



Comparison and validation of an *in vitro* skin sensitization strategy using a data set of 33 chemical references

Elodie Clouet^{a,b,*}, Saadia Kerdine-Römer^b, Pierre-Jacques Ferret^a

^a Pierre Fabre Dermo-Cosmetics Research & Development, Toxicology Division, Safety Department, Toulouse, France

^b UMR996 - Inflammation, Chemokines and Immunopathology, INSERM, Univ Paris-Sud, Université Paris-Saclay, 92296, Châtenay-Malabry, France

ARTICLE INFO

Keywords:

ITS
Skin sensitization
DPRA
h-CLAT
SENS-IS
AOP

ABSTRACT

Allergic contact dermatitis (ACD) is an adverse health effect that develops following repeated exposure to skin sensitizing chemicals. An animal testing ban has been applied in EU, leading to development of reliably predictive non-animal methods. Several *in vitro* methods have been developed as alternatives but one single non-animal test method is not been sufficient to fully address since the LLNA test ban.

Here, we have selected an ITS (Integrated Testing Strategy) for skin sensitization which focuses on three *in vitro* methods that covered the first three steps of the AOP (DPRA, SENS-IS or h-CLAT). The aim of this study was to compare these three methods due to the WoE approach based on a 2-out-of-3-assessment. The results of 33 references were compared to *in vivo* data (especially human). We have shown that tested firstly DPRA and SENS-IS have permitted to conclude on 29 of 33 chemicals, whereas DPRA and h-CLAT on 25, and SENS-IS and h-CLAT on 23. With this sequence, DPRA and SENS-IS and then h-CLAT in case of equivocal results, we conclude more quickly by performing fewer tests. Thereby, we have shown that it is better to follow a preferential sequence than testing chemicals simultaneously with these three methods.

1. Introduction

The principle of the “Three Rs” (Russell and Burch, 1959) has been present in spirit in European Union (EU) legislation from as early as 1986, when the first EU law for the protection of animals used for experimental and other scientific purposes was adopted. Russell and Burch proposed the principles of Replacement, Reduction and Refinement (the “Three Rs”) as key strategies of a systematic framework aimed to achieve the goal of humane experimental techniques. They saw replacement as the ultimate goal for laboratory animal based research, education and testing, with the other two, reduction and refinement, being more readily achievable in the short term. This concept, combined with societal change for 20 years, has led to a total animal testing ban in cosmetics which came into force with the

Cosmetic Regulation no. 1223/2009 on March 13th, 2016. Since 2004, EU law bans the sale of any cosmetic product in finished form tested on animals, as well as cosmetics containing ingredients subject to new animal testing since 2013.

Allergic Contact Dermatitis (ACD) represents a severe health problem with increasing worldwide prevalence, which is estimated at 15–20% in the world's population (Bruckner et al., 2000; Thyssen et al., 2007). ACD resulting from skin sensitization is provoked by repeated skin contact with low molecular weight chemicals (Kimber et al., 2011). Information on this toxicological endpoint is required in different legal frameworks, including the European Regulation of Registration, Evaluation, Authorization and Restriction of Chemical substances (REACH; EC 1907/2006) and under the United Nations Globally Harmonized System for Classification, Labeling and Packaging (GHS/CLP). One of

Abbreviations: Ac, Acylating agents; ACD, Allergic Contact Dermatitis; AOP, Adverse Outcome Pathway; ARE, Antioxidant Response Element; DC, dendritic cell; DMSO, dimethyl sulfoxide; DNCB, dinitrochlorobenzene; DPRA, Direct Peptide Reactivity Assay; ECVAM, European Centre for the Validation of Alternative Methods; FACS, Fluorescence-Activated Cell Sorting; FN, false negative; FP, false positive; GP, guinea pig; GPTs, guinea pig based tests; h-CLAT, human Cell Line Activation Test; HRIPT, Human Repeated Insult Patch Test; IATA, Integrated Approaches to Testing and Assessment; Ig, immunoglobulin; ITS, Integrated Testing Strategy; KE, Key Event; LLNA, local lymph node assay; MA, Michael acceptors; MFI, Mean Fluorescence Intensity; MIE, Molecular Initiating Event; MT1 and MT2, Metallothionein-1 and -2; MUSST, myeloid u937 skin sensitization test; OECD TG, Organisation Economic for Co-operation and Development Test Guidelines; PBS, phosphate-buffered saline; QRA, Quantitative Risk Assessment; QSAR, Quantitative structure-activity relationship; RFI, Relative Fluorescence Intensity; RhE, Reconstructed human Epidermis; ROS, Reactive Oxygen Species; RT-PCR, Reverse Transcription-Polymerase Chain Reaction; SB, Schiff base formers; SLS, sodium lauryl sulfate; SI, stimulation index; SN2, Substances reacting by nucleophilic substitutions of type 1 or 2; STS, Sequential Testing Strategy; TG, test guidelines; TMR, Test Methods Regulation; TN, true negative; TP, true positive; TNBS, 2,4,6-Trinitrobenzenesulfonic acid; WoE, Weight of Evidence

* Corresponding author at: Institut de Recherche Pierre Fabre, Recherche et Développement Pierre Fabre Dermo Cosmetique, 3 Avenue Hubert Curien, BP 13562, 31035 Toulouse Cedex, France.

E-mail address: elodie.clouet@pierre-fabre.com (E. Clouet).

<http://dx.doi.org/10.1016/j.tiv.2017.05.014>

Received 31 December 2016; Received in revised form 6 May 2017; Accepted 19 May 2017
0887-2333/ © 2017 Elsevier Ltd. All rights reserved.

the objectives of the REACH Regulation was to promote non-animal test methods. To better understand the impact of chemicals, reliable hazard and exposure information must be provided to demonstrate their safety. The hazard data has been traditionally generated through animal testing but now there are presently a number of other ways: by comparing substances with similar ones; by grouping them together into logical categories; by performing specialised computer modelling; by using Weight of Evidence; or by using a validated alternative test. The REACH Regulation is clear that every effort must be made and alternative methods used wherever possible and testing on animals is truly the last resort. To conform this, a revision of Annex VII of REACH was published at the end of 2016 (European Parliament, 2016). This text allows, if adequate, the use of *in vitro/in silico* methods in place of the local lymph node assay (LLNA), a sensitization test on mice for classification and labeling. Furthermore, under the EU animal testing Directive (2010/63/EU) companies must not test on animals when an alternative method is accepted under the EU legislation (European Union, 2010). In that vein, a revision of the ECHA R7a guidance on skin sensitisation was published recently at the end of 2016 (European Chemicals Agency, 2016).

To complete, information on skin sensitization is required in the safety assessment of EU Cosmetics Regulation (EC 1223/2009). Within these various legislative frameworks, information is necessary on the skin sensitization hazard potential. Furthermore, quantitative information that allows hazard assessment is needed to meet the regulatory requirements to sub-categorize skin sensitizers according to their potency class. Until recently, these legislative frameworks had been relying on information from animal models. Models accepted by the different regulations are the guinea pig (GP) models described in OECD TG406 (OECD, 1992) including the guinea pig maximization test (GPMT) and Buehler test, and the LLNA described in OECD TG429 (OECD, 2010a) and non-radioactive variants of the LLNA in OECD TG442A and 442B (OECD, 2010b, 2010c). From an animal welfare perspective, the LLNA has been a first refinement and reduction method compared to the classical guinea pig methods. Another advantage of the LLNA was generating dose-response data on lymphocyte proliferation. These quantitative data were used to derive the EC3 value, which was the Effective Concentration that induces a threefold increase in local lymph node proliferative activity. The EC3 correlates relatively well with known human thresholds for contact allergy (Basketter et al., 2005; Griem et al., 2003; Schneider and Akkan, 2004). This value is used for sub-categorizing of skin sensitizers, e.g. in the GHS 1A (EC3 ≤ 2%; strong skin sensitizers) and 1B (EC3 > 2%; other skin sensitizers) and allowing Quantitative Risk Assessment (QRA). Clearly, since the animal testing ban, novel predictive test methods or strategies for skin sensitization need to provide information that allow both hazard identification and potency sub-categorization of skin sensitizers.

