Human-Derived In Vitro Models Used for Skin Toxicity Testing Under REACh

Susanne N. Kolle and Robert Landsiedel

Contents

Abstract

In regulatory toxicology, in vivo studies are still prevailing, and human-derived in vitro models are mostly used in testing for local toxicity to the skin and the eyes. A single in vitro model may be limited to address one or few molecular or cellular events leading to adverse outcomes. Hence, in many instances their regulatory use involves the combination of several in vitro models to assess the hazard potential of test substance. A so-called defined approach combines different testing methods and a 'data interpretation procedure' to obtain a comprehensive overall assessment which is used for the regulatory hazard classification of the test substance.

Validation is a prerequisite of regulatory acceptance of new testing methods: This chapter provides an overview of the method development from an experi-

S. N. Kolle · R. Landsiedel (✉)
BASF SE, Experimental Toxicology and Ecology, Ludwigshafen am Rhein, Germany

mental method to a test guideline via application of GIVIMP (good in vitro method practice), standardization, validation to the regulatory adoption as an OECD test guidelines. Quandaries associated with the validation towards reference data from in vivo animal studies with limited accuracy and limited human relevance are discussed, as well as uncertainty and limitations arising from restricted applicability and technical and biological variance of the in vitro methods.

This chapter provides an overview of human-derived in vitro models currently adopted as OECD test guidelines: From the first skin corrosion tests utilizing reconstructed human epidermis models (RhE), to models to test for skin irritation, phototoxicity, eye irritation, and skin sensitization. The latter is using a battery of different methods and defined approaches which are still under discussion for their regulatory adoption. They will be a vanguard of future applications of human-derived models in regulatory toxicology. RhEs for testing of genotoxicity and of dermal penetration and absorption, have been developed, underwent validation studies and may soon be adopted for regulatory use; these are included in this chapter.

Abbreviations

AOP	Adverse outcome pathway
cat.	Category
DA	Defined approach
DIP	Data interpretation procedure
DPRA	Direct peptide reactivity assay
e.g.	exempli gratia
EC	European Council
ECHA	European Chemicals Agency
EIT	Eye irritation test
et al.	et alii, et aliae
EU	European Union
F/RAND	Fair, reasonable, and nondiscriminatory
FI	Fluorescence intensity
FP	False positives
GHS	Globally Harmonized System of Classification and Labelling of Chemicals
GIVIMP	Good in vitro method practice
GLP	Good laboratory practice
h-CLAT	Human cell line activation test
IL	Interleukin

ITS	Integrated testing strategy
KE	Key event
LLNA	Murine local lymph node assay
MIE	Molecular initiating event (KE1)
MPD	Mean peptide depletion
NPV	Negative prediction value
OECD	Organisation for Economic Co-operation and Development
PPV	Positive prediction value
RhCE	Reconstructed human cornea-like epithelium model
RhE	Reconstructed human epidermis model
SCT	Skin corrosion test
SI	Stimulation index
SIT	Skin irritation test
STS	Sequential testing strategy
TG	Test guideline no.
TN	True negatives
TP	True positives
UV	Ultraviolet light

1   Introduction

Traditionally, in toxicology, animal models are used, which require the extrapolation from observations in animals to humans. On the other hand, animal models provide the complexity of a whole organism which is not offered by cell or tissue cultures. Effects on the skin and eye are, however, local and may not involve complex interactions of different organs in the body. Therefore, human-derived in vitro models using cell or tissue cultures could avoid interspecies difference and yet be sufficient to address local effects. Even for local effects, a single in vitro model may be limited to one or few potential effects or biological events leading to adverse effects on human skin; this can be addressed by usage of testing batteries of several in vitro models: Regulatory use of in vitro models usually involves the combination of several models to assess the hazard potential of test substance. This requires a battery of testing methods and a data interpretation procedure (DIP) to combine the outcomes of these tests. A pre-defined set of testing methods and the fitting DIP is called defined approach.

The present chapter provides an overview of regulatorily accepted test methods based on human-derived in vitro models and defined approaches. The validation of testing methods is a prerequisite of their regulatory acceptance. Adopted test methods are available at the OECD's webpage.[1] OECD-adopted test methods are

---

[1]At https://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788.

usually taken up in the Annex to the EU Test Methods Regulation (Regulation (EC) No 440/2008 n.d.). Formal adoption of new methods or changes in existing methods by the OECD test guidelines and subsequently in the Annex to the EU Test Methods Regulation may take some time. However, "the latest version of an adopted test guideline should always be used when generating new data, independently of whether it is published by EU or OECD" (ECHA Endpoint Specific Guidance document. ECHA 2017a). In this chapter, we are referring to OECD test guideline methods.

2   Validation and Regulatory Acceptance

Since the early 2000s, several regulations include non-animal test methods (e.g., the European Chemicals Regulation REACH (Registration, Evaluation, Authorization and Restriction of Chemicals); (Regulation (EC) No 1907/2006 n.d.)) or even completely rely on in vitro (and other non-animal) testing methods, such as the EU Cosmetics Regulation (Regulation (EC) No 1223/2009 n.d.).

The validation of an in vitro method and its adoption in an OECD test guideline do not inevitably lead to its global regulatory acceptance and use, much less the complete replacement of the respective in vivo study. The regulatory acceptance of (in vitro) test data depends on regional authorities' regulations and may also be sector specific (e.g., different for pharmaceuticals, chemicals, cosmetics, and agrochemical formulations). The differences in regional regulatory needs to address skin sensitization have been exemplified in Daniel et al. (2018), and hurdles in regulatory acceptance of in vitro skin irritation and sensitization methods and use have been described by Sauer et al. (2016) and Eskes (2019).

The so-called mutual acceptance of data (MAD) avoids unnecessary repetition of tests for individual countries. Instead, all OECD member countries accept a study, which was performed according to an OECD test guideline and under good laboratory practice (GLP). Until writing of this chapter (January 2020), the MAD applies to individual test methods only, as there are no adopted guidelines for DA yet (see Sect. 3.4.4 Defined Approaches: Combination of in vitro methods to assess skin sensitization). The lack of mutual acceptance for defined approaches hampers the full regulatory acceptance of data obtained with human-derived in vitro models and hence the replacement of in vivo studies (Sauer et al. 2016); the OECD is currently working on validating and implementing DAs into its test guidelines.

The principles of the modular approach of validation have been described and evolved in several publications (Hartung et al. 2004, OECD guidance document no. 34 (OECD 2005), Zuang et al. 2015). A central part of method validation is the assessment of the method's reliability (i.e., to determine the test's intra- and interlaboratory variability and transferability) and its relevance (i.e., analyzing the test's predictive capacity as well as understanding its applicability domain) (Fig. 1).

The intra- (i.e., within a certain lab) and interlaboratory (i.e., between different labs) reproducibility is typically determined biostatistically using the data generated in ring trial studies with at least three participating labs which are blinded for test substances' identities. The assessment of the predictive capacity of a testing method
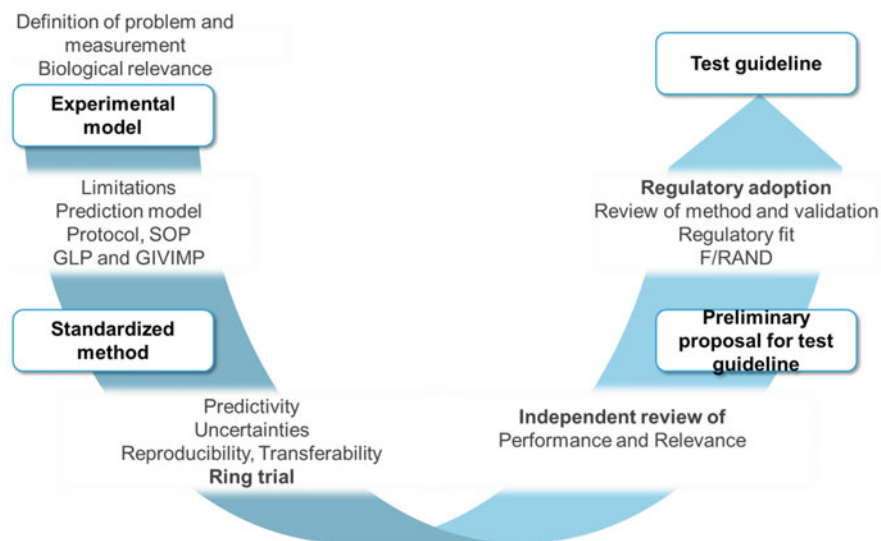
Fig. 1 Method development: From an experimental method to a test guideline via standardization, validation, and regulatory adoption. *F/RAND* fair, reasonable, and nondiscriminatory (ICH 2010; FDA 2011), *GIVIMP* good in vitro method practice (OECD 2018d), *SOP* standard operating procedure

or defined approach requires the testing of substances with well-known reference data (see also subchapter "Reference Data and Validation Sets"). The predictivity of the novel methods is assessed by comparing the results obtained with this method to the reference data. The predicitivity is described by the true positive rate (sensitivity), true negative rate (specificity), and the overall accuracy which are calculated according to Table 1 (Cooper et al. 1979).

