction rate of 37% and an under-prediction rate of 47%, no further test method refinement or validation was conducted.	*Test system:* This test method uses a reconstructed human epidermis derived from human keratinocytes cultured at the air-liquid interface, complete with an intact *s. corneum*. *Mode of action:* The RhE tissues model essential events in skin irritation (i.e. breach of the *s. corneum* barrier function and keratinocyte cytotoxicity, as well as cytokine release or expression), by measuring changes in cellular respiratory metabolism after a fixed exposure (e.g. MTT assay). The supplemental quantitation of cytokines in the culture medium by enzyme-linked immunoassays may be utilised where discrimination between mild and non-irritants is relevant. *Exposure time and washing:* Whereas the standard test method utilises a range of exposure times where the irritancy potential is unknown, single fixed exposure times of 24 hours continuous exposure for expected mild materials [61], or a 1 hour exposure, followed by 24 hours of a post-exposure [62] are conducted. After exposure, the RhE models are thoroughly rinsed. After 24 hours post-exposure, viability is assessed by the MTT assay. For optional cytokine analyses, the medium from under the tissues is collected and evaluated for cytokine content. *Test material limitations:* The same previously-mentioned limitations of the RhE-based test methods apply, and MTT interference can be overcome by using HPLC. *Human relevance*: The modes of action modelled for skin irritation are expected to be highly relevant. The model uses human-sourced cells to reconstruct human skin architecture. Although the skin models are generally more permeable than native human skin, and no adnexal structures are present, the model is highly human relevant. The concept of a time-to-toxicity is highly relevant as it recapitulates the exposure time-dependent chemical impacts observed *in vivo*, and is the basis for exposure time selections used subsequently in the RhE corrosivity and skin irritation test methods.
Phenion® Open-Source Reconstructed Epidermis (OS-Rep) *in vitro* Skin Irritation Test	*Overview:* The Phenion® OS-REep is a reconstructed epidermis method that has not been evaluated for inclusion in OECD Test Guidelines. *Classification:* According to GHS classification system, a chemical that reduces the relative tissue viability ≤ 50% is classified as skin irritant. If the relative tissue viability > 50%, a chemical is classified as a non-irritant [63,64]. *Performance:* The test method was assessed in a two-phase validation by multiple labs. Phase II validation study used 20 reference chemicals (10 non-irritants and 10 irritants) to evaluate the reliability and predictive ability of OS-REp test. The results revealed that the overall sensitivity of 90%, specificity of 70% and accuracy of 80%, which met the OECD PS criteria. The results of phase II validation study were comparable to the results of the phase I study [63,64].	*Test system:* This test method quantitatively measures a substance's ability to induce cytotoxicity in open-source reconstructed epidermis (OS-REp) model. The open-source model is cultured similarly to the commercially available RhE models defined in RhE Test Guidelines. *Mode of action:* The OS-REp models essential events in skin irritation (i.e. breach of the *s. corneum* barrier function and keratinocyte cytotoxicity), by measuring changes in cellular respiratory metabolism after a fixed exposure (e.g. MTT assay) where discrimination between skin irritants and non-irritants is relevant. *Exposure time and washing:* Test substances are applied to the epidermal tissue at room temperature for 35 minutes. After the exposure period, the tissue is rinsed 8 times and incubated at 37 °C/5% CO_2 for 42-hour post-exposure incubation before performing the MTT viability assessment. *Test material limitations:* This method relies on the RhE model and uses the same MTT endpoint as multiple OECD Test Guidelines for skin toxicology, so limitations of the RhE model will apply, and MTT interference can be overcome by using HPLC. *Human relevance*: The modes of action modelled for skin irritation are expected to be highly relevant. The model uses human-sourced cells to reconstruct human skin architecture. Although the skin models are generally more permeable than native human skin, and no adnexal structures are present, the model is highly human relevant. Furthermore, the combination of the initial chemical exposure time, and the subsequent post-treatment expression incubation time allows for multiple mechanisms of skin irritation to be modelled. This protocol lends itself well to adding downstream cell response endpoints such as induction of gene expression changes and protein / cytokine production.

(*Continued*)

Table 3. Continued.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
Phenion® FT *in vitro* Skin Irritation Test	*Overview:* The Phenion® FT is a reconstructed full-thickness skin model. The use of the full thickness model in the skin irritation test method has not been evaluated for inclusion in OECD Test Guidelines. *Classification:* According to a recent publication, this method was designed to classify antimicrobial mixtures according to the EPA classification system (> 70% relative tissue viability is classified EPA Category IV, ≤70% relative viability no category determination can be made). The typical GHS viability method can also be applied (≤ 50% relative tissue viability identifies a skin irritant, > 50% relative viability identifies a non-irritant) [65]. *Performance:* In a recent study using 24 antimicrobial formulations (6 non-irritants and 18 irritants), the predictivity of the Phenion® FT test was assessed. The results revealed an overall sensitivity of 78%, specificity of 83% and accuracy of 79%.	*Test system:* This test method quantitatively measures a substance's ability to induce cytotoxicity in a full-thickness (Phenion® FT) model. The FT model is cultured from keratinocytes and fibroblasts isolated from human skin samples, amplified in 2-D cell culture systems, and then subsequently seeded onto a collagen sponge. After a short interval under submersed conditions, the developing tissue equivalents are lifted to the air-liquid interface where they grow and differentiate into a 3-D reconstruction complete with a multi-layer epidermis and dermis [66]. *Mode of action:* The Phenion® FT models essential events in skin irritation (i.e. breach of the *s. corneum* barrier function and keratinocyte cytotoxicity), by measuring changes in cellular respiratory metabolism after a fixed exposure (e.g. MTT assay) where discrimination between skin irritants and non-irritants is relevant. *Exposure time and washing:* Test substances are applied to the epidermal tissue at room temperature for 15 minutes. After the exposure period, the tissue is rinsed and incubated at 37 °C/5% CO_2 for a 42-hour post-exposure incubation before cutting the tissue and performing the MTT viability assessment. *Test material limitations:* This method relies on the RhE model with minor differences, so limitations of the RhE model will apply, and MTT interference can be overcome by using HPLC. Additionally, the method has so far been tested with a limited number of substances. *Human relevance*: The model uses human-sourced cells to reconstruct human skin architecture. Although the skin models are generally more permeable than native human skin, the model is highly human relevant. Furthermore, the combination of the initial chemical exposure time, and the subsequent post-treatment expression incubation time allows for multiple mechanisms of skin irritation to be modelled. This protocol lends itself well to adding downstream cell response endpoints such as induction of gene expression changes and protein / cytokine production, especially with the presence of the dermis.
Neutral Red Uptake (NRU) Keratinocyte Cytotoxicity Test	*Overview:* The test quantitatively measures a substance's ability to induce damage to cell membranes in a monolayer of normal human epidermal keratinocytes (NHEK). NHEK cells are treated with a dilution series of test substance for 48 hours, followed by an assessment of viability using the neutral red uptake endpoint. Neutral red uptake can be used to determine the relative amount of viable cell-dependent accumulation of neutral red dye into lysosomes [67]. A dose-dependent relationship between test substance and relative neutral red uptake is used to rank order skin irritancy potential of aqueous-soluble chemicals. *Classification:* No classification schemes have been formally developed. Whereas the test method showed potential to provide a screening approach to rank ordering skin irritancy potential of ingredients, the test method performed poorly in relating human clinical data of personal care product formulations [67]. *Performance:* Information on reproducibility and repeatability is not currently available.	*Test System:* Human epidermal keratinocytes are cultured *in vitro* as a two-dimensional monolayer in multiwall tissue culture plates immersed in culture medium. Variations of this keratinocyte-based protocol include the use of normal human dermal fibroblasts. *Mode of action:* This assay is relevant to the epidermal cell death key event by directly assessing cell viability. Variations of this basic protocol have also included the collection of the culture medium and analyses of cytokines released from the keratinocytes. *Exposure time and washing:* Cells are cultured in 96-well plates and treated with a dilution series of test substance for 48 hours. Following the test substance treatments, cultures are rinsed and incubated with a neutral red solution for 3 hours. The amount of neutral red taken up by the cultures is determined spectrophotometrically [68]. *Test material limitations:* Aqueous-insoluble chemicals may not be compatible with the test method. Substances which absorb in the assay wavelength (540 nm) may interfere with the neutral red quantitation if test substance residues persist after rinsing. Spectrofluorimetric detection with excitation and emission at 530 and 645 nm can be used alternatively. Also, any chemical having a localised effect on the lysosomes can also cause non-cytotoxic effects on neutral red uptake [69]. *Human relevance*: Since the test method uses human-derived skin cells, the cellular responses to skin irritants are expected to be highly relevant. However, due to the lack of skin architecture and barrier function, the exposure kinetics are not human relevant, and thus may not be ideal for discriminating irritants from corrosives.

(*Continued*)

Table 3. Continued.