ACD is composed of two phases: the first called sensitization phase and the second one the elicitation phase. The chemical and biological pathways involved are relatively well characterized and have been simplified and structured into an Adverse Outcome Pathway (AOP). The KEs involved have now been formally described by the OECD (OECD, 2012). AOP represents existing knowledge concerning the causal linkages between initial molecular events and an adverse outcome at the individual and population levels. These events include skin bioavailability and metabolism, hapten formation (i.e., the ability of a chemical to react with skin proteins), epidermal inflammation via keratinocyte and/or dendritic cell (DC) signalling, DC activation, maturation and migration, and T-lymphocyte (T-cell) proliferation; meaning that several KEs are required for acquisition of skin sensitization. Thus, many *in vitro* methods including haptenisation with *in chemico* assays (Direct Peptide Reactivity Assay (DPRA), Peroxidase peptide reactivity assay (PPRA)), DC line activation (cell based assays (U-SENS™, mMUSST, h-CLAT)) and keratinocyte activation (Keratin-Sens™ and SENS-IS) have been developed as alternative methods and are currently used in different stages of formal validation (Reisinger

et al., 2015). Among those methods developed or under development, three *in vitro* methods have been recently validated:

- 1) The test method OECD 442C adopted in 2015 - Direct Peptide Reactivity Assay (DPRA) is an *in chemico* test method which addresses peptide reactivity, postulated to be the Molecular Initiating Event (MIE) (corresponding to the first Key Event (KE1)) of the skin sensitization AOP (OECD, 2012). Reactivity is measured by quantifying depletion of synthetic peptides containing either cysteine or lysine.
- 2) The test method OECD 442D adopted in 2015- ARE-Nrf2 Luciferase Test Method (KeratinSens™) is an *in vitro* test method which addresses keratinocyte induction of a cyto-protective gene pathway linked to skin sensitization, i.e. KE2 of the AOP (OECD, 2012). The test method uses luminescence detection to measure gene expression of the Antioxidant/electrophile Response Element (ARE)-dependent pathway.
- 3) The test method OECD 442E adopted in 2016 - human Cell Line Activation Test (h-CLAT) is an *in vitro* method which addresses activation of the dendritic cells (DCs), i.e. KE3 of the AOP (OECD, 2012). The assay quantitatively measures changes in the expression of cell surface markers (CD86 and CD54), associated with the activation of DCs, in a human monocytic leukemia cell line by flow cytometry analysis.

These tests cover specific KEs within the skin sensitization AOP, previously outlined (OECD, 2012).

Besides these, many other assays are in development and at different stages of the evaluation process (EURL-ECVAM, 2016). Out of these, two *in vitro* keratinocyte assays are in development: the LuSens assay (Bauch et al., 2012; Ramirez et al., 2016; Ramirez et al., 2014; Urbisch et al., 2015), a test method similar to KeratinSens™ and the SENS-IS assay (Cottrez et al., 2016b; Cottrez et al., 2015), a test determining the sensitizing potential and potency of chemicals using a panel of genes. Moreover, data on over a hundred substances is available for the myeloid U937-based skin sensitization assays (U-SENS™, formerly known as MUSST, myeloid U937 skin sensitization test) (Ade et al., 2006; Alépée et al., 2015; Bauch et al., 2011; Piroird et al., 2015; Urbisch et al., 2015), which is similarly to the h-CLAT. The U-SENS™ test is currently being peer-reviewed by EURL-ECVAM and a proposal for the development of an OECD Test Guideline has been included in the OECD work program (EURL-ECVAM, 2016; Ezendam et al., 2016).

However, there is a current scientific consensus that one single non-animal test method will not be able to provide sufficient information to fully address this AOP and replace the use of animals, compared to LLNA (Adler et al., 2011). In many cases, the use of a battery of tests in combination and/or with other information in the context of integrated approaches such as Weight of Evidence (WoE) or testing and assessment strategies have to be considered (European Parliament, 2016; Patlewicz et al., 2014; Tollefsen et al., 2014). There are several testing strategies proposed that combine *in silico* (QSARs), *in chemico* (OECD TG 442C) and/or *in vitro* tests (e.g. OECD TG 442D and 442E) to distinguish skin sensitizers from non-sensitizers.

In addition, the new REACH consolidation as well as ECHA guidance R7a or OECD EU test guidelines updated agree that these methods can be used within a WoE assessment. Recently, the OECD has also developed two guidance documents on defined approaches to testing and assessment. The first is a general document that provides a set of principles and how to use them within an Integrated Approaches to Testing and Assessment (IATA) (OECD, 2016a); the second document specifically focuses on case studies (OECD, 2016b).

Therefore, all of these tests (except for the SENS-IS) are already being successfully used in combination in a number of Integrated Testing Strategies (ITS). Some of them, which include, but are not limited to, the “2 out of 3” approach (Bauch et al., 2012; Natsch et al.,

2013; van der Veen et al., 2014a), an ITS based on Bayesian Networks (Jaworska et al., 2013) or an ITS which includes an assessment of potency (Nukada et al., 2013; Tsujita-Inoue et al., 2014). Some approaches are constructed on a mechanistic base using parts of the AOP, e.g. the “2 out of 3” ITS (Urbisch et al., 2015) or the Sequential Testing Sequence (STS) of the RIVM (van der Veen et al., 2014a, 2014b). Others apply statistical approaches to conduct their strategy (e.g. the integrated decision strategy of Strickland et al. (2016)) or use a combination of both (e.g. the Bayesian ITS of Jaworska et al. (2015)). The skin sensitization ITS proposed by BASF (Bauch et al., 2012) and later updated by Natsch et al. (2013) and Urbisch et al. (2015), is based on assays reflecting the first 3 KEs: KE1 with DPRA, KE2 with LuSens assay or the KeratinoSens™ and KE3 with the mMUSST or the h-CLAT. Their initial approach was developed using assays that reflect the first three KEs of the AOP. The preliminary prediction model was based on a dataset of 54 substances. Application of this prediction model resulted in an accuracy of 83–94%, a sensitivity of 81–93% and a specificity of 88–95% compared to LLNA or human data, respectively (Bauch et al., 2012). This ITS further evolved with a larger dataset of 145 substances (102 skin sensitizers and 43 non-sensitizers). The data interpretation procedure used was simplified compared to that used in the initial study of Bauch et al. (2012). Thus, an ITS combining data from the regulatory accepted test methods (DPRA, KeratinoSens™ and h-CLAT) resulted in the highest accuracy of 90%; sensitivity and specificity both were 90% as well, compared to human data. This predictivity was more accurate compared to predictions based on the LLNA for the same dataset, for which the accuracy, sensitivity and specificity were 82%, 91% and 64%, respectively (Urbisch et al., 2015). Many AOP-based testing strategies are focused mainly on MIE (KE1), KE2 and KE3, meaning that the AOP is not yet totally covered. Informations on these last events seem to be sufficient to predict skin sensitization with a high certainty. Furthermore, the “2 out of 3” approach shows that it is possible to identify skin sensitizers with a high accuracy based on only two KEs (Ezendam et al., 2016). But, the majority of predictive *in vitro* or *in chemico* assays cannot estimate potency on their own. A possible explanation may be that the dynamic range of *in vitro* and *in chemico* assays is too limited to detect the broad range of skin-sensitizing potencies (e.g. lack of bioactivation of prohaptens, water solubility of the test substances, ...).