It shall be mentioned that this comparison of new in vitro models with traditional in vivo models is questionable: The identification of hazard properties of a test substance and the classification and labelling criteria were defined according to the animal models, e.g., and it is grueling to try to accurately reproduce results defined by the parameters of an in vivo method by in vitro models. Moreover, the reproducibility of in vivo test results is limited, even though the test methods are highly standardized: The reproducibility of the refinement in vivo study, the murine local lymph node assay (OECD test guideline 429, OECD 2010), was found to be 89% (based on 296 test substances) and 73% for seriously eye-damaging findings in the Draize rabbit eye irritation (based on 46 test substances) (Luechtefeld et al. 2016a, b). As early as 1971, Weil and Scala reported on the intra- and interlaboratory variability of rabbit eye and skin irritation tests in 25 different laboratories and concluded that "The all-or-none, irritant or nonirritant, eye or skin rating of the reference samples was determined quite differently in different laboratories" (Weil and Scala 1971). In other words, it is a forlorn task to exactly reproduce the results of imperfect in vivo animal methods by in vitro models. Instead, we should strive for

Table 1 Calculation of a test method's predictive performance (confusion matrix)

		"True" condition		
		Reference data POSITIVE	Reference data NEGATIVE	
Predicted	Test data POSITIVE	TRUE POSITIVE	FALSE POSITIVE (Type I error)	$PPV = \dfrac{\sum TP}{\sum TP + \sum FP}$
	Test data NEGATIVE	FALSE NEGATIVE (Type II error)	TRUE NEGATIVE	$NPV = \dfrac{\sum TN}{\sum TN + \sum FN}$
		$Sensitivity = \dfrac{\sum TP}{\sum TP + \sum FN}$	$Specificity = \dfrac{\sum TN}{\sum TN + \sum FP}$	
		$Accuracy = \dfrac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} = \dfrac{\sum TP + \sum TN}{\sum Reference\ data}$		

With *FN* false negatives, *FP* false positives, *TN* true negatives, *TP* true positives, *PPV* positive prediction value, *NPV* negative prediction value

human relevance and address disturbance of relevant physiological processes in humans.

3 Regulatory-Accepted Human-Derived In Vitro Models

3.1 Skin Irritation and Corrosion

3.1.1 Testing Methods: Reconstructed Human Epidermis (RhE) Used in OECD Test Guidelines No. 431 and 439

The first regulatory-accepted non-animal method using a human-derived model is the in vitro skin corrosion test utilizing *reconstructed human epidermis model (RhE)*. Typically, the RhE are generated from non-transformed human epidermal keratinocytes forming a multilayered, highly differentiated model of the human epidermis. They consist of organized basal, spinous, and granular layers and a multilayered stratum corneum containing intercellular lamellar lipid layers arranged in patterns analogous to that found in vivo resulting in similar biochemical and physiological properties to human epidermis.

The skin corrosion test assay was first adopted by OECD in 2004 as OECD test guideline no. 431 (OECD 2019a). The corresponding skin irritation test was first adopted in 2010 (and revised several times since then) as OECD test guideline no. 439 (OECD 2019b). Skin irritation and corrosion tests using RhE are based on the experience that skin irritant and corrosive substances induce localized trauma as the underlying mechanism of skin irritation in vivo. The RhE-based tests are designed to predict a skin corrosion or irritation potential of a test substance after exposure on a RhE. Testing according to both OECD test guidelines can be conducted with several commercially available tissues (with similar but distinct exposure protocols and prediction models for each of the different models and irritation and corrosion endpoints).

After application of the test material to the *stratum corneum* surface of the reconstructed tissue, the induced cytotoxicity is measured by a colorimetric assay. Cytotoxicity is expressed as the reduction of mitochondrial dehydrogenase activity measured by the amount of reduced tetrazolium dye. After isopropanol extraction of the formazan from the tissues, the optical density of the extract is determined spectrophotometrically and compared to negative control values[2] to express relative tissue viability. Test substances reducing viability below certain cutoffs are then identified as skin corrosive or irritant according to the prediction models described in OECD test guidelines no. 431 and 439, respectively.

---

[2]Positive and negative controls are typically included in all in vitro assays (but not in all traditional in vivo assays) as well as lists of reference substances to demonstrate proficiency in a method ("proficiency chemicals") and a list of substances to demonstrate a novel assay performs similar ("performance standards").

The prediction models of the skin corrosion test according to *OECD test guideline no. 431* have been initially developed and adopted to identify substances not corrosive to the skin and those corrosive to the skin. In the EU, evidence of toxicological effects (at the time of writing this chapter, January 2020, mostly results of animal studies) trigger classification (and labelling) of substances (Regulation (EC) No 1272/2008 n.d.). The classification criteria were agreed at the UN level, the so-called Globally Harmonized System of Classification and Labelling of Chemicals, GHS (United Nations 2007). When toxicological data on a substance meet the classification criteria, the hazards of the substance are identified by assigning a certain hazard category; i.e., a substance is classified in skin corrosion category 1 if "Destruction of skin tissue, namely, visible necrosis through the epidermis and into the dermis, in at least one tested animal after exposure ≤ 4 h" is observed in rabbits tested according to OECD test guideline no. 405 (of note, the classification criteria are defined based on results of the animal studies). This can further be subclassified into subcategories 1A, 1B, and 1C.[3] This subcategorization of skin corrosion was initially not addressed by the OECD test guideline no. 431, but corrosives were distinguished from non-corrosives, only. Since the year 2015, the OECD test guideline no. 431 now supports the subcategorization into skin corrosives 1A and a combined 1B/C. Of note, an overprediction rate of approximately 30% for substances identified as UN GHS subcategory 1A actually belonging to subcategories 1B or C has been reported (OECD 2019a). In case subcategorization of the corrosive classes is needed and in particular in cases where UN GHS subcategories 1B and 1C have to be differentiated, the biomembrane-based *Corrositex* assay (OECD test guideline no. 435, OECD 2015a) can be conducted (as this assay is not using a human-derived model, it is not discussed further here).

OECD test guideline no. 439 provides a prediction model to identify substances nonirritant to the skin. In case the test is positive, additional testing is required to provide information whether a substance should be classified as skin irritant (UN GHS category 2) or skin corrosive (UN GHS category 1).

3.1.2 Combination of Methods to Assess Skin Irritation and Skin Corrosion

As can be concluded from the predicted UN GHS categories from OECD test guidelines nos. 431 and 439 and summarized in Fig. 2, in many cases a combination of assays (OECD guidance document no. 203, OECD 2014) is needed to cover the full irritation scale that was covered by the in vivo skin irritation test (OECD test guideline no. 404, OECD 2015b).