Test Method	Description and Applications i. *Overview* ii. *Classification* iii. *Performance*	Characteristics i. *Test system* ii. *Mode of action* iii. *Exposure time and washing* iv. *Test material limitations* v. *Human relevance*
Cytosensor Microphysiometer (CM) Test	*Overview:* The Cytosensor Microphysiometer (CM) test quantitatively measures the concentration of test material that causes a 50% decrease in the acidification rate (MRD_{50}) in a sub-confluent monolayer of mouse L929 fibroblasts, using a pH metre to detect changes in acidity. The MRD_{50} concentrations are used to compare and rank order skin irritancy potential of chemicals and ingredients, and the inclusion of benchmark materials can aid in irritation categorisation. This test method requires specialised equipment that currently is not readily available. *Classification:* Whereas the essential CM procedures were used by industry for both skin and eye irritation evaluations, only the eye irritation test method has undergone validation for identification of GHS category 1 substances using a top-down approach, and aqueous-soluble substances not classified by GHS using a bottom-up approach [70]. No comparable skin irritation classification scheme has been developed such that the method would not be useful for regulatory classification. *Performance:* A specific application compared the MRD50 values from testing surfactant containing products to the chronic toxicity in the 21-day cumulative irritancy patch test (CIPT). When predicting 21-day CIPT scores for surfactant containing products, MRD_{50} values of 50 mg/mL were found to be associated with acceptable market histories, while MRD_{50} values of greater than 78 mg/mL were found to have no positive reactions (NOEL) in the clinical test [71]. Performance parameters of the eye irritation application: inter-laboratory reproducibility for bottom-up assessments were 100% for GHS classification and 94.44% for EPA, and, for top-down assessments, were 87.62% for GHS with no values provided for EPA classification. Intra-laboratory repeatability was assessed based on calculated CVs for MRD_{50} scores from two different studies. Mean CVs tended to be higher for surfactant substances than non-surfactant substances, ranging from 10% to 24% [70].	*Test system:* Mouse L929 fibroblasts are cultured *in vitro* as a confluent monolayer on permeable tissue culture inserts immersed in culture medium. The non-human cell line is used as a model to evaluate non-specific cytotoxic events in epidermal keratinocytes and fibroblasts. *Mode of action:* This assay is relevant to the epidermal cell death key event by assessing changes in cellular metabolic rate (e.g. by changes in release of acidic metabolites). Since the test method does not include a skin-relevant barrier function, it is limited mechanistically in its ability to apply skin permeation kinetics to fully discriminate among the spectrum of skin irritants and corrosives. *Exposure time and washing:* Increasing concentrations of the substance are introduced, via flow-through, to the cells over 13.5 minutes, and then the cells are washed [70]. *Test material limitations:* The principle of the CM relies on measurable changes in pH to infer interference with metabolic activity, which limits this test to aqueous-soluble materials and stable aqueous suspensions that do not directly affect the pH of the medium. *Human relevance*: The test method uses non-human mammalian cells, such that the cellular responses to skin irritants may not be highly relevant. Furthermore, due to the lack of skin architecture and barrier function, the exposure kinetics are not human relevant, and thus may not be ideal for discriminating irritants from corrosives.

aforementioned mammalian skin models tend to be thinner than that of human skin and allow more rapid chemical permeation into the epidermal and dermal tissue. Skin permeation rates are directly related to the amount of chemical able to target viable tissues, and thus also directly related to skin irritation and corrosion potential.

Structure and function of in chemico and in vitro models of skin

In recent decades, several *in chemico* and *in vitro* models have been developed and implemented for irritation and corrosion testing, each with differing structural characteristics and mechanistically-based endpoints which can inform on the skin irritation and corrosion potential of chemicals and products [44,67,80–82]. The need for test methods that are more relevant to human biology and more reliable in performance has led to the refinement and optimisation efforts such that several of these methods have been validated and accepted within the OECD Guidelines for the Testing of Chemicals [50, 55,56, 83–87]. In contrast to the *in vivo* test method, in which a full spectrum of irritant and corrosive apical outcomes can be observed, *in chemico* and *in vitro* protocols address either corrosion or irritation endpoints separately by focusing on specific key events in the aetiology of skin irritation/corrosion. The *in chemico* models are characterised as biochemically relevant synthetic membranes or 'biobarriers' comprised of macromolecules similar or analogous to those found in human skin, and currently there are only two vendors providing these test methods (Table 2). In contrast, the *in vitro* methods are more diverse with human and mammalian cell-based 2-dimensional (monolayer) cell culture systems and 3-dimensional reconstructed skin models using human-derived cells (Table 3).

Discussion

Human relevance of skin irritation and corrosion test methods

Within global regulatory categorisation systems, substances are classified based upon the severity and persistence of skin reactions observed in rabbits. Whereas the specific responses observed in the rabbit test may not be highly relevant to predicting human responses, a framework which categorises substances based upon the likely severity and persistence of skin reactions in humans is a universal goal. As an example of a practical application, such a framework may discriminate between substances that are corrosive to skin and those that may cause reversible skin irritation, to inform the levels of

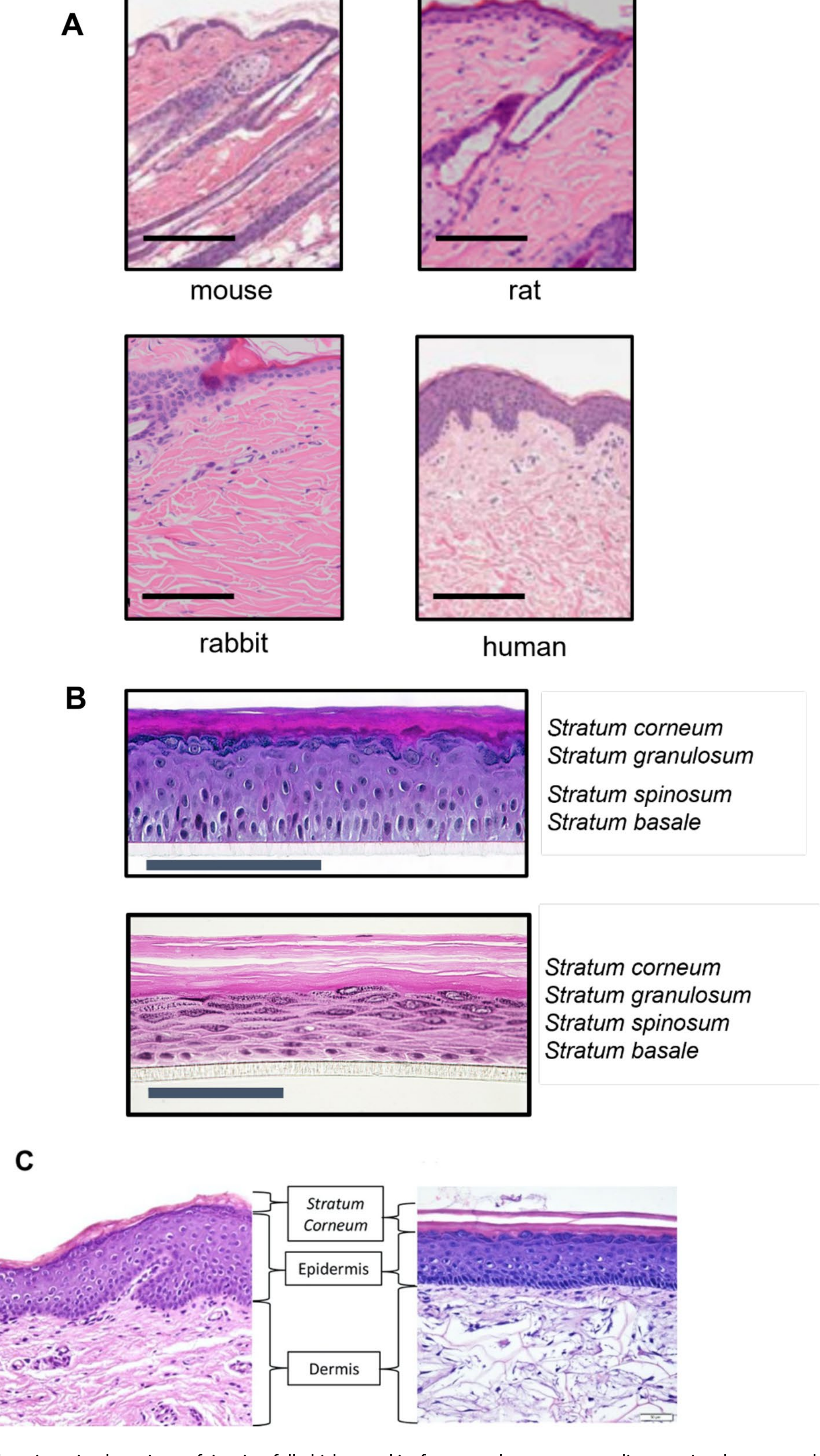


Figure 3. Haematoxylin and eosin-stained sections of *in vivo* full thickness skin from non-human mammalian species, human, and *in vitro* human epidermal cultures. **(A)** Comparative histology of mouse, rat, rabbit, and human skin. Epidermal layers (purple) and the dermis (pink) differ in structure and thickness across species. Modified with permission [72,73] **(B)** Histology of reconstructed human epidermis (RhE) cultures showing differentiation of epidermal strata. SkinEthic™ RHE model (EpiSkin, Lyon, France) (upper) and EpiDerm™ model (MatTek, Ashland, MA) (lower). Scale bars indicate 100 µm. **(C)** Cross section through human skin (left) and Phenion ® FT Full Thickness Skin Model (right) [65].

personal protective equipment needed in an industrial hygiene setting. Depending upon the regulatory purpose, further discrimination amongst moderate and mild skin irritants may be desired. Although the available test methods evaluated for use within the GHS classification scheme were evaluated relative to the rabbit reference data, the utility and

applications of the various test methods are presented here relative to the mechanistic bases of the methods and their relevance to predicting human outcomes.