In this study, we propose an in-house IATA selecting three *in vitro* tests that cover main KEs defined by the AOP (Fig. 1): DPRA (MIE), SENS-IS (KE2) and h-CLAT (KE3). According to the previous explanation, the originality of our ITS has been to integrate the SENS-IS which measures the gene expression of irritation and sensitization biomarkers at the KE2 level. The SENS-IS assay was submitted to ECVAM in September 2011 and was enrolled in the OECD guideline work in 2015. Despite the fact that this assay is not yet approved, it can provide an accurate prediction of potency categories (Cottrez et al., 2015). Moreover, the use of 3D model allows us to avoid any solubility limitation of the test item. In addition, several studies have shown that regulation of gene expression correlates with potency (McKim et al., 2012; van der Veen et al., 2013). Thus, the different read-out(s) issued from the SENS-IS assay seems to be more capable of distinguishing sensitizer, irritant or antioxidant and assigning potencies. Thereby, SENS-IS provide a possible way to encompass metabolism as well as bioavailability limitations of ingredients and come closer to potency assessment (Cottrez et al., 2015). Therefore, the aim of this study was to compare the available data, previously published (Cottrez et al., 2015; Urbisch et al., 2015) with the above mentioned *in vitro* methods (DPRA, SENS-IS and h-CLAT). All results obtained within this ITS was compared to LLNA and human data from literature. The predictivities of the individual methods as well as their combinations (2 or 3 methods) are evaluated. A pragmatic “WoE” approach is used: Any two of the three tests decide the overall result (any two assays must be positive to rate the substance as a skin sensitizer and any two assays must be negative to rate the substance as a nonsensitizer). The final aim of our

ITS is to give reliable information about the potential sensitization risk induced by a chemical.

2. Material and methods

2.1. Strategy for selecting tests

Each AOP starts with a Molecular Initiating Event (MIE), consisting in the first step that initiates the perturbation caused by a chemical, followed by several KEs at the cellular, organ or organism level, eventually resulting in the adverse outcome. For skin sensitization, the MIE is the covalent binding of a chemical or its metabolites to skin proteins, resulting in the formation of an immunogenic conjugate. This triggers two cellular events, KE2 and KE3. The electrophilic properties of skin sensitizers cause an activation of keratinocytes (KE2), resulting in the release of pro-inflammatory cytokines and chemokines, and activation of antioxidant responses. These mediators are essential components for the generation of an adaptive immune response (Martin et al., 2011; van der Veen et al., 2011). The immunogenic conjugate is taken up by dermal DCs and/or Langerhans cells, resulting in activation and migration of these cells to the draining lymph nodes (KE3). The DC present epitopes of the conjugate to T cells, resulting in activation and proliferation of T cells (KE4), occurring in sensitization phase of the immune system. The AOP is a simplification of a complex biological process, which offers the opportunity to anchor existing test methods or *in silico* tools that measure the MIE or other KEs. This AOP can be used to mechanistically underpin the development of strategies for the testing and assessment of chemicals for a specific regulatory need (MacKay et al., 2013; Patlewicz et al., 2014). Considering this AOP, ITS used within the IATA consists of methods with a mechanistic relevance for the endpoint that must be assessed. Placed in the context, we have selected validated or under-validation *in vitro* methods that covered three main KEs.

The DPRA evaluates KE1 and the protein/peptide reactivity of a substance (covalent interaction with cellular proteins). The SENS-IS represents KE2 and addresses keratinocyte activation (through Pathways-associated gene expression). And the h-CLAT describes KE3 and DCs activation (expression of co-stimulatory and adhesion cell surface markers). When used together, they cover the first three of the four KEs of the sensitization process, thus being of mechanistic relevance and supporting the scientific rationale for using a combination of these methods in an AOP-based ITS.

2.2. Selection of chemicals

The main selection criterion for the substances was the availability of both human and LLNA reference data. Only molecules having results in each of 3 *in vitro* methods were retained: DPRA, h-CLAT and SENS-IS as well as LLNA and human data, taken from two publications (Cottrez et al., 2015; Urbisch et al., 2015) indicated in Table 1. These chemicals are used in cosmetics as well as fragrance or pharmaceutical use. Among the 33 selected, 23 (70%) are classified as sensitizers according to LLNA and 19 (58%) according to human data. In Table 1, all results are combined per test to each molecule, and compared to both human and LLNA data.

For each chemical and per test, we attributed a simple number: “0” for negative test results, “1” indicates positive test results for all different tests and “0/1” for equivocal results. However, for equivocal results, and during combination process, the worst-case will be retained (e.g. “1” for salicylic acid which is quoted in DPRA (“0/1”).