---

[3]Subcategory 1A "Corrosive responses in at least one animal following exposure ≤ 3 min during an observation period ≤ 1 h," Subcategory 1B "Corrosive responses in at least one animal following exposure >3 min and ≤ 1 h and observations ≤ 14 days," and Subcategory 1C "Corrosive responses in at least one animal after exposures >1 h and ≤ 4 h and observations ≤ 14 days" (ECHA 2017b).
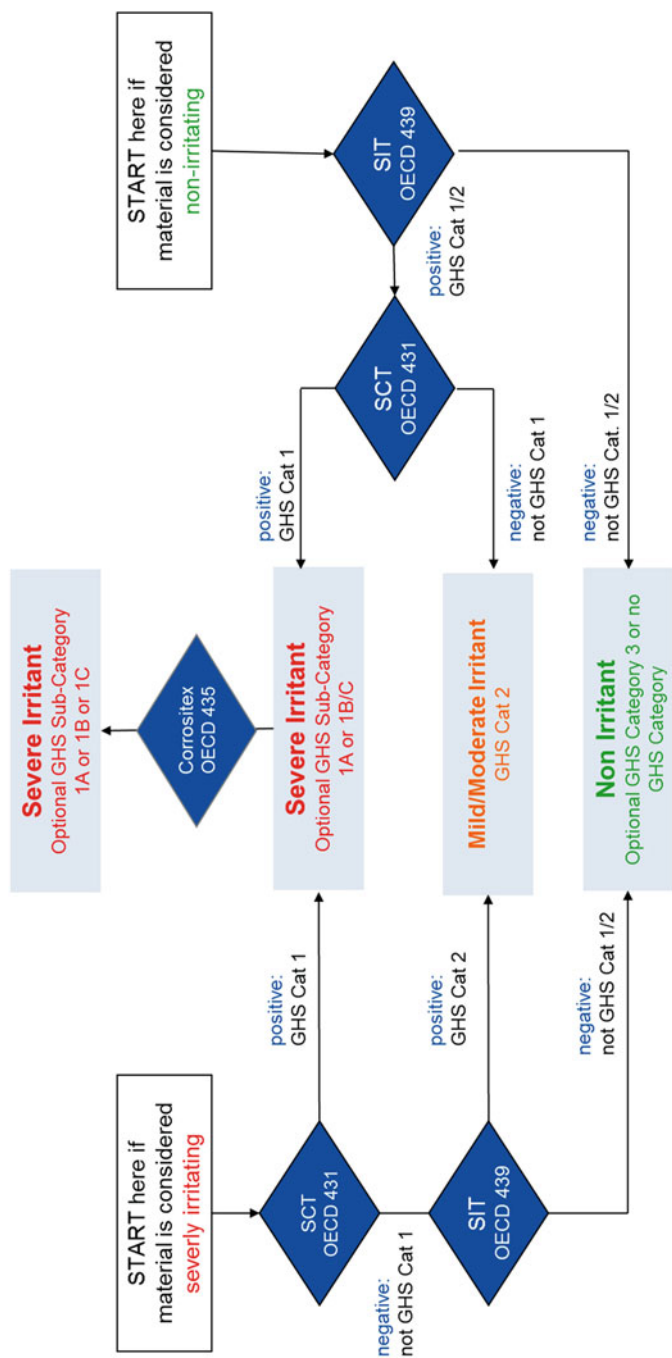
Fig. 2 Combination of in vitro methods to assess skin irritation/corrosion. Depending on the expected effects (severely irritating or nonirritating), the first test conducted is chosen. *Cat* category, *OECD* OECD test guideline, *SCT* skin corrosion test, *SIT* skin irritation test

3.2   Phototoxicity

Substances applied to the skin may form active substances by sunlight irradiation causing phototoxic (irritating) effects. Standardized and internationally harmonized in vitro methods and a tiered testing strategy are available, to test for these effects (Kolle et al. 2018). The 3T3 Neutral Red Uptake (NRU) phototoxicity test method (OECD test guideline no. 432, OECD 2019i) is using a mouse fibroblast line (BALB/3T3). Human-derived methods using RhE were developed and successfully pre-validated for phototoxicity assessment (Liebsch et al. 1999) and have been added to the OECD work plan in 2019. Both models were found to be overpredictive (Jirova et al. 2007), and today testing is usually performed according to a tiered testing strategy including light absorption, photoreactivity (formation of reactive oxygen species, ROS; *OECD test guideline no. 495*, OECD 2019j) of the test substance, as well as its distribution to the human skin (ICH 2013).

3.3   Eye Irritation

3.3.1   Testing Methods: Reconstructed Human Cornea-Like Epithelium Models (RhE) Used in OECD Test Guideline No. 492

The eye irritation test (EIT) based on *reconstructed human cornea-like epithelium models (RhCE)* was first adopted as *OECD test guideline no. 492* in 2015 (OECD 2019c). The RhCE tissue models are three-dimensional, non-keratinized tissue constructs composed of normal human-derived epidermal keratinocytes used to model the human corneal epithelium. RhCE have similar biochemical and physiological properties to human cornea epithelium.

After application of the test material to the surface of the RhCE, the induced cytotoxicity (= loss of viability, specifically of mitochondrial dehydrogenase activity) is measured by a colorimetric assay. Test substances that do not reduce viability below certain cutoffs are then identified as nonirritant to the eye according to the prediction model described in OECD test guideline no. 492. OECD test guideline no. 492 can be conducted with several commercially available tissues (with similar but distinct exposure protocols and prediction models for each of the different models and irritation and corrosion endpoints; OECD 2019c).

3.3.2   Testing Methods: Immortalized Corneal Epithelial Cells Used in OECD Test Guideline No. 494

In the *Vitrigel*-EIT, immortalized corneal epithelial cells are fabricated in a collagen *Vitrigel* membrane chamber. In this assay the time-dependent change in transepithelial electrical resistance is used to monitor the disruption of the barrier function. The *Vitrigel*-EIT assay has been adopted as *OECD test guideline no. 494* in 2019 for the identification of ocular nonirritants and seriously eye-damaging substances (UN GHS category 1) (OECD test guideline no. 494, OECD 2019d).

3.3.3 Defined Approaches: Combination of Methods to Assess Eye Irritation and Serious Eye Damage

In 2010 the concept of top-down and bottom-up approaches has been described for eye irritation (Scott et al. 2010) for the replacement of the in vivo eye irritation test (OECD test guideline no. 405, OECD 2017a). Like with skin irritation and corrosion testing, the first test to be conducted is selected based on the expected ocular irritant potential (Fig. 3, OECD guidance document no. 263; OECD 2019e). Both human-derived eye irritation test methods presented above could be employed to identify ocular nonirritants, while at least an additional method is needed to identify UN GHS Cat 1 seriously eye-damaging substances. Meanwhile several in vitro methods have been adopted to identify seriously eye-damaging substances (UN GHS category 1) by the OECD: the bovine corneal opacity and permeability test using bovine corneas (OECD test guideline no. 437 (OECD 2017b)), the isolated chicken eye test using chicken eyes (OECD test guideline no. 438 (OECD 2018c)), the fluorescein leakage test method using Madin-Darby canine kidney cells (OECD test guideline no. 460 (OECD 2017c)), the short-term exposure test method using *Statens Serum Institut* rabbit cornea cells (OECD test guideline no. 491 (OECD 2018e)), or the Ocular Irritection test method using a complex macromolecular matrix (OECD test guideline no. 496 (OECD 2019k)). As none of these assays is using a human-derived model, they are not discussed further here. Two defined approaches based on in vitro bottom-up approaches combined with physiochemical properties for ocular toxicity have been added to the OECD work plan in 2019.

3.4 Skin Sensitization

The underlying mechanism of skin sensitization is quite well understood and has been broken down into an adverse outcome pathway (OECD 2012a, b). Three of the key events can be assessed experimentally using non-animal methods (OECD 2018a, b, 2019f). Chemical reactivity has been shown to be well associated with allergenic potency and has been described as the molecular initiating event in the adverse outcome pathway. As a second key event of the skin sensitization adverse outcome pathway, keratinocytes must be activated to induce essential ("danger") signalling molecules. The third key event is the activation of the skin dendritic cells as antigen-presenting cells must upregulate cell surface markers to interact with T cells.