The *in vivo* and *ex vivo* test methods described in Table 1 utilise responses observed in rabbit, rat, and mouse skin to inform hazards to humans from chemical exposures. Regarding animal welfare, it is important to note that live animals are used for both *in vivo* and *ex vivo* methods at some point within the test procedures; the former utilises live animals throughout the testing procedures, while the latter requires the sacrifice of animals (albeit typically fewer in number) to obtain sufficient skin for subsequent testing. Whereas the endpoints for the rabbit test include a range of apical outcomes, the endpoints for the *ex vivo* rodent skin test methods focus on the single upstream key event of barrier function disruption that is frequently associated with the aetiology of skin irritation and corrosion. The *in chemico* membrane test methods described in Table 2 model the essential key event of barrier function disruption by irritant or corrosive substances in biochemically relevant synthetic membranes. All of the *in vitro* cell and tissue culture methods described in Table 3 utilise the common key event of cytotoxicity to approximate the response of human skin to irritant or corrosive substances. The cell-based methods are further subdivided into either 2-D monolayer cell cultures submerged in aqueous media or as complex 3-D reconstructions of skin tissue. The 2-D monolayer systems typically lack a functional barrier and require that chemicals and products are diluted in the aqueous medium prior to dosing, which in particular can be problematic for testing hydrophobic substances and complex formulations. In contrast, owing to the presence of an *s. corneum* barrier, the 3-D reconstructed human epidermis (RhE)-based methods model highly relevant chemical exposure and permeation kinetics, and thus allow substances and formulations to be applied topically on the model just as they occur *in vivo*.

Use and performance of rabbit and rodent skin in skin irritation and corrosion testing

The Draize method (Table 1) was adopted as Organisation for Economic Co-operation and Development (OECD) Test Guideline 404 (TG 404) in 1981 [34,35] and has been used to meet the various classification and labelling needs for industrial and regulatory purposes. Of the available test methods for skin irritation, only the Draize rabbit test shares the same observable apical endpoints of erythema and edoema that are included in human patch tests. Furthermore, only the Draize rabbit test has the capacity to demonstrate reversal of chemically-induced irritation, thus providing a basis for discriminating between recoverable skin irritation and irreversible corrosive outcomes.

The Draize test has long been considered to be a protective method for identifying irritation and corrosion hazards due to the increased sensitivity to irritants in rabbits relative to humans [80, 88]. Indeed, the over-prediction rate of the rabbit test when compared to human patch test data has been shown to be approximately two-fold for many irritation classifications. For example, in one study of 40 common cosmetic ingredients (mostly oils and surfactants), the rabbit test showed a positive predictive value of only 57%. [89]. However, Nixon et al. [90] and Marzulli and Maibach [91] (1975) reported early on in the implementation of the Draize test that some ingredients in household products [90] and sunscreen products [91], respectively, were found to be more irritating in human skin irritation tests than in the rabbit test.

Further, the low reproducibility of the rabbit test has been recently highlighted. When comparing results of monoconstituent substances that were tested *in vivo* more than once, a surprising lack of concordance was observed for corrosion and irritation categories using both EPA and GHS classifications. For both systems, the middle categories (EPA Category II and III, GHS Category 2 and 3) were only able to be reproduced in a second test approximately half the time. For example, only 64% of mono-constituent substances identified as GHS Category 2 by the rabbit test once were identified as such in the subsequent *in vivo* tests. When using the EPA classification system, the observed reproducibility was even poorer, with less than half of Category II substances being classified in Category II again [40]. For these substances, EPA Category IV/GHS No Category determinations were of similar likelihood as EPA Category III/GHS Category 3 determinations when evaluating the substance data from subsequent studies. Similarly, only 45% and 54% of irritants classified as GHS Category 3 and EPA Category III, respectively, maintained that classification upon retesting *in vivo* [40]. Most concerning, substances were often classified into a non-adjacent category, including a considerable proportion of inconsistent classifications into either corrosive or mild or non-irritant categories. In both datasets, substances first labelled as EPA Category II/ GHS Category 2 had an approximately 1 in 5 chance of being labelled as EPA Category IV/GHS No Category.

In all of the alternatives to the *in vivo* rabbit test, quantitative endpoints measure key events upstream of the apical outcomes of erythema, edoema, and eschar formation. As such, these methods have the minimum complexity necessary to assess a specified key event using a given protocol. For the *ex vivo* test methods such as the rat TER and mouse SIFT, the endpoints inform on the loss of tight junction integrity and disruption in skin barrier function, by objectively measuring reductions in trans-epidermal resistance and increases in trans-epidermal water loss shortly after chemical exposure (Table 1). These upstream endpoints are generally relevant to the tissue barrier-disrupting effects of many skin irritants and corrosives, but given this common mode of action, the methods may not provide a mechanistic basis for discriminating between skin irritants limited to inducing epidermal tissue damage versus those corrosives inducing damage into the dermis, nor do they identify skin irritant or corrosive effects due to disruption of cellular functions in the absence of necrotic tissue destruction. As a consequence, in the absence of other data, positive results from these latter methods may only be used to conservatively classify materials as corrosives.

In practice, the *ex vivo* Rat Skin TER test for skin corrosion was approved based on an acceptably high 94% sensitivity and otherwise general alignment with the rabbit corrosion reference data [43, 48]. The high concordance of the rat skin TER with corrosive chemicals is likely based on both the rat skin TER's ability to quantitatively measure the rapid

breakdown of the skin barrier function as an initial requisite key event in skin corrosion, and the generally high reliability of the Draize reference test method in identifying corrosive materials (EPA Category I/GHS Category 1).

The *ex vivo* mouse skin test was not found to be sufficiently concordant with European Risk phrase R38 skin irritants[1], and those chemicals not requiring classification in the initial phase of a multiphase validation study by the European Centre for the Validation of Alternative Methods (ECVAM) [47,48].

In chemico models used in skin irritation and corrosion testing

There are currently two commercially available *in chemico* test methods; the Corrositex® skin corrosivity test method and the Dermal Irritection® Assay System for skin irritation (In Vitro International, Placentia, CA) (Table 2). While not a living system, the Corrositex® test system is a synthetic biobarrier membrane comprised of a proprietary homogeneous proteinaceous mixture of keratins, collagen, and lipids, in a stable gel that serves as an analogous model of the organised biochemistry of the cell layers in the epidermis [49] and which provides for a standardised barrier function model. In this method, chemical disruption of the biobarrier serves as a model of membrane disruption, a cellular key event that would necessarily result in severe damage to epithelial cells. Otherwise, the homogeneous acellular biobarrier membrane does not resemble the stratified structure of human or mammalian skin. The test method determines the corrosive potential of a chemical based upon the chemical's potential to degrade the biobarrier sufficiently to penetrate the model. Furthermore, the test method can be used to subcategorise corrosive chemicals based upon the elapsed time for biobarrier breakthrough. Thus, the relevance of the model is based upon the analogous barrier function to that of skin *in vivo*. This *in chemico* test method has been validated and approved for the identification and subcategorization of corrosive substances as described in OECD TG 435 [49]. Similar to the proposed reasons the rat skin TER method correlates well to the *in vivo* corrosivity data, the Corrositex® also quantitatively measures the rapid breakdown of the skin barrier function as the requisite key event in skin corrosion. Based upon the endpoint method, the applicability domain of the test method is generally limited to chemicals that can induce a colour change in a buffered pH indicator solution. It should be noted that Corrositex® was calibrated to available rabbit-derived corrosivity data during the development and validation of the method.

The Dermal Irritection® test system is similar in composition to the Corrositex® test system which includes a membrane matrix of keratins and collagens (and an indicator dye) which mimics the epidermal barrier function, as well as a reagent substrate consisting of a highly organised globulin/protein macromolecular solution [53]. The Dermal Irritection® test has currently not undergone validation for regulatory application, however, the mechanistically-similar method for eye irritation (Ocular Irritection®) by the same test method developer has been validated and approved for identifying chemicals that can induce serious eye damage and those not requiring classification within the United Nations Globally Harmonised System (GHS) [54]. Irritants are determined by the ability to degrade the membrane matrix and induce conformational changes in the organised globulin/protein matrix. Although neither the membrane matrix nor the reagent substrate resemble the stratified structure of human or mammalian skin, the relevance of the model is based upon the modelling of similar chemical-induced changes in skin proteins as occurs *in vivo*. Although the relevance of the acellular test method is mechanistically based on the ability to quantitatively measure the chemical disruption of tissue and cellular proteins and lipids, that key event may not be sufficient for modelling the various cellular responses involved in skin irritation, and in particular for those non-lethal, non-necrotic cellular responses such as changes in gene and cytokine expression at the milder end of the irritation continuum.

In vitro models used in skin irritation and corrosion testing

Cell membrane perturbations, altered cell metabolism, and cell death are generally accepted as key cellular mechanisms and events (Figure 2) in skin irritation and corrosion after acute exposure and skin penetration [22], and accordingly, early *in vitro* cell-based cytotoxicity test methods were developed to measure increases in cytotoxicity (i.e. reductions in cell viability). Initial cell-based methods utilised available mammalian cell lines cultured as 2-dimensional (2-D) monolayers immersed in aqueous medium in culture dishes or multi-well plates and compared the cytotoxicity of serial dilutions of test chemical to control values. Various cell death (e.g. lactate dehydrogenase release) and viability (e.g. neutral red uptake, succinate dehydrogenase activity, glucose metabolism acidification rate) biomarkers can be measured quantitatively to determine cytotoxicity. Based upon the successful validation of the Cytosensor Microphysiometer test method for identifying ocular irritants by measuring changes in the baseline acidification rates in murine L929 cells, a modification of the test method was applied to provide rank order skin irritation characterisation of raw materials for personal care paper products [71]. To increase the relevance of these 2-D systems for human skin irritation predictions, cytotoxicity assays using human-derived epidermal keratinocytes and dermal fibroblasts were utilised, each employing a neutral red uptake (NRU) endpoint [81]. Although some of the test methods used human-derived cells, the 2-D monolayer cell culture systems do not reflect the stratified structure of human or mammalian skin, nor do they provide a skin-relevant barrier function, and thus are limited in their relevance to skin irritation characterisation in humans. Accordingly, most of the 2-D methods were generally found to be limited to rank ordering the irritation potential of individual chemicals [92].