2.3. Non-animal tests methods selected: *in vitro* skin sensitization studies according to AOP Skin Sensitization

2.3.1. Direct peptide reactivity assay (DPRA) for the Key Event 1

The Molecular Initiating Event leading to skin sensitization is

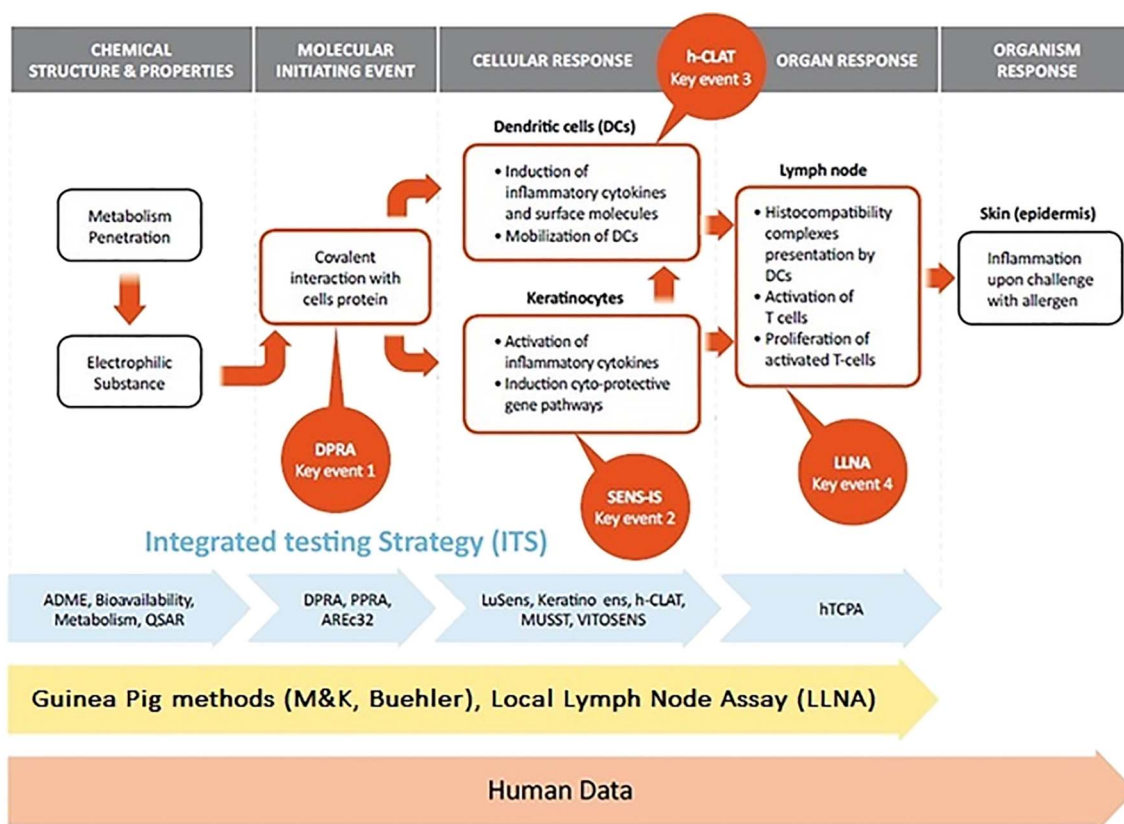


Fig. 1. Schematic representation of Adverse Outcome Pathway for Skin Sensitization and necessity to follow an Integrated Testing Strategy (ITS) since the LLNA ban to cover the different keys events.

This figure is inspired by the AOP for skin sensitization (OECD, 2012). It shows the *in vitro* methods selected in our prediction model to cover the different key events of the AOP (Adverse Outcome Pathway). All results obtained were compared to LLNA and human data.

Table 1

Chemical references selected including available data for *in vitro* methods as well as *in vivo* tests (LLNA and human data). The references have been classified according to their potency in LLNA.

Chemical name	CAS number	LLNA EC3 (%)	LLNA category / Potency	Proposed protein-binding mechanism	Human Data	LLNA	DPRA	SENS-IS	h-CLAT
Oxazolone	15646-46-5	0,003	Extreme	MA / Ac	1	1	1	1	1
Diphenylcyclopropanone	886-38-4	0,003	Extreme	Ac	1	1	1	1	1
p-Benzoquinone	106-51-4	0,01	Extreme	MA	1	1	1	1	1
1,4-Phenylenediamine	106-50-3	0,16	Strong	quinone precursor	1	1	1	1	1
Phthalic anhydride	85-44-9	0,16	Strong	Ac	0	1	1	1	0
Propyl gallate	121-79-9	0,32	Strong	quinone precursor	1	1	1	1	1
2-Amino-phenol	95-55-6	0,4	Strong	quinone precursor	1	1	1	1	1
Formaldehyde	50-00-0	0,7	Strong	SB	1	1	1	1	1
Methyldibromo glutaronitrile	35691-65-7	0,9	Strong	S _N 2	1	1	1	1	1
Glyoxal	107-22-2	1,4	Moderate	SB	1	1	1	1	1
2-Mercaptobenzothiazole	149-30-4	1,7	Moderate	S _N 2	1	1	1	1	1
Isoeugenol	97-54-1	1,8	Moderate	quinone precursor	1	1	1	1	0
1,2-Benzisothiazolin-3-one	2634-33-5	2,4	Moderate	S _N 2	1	1	1	1	1
Cinnamic aldehyde	104-55-2	3,1	Moderate	MA	1	1	1	1	1
Citral	5392-40-5	5	Moderate	MA	1	1	1	1	1
Tetramethylthiuram disulfide	137-26-8	5,2	Moderate	S _N 2	1	1	1	1	1
Diethyl maleate	141-05-9	5,8	Moderate	MA	1	1	1	1	1
Alpha hexyl cinnamic aldehyde	101-86-0	11	Weak	MA	0	1	0	1	0
Eugenol	97-53-0	13	Weak	quinone precursor	1	1	1	1	1
Benzyl benzoate	120-51-4	17	Weak	S _N 2	0	1	0	1	0
Hydroxycitronellal	107-75-5	23	Weak	SB	1	1	1	1	1
Butyl glycidil ether	2426-08-6	31	Weak	S _N 2	1	1	1	1	0
Salicylic acid	69-72-7	>50	None	no alert	0	0	0/1	0	1
Benzalkonium chloride	8001-54-5	False positive	None	no alert	0	0	0	0	0
Propyl paraben	94-13-3	>50	None	no alert	0	0	0	0	1
Sodium laurylsulfate	151-21-3	False positive	None	no alert	0	0	0/1	0	0
Octanoic acid/caprylic acid	124-07-2	>50	None	no alert	0	0	0	0	1
1-Butanol	71-36-3	>50	None	no alert	0	0	0	0	0
Hexyl salicylate	6259-76-3	>50	None	no alert	0	1	0	0	1
Glycerol	56-81-5	>50	None	no alert	0	0	0	0	0
Lactic acid	50-21-5	>50	None	no alert	0	0	0	0	0
Propylene glycol	57-55-6	>50	None	no alert	0	0	0	0	0
Diethyl phthalate	84-66-2	>50	None	no alert	0	0	0	0	1
Sensitizer (S)					19	23	22	22	22
Not sensitizer (NS)					14	10	11	11	11

* For Salicylic acid and Sodium Lauryl Sulfate, the results were equivocal, so we considered them ultimately as positives.

postulated in the AOP to be covalent binding of electrophilic chemical species with selected nucleophilic molecular sites of action in skin proteins (Gerberick et al., 2008; Karlberg et al., 2008).

The correlation of skin reactivity and skin sensitization is well established (Gerberick et al., 2008; Jordan, 1983; Landsteiner and Jacobs, 1936; Rastogi et al., 1998). Gerberick et al. (Gerberick et al., 2009; Gerberick et al., 2004) have developed a Direct Peptide Reactivity assay (DPRA) for screening the skin sensitization potential of chemicals by measuring peptide depletion following incubation with allergens and non-allergens.

The DPRA was performed according to the OECD TG442C (OECD, 2015a) and describing into the publication of Urbisch et al. (2015). Peptide reactivity data were generated using a method to measure reactivity of a test chemical with model heptapeptides containing lysine (Ac-RFAAKAA-COOH) or cysteine (Ac-RFAACAA-COOH) (Gerberick et al., 2004). Peptide reactivity was reported as percent depletion based on the decrease in non-reacted peptide concentration in the sample relative to the average concentration measured in the control. Peptides were prepared and purified by the SynPep Corporation (Dublin, CA, USA) to > 90% purity as measured by HPLC, and molecular weight confirmation was determined by flow injection positive-ion electrospray mass spectrometry. Briefly, 400 μ L of a 1.25 mM peptide stock solution prepared in buffer and a 100 mM test chemical stock solution prepared in either acetonitrile or DMSO/acetonitrile were added to 100 mM ammonium acetate buffer (pH 10.2) for the lysine peptide or 100 mM sodium phosphate buffer (pH 7.5) for the cysteine peptide. The final reaction, containing 0.5 mM of the peptide and 5 or 25 mM of the test chemical, representing 1:10 and 1:50 M ratios, was mixed and incubated in the dark for 24 h at 25 °C. Control samples and standards used for defining the calibration curve for each analysis were prepared without test chemical for each peptide and ranged from 0.0156 to 1.0 mM. All samples were prepared in triplicate. Following incubation, the peptide was quantified by reverse-phase HPLC (Waters 2695 Alliance) on a Zorbax SB-C18 column (2.1 $\times$ 100 mm, 3.5 μ m) with UV detection at 220 nm (Waters 996 PDA detector) using an external standard linear calibration curve. The UV spectrum was collected from 210 to 400 nm to allow verification of the peptide peak identity. For salicylic acid and sodium lauryl sulfate, the results are equivocal, so we considered these results as positive.