3.4.1 Testing Methods: Synthetic Peptides Used in OECD Test Guideline No. 442C

In the direct peptide reactivity assay (DPRA), the reactivity of a test substance towards synthetic cysteine- and lysine-containing peptides is addressed. For this purpose, a single test substance concentration is incubated with synthetic peptides for ca. 24 h at ca. 25 °C, and the remaining non-depleted peptide concentrations are determined by high-performance liquid chromatography (HPLC) with gradient elution and UV detection at 220 nm.
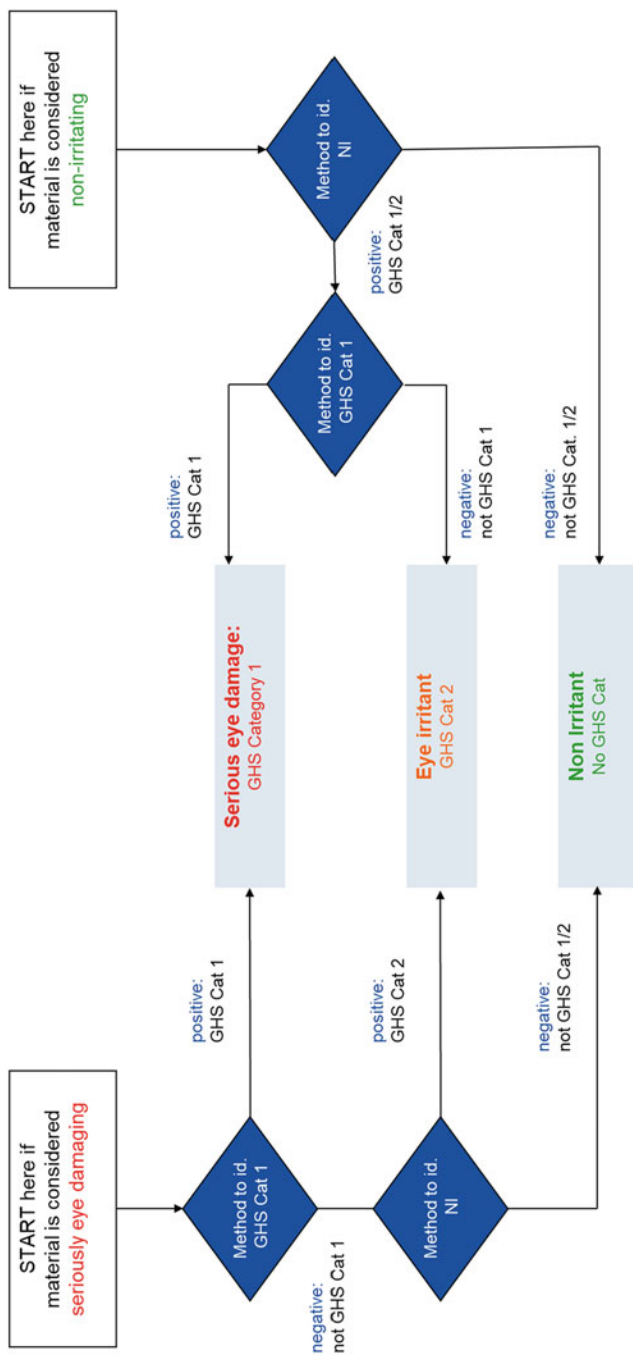
Fig. 3 Combination of in vitro methods to assess eye irritation/serious eye damage. Depending on the expected effects (seriously eye damaging or nonirritating), the first test conducted is chosen. Cat, category; NI, nonirritant; id., identify. Please note that a default GHS Cat 2 identification by exclusion of "nonirritant" and exclusion of "serious eye damage" may not be accepted by all regulatory bodies and additional information such as in vivo data may be required (OECD 2019e)

The peptide depletion of test substance incubated samples is compared to the peptide depletion of the negative control samples and expressed as relative peptide depletion. The DPRA has been first adopted as *OECD test guideline no. 442C* in 2015 (OECD 2019f).

The DPRA is not using a human-derived cell- or tissue model nor a biomacromolecule as test system, but rather a synthetic heptapeptide. The assay is, however, complementing human-derived models (described below, Sects. 3.4.2 and 3.4.3) in testing batteries to predict a skin sensitization potential in humans.

As the DPRA is not using biological systems, but rather a synthetic heptapeptide, it is often termed an *in chemico* rather than "in vitro" assay. Information on the reactivity of a test substance towards a peptide (as a proxy for skin proteins) can also be obtained by in silico methods. Several commercial and not-for-profit models have been evaluated (Teubner et al. 2013; Urbisch et al. 2016b; Fitzpatrick et al. 2018) and provided a lower overall predictivity, but good concordance with experimental results with specific models within their applicability domain. So far, peptide reactivity is used to predict a sensitization potential (presence or absence of hazard). Recently, the DPRA was extended to also predict potency classes (*kDPRA*). In the kDPRA several test substance concentrations are assessed after several incubation times to determine reaction rate constants which are then used to identify strong sensitizers (UN GHS category 1A) (Wareing et al. 2017).

3.4.2 Testing Methods: Human-Derived Keratinocytes Used in OECD Test Guideline No. 442D

As a second key event in the adverse outcome pathway for skin sensitization, keratinocyte activation can be assessed by the *KeratinoSens* and *LuSens* assays using the genetically modified human keratinocyte cell lines. Both assays employ the reporter gene for luciferase under the control of an antioxidant response element and hence monitor Nrf-2 transcription factor activity. The endpoint measurement is the upregulation of the luciferase activity after incubation with test substances. This upregulation is an indicator for the activation of the Keap1/Nrf2/ARE signalling pathway. The ARE-Nrf2 luciferase test methods have been first adopted in 2015 as *OECD test guideline no. 442D* (OECD 2018a).

3.4.3 Testing Methods: Human-Derived Dendritic-Like Cells Used in OECD Test Guideline No. 442E

Dendritic cell activation, the third key event in the adverse outcome pathway for skin sensitization, is addressed by the test methods described in *OECD test guideline no. 442E* first adopted in 2016 (OECD 2018b). The assays evaluate the potential to activate dendritic cells either by measuring changes in the cell surface marker expression (human cell line activation test (*h-CLAT*) and the U937 Cell Line Activation Test (*U-SENS*)) or by means of inducing the cytokine IL-8 in the interleukin-8 reporter gene assay (*IL8LUC*). The h-CLAT is performed using the human monocytic leukemia cell line THP-1 as surrogate for dendritic cells. As readout, the change in the expression of the cell membrane markers CD 54 and CD 86 is determined by flow cytometry after test substance exposure. Similarly, in

the U-SENS the change in the expression of the cell membrane marker CD 86 measured by flow cytometry after test substance exposure of U937 cells is determined. In the IL8LUC a THP-1 derived IL-8-reporter cell line, IL-8 dependent luciferase activity is determined after test substance exposure (OECD 2018b).

3.4.4 Defined Approaches: Combination of In Vitro Methods to Assess Skin Sensitization

Although non-animal methods addressing individual key events of the skin sensitization adverse outcome pathway are available as OECD-adopted test methods, none of the available methods should be considered as a stand-alone method to address the endpoint of skin sensitization, but rather the methods have to be combined in defined approaches. To conclude on the sensitizing potential of a test substance, the data from several methods are combined to a defined approach in which a fixed data interpretation procedure serves as prediction model for the combination of results. Several defined approaches have been described (Table 2, OECD 2016a, b), and in the following, we briefly describe one of the less complex defined approaches for the identification of the skin sensitization hazard (Fig. 4, Bauch et al. 2012; Urbisch et al. 2015).

In the *2 out of 3 approach* (Case study 1 in Table 2), assays addressing three of the key events of the skin sensitization adverse outcome pathway are conducted, and two concordant results determine the overall hazard prediction (i.e., if a test substance is positive in any two of the three assays, it is predicted to be a sensitizer).

This defined approach (and indeed most defined approaches) is combining the data of several test methods adopted by OECD. Yet, the adoption of defined approaches into OECD test guideline is still pending at the time of writing of this chapter (January 2020) and this is prohibiting defined approaches from providing the same regulatory recognition and mutual acceptance of data as the animal tests. Since

Table 2 Case studies of defined approaches described in OECD GD 256 (OECD 2016a, b)