The reconstructed human epidermis (RhE)-based test methods utilise a tissue model of epidermal function to predict human corrosion responses directly based on cytotoxicity [43]. RhE models are comprised of human-derived epidermal keratinocytes cultured at the air-liquid interface

to replicate the differentiated epidermal keratinocyte layers complete with viable cells and an effective barrier. The same process of terminal differentiation described for human skin *in vivo* is used to recapitulate human epidermal tissue *in vitro*, starting with the culturing of a proliferative basal layer of keratinocytes at an air-liquid interface and subsequent upward displacement and differentiation into distinct layers. In RhE models, the full differentiation of epidermal layers from the *s. basale* to a functional *s. corneum* layer is histologically evident (Figure 3B). Just as occurs *in vivo*, it is the barrier function of the RhE models that provide the barrier to chemical insults, such that individual chemicals as well as complex formulations can be applied topically onto the apical surfaces to model *in vivo* exposures. As the composition of the RhE barrier layers more closely models that of human skin, the diffusion kinetics of individual ingredients out of a formulation and into the skin would be expected to be more similar between human skin and the reconstructed models than for any of the aforementioned *in vivo*, *ex vivo* or *in vitro* models.

There are several commercially available RhE models globally, with most based upon a relatively simple differentiated keratinocyte architecture. Bespoke epidermal models are available with specific accessory cells, and some are adapted to specialised tissue culture support systems, and other more complex full thickness reconstructed skin models incorporating a fibroblast-based dermis are also available (Table 3). The RhE models validated for use in skin irritation and corrosion test guidelines lack other accessory cell types like melanocytes, Langerhans cells, and Merkel cells, as well as innervation, vasculature, and other appendages that originate in the dermis. However, these adnexal features are not needed to model epidermal cell death following *s. corneum* penetration, and thus the basic RhE models are highly relevant to addressing the key events relevant to skin irritation and corrosion. Lastly, although the commercially available RhE models are human relevant in terms of cell sourcing and general tissue architecture, the models are generally more permeable to a range of chemicals relative to excised human skin [93]. To compensate for the less robust barrier function, the specific skin irritation test method exposure kinetics (i.e. dose volume and exposure times) were optimised for each tissue model during test method development to calibrate the test system responses to the reference skin irritation data.

In vitro test methods for skin irritation and corrosion performance. Several commercially available RhE models have undergone formal validation for use in either skin corrosion and/or skin irritation applications. Test methods included in OECD TG 431 can discriminate among corrosives (GHS 1) and non-corrosives and provide subcategorization to discriminate between GHS 1A vs GHS 1B/1C combined, and those test methods included in OECD TG 439 can discriminate GHS Category 2 and No Category across a variety of materials, including organic acids, surfactants, and electrophiles, oily and aqueous chemicals, as well as solids and liquids. Performance standards for within-laboratory and between-laboratory reproducibility as well as sensitivity, specificity, and accuracy have been met or exceeded using defined reference data [52]. Since the mandate for the performance of alternative methods was to be as good as or better than the Draize rabbit test (i.e. at least as protective), the current regulatory-accepted methods are more likely to be over-predictive than under-predictive relative to human hazard potential.

A complete evaluation of the RhE method performance for identifying corrosives within the agrochemical sector is not practical given that very few products are indeed corrosive. For example, a retrospective analysis of one company's agrochemical formulations revealed only two corrosive (GHS Category 1) materials [94] out of 207. In a separate analysis of 81 agrochemical formulations, four corrosives (GHS Category 1) were identified by the *in vitro* RhE method that were not identified by the *in vivo* method. Although this is not a sufficient number of positives to draw a conclusion on overall method performance, these data suggest that the RhE method performance is capable of identifying corrosive agrochemical formulations.

There are also limited published data assessing the skin irritation potential of pesticide formulations with *in vitro* methods. In a comparison of test results from 25 irritating (GHS Category 2) agrochemical formulations, 44% (11/25) of formulations which were found to be severely irritating in the rabbit test were also found to be irritating in the *in vitro* RhE test, using the OECD TG 439 protocol [58]. Further analyses reveal that of the 14 formulations classified as Cat 2 by the rabbit test, but predicted as non-irritant by the *in vitro* SIT, 7 formulations resulted in <85% relative viability demonstrating the ability of the test method to detect the induction of cytotoxic effects, albeit above the prediction threshold. Additionally, of the 8 formulations classified by the animal test to be mildly irritating (GHS Category 3), 6 resulted in positive predictions in the *in vitro* test. Thus, it should not be immediately inferred that agrochemical formulations are outside of the applicability domain of the *in vitro* RhE test methods, without questioning both the relevance and reliability of the *in vivo* reference data, as it is unclear whether the low apparent sensitivity of the RhE method more reasonably implies the presence of false positives in the rabbit test or false negatives in the RhE method, as the aforementioned sensitivity rate is in proportion with observed false positive rates in the rabbit test [89]. It should also be noted that this observed sensitivity is not unexpected given the low level of reproducibility of the rabbit test for this same category, including significant re-classifications between GHS Category 2 and No Category [40]. Even among the reference chemicals used to validate new RhE methods are two chemicals considered to represent false-positive Category 2 classifications in rabbits [56]. Another recent study used 24 antimicrobial formulations (6 non-irritants and 18 irritants) to evaluate the predictivity of the Phenion® FT test. The results revealed an overall sensitivity of 78%, specificity of 83% and accuracy of 79% [65].

Pesticidal formulations represent a broad class of chemistries and complex mixtures, including mineral clays, wetting agents, foaming agents, and/or dispersing agents designed to optimise the activity of the active ingredients. The great majority of formulations are broadly grouped as water- or

organic solvent-based, or solid formulations, which generally fall within the RhE-based test method applicability domains [41]. Taking the above into account, in general RhE-based methods not only provide human relevant structure and function of the endpoint-relative features of human skin, they also exhibit good technical performance. While published comparative data for pesticide formulations are currently limited, quantitative RhE methods have been demonstrated to be reliable, reproducible, and protective of human health for substances representing a diversity of physicochemical properties [89,95].

Applicability of current methods for decision making

An evaluation of methods for regulatory application should consider their fitness for the intended purpose, assess human biological and mechanistic relevance, and ensure appropriate technical performance [95]. With this in mind, there are several non-animal test methods and testing strategies that can be utilised without further development or validation to address a portion of the skin irritation/corrosion continuum. By applying a top-down testing strategy within an Integrated Approach to Testing and Assessment (IATA) (Figure 4A), the corrosive potential of test substances can readily be evaluated using either the *in chemico* Corrositex® test method following procedures described in OECD TG 435, or using any of the available *in vitro* RhE models following procedures described in OECD TG 431 [41,96]. The *ex vivo* TER assay (OECD TG 430) [42] could also be used but it is not based on human skin. Selection of the appropriate platform may be dictated in part by compatibility of the test substance with the specific corrosivity test method; for example, substances or components of mixtures tested in Corrositex® must be outside the pH range 4.5 – 8.5. These methods can be used to identify a corrosive, and further subcategorise if needed, or can be used to rule out corrosive potential. If the test substance is not found to be corrosive, *in vitro* RhE models may be used to identify whether the substance is likely to be a GHS 2 or EPA Category II skin irritant, following the SIT procedures described in OECD TG 439. The SIT was validated to discriminate between GHS category 2 skin irritants, and those substances that do not require classification within the GHS,