2.3.2. Explorative sensitization prediction test (SENS-IS test) for the Key Event 2

Haptens can also react with cell surface proteins and activate response pathways in keratinocytes (Weltzien et al., 2009). Uptake of the hapten by keratinocytes activates multiple events, including the release of pro-inflammatory cytokines and the induction of cytoprotective cellular pathways. Keratinocyte exposure to sensitizers also results in induction of ARE/RhE-dependent pathways (Natsch and Emter, 2008). Expression of ARE/RhE-dependent genes and other cytoprotective genes (including CYP1A1, MT1 and MT2) in HaCaT cells are part of a proprietary *in vitro* battery approach to determine sensitization potency (McKim et al., 2010).

To address that complexity and this second key event described in the AOP, the SENS-IS model was developed by Immunosearch Company to assess the potential of the test products to specifically induce the expression of irritation and/or sensitization biomarkers in an *in vitro* 3D skin model. Immunosearch developed this assay after analyzing genes modulated during the sensitization process either on mice (LLNA) or human (blisters). The panel of sensitization biomarkers defined includes genes already identified as markers, such as the ARE family, and others not yet associated with the sensitization process (called SENS-IS gene subset).

The SENS-IS is based on the overexpression analysis of three sets of genes (62 genes) defined by Cottrez et al. (2015), one specifically reflecting the irritant potential (IRR genes) and the other two set of genes (ARE and SENS-IS genes) correlating with the sensitization

potential. ARE genes control the expression of proteins critical in the detoxification and elimination of Reactive Oxygen Species (ROS) and electrophiles (Nguyen et al., 2009). The SENS-IS assay was submitted to ECVAM in September 2011 and was enrolled in the OECD guideline work in 2015. SENS-IS expose 3D epidermal skin tissues addressing substance activation by keratinocytes as well as the cytotoxicity of a substance. This assay was primarily developed for simple ingredients and has been tested and validated with around 180 chemicals (100 from phase II, + 30 RIFM + 50 Phase III). Furthermore, its prediction model gives a quantitative response (potency) for ingredient that correlate to the LLNA. EpiSkin® tissues (L'Oréal, France) are exposed with the test substances using the protocol to determine *in vitro* skin irritation described in OECD TG439 (OECD, 2015b). Chemical references were used diluted in PBS, Olive oil or DMSO, at the concentrations of 50% to 0.1% depends on the LLNA category. 30 μ L (26.3 μ L/cm²) of each dilution was applied on the top of each reconstituted epidermis. The test product was gently spread on the epidermis surfaces to ensure it covers the entire surface. After 15 min exposure, the EpiSkin® were rinsed with PBS. Epidermis were then incubated at 37 °C for 6 h and then samples were harvested and frozen in liquid nitrogen before tissue lysing and RNA extraction. The complete epidermis was collected, placed in a RNazol® solution and the total RNA was isolated. After reverse transcription, quantitative gene expression was measured by RT-PCR with a SYBR green buffer using the LC480 Roche's apparatus and specific biomarkers primers defined for the SENS-IS test. Three negative controls (PBS, Olive oil and DMSO treated skins), a positive irritation control (5% SLS) and two positive sensitization controls (Hexylcinnamaldehyde (HCA) at 50% in Olive oil and TNBS at 1% in PBS) were performed.

A test substance is considered to be sensitizer if it increases the expression (compared to the solvent control) of at least 7 genes measured by qPCR in either the "SENS-IS" (7/21) or the "ARE" (7/17) gene sets. To take into account non-specific gene over-expression due to cell stress, the induction of > 15 genes (on 23) in the irritation gene set, classifies a result as inconclusive and the test substance is re-analyzed at a lower concentration. Similarly to the LLNA, potency is classified according to the concentration of test material needed to induce a positive response: positive at 0.1%: extreme; positive at 1%: strong; positive at 10%: moderate; positive at 50%: weak. SENS-IS is considered to mainly address KE2 from the AOP of skin sensitization, but may, as ARE-activated genes are included, also provide information on protein reactivity of a test chemical. Furthermore, up to today, product safety evaluation in the EU is predominantly based on data/information on their individual ingredients. However, for skin sensitization, mixture component effects have been demonstrated in numerous studies either by the presence of irritating chemicals or by the modification of dermal absorption. SENS-IS assay has evaluate the ability of mixture component effects on the potency of skin sensitization of chemicals and seems to be efficient to measure the component effects and assess mixtures or final products. Influence of formulations (oil-in-water, emulsion, cleansing water, oil, water-in-oil emulsion) and state of the skin (irritated or not) on the sensitizing potential of some components has been measured (Cottrez et al., 2016a). SENS-IS assay could be used in a near future to test complex mixtures and final products.

2.3.3. Human Cell Line Activation Test (h-CLAT) for the Key Event 3

As noted in the AOP during allergen contact with the skin, immature epidermal DCs (known as Langerhans cells and dermal DCs) serve as antigen-presenting cells (Kimber et al., 2011; Ryan et al., 2007). In this role, they recognize and internalize the hapten-protein complex formed during covalent binding. Subsequently, the DC loses its ability to seize new hapten-protein complex and gains the potential to display the allergen-MHC-complex to naive T-cells; this process is often referred to as DC maturation. The h-CLAT test is an *in vitro* method measuring by flow cytometry the expression of CD86 and CD54 protein markers on

the cell surface of the human monocytic leukemia cell line THP-1 (American Type Culture Collection, Manassas, VA, USA) used as a surrogate for human DC (Ashikaga et al., 2010; Ashikaga et al., 2006; Sakaguchi et al., 2010; Sakaguchi et al., 2009; Sakaguchi et al., 2006). The h-CLAT has been validated by ECVAM and OECD (Aeby et al., 2010; European Commission, 2015; OECD, 2016c).

For dose finding, cytotoxicity tests were conducted and the concentration resulting in 75% cell viability, termed CV75, was calculated based on the analysis of viable cells. THP-1 cells were treated with eight different concentrations, decided based on dose finding cytotoxicity test, for 24 h. After removing the test substance, the expression of CD86 and CD54 on the cell surface was measured by flow cytometry. The Relative Fluorescence Intensity (RFI) was used as an indicator of CD86 and CD54 expression. If the RFI of CD86 or CD54 was > 150% or 200% at any dose in at least two out of three experiments, the substance was judged as a sensitizer. Otherwise, it was considered a non-sensitizer (Ashikaga et al., 2006).

The Relative Fluorescence Intensity (RFI) is used as an indicator of CD86 and CD54 expression. RFI is calculated using the following formula:

$$\text{RFI} = \text{MFI of test item-treated cells} - \text{MFI of test item} \\ - \text{treated isotype control cells}$$

$$\text{MFI of vehicle control cells} - \text{MFI of vehicle isotype control cells}$$

2.4. In vivo study and human data

The predictivities of the non-animal test methods listed previously were compared to LLNA and human data.