|	
1	An adverse outcome pathway-based "2 out of 3" integrated testing strategy approach to skin hazard identification
2	Sequential testing strategy (STS) for hazard identification of skin sensitizers
3	A non-testing pipeline approach for skin sensitization
4	Stacking meta-model for skin sensitization hazard identification
5	Integrated decision strategy for skin sensitization hazard
6	Consensus of classification trees for skin sensitization hazard prediction
7	Sensitizer potency prediction based on key event 1 + 2: Combination of kinetic peptide reactivity data and *KeratinoSens*® data
8	The artificial neural network model for predicting LLNA EC3
9	Bayesian network DIP (BN-ITS-3) for hazard and potency identification of skin sensitizers
10	Sequential testing strategy (STS) for sensitizing potency classification based on in chemico and in vitro data
11	Integrated testing strategy (ITS) for sensitizing potency classification based on in silico, in chemico, and in vitro data
12	DIP for skin allergy risk assessment (SARA)
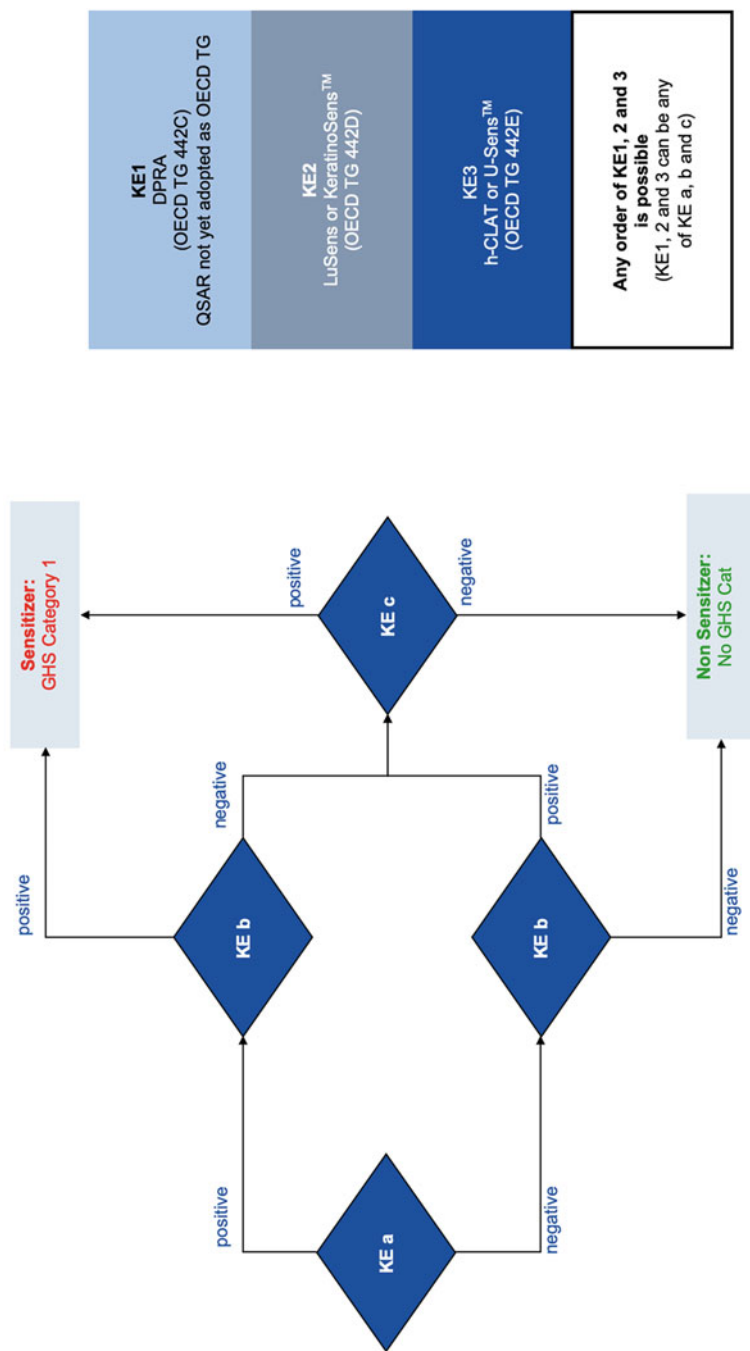
Fig. 4 The 2 out of 3 defined approach (according to Urbisch et al. 2015). KE, key event in the adverse outcome pathway for skin sensitization with key events (KE) A, KE B, and KE C being key events 1, 2, and 3 in an arbitrary sequence. Any two concordant results determine the overall prediction with at least 2 out the 3 KE addressed. KE1 is addressed by the DPRA (OECD test guideline no. 442C), KE2 by the LuSens/KeratinoSens™ assays (OECD test guideline no. 442D), and KE3 by the h-CLAT (OECD test guideline no. 442E)

the "2 out of 3" approach has been first submitted for regulatory acceptance to the European Centre for the Validation of Alternative Methods in 2011, a lot of progress has been made. A project to draft a guideline has been added to the OECD work plan in 2017, and a second draft guideline and supporting documents became available in September 2019 (OECD 2019g, h). The work undertaken to draft these documents included an extensive review of the human and mouse skin sensitization reference data (see also Reference Data and Validation Sets for a discussion of reference data in more general) as well as a discussion about the applicability domains (extending on the applicability domains of the individual assays).

3.5   Genotoxicity

There is a variety of in vitro genotoxicity and mutagenicity models available and used within testing batteries (Kirkland et al. 2006; ICH 2011; SCCS 2014). Interestingly none of the regulatorily accepted models is using human-derived models except for the so-called HuLy assay which is utilizing primary human lymphocytes (OECD test guideline no. 487, OECD 2016c). Recently, methods to detect genotoxic and mutagenic effects in human-derived reconstructed epidermal models were developed and validated in ring trials to detect genotoxicity (Reisinger et al. 2018) and chromosomal aberrations (Curren et al. 2006; Aardema et al. 2010). These methods await the finalization of their validation processes and inclusion in OECD test guidelines to be used within genotoxicity and mutagenicity in vitro testing batteries with the in vitro genotoxicity test for dermal exposure using 3D models added to the OECD work plan.

3.6   Dermal Penetration and Absorption

Dermal penetration and absorption methods are not testing for adverse effects on human skin but are assessing the penetration of dermally applied substance into the skin and through the skin to become systemically available in the human body. The OECD-adopted in vitro method (OECD test guideline no. 428; OECD 2004; Fabian et al. 2017) is utilizing human skin preparations. Non-viable skin can also be used provided that the integrity of the skin can be demonstrated. Either epidermal membranes or split-thickness skin (typically 200–400 μm thick) prepared with a dermatome are acceptable. The principal diffusion barrier of the skin is the non-viable stratum corneum; active transport of chemicals through the skin has not been identified, and dermal metabolism (Bätz et al. 2013; Oesch et al. 2018) is not rate limiting in terms of actual absorbed dose (OECD 2004). Methods to utilize human reconstructed epidermal or full-thickness skin models have been developed and pre-validated (Schäfer-Korting et al. 2008; Ackermann et al. 2010) but are not yet regulatorily accepted.

4 Limitations

4.1 Technical Limitations

Specific test substances may not be applicable to certain test systems. Some examples of these technical limitations are listed below (using the example of the in vitro skin sensitization methods).

In the DPRA (OECD test guideline no. 442C), the depletion of the synthetic heptapeptide is quantified by its UV light absorption after HPLC elution. Test substances co-eluting at the same retention times as the model peptides may hamper the peptide quantification.

The cell-based assay methods (OECD test guidelines no. 442D and 442E) use luciferase-generated bioluminescence or fluorescence of fluorochrome-labelled antibodies as detection methods. Test substances quenching the fluorescence or otherwise interfering with the optical detection may hamper the quantification of the luciferase induction or the identification of labelled cells.

Dendritic cell activation (OECD test guideline no. 442E) is analyzed by flow cytometric determination of the cell surface marker expression. Insoluble particles and polymers may however limit the technical applicability by clogging the flow cytometer.

4.2 Predictive Limitations

4.2.1 Mechanistic Limitations

In the following we present (and partially discuss) known mechanism limitation of in vitro assays. This is not to be understood as a comprehensive list; we'd like to present a few examples based on the in vitro skin sensitization assays.

OECD Test Guideline No. 442C The DPRA is based on reactivity of a test substance with cysteine and lysine residues. Metals do not form covalent bonds with those two amino acid residues and hence are out of the applicability domain of the assay. Also, test substance reacting with amino acid residues different from cysteine and lysine will not be detected in the DPRA. It has been described that some test substances favor the dimerization or oxidation of the peptide leading to an overestimation of a true peptide depletion (or non-covalent, specific binding, e.g., Roediger and Weninger 2011).

OECD Test Guideline No. 442D The underlying mechanism for the antioxidant response element pathway addressed in both the *KeratinoSens* and *LuSens* assays is closely linked to cysteine reactivity. Therefore, test substance primarily reacting with other amino acid residues (such as acylating agents reacting with lysine) would be expected to be underpredicted in the *KeratinoSens* and *LuSens* assays.

Metabolic Capacity

The three in vitro skin sensitization tests described above do not contain any external source of metabolic capacity. Nevertheless, the test systems can detect most pre- and pro-haptens. In vitro investigations (Urbisch et al. 2016a; Patlewicz et al. 2016) using test substances requiring molecular transformation to attain a sensitizing potential have shown that pre-haptens can readily be detected in the DPRA, many of which involve autoxidation processes. Moreover, many pro-haptens are also activated by nonenzymatic oxidation (and therefore are pre- *and* pro-haptens). The cellular models h-CLAT and *LuSens* have been shown to detect pro-haptens more efficiently; respective enzyme activities were detected in the cell lines (Fabian et al. 2013; Oesch et al. 2018). Thus, it can be concluded that potentially relevant molecular transformations are generally sufficiently considered using the in vitro skin sensitization tests DPRA, *LuSens* or *KeratinoSens*, and h-CLAT.