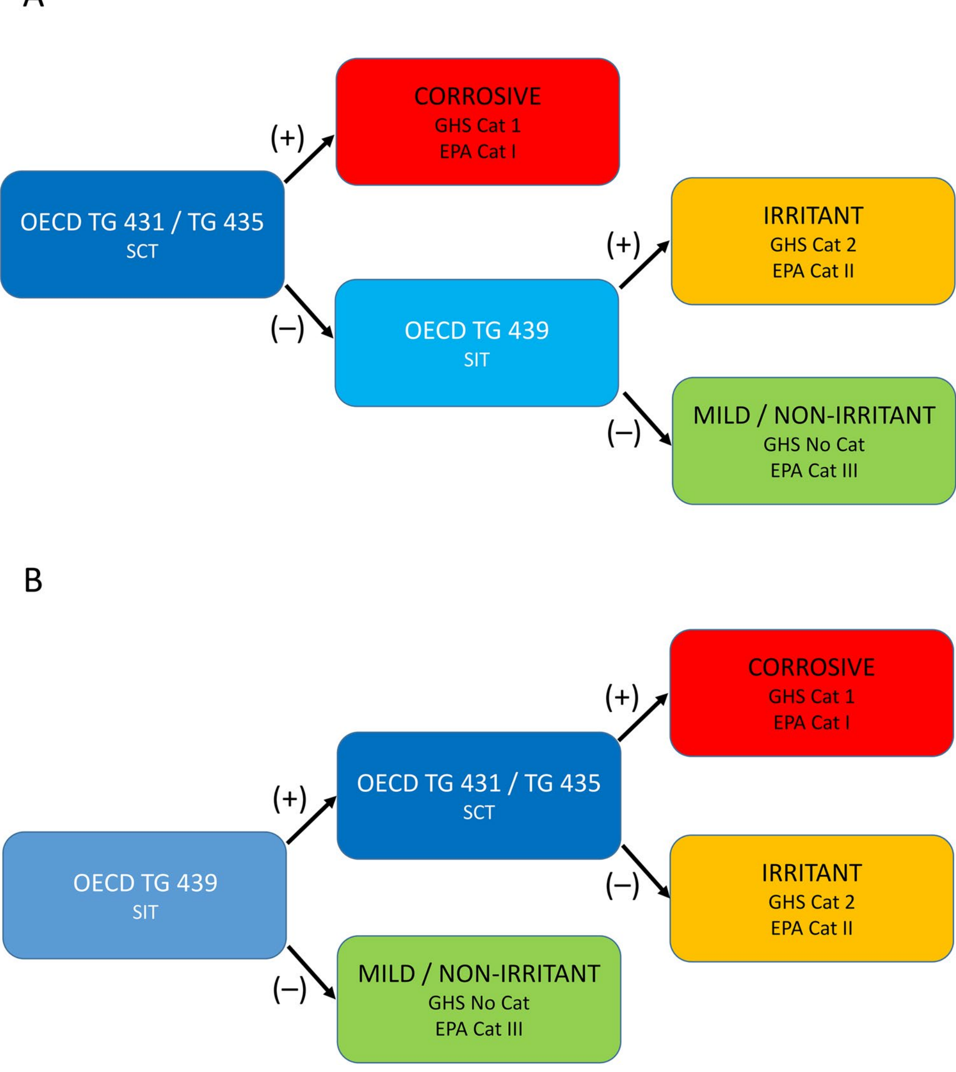


Figure 4. Testing strategies for irritants and corrosives using currently available OECD-approved *in chemico* and *in vitro* test methods. **(A)** Top-down testing strategy. **(B)** Bottom-up testing strategy.

the latter of which include minimally or mildly irritating substances. Consequently, the SIT has not been validated to discriminate between EPA Categories II, III and IV, such that a positive outcome would conservatively drive an EPA Category II classification, at the risk of potential over-prediction of some mildly irritating substances. Similarly, a negative outcome in the SIT without any further evidence may drive a conservative EPA Cat III classification.

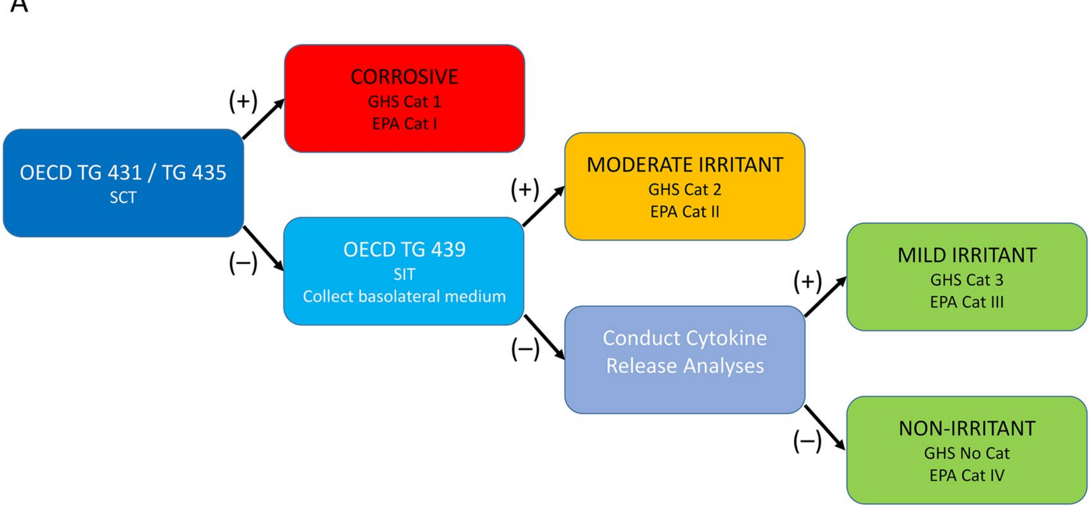
Alternatively, a bottom-up testing strategy may be employed if the substance is not expected to be highly irritating (Figure 4B), wherein the irritant potential of the test substance is first evaluated in the SIT, following the procedures described in OECD TG 439, as described above. A positive prediction in the SIT would require conducting one of the corrosion test methods to determine if the irritant substance also has the potential to be corrosive.

Next steps: Approaches to utilize non-animal test methods for predicting mild skin irritation

The RhE corrosion and SIT test endpoints rely on measuring overt cytotoxic effects in the tissue models, which is mechanistically relevant for the vast majority of highly irritant and corrosive substances. However, for milder irritants the cytotoxicity endpoint may not be fully useful for identifying the majority of mild irritants where comparatively fewer cells are damaged or lysed upon skin exposure. Fortunately given the biological relevance of the RhE models, various approaches and endpoints can be applied to allow for improvements in the prediction of mild and moderate irritants. For example, sensitivity to milder materials can be enhanced in the cytotoxicity-based assays by increasing the duration of the exposure times, by enhancing the exposure kinetics, and by modifying the positivity prediction thresholds. In cases where cytotoxicity is not prominent, and the aforementioned modifications to the exposure and expression kinetics are not desired, the release of inflammatory cytokines can be evaluated.

The upregulation and release of pro-inflammatory cytokines in epidermal cells has been implemented in non-regulatory testing especially in the consumer product arena as an additional endpoint for identifying milder irritants [62,85,97], and is recognised as one of the key cellular events that occur upstream of overt cytotoxicity. Specifically, the release of the primary cytokine IL-1α by epidermal cells after chemical insult *in vivo* can be modelled in the *in vitro* RhE and reconstructed full thickness skin models by quantifying the amount of IL-1α released into the culture medium using available ELISA technologies. IL-1α released into the medium has shown utility in the identification of mild irritants and confirmation of non-irritant results [85,98]. Correlation to measured IL-1α has also been observed for both clinical trans-epidermal water loss measurements and classification of commercial cleansers [62] and detergents [99], and because the IL-1α endpoint has shown to be highly reproducible in routine use, it was established as the primary criterion for discriminating amongst candidate ingredients [62]. During the optimisation of the SIT test method IL-1α release was evaluated for consideration as an additional endpoint to the MTT viability assessment to identify R38 skin irritants but was not included as it did not further add to identifying R38 irritants [85], and in fact was found to be released by milder non-R38 reference substances in addition to the R38 reference materials[2]. This finding further supports the hypothesis that IL-1α release is an upstream key event in common with both mild and more severe irritants and could be used to discriminate mild irritants from non-irritants in the absence of overt cytotoxicity. Lastly, the full range of erythema scores *in vivo* in hairless rats were highly correlated to increasing levels of expression of the secondary cytokines IL-6 and IL-8 in a full thickness skin model after exposure to aliphatic hydrocarbons, thus demonstrating the relevance of cytokine signalling events *in vitro* to apical outcomes *in vivo* [87]. While the data cited here are not from testing pesticide formulations, the breadth of types of products and ingredients (surfactants, detergents, neat chemicals, and medical device extracts) used indicates utility of the marker across product types and demonstrates that test protocols can be modified to fit the relevant physicochemical properties of test materials. Indeed many of the same aforementioned classes of chemicals such as surfactants and solvents may also be used in agrochemical formulations to improve the dissolution/suspension, dispersal, and adhesion properties of active ingredients in the formulation [58].

Based upon the endpoints and human-relevant mechanistic justifications presented above an approach for consideration would be to include measurement of IL-1α released into the culture medium during the post-treatment expression incubation of the SIT assay. An envisioned prediction model would allow for categorisation into three categories allowing discrimination between moderate skin irritants (consistent with GHS 2 criteria), milder irritants likely inducing only transient erythema and/or edoema, and those substances not likely to induce notable clinical effects. Upon testing, those substances which result in a cytotoxic response as determined by the MTT viability assay would be categorised as moderate skin irritants, regardless of cytokine release; those substances which result in a positive response in the IL-1α release endpoint in the absence of overt cytotoxicity would be categorised as mild skin irritants, and those substances which result in negative responses in both the MTT viability and IL-1α release assays would be characterised as non-irritants to skin. Establishing the appropriate thresholds for cytokine release could be done based upon available human clinical data to better discriminate between non-irritants and those that showed positive reactions in humans. This approach, where feasible, would apply human-relevant mechanistic-based methodologies to fit the human-derived data in establishing useful EPA category III and IV criteria. Accordingly, in a top-down testing strategy, the corrosive potential of test substances would be evaluated following the procedures in OECD TG 435 or TG 431, as described previously (Figure 5A), or by applying a bottom-up testing strategy utilising a modified TG 439 with cytokine analyses (Figure 5B**)** to allow for further discrimination between mild and non-irritants; an enhancement that is not provided in the first strategy described above. Regardless of whether a top-down or