2.4.1. Guinea pig assays and local lymph node assay (LLNA)

In the regulatory context, the skin sensitization potential of a chemical has traditionally been evaluated using animal tests, in particular the guinea pig based tests (GPTs) including guinea pig maximization test (GPMT) and Buehler test described in OECD TG406 (OECD, 1992) or the mouse-based local lymph node assay (LLNA) described in OECD TG429 (OECD, 2010a) (or OECD TG442 A + B (OECD, 2010b, 2010c)). In the GPMT, guinea pigs are exposed to the test substance by intradermal injection and topical application by occlusion. Following a rest period of ten to fourteen days, the challenge dose is introduced dermally. The extent and degree of skin reactions to this challenge exposure are then compared with control animals. In the Buehler test, guinea pigs are repeatedly exposed to the test substance by topical application under occlusion. Following a rest period of twelve days, a dermal challenge treatment is performed under occlusive conditions. The LLNA is the gold standard regulatory toxicology test for skin sensitization along with GPMT. Compared with the GPMT, LLNA uses fewer animals, it is quantitative, and it gives a numerical prediction of potency. In the LLNA, the test substance is applied to the ears of mice for three days and later tritiated radioactive thymidine is introduced to measure cell proliferation in auricular lymph nodes. An increase in lymph node cell proliferation compared to control animals indicates sensitization.

In the main study, a minimum of four animals per dose group is recommended. Test substance treatment dose levels were based on the recommendations given by Kimber and Basketter (1992) and in the ICCVAM Panel Report (ICCVAM, 1999). Dose levels selected from the concentration series 100%, 50%, 25%, 10%, 5%, 2.5%, 1%, 0.5%, etc. The maximum concentration tested should be the highest achievable level while avoiding excessive local irritation and overt systemic toxicity.

The LLNA experimental procedure is performed as follows:

In the main study, groups of 4 or 5 mice were treated by topical applications of 25 μL test or control solutions to the entire dorsal surface of each ear. The topical applications were performed once daily

over three consecutive days. In each study, a positive control was included. The negative control group was treated with the vehicle alone. On day 6, record the weight of each mouse and inject 250 μL of sterile phosphate-buffered saline (PBS) containing 20 μCi of (^3H)-methyl thymidine or 250 μL PBS containing 2 μCi of (^{125}I)-iododeoxyuridine (^{125}IU) and 10^{-5} M fluorodeoxyuridine into each mouse via the tail vein (Kimber, 1995; Loveless et al., 1996). Five hours later, each mouse is euthanized and the draining (“auricular”) lymph nodes of both ears are collected and placed in PBS (one container per mouse). Both bilateral draining lymph nodes must be collected in PBS and individually pooled for each animal. Following further processing steps, the incorporation of ^3H -methyl thymidine into the lymph nodes was measured in a scintillation counter and expressed as the number of disintegrations per minute (DPM). The proliferative response of lymph node cells was expressed as the number of radioactive disintegrations per minute per lymph node (i.e., DPM/lymph node). For each dose group, the stimulation index (SI) was calculated as the ratio of the arithmetic mean DPM/lymph node values (corrected for background) of the test and negative control groups. The EC3 value (theoretical concentration resulting in a SI value of 3) was determined by linear interpolation of points on the dose-response curve, immediately above and below the 3-fold threshold.

2.4.2. Human data/clinical sensitization exposure

Human data were available for all 33 substances selected; we retained only the column that shown the final classification of human data mainly issued of Urbisch et al. (2015) board (“1” indicates positive test results and “0” for negative test results).

2.5. Cooper statistics to compare in vitro assays as stand-alone or in combination

The Cooper statistics of predictivity were calculated for the single assays and their combinations as well as for the “2 out of 3” prediction model (Cooper et al., 1979). All parameters are based on a 2×2 contingency table counting the number of compounds that are “true positive” (TP), “false positive” (FP), “true negative” (TN) and “false negative” (FN):

$$\text{Sensitivity}[\%] = (\text{TP}/(\text{TP} + \text{FN})) \times 100 \text{ and } \text{Specificity}[\%] = (\text{TN}/(\text{TN} + \text{FP})) \times 100$$

$$\text{Positive predictive value}[\%] = (\text{TP}/(\text{TP} + \text{FP})) \times 100$$

$$\text{Negative predictive value}[\%] = (\text{TN}/(\text{TN} + \text{FN})) \times 100$$

$$\text{Accuracy}[\%] = ((\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{TN} + \text{FN})) \times 100$$

3. Results

3.1. Comparison of single assays compared to LLNA and human data

After attributing a positive (“1”) or negative (“0”) results to each chemical reference (Table 1) in each alternative method as well as LLNA and human data as described in Mat & Med, we compared results of single assays in terms of sensitivity, specificity and accuracy using Cooper statistics.

To determine the performance of the LLNA to predict skin sensitization hazard of this data set, LLNA data were directly compared to the available human data for all substances. A total of 4 (12%) of the substances (phthalic anhydride, alpha hexyl cinnamic aldehyde, benzyl benzoate and hexyl salicylate) were incorrectly identified as FP by the LLNA. In total, comparison to human data revealed a sensitivity of 100%, a specificity of 71% and an overall accuracy of 88% of the LLNA (Table 3).

For salicylic acid and sodium lauryl sulfate, the results were

Table 2
Number attribution system for combination of different *in vitro* methods and for our prediction model (DPRA, h-CLAT and SENS-IS).

Chemical name	LLNA EC3 (%)	LLNA category / potency	Use	Proposed protein-binding mechanism	Human Data	LLNA	DPRA	SENS-IS	h-CLAT	DPRA and h-CLAT	DPRA and SENS-IS	h-CLAT and SENS-IS	DPRA, h-CLAT and SENS-IS
Oxazolone	0.003	Extreme	Other	MA/Ac	1	1	1	1	1	1	1	1	1
Diphenylcyclopropenone	0.003	Extreme	pharmaceutical use	Ac	1	1	1	1	1	1	1	1	1
p-Benzoquinone	0.01	Extreme	pharmaceutical use	MA	1	1	1	1	1	1	1	1	1
1,4-Phenylenediamine	0.16	Strong	dye (production)	quinone precursor	1	1	1	1	1	1	1	1	1
Phthalic anhydride	0.16	Strong	dye (production)	Ac	0	1	1	1	0	0/1	1	0/1	1
Propyl gallate	0.32	Strong	preservative/disinfectant	quinone precursor	1	1	1	1	1	1	1	1	1
2-Amino-phenol	0.4	Strong	cosmetics (other use)	quinone precursor	1	1	1	1	1	1	1	1	1
Formaldehyde	0.7	Strong	preservative/disinfectant	SB	1	1	1	1	1	1	1	1	1
Methyldibromo glutaronitrile	0.9	Strong	preservative/disinfectant	S _N 2	1	1	1	1	1	1	1	1	1
Glyoxal	1.4	Moderate	Other	SB	1	1	1	1	1	1	1	1	1
2-Mercaptothiazole	1.7	Moderate	pesticide	S _N 2	1	1	1	1	1	1	1	1	1
Isoeugenol	1.8	Moderate	fragrance	quinone precursor	1	1	1	1	0	0/1	1	0/1	1
1,2-Benzisothiazolin-3-one	2.4	Moderate	preservative/disinfectant	S _N 2	1	1	1	1	1	1	1	1	1
Cinnamic aldehyde	3.1	Moderate	fragrance	MA	1	1	1	1	1	1	1	1	1
Citral	5	Moderate	fragrance	MA	1	1	1	1	1	1	1	1	1
Tetramethylthiuram disulfide	5.2	Moderate	pesticide	S _N 2	1	1	1	1	1	1	1	1	1
Diethyl maleate	5.8	Moderate	Other	MA	1	1	1	1	1	1	1	1	1
Alpha hexyl cinnamic aldehyde	11	Weak	fragrance	MA	0	1	0	1	0	0	0/1	0/1	0
Eugenol	13	Weak	fragrance	quinone precursor	1	1	1	1	1	1	1	1	1
Benzyl benzoate	17	Weak	preservative/disinfectant	S _N 2	0	1	0	1	0	0	0/1	0/1	0
Hydroxycitronellal	23	Weak	fragrance	SB	1	1	1	1	1	1	1	1	1
Buyl glycidyl ether	31	Weak	Other	S _N 2	1	1	1	1	0	0/1	1	0/1	1
Salicylic acid	>50	None	preservative/disinfectant	no alert	0	0	0	0/1	0	1	0	0	0
Benzalkonium chloride	False positive	None	preservative/disinfectant	no alert	0	0	0	0	0	0	0	0	0
Propyl paraben	>50	None	preservative/disinfectant	no alert	0	0	0	0	1	0/1	0	0/1	0
Sodium lauryl sulfate	False positive	None	surfactant	no alert	0	0	0	0/1	0	0/1	0/1	0	0
Octanoic acid/caprylic acid	>50	None	dye (production)	no alert	0	0	0	0	1	0/1	0	0/1	0
1-Butanol	>50	None	solvent	no alert	0	0	0	0	0	0	0	0	0
Hexyl salicylate	>50	None	fragrance	no alert	0	1	0	0	1	0/1	0	0/1	0
Glycerol	>50	None	cosmetics (other use)	no alert	0	0	0	0	0	0	0	0	0
Lactic acid	>50	None	preservative/disinfectant	no alert	0	0	0	0	0	0	0	0	0
Propylene glycol	>50	None	solvent	no alert	0	0	0	0	0	0	0	0	0
Diethyl phthalate	>50	None	Other	no alert	0	0	0	0	1	0/1	0	0/1	0
Sensitizer (S)					19	23	22	11	22	26	24	27	21
Not sensitizer (NS)					14	10	11	11	11	7	9	6	12