Water Solubility and Lipophilicity

OECD test guideline no. 442E describes that the h-CLAT result may be underpredictive (false negative) for test substances with log $K_{OW} > 3.5$. In this case a "negative" result is interpreted as "inconclusive." However, a "positive" result will be accepted. While with increasing log K_{ow}, less solubility is expected, it should however not be neglected that OECD test guideline no. 442E (OECD 2018b) and OECD test guideline no. 442D (OECD 2018a) allow the testing of homogenous but non-completely dissolved test substances.

4.2.2 Agrochemical Formulation in In Vitro Skin and Eye Irritation Tests

(Non-animal) tests are typically validated against well-characterized individual reference chemicals. Different chemistries, use classes, and, e.g., special types of mixtures can, however, not always (comprehensively) be included in validation exercises. Therefore, important post-validation experience is gained during routine testing (frequently after test guideline adoption and regulatory acceptance). We present here two examples of test methods both based on reconstructed human tissues. Agrochemical formulations are a special type of mixture (which as such fall under the GHS mixture definition and could generally be considered within the applicability domain of the OECD test guidelines) containing a variety of ingredients to alter the properties of the active ingredient. Oftentimes thereby toxicity is also affected, and rules of additivity do not simply apply. In 2015 Kolle et al., based on a comparative dataset of 97 agrochemical formulations, have reported excellent sensitivity of the *EpiOcular* eye irritation test assay to predict agrochemical formulations nonirritant to the eye (until the writing of this chapter (January 2020), there is still no non-animal method to reliably predict seriously eye-damaging agrochemical formulations) (Kolle et al. 2015, 2017a). This would lead to the notion that maybe reconstructed human tissues work well for the lower end of the irritation scale also for skin. This could unfortunately not be confirmed for a comparative dataset of 65 agrochemical formulations which showed that the in vitro skin irritation test was

neither sufficiently sensitive nor specific (Kolle et al. 2017b). Also proof of concept with five formulations assessed using a protocol modification of the SIT (using a 15 min exposure followed by a 24 or 42 h post-exposure instead of a 60 min exposure followed by a 42 h post-exposure in the EpiDerm SIT according to OECD test guideline no. 439) was not successful (unpublished data). Therefore, most unfortunately, the in vivo assay (OECD test guideline no. 404) is still needed to evaluate the skin irritation potential of agrochemical formulations.

At the time of writing of this chapter (January 2020), there is no regulatory-accepted method available to reliably predict the skin irritation potential of agrochemical formulations; the development and validation of such methods should be fostered.

4.3   Uncertainty

4.3.1   Reference Data and Validation Sets

When evaluating or implementing novel (non-animal) methods, it is of upmost importance to use substance with critically reviewed reference data. There have been reports on cases where method implementation based on so-called proficiency chemicals provided in OECD test guidelines has been hampered (Kolle et al. 2019a, b). Another example is the evaluation of the defined approach for skin sensitization which started with 128 substances for which human as well as local lymph node assay data was available. During the review of the guideline, the reference data was extensively reviewed and resulted in a reduced curated dataset of 105 substances with local lymph node assay data and 76 substances with human reference data (OECD 2019h).

4.3.2   Borderline Range: Uncertainty Arising from Technical and Biological Variance

The borderline range depicts the variance of the individual test methods, including technical and biological variability (Leontaridou et al. 2017). It addresses the uncertainty of the three assays around their respective classification thresholds and represents a range in which the likelihood to obtain a positive or negative result just below or above the classification threshold is equal (Fig. 5).

The borderline range can be determined statistically (e.g., using pooled standard deviations), using historical intra-laboratory data (Leontaridou et al. 2017). It is useful especially for assays for which no individual statistical analysis is possible due to low number of replicates per treatment (e.g., h-CLAT and DPRA). This evaluation is an amendment to the evaluation given in the respective OECD test guidelines, and it also influences a method's precision (Leontaridou et al. 2019) (Table 3).

The definition of a borderline range allows the possible prediction as "ambiguous" underlining the fact that a result close to a classification cutoff is rather random.

Table 3 summarizes the borderline range for the in vitro skin sensitization test methods. Borderline ranges rather than discrete cutoff values should be used in
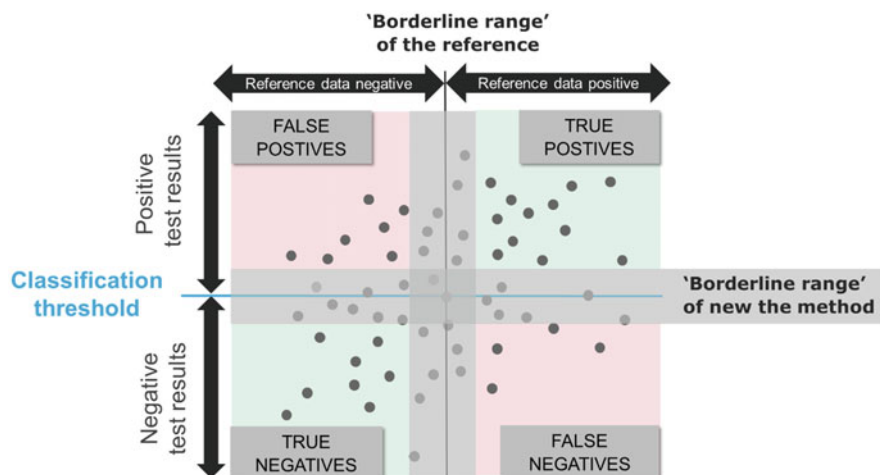
Fig. 5 The borderline range of new and reference methods

Table 3 Pooled standard deviations and borderline ranges of in vitro methods to predict skin sensitization in humans (Leontaridou et al. 2017)

Testing method	Borderline range
DPRA (OECD test guideline no. 442C)	MPD = {1.35%, 11.41%}
LuSens (OECD test guideline no. 442D)	FI = {1.26, 1.74}
h-CLAT (OECD test guideline no. 442E)	CD54 FI = {1.74, 2.26}, CD86 FI = {1.2, 1.81}
LLNA (OECD test guideline no. 429) for comparison	SI = {2.20, 3.71}

Test results in the range of $3.38\% \leq MPD \leq 9.38\%$ (DPRA), $FI \leq 3$ (LuSens), and CD86 $FI \leq 3$ as well as CD54 $FI \leq 3$ (h-CLAT) were considered for quantifying the pooled standard deviation (Leontaridou et al. 2017)

DPRA direct peptide reactivity assay, *h-CLAT* human cell line activation test, *FI* fluorescence intensity, *MPD* mean peptide depletion, *SI* stimulation index

prediction models (or data interpretation procedures, DIP), and potential outcomes of studies which dichotomize continuous data into classes are really "positive," "negative," and "inconclusive."

References

Aardema MJ, Barnett BC, Khambatta Z, Reisinger K, Ouedraogo-Arras G, Faquet B et al (2010) International pre-validation studies of the EpiDerm™ 3D human reconstructed skin micronucleus (RSMN) assay: transferability and reproducibility. Mutat Res Genet Toxicol Environ Mutagen 701(2):123–131

Ackermann K, Borgia SL, Korting HC, Mewes KR, Schäfer-Korting M (2010) The Phenion® full-thickness skin model for percutaneous absorption testing. Skin Pharmacol Physiol 23 (2):105–112

Bätz FM, Klipper W, Korting HC, Henkler F, Landsiedel R, Luch A et al (2013) Esterase activity in excised and reconstructed human skin–biotransformation of prednicarbate and the model dye fluorescein diacetate. Eur J Pharm Biopharm 84(2):374–385

Bauch C, Kolle SN, Ramirez T, Eltze T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B, Landsiedel R (2012) Putting the parts together: combining in vitro methods to test for skin sensitizing potentials. Regul Toxicol Pharmacol 63(3):489–504

Cooper J, Saracci R, Cole P (1979) Describing the validity of carcinogen screening tests. Br J Cancer 39:87–89