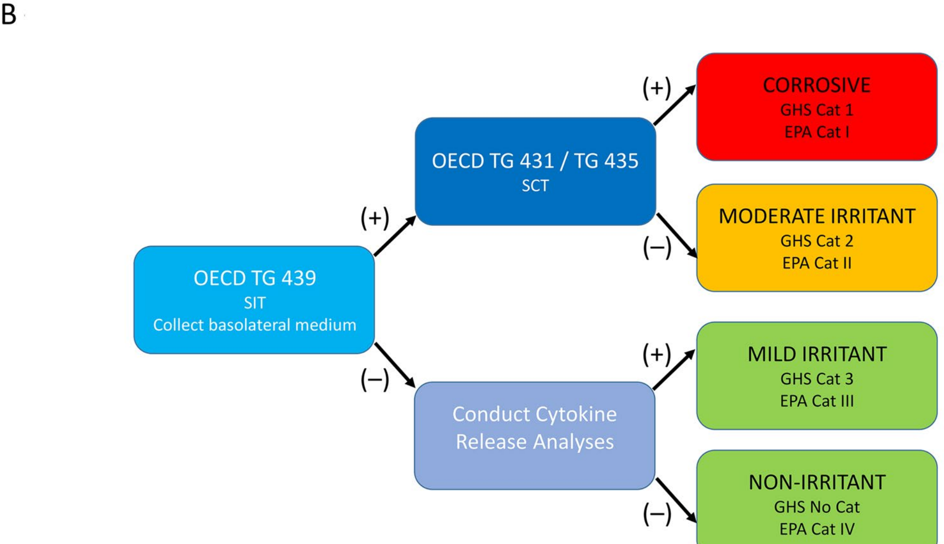


Figure 5. Potential testing strategies using cytokine release in conjunction with currently available test methods to allow for additional discrimination between mild and non-irritants. **(A)** Top-down testing strategy. **(B)** Bottom-up testing strategy.

bottom up strategy is utilised, the testing results should give rise to the same predictions.

Conclusions

Several considerations, including major species differences in the structure and function of the epidermis relative to humans, high subjectivity in evaluation, and concomitant poor reproducibility especially in the mild to moderate irritation range, have illustrated the limitations in the Draize rabbit test for the categorisation of substances for skin irritation. In contrast, several non-animal test methods, and in particular RhE-based test methods, have a distinct advantage in the ability to monitor human tissue responses *in vitro* with mechanistically based, quantitative protocols designed to reflect the expected responses in humans. The current regulatory-approved *in chemico* and *in vitro* tests for identifying moderate to severe skin irritation and corrosion can be used today to make human-relevant regulatory decisions for substances and mixtures, including pesticidal formulations, following current OECD Test Guidelines and IATA guidance documents [41,55,56]. Several *in vitro* methods are accepted for many applications for hazard classification and have been included in the GHS Revision 8, conferring acceptance by countries using the GHS [100]. Further, in the future, RhE-based test methods can improve prediction of relevant endpoints by providing quantitative hazard classification for mild irritants based on cytokine release. Additional investigations into the dynamics of cytokine signalling could provide mechanistic insight that could increase confidence in delineating mild and moderate irritants.

Notes

1. The authors note that risk-phrase categories like R38 (irritating to skin) are no longer used in the EU.
2. The authors note that risk-phrase categories like R38 (irritating to skin) are no longer used in the EU.

Acknowledgements

The authors would like to thank Nicole Kleinstreuer, Deborah Ramsingh, Lisa Woods, Yavinder Bhuller, Xiyuan Cheng, Erin Kollitz, Amy Clippinger, and Jet Gijbrechts for their input in discussions and/or review of this paper. The authors would like to thank EPISKIN, Henkel, and the Institite for In Vitro Sciences for providing histology images.

Author contributions statement

All authors have approved the final version for publication and agree to be accountable for all aspects of the work. HAR, DGA, AL, MC, KP, EB, and KS were involved in the conception and design of the article; HAR, GEC, DGA, AL, LO, and KS were involved in drafting of the paper, and HAR, GEC, DGA, AL, MC, LO, JBA, KP, MP, TFS, and WW provided critical revisions.

Disclosure statement

Marco Corvaro works for Corteva Agriscience and Kathryn Page works for The Clorox Company. All other authors declare no competing interests to declare. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency or other regulatory authorities.

Funding

This project was funded in part with federal funds from the NIEHS, NIH under Contract No. HHSN273201500010C. This work was supported by a grant from the Carroll Petrie Foundation.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

References

1. Kleinstreuer NC, Hoffmann S, Alépée N, et al. Non-animal methods to predict skin sensitization (II): an assessment of defined approaches (*). Crit Rev Toxicol. 2018;48(5):359–374. doi:10.1080/10408444.2018.1429386.
2. Li H, Bai J, Zhong G, et al. Improved defined approaches for predicting skin sensitization hazard and potency in humans. ALTEX - Altern Anim Experiment. 2019;36(3):363–372. doi:10.14573/altex.1809191.
3. Guideline OECD. No. 497: Defined Approaches on Skin Sensitisation [Internet]. 2023. Available from: https://www.oecd-ilibrary.org/content/publication/b92879a4-en.
4. EPA. Interim science policy: use of alternative approaches for skin sensitization as a replacement for laboratory animal testing draft for public comment [Internet]. 2018. Available from: https://www.regulations.gov/document/EPA-HQ-OPP-2016-0093-0090.
5. Hamm J, Allen D, Ceger P, et al. Performance of the GHS mixtures equation for predicting acute oral toxicity. Regul Toxicol Pharmacol. 2021;125:105007. doi:10.1016/j.yrtph.2021.105007.
6. Latorre AO, Floresta PVM, Boff MM, et al. Non-relevance of acute dermal toxicity testing for assessing human health protection in the regulatory decision-making for agrochemical formulated products. Regul Toxicol Pharmacol. 2019;106:105–110. doi:10.1016/j.yrtph.2019.04.014.
7. OECD. Guidance document on considerations for waiving or bridging of mammalian acute toxicity tests. Series on Testing And Assessment No. 237. 2017.
8. PMRA. Science Policy Note SPN2017-03, Acute Dermal Toxicity Study Waiver [Internet]. 2017. Available from: https://www.canada.ca/en/health-canada/services/consumer-product-safety/reports-publications/pesticides-pest-management/policies-guidelines/science-policy-notes/2017/acute-dermal-toxicity-waiver-spn2017-03.html.
9. EPA. Guidance for waiving acute dermal toxicity tests for pesticide technical chemicals & supporting retrospective analysis [Internet]. Washington DC, USA: Office of Pesticide Programs, Office of Chemical Safety and Pollution Prevention, United States Environmental Protection Agency; 2020. [cited 2021 Nov 13]. p. 8. Report No.: EPA 705-G-2020-3722. Available from: https://www.epa.gov/sites/default/files/2021-01/documents/guidance-for-waiving-acute-dermal-toxicity.pdf.
10. EPA. Use of an alternate testing framework for classification of eye irritation potential of EPA pesticide products [Internet]. 2015. Available from: https://www.epa.gov/sites/default/files/2015-05/documents/eye_policy2015update.pdf.
11. Clippinger AJ, Raabe HA, Allen DG, et al. Human-relevant approaches to assess eye corrosion/irritation potential of agrochemical formulations. Cutan Ocul Toxicol. 2021;40(2):145–167. doi:10.1080/15569527.2021.1910291.
12. Monteiro-Riviere NA, Bristol DG, Manning TO, et al. Interspecies and interregional analysis of the comparative histologic thickness and laser doppler blood flow measurements at five cutaneous sites in nine species. J Invest Dermatol. 1990;95(5):582–586. doi:10.1111/1523-1747.ep12505567.
13. Jung EC, Maibach HI. Animal models for percutaneous absorption. J Appl Toxicol. 2015;35(1):1–10. doi:10.1002/jat.3004.
14. DeWever B, Rheins L. An in vitro human skin analog. In Vitro skin toxicology: Irritation, phototoxicity, sensitization. New York: Mary Ann Liebert; 1994. p. 121–131.
15. Leeson. Histology: the skin and its appendages. 4th ed. Pennsylvania: W.B. Saunders Company; 1991.
16. Otberg N, Richter H, Schaefer H, et al. Variations of hair follicle size and distribution in different body sites. J Invest Dermatol. 2004;122(1):14–19. doi:10.1046/j.0022-202X.2003.22110.x.
17. Elmahdy A, Cao Y, Hui X, et al. Follicular pathway role in chemical warfare simulants percutaneous penetration. J Appl Toxicol. 2021;41(6):964–971. doi:10.1002/jat.4081.
18. Otberg N, Teichmann A, Rasuljev U, et al. Follicular penetration of topically applied caffeine via a shampoo formulation. Skin Pharmacol Physiol. 2007;20(4):195–198. doi:10.1159/000101389.
19. Chu I, Dick D, Bronaugh R, et al. Skin reservoir formation and bioavailability of dermally administered chemicals in hairless Guinea pigs. Food Chem Toxicol. 1996;34(3):267–276. doi:10.1016/0278-6915(95)00112-3.
20. Hayes B, Patrick E, Maibach H. Dermatotoxicology. In: Hayes A, editor. Principle and methods of toxicology. 5th ed. Boca Raton: CRC Press, Taylor & Francis Group; 2008.
21. Harbell J, Raabe H. Chapter 5: in vitro methods for the prediction of ocular and dermal toxicity. Handbook of toxicology. 3rd ed. Florida: CRC Press; 2014. p. 197–231.
22. Patil S, Patrick E, Maibach H. Animal, human, and in vitro test methods for predicting skin irritation. In: Marzulli M, editors. Dermatotoxicology. 5th ed. Boca Raton: Taylor & Francis; 1996.
23. Patrick E, Maibach HI, Burkhalter A. Mechanisms of chemically induced skin irritation. I. Studies of time course, dose response, and components of inflammation in the laboratory mouse. Toxicol Appl Pharmacol. 1985;81(3 Pt 1):476–490. doi:10.1016/0041-008x(85)90419-3.
24. Welss T, Basketter DA, Schröder KR. In vitro skin irritation: facts and future. State of the art review of mechanisms and models. Toxicol In Vitro. 2004;18(3):231–243. doi:10.1016/j.tiv.2003.09.009.
25. Berardesca E, Distante F. The modulation of skin irritation. Contact Dermatitis. 1994;31(5):281–287. doi:10.1111/j.1600-0536.1994.tb02019.x.
26. Lee HY, Stieger M, Yawalkar N, et al. Cytokines and chemokines in irritant contact dermatitis. Mediators Inflamm. 2013;2013:916497–916497. doi:10.1155/2013/916497.