TP	True positive
FP	False positive
FN	False negative
TN	True negative

equivocal in DPRA, so we considered them finally as positive to get into the worst-case approach.

When calculating the Cooper statistics, the assessments were reduced to Yes/No (“1”/“0”) ratings according to the prediction models of the individual assays. Using the data, a direct comparison between the predictivities of the non-animal test methods and human is possible (Table 3). Accuracies of the single assays compared to LLNA data range from 79 to 91% for these 33 chemicals. Interestingly, when comparing the positive predictive values of almost all single non-animal test methods with human data, the percentages were generally higher than comparison with LLNA data (except for h-CLAT in that case).

3.2. Comparison of combination assays compared to LLNA and human data

We have performed an attribution number system combining two *in vitro* methods among those selected, as shown in Table 2. To attribute the final number, we have classified the chemicals as a sensitizer or “1” when for the both tests the score is 1, non-sensitizer or “0” when a score is 0 for both tests. For ambiguous result, positive in one test and negative in another one, we have classified as an equivocal result or “0/1”. To realize Cooper statistics, we considered them finally as positive to get into the worst-case. Using all the data from Table 2, a direct comparison between the predictivities of the non-animal test methods and human and LLNA data has been possible thanks to Cooper statistics (Table 3). Thus, in combining two alternatives methods, we didn't obtain better result than with LLNA. Nevertheless, the best two *in vitro* methods combination seems to be DPRA and SENS-IS compared to the two others, DPRA and h-CLAT, and h-CLAT and SENS-IS.

3.3. “2 out of 3” prediction model and order to carried out the testing sequence.

We realized the same process to combine our three selected *in vitro* methods (Table 2). The least complicated way to assess the skin sensitization hazard potential of a substance is to use the results of single assays which reflect key steps of the AOP within a “2 out of 3” prediction model. For the assays addressing the three KE described in the OECD AOP on skin sensitization mentioned above (OECD, 2012), a “2 out of 3” assessment was introduced for the first time by Bauch et al. (2012). In our study, this prediction model was applied using DPRA, SENS-IS and h-CLAT data. We realized the previous attribution number system combining these three *in vitro* methods, shown in Table 2. Any two congruent results of the three tests rule the overall assessment: If at least two of the three assays were positive, the substance was rated to be a skin sensitizer (“1”). If at least two of the three assays were negative, the substance was rated to be a non-sensitizer (“0”). The classification as a sensitizer or non-sensitizer is therefore based on a WoE pertaining to key events of the AOP. Cooper statistics for our final prediction model were determined in comparison to human data (and benchmarked to the previous gold standard test, the LLNA). Results for 33 substances, respectively, were obtained using this prediction model in Table 3.

Our prediction model, DPRA, SENS-IS and h-CLAT, provides some-

what higher accuracies compared to the predictivities of the single assays, whereby the accuracies compared to human or LLNA data were 94% or 88%, respectively. These results indicate that the investigated non-animal methods generally predict human data more accurately than LLNA data. To determine the performance of the LLNA to predict skin sensitization hazard of this data set, LLNA data were directly compared to the available human data. For comparison, our prediction model revealed a sensitivity of 100%, a specificity of 86% and an overall accuracy of 94% for the same set of substances suggesting predictions to be more relevant for humans. The analysis shows that combination of two *in vitro* methods is not enough to obtain better predictivity than LLNA (Fig. 2). Finally, the results indicate that the investigated non-animal methods generally predict human data more accurately than LLNA data.

After determine our prediction model in a “2 out of 3” strategy in order to predict better human sensitization than LLNA, we wondered whether these three tests should be performed at the same time or sequentially. According to the Fig. 2, three different “2 out of 3” strategy are conceivable. With the first strategy, DPRA and SENS-IS combined, we are able to classify 29 out of 33 chemicals and for 4 others, the h-CLAT test is added to draw a conclusion from the WoE. For the second strategy, combining DPRA and h-CLAT, we can classify 25 out of 33 chemicals and for 8 others, the SENS-IS test is added to draw conclusion from the WoE. Within the third strategy, SENS-IS and h-CLAT combined, we can classify for fewer chemicals, 23 out of 33, and for 10 others, the DPRA is added to draw a conclusion from the WoE. Finally, the first strategy permits us to classify a larger number of molecules, associating fewer tests than the two other strategies.

It is interesting to note that the number of tests performed is increased by +5% and +8% respectively with strategies 2 and 3 compared to the first. So, the strategy 1 appears to be the best one.

4. Discussion

4.1. Discordant results between LLNA and human data

The arbitrary assignment of “1” to equivocal results in LLNA and human data could be questionable because that triggers false results and particularly a final positive result. Regarding the difference between human data and LLNA results obtained, we can discuss as follows for the 4 molecules, phthalic anhydride, α -hexyl cinnamic aldehyde, benzyl benzoate and hexyl salicylate.

For example in Table 1, phthalic anhydride is positive in DPRA and SENS-IS (classified as a strong sensitizer). Moreover, phthalic anhydride is classified in the Annex VI of CLP as a sensitizer (H317) because positive in LLNA but it appears negative in humans (Urbisch et al., 2015). According to Urbisch et al. (2015), phthalic anhydride is known to be a respiratory sensitizer, but not a skin sensitizer (Dearman et al., 2013). Phthalic anhydride led to complete peptide depletion and adduct formation at Lys-residues of the peptide, due to its amine reactive function (Natsch et al., 2015). Phthalic anhydride appears negative in human population, but it is considered a sensitizer on specific workers exposed daily to this molecule. Their sensitivity was confirmed by

Table 3
Cooper statistics results for single assays as well as their combination compared to human data.