Curren RD, Mun GC, Gibson DP, Aardema MJ (2006) Development of a method for assessing micronucleus induction in a 3D human skin model (EpiDerm). Mutat Res 607(2):192–204

Daniel AB, Strickland J, Allen D, Casati S, Zuang V, Barroso J, Whelan M, Régimbald-Krnel MJ, Kojima H, Nishikawa A, Park HK, Lee JK, Kim TS, Delgado I, Rios L, Yang Y, Wang G, Kleinstreuer N (2018) International regulatory requirements for skin sensitization testing. Regul Toxicol Pharmacol 95:52–65

ECHA (2017a) Guidance on information requirements and chemical safety assessment chapter R.7a: endpoint specific guidance version 6.0. https://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf

ECHA (2017b) Guidance to regulation (EC) no 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. Version 5.0, July 2017. https://echa.europa.eu/documents/10162/23036412/clp_en.pdf

Eskes C (2019) The usefulness of integrated strategy approaches in replacing animal experimentation. Ann Ist Super Sanita 55(4):400–404

Fabian E, Vogel D, Blatz V, Ramirez T, Kolle S, Eltze T, van Ravenzwaay B, Oesch F, Landsiedel R (2013) Xenobiotic metabolizing enzyme activities in cells used for testing skin sensitization in vitro. Arch Toxicol 87(9):1683–1696

Fabian E, Oesch F, Ott K, Landsiedel R, van Ravenzwaay B (2017) A protocol to determine dermal absorption of xenobiotica through human skin in vitro. Arch Toxicol 91(3):1497–1511

FDA (2011) Guidance for industry process validation: general principles and practices. https://www.fda.gov/media/71021/download

Fitzpatrick JM, Roberts DW, Patlewicz G (2018) An evaluation of selected (Q) SARs/expert systems for predicting skin sensitization potential. SAR QSAR Environ Res 29(6):439–468

Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Hoffmann S, Roi AJ, Prieto P, Sabbioni E, Scott L, Worth A, Zuang V (2004) A modular approach to the ECVAM principles on test validity. Altern Lab Anim 32(5):467–472

ICH (2010) ICH harmonised tripartite guideline Q7: good manufacturing practice guide for active pharmaceutical ingredients. https://database.ich.org/sites/default/files/Q7_Guideline.pdf

ICH (2011) ICH guideline guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use s2(r1) (2011) approval by the steering committee of S2(R1) under step 4 and recommendation for adoption to the three ICH regulatory bodies (9 November 2011)

ICH (2013) International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. In: ICH harmonised tripartite guideline: photosafety evaluation of pharmaceuticals. Current step 4 version, 13, 2013

Jirova D, Liebsch M, Basketter D, Spiller E, Kejlova K, Bendova H et al (2007) Comparison of human skin irritation and photo-irritation patch test data with cellular in vitro assays and animal in vivo data. ALTEX 14:359–365

Kirkland D, Aardema M, Müller L, Makoto H (2006) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens II. Further analysis of mammalian cell results, relative predictivity and tumour profiles. Mutat Res 608 (1):29–42

Kolle SN, Rey Moreno MC, Mayer W, van Cott A, van Ravenzwaay B, Landsiedel R (2015) The EpiOcular™ eye irritation test is the method of choice for the *In vitro* eye irritation testing of agrochemical formulations: correlation analysis of EpiOcular eye irritation test and BCOP test data according to the UN GHS, US EPA and Brazil ANVISA classification schemes. ATLA Altern Lab Anim 43(3):181–198

Kolle SN, Van Cott A, van Ravenzwaay B, Landsiedel R (2017a) Lacking applicability of *in vitro* eye irritation methods to identify seriously eye irritating agrochemical formulations: Results of bovine cornea opacity and permeability assay, isolated chicken eye test and the EpiOcular™ ET-50 method to classify according to UN GHS. Regul Toxicol Pharmacol 85:33–47

Kolle SN, van Ravenzwaay B, Landsiedel R (2017b) Regulatory accepted but out of domain: in vitro skin irritation tests for agrochemical formulations. Regul Toxicol Pharmacol 89:125–130

Kolle SN, Teubner W, Landsiedel R (2018) Modern skin toxicity testing strategies. In: Environment and skin. Springer, Cham, pp 27–40

Kolle SN, Hill E, Raabe H, Landsiedel R, Curren R (2019a) Regarding the references for reference chemicals of alternative methods. Toxicol In Vitro 57:48–53

Kolle SN, Natsch A, Gerberick GF, Landsiedel R (2019b) A review of substances found positive in 1 of 3 *in vitro* tests for skin sensitization. Regul Toxicol Pharmacol 106:352–368

Leontaridou M, Urbisch D, Kolle SN, Ott K, Mulliner DS, Gabbert S, Landsiedel R (2017) The borderline range of toxicological methods: quantification and implications for evaluating precision. ALTEX 34(4):525–538

Leontaridou M, Gabbert S, Landsiedel R (2019) The impact of precision uncertainty on predictive accuracy metrics of non-animal testing methods. ALTEX-alternatives to animal experimentation

Liebsch M, Traue D, Barrabas C, Spielmann H, Gerberick F, Cruse L et al (1999) Prevalidation of the Epiderm™ phototoxicity test. In: Clark D, Lisansky S, Macmillan R (eds) Alternatives to animal testing II: proceedings of 2nd international science conference organised by the European cosmetic industry, Brussels. Belgium CPL Press, Newbury, pp 160–166

Luechtefeld T, Maertens A, Russo DP, Rovida C, Zhu H, Hartung T (2016a) Analysis of publically available skin sensitization data from REACH registrations 2008–2014. ALTEX 33(2):135–148

Luechtefeld T, Maertens A, Russo DP, Rovida C, Zhu H, Hartung T (2016b) Analysis of Draize eye irritation testing and its prediction by mining publicly available 2008–2014 REACH data. ALTEX 33(2):123–134

OECD (2004) OECD guideline for testing of chemicals test no. 428: skin absorption: in vitro method, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264071087-en

OECD (2005) OECD series on testing and assessment no 34 guidance document on the validation and international acceptance of new or updated test methods for Hazard assessment. OECD Publishing, Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2005)14&doclanguage=en

OECD (2010) OECD guideline for testing of chemicals test no. 429: skin sensitisation: local lymph node assay. OECD Publishing, Paris. https://doi.org/10.1787/9789264071100-en

OECD (2012a) OECD series on testing and assessment no. 168: the adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins, part 1: scientific evidence. OECD Publishing, Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2012)10/part1&doclanguage=en

OECD (2012b) OECD series on testing and assessment no. 168: the adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins, part 2: use of the AOP to develop chemical categories and integrated assessment and testing. OECD Publishing, Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2012)10/part2&doclanguage=en

OECD (2014) OECD series on testing and assessment no. 203: guidance document on an integrated approach on testing and assessment (IATA) for skin corrosion and irritation. OECD Publishing,

Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2014)19&doclanguage=en

OECD (2015a) OECD guideline for testing of chemicals test no. 435: *In vitro* membrane barrier test method for skin corrosion. OECD Publishing, Paris. https://doi.org/10.1787/9789264242791-en

OECD (2015b) OECD guideline for testing of chemicals test no. 404: acute dermal irritation/corrosion. OECD Publishing, Paris. https://doi.org/10.1787/9789264242678-en

OECD (2016a) OECD series on testing and assessment no.256 guidance document on the reporting of defined approaches to be used within integrated approaches to testing and assessment annex 1, OECD Publishing, Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)28&doclanguage=en

OECD (2016b) OECD series on testing and assessment no.256 guidance document on the reporting of defined approaches and individual information sources to be used within integrated approaches to testing and assessment (IATA) for skin sensitization, annex 2. OECD Publishing, Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)29&doclanguage=en

OECD (2016c) OECD guideline for testing of chemicals test no. 487: in vitro mammalian cell micronucleus test, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264264861-en

OECD (2017a) OECD guideline for testing of chemicals test no. 405: acute eye irritation/corrosion OECD. OECD Publishing, Paris. https://doi.org/10.1787/9789264185333-en

OECD (2017b) OECD guideline for testing of chemicals test no. 437: bovine corneal opacity and permeability test method for identifying (i) chemicals inducing serious eye damage and (ii) chemicals not requiring classification for eye irritation or serious eye damage, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264203846-en

OECD (2017c) Test No. 460: fluorescein leakage test method for identifying ocular corrosives and severe irritants, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264185401-en