27. Keppel Hesselink JM. Fundamentals of and critical issues in lipid autacoid medicine: a review. Pain Ther. 2017;6(2):153–164. doi:10.1007/s40122-017-0075-4.
28. Werman A, Werman-Venkert R, White R, et al. The precursor form of IL-1alpha is an intracrine proinflammatory activator of transcription. Proc Natl Acad Sci U S A. 2004;101(8):2434–2439. doi:10.1073/pnas.0308705101.
29. Dinarello CA. Overview of the IL-1 family in innate inflammation and acquired immunity. Immunol Rev. 2018;281(1):8–27. doi:10.1111/imr.12621.
30. Cohen I, Rider P, Carmi Y, et al. Differential release of chromatin-bound IL-1alpha discriminates between necrotic and apoptotic cell death by the ability to induce sterile inflammation. Proc Natl Acad Sci U S A. 2010;107(6):2574–2579. doi:10.1073/pnas.0915018107.
31. Di Paolo NC, Shayakhmetov DM. Interleukin 1α and the inflammatory process. Nat Immunol. 2016;17(8):906–913. doi:10.1038/ni.3503.
32. Barker JN, Mitra RS, Griffiths CE, et al. Keratinocytes as initiators of inflammation. Lancet. 1991;337(8735):211–214. doi:10.1016/0140-6736(91)92168-2.
33. McKenzie R, Sauder D. The role of keratinocyte cytokines in inflammation and immunity. J Invest Dermatol. 1990;95(6 Suppl):105S–107S. doi:10.1111/1523-1747.ep12874955.
34. Draize JH, Woodard G, Calvery HO. Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J Pharmacol Exp Ther. 1944;82:377.
35. OECD. Test no. 404: acute dermal irritation/corrosion [Internet]. 2015. Available from: https://www.oecd-ilibrary.org/content/publication/9789264242678-en.
36. EPA. Chemical hazard classification and labeling : Comparison of OPP requirements and the GHS [Internet]. 2004. Available from: https://www.epa.gov/sites/default/files/2015-09/documents/ghscriteria-summary.pdf.
37. EPA. Federal Insecticide, fungicide, rodenticide act, Pesticide assessment guidelines. 1984. 55e.
38. EPA. OPPTS 870.2500: Acute Dermal Irritation [Internet]. 1998 [cited 2023 Dec 22]. Available from: https://ntp.niehs.nih.gov/sites/default/files/iccvam/suppdocs/feddocs/epa/epa_870_2500.pdf.
39. EEC. Part B: Methods for the Determination of Toxicity, No. L 141/142, B.4,Acute Toxicity (Skin Irritation). 2008.
40. Rooney JP, Choksi NY, Ceger P, et al. Analysis of variability in the rabbit skin irritation assay. Regul Toxicol Pharmacol. 2021;122:104920. doi:10.1016/j.yrtph.2021.104920.
41. OECD. Guidance document on an integrated approach on testing and assessment (IATA) for skin corrosion and irritation [Internet]. 2017. Available from: https://www.oecd-ilibrary.org/content/publication/9789264274693-en.
42. OECD. Test no. 430: in vitro skin corrosion: transcutaneous electrical resistance test method (TER) [Internet]. 2015. Available from: https://www.oecd-ilibrary.org/content/publication/9789264242739-en.
43. Fentem JH, Archer GE, Balls M, et al. The ECVAM international validation study on in vitro tests for skin corrosivity. 2. Results and evaluation by the management team. Toxicol In Vitro. 1998;12(4):483–524. doi:10.1016/s0887-2333(98)00019-8.
44. Oliver GJ, Pemberton MA, Rhodes C. An in vitro skin corrosivity test–modifications and validation. Food Chem Toxicol. 1986;24(6-7):507–512. doi:10.1016/0278-6915(86)90102-x.
45. Botham PA, Hall TJ, Dennett R, et al. The skin corrosivity test in vitro. Results of an inter-laboratory trial. Toxicol In Vitro. 1992;6(3):191–194. doi:10.1016/0887-2333(92)90031-l.
46. Heylings JR, Diot S, Esdaile DJ, et al. A prevalidation study on the in vitro skin irritation function test (SIFT) for prediction of acute skin irritation in vivo: results and evaluation of ECVAM phase III. Toxicol In Vitro. 2003;17(2):123–138. doi:10.1016/s0887-2333(02)00130-3.
47. Fentem JH, Briggs D, Chesné C, et al. A prevalidation study on in vitro tests for acute skin irritation. results and evaluation by the management team. Toxicol In Vitro. 2001;15(1):57–93. doi:10.1016/s0887-2333(01)00002-9.
48. Spielmann H, Hoffmann S, Liebsch M, et al. The ECVAM international validation study on in vitro tests for acute skin irritation: report on the validity of the EPISKIN and EpiDerm assays and on the skin integrity function test. Altern Lab Anim. 2007;35(6):559–601. doi:10.1177/026119290703500614.
49. OECD. Test no. 435: in vitro membrane barrier test method for skin corrosion [internet]. 2015. Available from: https://www.oecd-ilibrary.org/content/publication/9789264242791-en.
50. Kandárová H, Liebsch M, Spielmann H, et al. Assessment of the human epidermis model SkinEthic RHE for in vitro skin corrosion testing of chemicals according to new OECD TG 431. Toxicol In Vitro. 2006;20(5):547–559. doi:10.1016/j.tiv.2005.11.008.
51. NICEATM. Corrositex®: An In Vitro Test Method for Assessing Dermal Corrosivity Potential of Chemicals. The Results of an Independent Peer Review Evaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) [Internet]. National Institutes of Health; 1999. [cited 2023 Dec 22]. Available from: https://ntp.niehs.nih.gov/sites/default/files/iccvam/docs/dermal_docs/corprrep.pdf.
52. ICCVAM. Recommended Performance Standards for In Vitro Test Methods for Skin Corrosion [Internet]. National Institutes of Health; 2004. Available from: https://ntp.niehs.nih.gov/sites/default/files/iccvam/docs/dermal_docs/ps/ps044510.pdf.
53. Hassan Z, Ismail R, Ahmad S. Safety evaluation for dermal and ocular irritation of palm dihidroxystearic acid as a cosmetics ingredient. J Palm Oil Res. 2005;17:160–167.
54. OECD. Test No. 496: In vitro Macromolecular Test Method for Identifying Chemicals Inducing Serious Eye Damage and Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage [Internet]. 2023. Available from: https://www.oecd-ilibrary.org/content/publication/970e5cd9-en.
55. OECD. Test no. 431: in vitro skin corrosion: reconstructed human epidermis (RHE) test method [Internet]. 2019. Available from: https://www.oecd-ilibrary.org/content/publication/9789264264618-en.
56. OECD. Test no. 439: in vitro skin irritation: reconstructed human epidermis test method [Internet]. 2021. Available from: https://www.oecd-ilibrary.org/content/publication/9789264242845-en.
57. Sheehan D, Costin G-E, Diersen V, et al. Evaluation of the validated in vitro skin irritation test (OECD TG 439) for the assignment of EPA hazard categories. Washington, DC; 2016. [cited 2023 Dec 23]. Available from: http://iivs.org/wp-content/uploads/2016/09/poster_2016_final.pdf.
58. Kolle SN, van Ravenzwaay B, Landsiedel R. Regulatory accepted but out of domain: in vitro skin irritation tests for agrochemical formulations. Regul Toxicol Pharmacol. 2017;89:125–130. doi:10.1016/j.yrtph.2017.07.016.
59. Robinson MK, Cohen C, de Fraissinette A. D B, et al. Non-animal testing strategies for assessment of the skin corrosion and skin irritation potential of ingredients and finished products. Food Chem Toxicol. 2002;40(5):573–592. doi:10.1016/s0278-6915(02)00005-4.
60. Faller C, Bracher M, Dami N, et al. Predictive ability of reconstructed human epidermis equivalents for the assessment of skin irritation of cosmetics. Toxicol In Vitro. 2002;16(5):557–572. doi:10.1016/s0887-2333(02)00053-x.
61. Perkins MA, Osborne R, Rana FR, et al. Comparison of in vitro and in vivo human skin responses to consumer products and ingredients with a range of irritancy potential. Toxicol Sci. 1999;48(2):218–229. doi:10.1093/toxsci/48.2.218.
62. Walters RM, Gandolfi L, Mack MC, et al. In vitro assessment of skin irritation potential of surfactant-based formulations by using a 3-D skin reconstructed tissue model and cytokine response - PubMed. Altern Lab Anim. 2016;44(6):523–532. doi:10.1177/026119291604400611.

63. Mewes KR, Fischer A, Zöller NN, et al. Catch-up validation study of an in vitro skin irritation test method based on an open source reconstructed epidermis (phase I). Toxicol In Vitro. 2016;36:238–253. doi:10.1016/j.tiv.2016.07.007.
64. Groeber F, Schober L, Schmid FF, et al. Catch-up validation study of an in vitro skin irritation test method based on an open source reconstructed epidermis (phase II). Toxicol In Vitro. 2016;36:254–261. doi:10.1016/j.tiv.2016.07.008.
65. Page K, Westerink W, Sullivan K, et al. Assessment of the utility of the novel phenion® full thickness human skin model for detecting the skin irritation potential of antimicrobial cleaning products. Toxicol In Vitro. 2024;94:105726. doi:10.1016/j.tiv.2023.105726.
66. Mewes KR, Raus M, Bernd A, et al. Elastin expression in a newly developed full-thickness skin equivalent. Skin Pharmacol Physiol. 2007;20(2):85–95. doi:10.1159/000097655.
67. Osborne R, Perkins MA. An approach for development of alternative test methods based on mechanisms of skin irritation. Food Chem Toxicol. 1994;32(2):133–142. doi:10.1016/0278-6915(94)90174-0.
68. Borenfreund E, Puerner JA. Toxicity determined in vitro by morphological alterations and neutral red absorption. Toxicol Lett. 1985;24(2-3):119–124. doi:10.1016/0378-4274(85)90046-3.
69. Repetto G, del Peso A, Zurita JL. Neutral red uptake assay for the estimation of cell viability/cytotoxicity. Nat Protoc. 2008;3(7):1125–1131. doi:10.1038/nprot.2008.75.
70. Nash JR, Mun G, Raabe HA, et al. Using the cytosensor microphysiometer to assess ocular toxicity. Curr Protoc Toxicol. 2014;61:1–11. 1.13.
71. Landin WE, Mun GC, Nims RW, et al. Use of the cytosensor microphysiometer to predict results of a 21-day cumulative irritation patch test in humans. Toxicol In Vitro. 2007;21(6):1165–1173. doi:10.1016/j.tiv.2007.03.006.
72. Gudjonsson JE, Johnston A, Dyson M, et al. Mouse models of psoriasis. J Invest Dermatol. 2007;127(6):1292–1308. doi:10.1038/sj.jid.5700807.
73. Mitsuishi M, Oshikata T, Kumabe S, et al. Histological dermal changes caused by preparation and application procedures in percutaneous dose toxicity studies in dogs, rabbits and rats. J Toxicol Pathol. 2015;28(1):1–9. doi:10.1293/tox.2014-0021.
74. Lucy K, Ashok N, Maya S, et al. Histomorphological comparison of dermis in different breeds of rabbits. Pharma Innovation. 2021;10:551–555.
75. Todo H. Transdermal permeation of drugs in various animal species. Pharmaceutics. 2017;9(3):33. doi:10.3390/pharmaceutics9030033.
76. Wei JCJ, Edwards GA, Martin DJ, et al. Allometric scaling of skin thickness, elasticity, viscoelasticity to mass for micro-medical device translation: from mice, rats, rabbits, pigs to humans. Sci Rep. 2017;7(1):15885. doi:10.1038/s41598-017-15830-7.
77. Bronaugh R, Hood H, Kraeling M, et al. Determination of percutaneous absorption by in vitro techniques. J Toxicol Cutaneous Ocular Toxicol. 3rd ed. 2001;20(4):423–427. doi:10.1081/CUS-120001867.
78. Chowhan ZT, Pritchard R. Effect of surfactants on percutaneous absorption of naproxen I: Comparisons of rabbit, rat, and human excised skin. J Pharm Sci. 1978;67(9):1272–1274. doi:10.1002/jps.2600670921.
79. Walker M, Dugard P, Scott R. In vitro percutaneous absorption studies: a comparison of human and laboratory species. Human Toxicol. 1983;2:561.
80. Nixon GA, Bannan EA, Gaynor TW, et al. Evaluation of modified methods for determining skin irritation. Regul Toxicol Pharmacol. 1990;12(2):127–136. doi:10.1016/s0273-2300(05)80054-6.
81. Osborne R, Perkins MA. In vitro skin irritation testing with human skin cell cultures. Toxicol In Vitro. 1991;5(5-6):563–567. doi:10.1016/0887-2333(91)90094-t.
82. Régnier M, Caron D, Reichert U, et al. Reconstructed human epidermis: a model to study in vitro the barrier function of the skin. Skin Pharmacol. 1992;5(1):49–56. doi:10.1159/000211017.
83. Jung K-M, Lee S-H, Jang W-H, et al. KeraSkin™-VM: a novel reconstructed human epidermis model for skin irritation tests. Toxicol In Vitro. 2014;28(5):742–750. doi:10.1016/j.tiv.2014.02.014.
84. Hoffmann J, Heisler E, Karpinski S, et al. Epidermal-skin-test 1000 (Est-1000)—a new reconstructed epidermis for in vitro skin corrosivity testing. Toxicol In Vitro. 2005;19(7):925–929. doi:10.1016/j.tiv.2005.06.010.
85. Cotovio J, Grandidier M-H, Lelièvre D, et al. In vitro acute skin irritancy of chemicals using the validated EPISKIN model in a tiered strategy results and performances with 184 cosmetic ingredients. AATEX. 2007;14:S351–S358.
86. Katoh M, Hamajima F, Ogasawara T, et al. Assessment of human epidermal model LabCyte EPI-MODEL for in vitro skin irritation testing according to european Centre for the validation of alternative methods (ECVAM)-validated protocol. J Toxicol Sci. 2009;34(3):327–334. doi:10.2131/jts.34.327.
87. Mallampati R, Patlolla RR, Agarwal S, et al. Evaluation of EpiDerm full thickness-300 (EFT-300) as an in vitro model for skin irritation: Studies on aliphatic hydrocarbons. Toxicol In Vitro. 2010;24(2):669–676. doi:10.1016/j.tiv.2009.08.019.
88. Phillips L, Steinberg M, Maibach H, et al. A comparison of rabbit and human skin response to certain irritants - PubMed. Toxicol Appl Pharmacol. 1972;21(3):369–382. doi:10.1016/0041-008x(72)90157-3.
89. Sugiyama M, Akita M, Alépée N, et al. Comparative assessment of 24-hr primary skin irritation test and human patch test data with in vitro skin irritation tests according to OECD test guideline 439 (for quasi-drugs in Japan). J Toxicol Sci. 2018;43(12):751–768. doi:10.2131/jts.43.751.
90. Nixon GA, Tyson CA, Wertz WC. Interspecies comparisons of skin irritancy. Toxicol Appl Pharmacol. 1975;31(3):481–490. doi:10.1016/0041-008x(75)90272-0.
91. Marzulli FN, Maibach HI. The rabbit as a model for evaluating skin irritants: a comparison of results obtained on animals and man using repeated skin exposures. Food Cosmet Toxicol. 1975;13(5):533–540. doi:10.1016/0015-6264(75)90008-5.
92. Hobson D, Blank J. In vitro alternative methods for the assessment of dermal irritation and inflammation. In: Dermal and ocular toxicology: fundamentals and methods. Boca Raton: CRC Press; 1991.
93. Schäfer-Korting M, Bock U, Diembeck W, et al. The use of reconstructed human epidermis for skin absorption testing: results of the validation study. Altern Lab Anim. 2008;36(2):161–187. doi:10.1177/026119290803600207.
94. Corvaro M, Gehen S, Andrews K, et al. A retrospective analysis of in vivo eye irritation, skin irritation and skin sensitisation studies with agrochemical formulations: Setting the scene for development of alternative strategies. Regul Toxicol Pharmacol. 2017;89:131–147. doi:10.1016/j.yrtph.2017.06.014.
95. van der Zalm AJ, Barroso J, Browne P, et al. A framework for establishing scientific confidence in new approach methodologies. Arch Toxicol. 2022;96(11):2865–2879. doi:10.1007/s00204-022-03365-4.
96. Alépée N, Grandidier M-H, Tornier C, et al. An integrated testing strategy for in vitro skin corrosion and irritation assessment using SkinEthic™ reconstructed human epidermis. Toxicol In Vitro. 2015;29(7):1779–1792. doi:10.1016/j.tiv.2015.07.012.
97. Coquette A, Berna N, Vandenbosch A, et al. Analysis of interleukin-1alpha (IL-1alpha) and interleukin-8 (IL-8) expression and release in in vitro reconstructed human epidermis for the prediction of in vivo skin irritation and/or sensitization. Toxicol In Vitro. 2003;17(3):311–321. doi:10.1016/s0887-2333(03)00019-5.
98. Olsen DS, Lee M, Turley AP. Assessment of test method variables for in vitro skin irritation testing of medical device extracts. Toxicol In Vitro. 2018;50:426–432. doi:10.1016/j.tiv.2017.11.012.
99. Fowler JF, Zirwas MJZJ, Napolitano L, et al. A novel multifactorial approach to developing mild laundry detergents and assessing their relative mildness. J Drugs Dermatol. 2017;16:1235–1239.
100. Nations Unies. Commission économique pour l'Europe. Globally harmonized system of classification and labelling of chemicals (GHS). New York; Geneva: United Nations; 2005.