OECD (2018a) OECD guideline for testing of chemicals test no. 442D: *in vitro* skin sensitisation: ARE-Nrf2 luciferase test method. OECD Publishing, Paris. https://doi.org/10.1787/9789264229822-en

OECD (2018b) OECD guideline for testing of chemicals test no. 442E: *in vitro* skin sensitisation: *in vitro* skin sensitisation assays addressing the key event on activation of dendritic cells on the adverse outcome pathway for skin sensitisation. OECD Publishing, Paris. https://doi.org/10.1787/9789264264359-en

OECD (2018c) OECD guideline for testing of chemicals test no. 438: isolated chicken eye test method for identifying (i) chemicals inducing serious eye damage and (ii) chemicals not requiring classification for eye irritation or serious eye damage, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264203860-en

OECD (2018d) Guidance document on good in vitro method practices (GIVIMP), OECD series on testing and assessment, no. 286. OECD Publishing, Paris. https://doi.org/10.1787/9789264304796-en

OECD (2018e) Test No. 491: short time exposure in vitro test method for identifying (i) chemicals inducing serious eye damage and (ii) chemicals not requiring classification for eye irritation or serious eye damage, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264242432-en

OECD (2019a) OECD guideline for testing of chemicals test no. 431: In vitro skin corrosion: reconstructed human epidermis (RHE) test method. OECD Publishing, Paris. https://doi.org/10.1787/9789264264618-en

OECD (2019b) OECD guideline for testing of chemicals test no. 439: in vitro skin irritation: reconstructed human epidermis test method. OECD Publishing, Paris. https://doi.org/10.1787/9789264242845-en

OECD (2019c) OECD guideline for testing of chemicals test no. 492: reconstructed human cornea-like epithelium (RhCE) test method for identifying chemicals not requiring classification and labelling for eye irritation or serious eye damage. OECD Publishing, Paris. https://doi.org/10.1787/9789264242548-en

OECD (2019d) OECD guidelines for the testing of chemicals test no. 494: Vitrigel-eye irritancy test method for identifying chemicals not requiring classification and labelling for eye irritation or serious eye damage. OECD Publishing, Paris. https://doi.org/10.1787/9f20068a-en

OECD (2019e) OECD series on testing and assessment no. 263 guidance document on integrated approaches to testing and assessment (IATA) for serious eye damage and eye irritation. OECD Publishing, Paris. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2014)19&doclanguage=en

OECD (2019f) OECD guideline for testing of chemicals test no. 442C: in chemico skin sensitisation: assays addressing the adverse outcome pathway key event on covalent binding to proteins. OECD Publishing, Paris. https://doi.org/10.1787/9789264229709-en

OECD (2019g) Draft OECD guideline defined approaches for skin sensitisation. OECD Publishing, Paris. https://www.oecd.org/env/ehs/testing/GL%20DASS_22Sep2019v2.pdf

OECD (2019h) Supporting document for evaluation and review of draft guideline (GL) for defined approaches (DAs) for skin sensitisation. OECD Publishing, Paris. https://www.oecd.org/env/ehs/testing/DAGL%20supporting%20document_23%20Sep2019.pdf

OECD (2019i) OECD guideline for testing of chemicals test no. 432: in vitro 3T3 NRU phototoxicity test, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/9789264071162-en

OECD (2019j) OECD guideline for testing of chemicals test no. 495: Ros (reactive oxygen species) assay for Photoreactivity. OECD Publishing, Paris. https://www.oecd-ilibrary.org/docserver/915e00ac-en.pdf?expires=1578570534&id=id&accname=guest&checksum=C7B321D86BFA753562D53BDAC49F6F0A

OECD (2019k) 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, OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. https://doi.org/10.1787/970e5cd9-en

Oesch F, Fabian E, Landsiedel R (2018) Xenobiotica-metabolizing enzymes in the skin of rat, mouse, pig, Guinea pig, man, and in human skin models. Arch Toxicol 92(8):2411–2456

Patlewicz G, Casati S, Basketter DA, Asturiol D, Roberts DW, Lepoittevin JP, Worth AP, Aschberger K (2016) Can currently available non-animal methods detect pre and pro-haptens relevant for skin sensitization? Regul Toxicol Pharmacol 82:147–155

Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products (n.d.). http://data.europa.eu/eli/reg/2009/1223/2019-11-27

Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (n.d.). http://data.europa.eu/eli/reg/2008/1272/oj

Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC (n.d.). http://data.europa.eu/eli/reg/2006/1907/2014-04-10

Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) (n.d.). http://data.europa.eu/eli/reg/2008/440/oj

Reisinger K, Blatz V, Brinkmann J, Downs TR, Fischer A, Henkler F et al (2018) Validation of the 3D skin comet assay using full thickness skin models: transferability and reproducibility. Mutat Res Genet Toxicol Environ Mutagen 827:27–41

Roediger B, Weninger W (2011) How nickel turns on innate immune cells. Immunol Cell Biol 89 (1):1–2

Sauer UG, Hill EH, Curren RD, Raabe HA, Kolle SN, Teubner W, Mehling A, Landsiedel R (2016) Local tolerance testing under REACH: accepted non-animal methods are not on equal footing with animal tests. ATLA Altern Lab Anim 44(3):281–299

SCCS (2014) SCCS ADDENDUM to the SCCS's notes of guidance (NoG) for the testing of cosmetic ingredients and their safety evaluation 8th revision (SCCS/1501/12). 2014. http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_156.pdf

Schäfer-Korting M, Bock U, Diembeck W, Düsing HJ, Gamer A, Haltner-Ukomadu E et al (2008) The use of reconstructed human epidermis for skin absorption testing: results of the validation study. Altern Lab Anim 36(2):161–187

Scott L, Eskes C, Hoffmann S, Adriaens E, Alepée N, Bufo M, Clothier R, Facchini D, Faller C, Guest R, Harbell J, Hartung T, Kamp H, Varlet BL, Meloni M, McNamee P, Osborne R, Pape W, Pfannenbecker U, Prinsen M, Seaman C, Spielmann H, Stokes W, Trouba K, Berghe CV, Goethem FV, Vassallo M, Vinardell P, Zuang V (2010) A proposed eye irritation testing strategy to reduce and replace in vivo studies using bottom-up and top-down approaches. Toxicol In Vitro 24(1):1–9

Teubner W, Mehling A, Schuster PX, Guth K, Worth A, Burton J et al (2013) Computer models versus reality: how well do in silico models currently predict the sensitization potential of a substance. Regul Toxicol Pharmacol 67(3):468–485

United Nations (2007) Globally harmonized system of classification and labelling of chemicals (GHS). Part 3, Health Hazards. http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev02/English/03e_part3.pdf

Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle SN, Landsiedel R, Jaworska J, Kern P, Gerberick F, Natsch A, Emter R, Ashikaga T, Miyazawa M, Sakaguchi H (2015) Assessing skin sensitization hazard in mice and men using non-animal test methods. Regul Toxicol Pharmacol 71(2):337–351

Urbisch D, Becker M, Honarvar N, Kolle SN, Mehling A, Teubner W et al (2016a) Assessment of pre-and pro-haptens using nonanimal test methods for skin sensitization. Chem Res Toxicol 29 (5):901–913

Urbisch D, Honarvar N, Kolle SN, Mehling A, Ramirez T, Teubner W, Landsiedel R (2016b) Peptide reactivity associated with skin sensitization: the QSAR toolbox and TIMES compared to the DPRA. Toxicol In Vitro 34:194–203

Wareing B, Urbisch D, Kolle SN, Honarvar N, Sauer UG, Mehling A, Landsiedel R (2017) Prediction of skin sensitization potency sub-categories using peptide reactivity data. Toxicol In Vitro 45:134–145

Weil CS, Scala RA (1971) Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. Toxicol Appl Pharmacol 19(2):276–360

Zuang V, Desprez B, Barroso J, Belz S, Berggren E, Bernasconi C, Bessems J, Bopp S, Casati S, Coecke S, Corvi R, Dumont C, Gouliarmou V, Griesinger C, Halder M, Janusch-Roi A, Kienzler A, Landesmann B, Madia F, Mil-camps A, Munn S, Price A, Prieto P, Schäffer M, Triebe J, Wittwehr C, Worth A, Whelan M (2015) EURL ECVAM status report on the development, validation and regulatory acceptance of alternative methods and approaches JRC97811 — EUR 27474 EN, September 2015, 114pp. European Union, Brussels