«Compared to human»			Positive predictive value	Negative predictive value	Sensitivity	Specificity	Accuracy
<i>In vitro</i> assays	Individual assays	DPRA	86	100	100	79	91
		h-CLAT	77	82	89	64	79
		SENS-IS	86	100	100	79	91
	Combinations	DPRA and h-CLAT	73	100	100	50	79
		DPRA and SENS-IS	79	100	100	64	85
		hCLAT and SENS-IS	70	100	100	43	76
	Prediction model	DPRA, h-CLAT and SENS-IS	90	100	100	86	94
	<i>In vivo</i> standard	LLNA	83	100	100	71	88

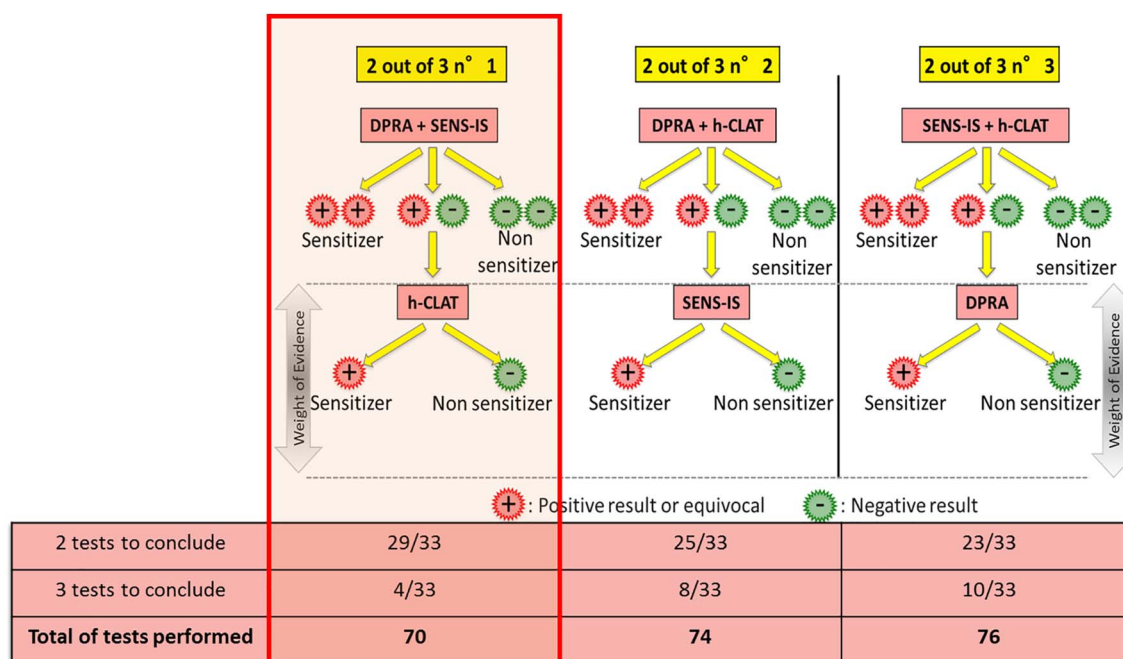


Fig. 2. Different «2 out of 3» strategies and preferential sequence selected.

direct bronchial tests, positive prick tests and positive RAST (Newcombe, 1992).

For α -hexyl cinnamic aldehyde, Urbisch et al. (2015) discuss the final inconclusive human data but we gave a final negative result in humans according to the RIFM database and Basketter et al. (2014). For benzyl benzoate, it is probably not or only a weak sensitizer in humans (Basketter et al., 2014) (class 5 of potency) with no evidence of induction in HRIPT at 59000 $\mu\text{g}/\text{cm}^2$.

Finally, concerning the hexyl salicylate, we observed no induction in HRIPT at 35433 $\mu\text{g}/\text{cm}^2$ and in HMT (Human Maximization Test) at 20654 $\mu\text{g}/\text{cm}^2$. Therefore, the very low EC3 (0.18%) may be due to its irritating properties (possibly sensitizing impurities).

4.2. Misclassified substances and applicability domain in single assays

Concerning equivocal results obtained in stand-alone assays in Table 1, two molecules, salicylic acid and sodium lauryl sulfate, had inconclusive results.

Salicylic acid appears positive (96.5% Cysteine/78.9% Lysine) in DPRA of Natsch et al. (2013) and negative (91.3% Cys/99% Lys) in Bauch et al. (2012). Similarly, sodium lauryl sulfate (SLS) appears negative (100% Cys/100% Lys) in DPRA of Natsch et al. (2013) and positive (100% Cys/2.9% Lys) in Bauch et al. (2012).

According to Gerberick et al. (2007), salicylic acid co-eluted with lysine peptide, indicating that it could not be evaluated using DPRA. However, salicylic acid could be easily separated from NAL using ADRA (Amino acid Derivative Reactivity Assay), and its depletion could hence be determined (Fujita et al., 2014). Finally, in case of an inconclusive result, we assumed a worst-case scenario by assigning “1” for individual method, without taking into account the applicability domain or the physical-properties of the test item.

The discussion of false predictions provided by the single non-animal test methods would be beyond the scope of this study. Hence, the analysis was restricted to the detailed investigation of false positive and false negative predictions by the “2 out of 3” approach, when compared to LLNA and human data.

4.3. Integrated Testing Strategy results

The “sensitization testing strategy” is an approach, which incorporates the latest validated (or under validation) methods. The safety assessor evaluating or analyzing simple cosmetic ingredients can apply it. Before defining its ITS, it is important to have an understanding on each test selected and especially to know their respective applicability domain as well as the strategy to follow. Various strategies are presently proposed (van der Veen et al., 2014a). The outcome of the approach suggested here meets the requirements of the EU Cosmetics Regulation by providing a sensitization profile, which covers all available sensitization endpoints, required to make a safety decision. In addition, this approach, to an extent, overcomes the challenges of the animal testing ban. Finally, this approach was confirmed by previous *in vivo* tests and human data. In order to facilitate acceptance of *in vitro* sensitization methods by regulatory committees, results of the single test methods as well as simple “2 out of 3” ITS have been compiled and qualitatively compared to both animal and human data. The sensitizing potential of the chemical reference was investigated in three explorative *in vitro* tests, DPRA, SENS-IS and h-CLAT validated or under validation by OECD (OECD, 2016c; 2015a) or ECVAM (under submission and external validation study). These three methods covered three out of four KEs defined by AOP for skin sensitization (OECD, 2012). Unfortunately, despite a good correlation between the proposed ITS and the results obtained by LLNA and human data, it is presently not yet approved to draw conclusions on potency by using only *in vitro* methods (Bauch et al., 2011). Regarding the results obtained from this prediction model, we achieved a good correlation between our final prediction and human data for 31 chemicals, except for phthalic anhydride and salicylic acid. As explained above, phthalic anhydride is known to be a respiratory allergen and supposed to be a cutaneous sensitizer in some specific populations. Regarding salicylic acid, its inconclusive result comes from the co-elution obtained with the lysine peptide, whereas it is a non-sensitizer in humans. In h-CLAT, salicylic acid was false negative in all labs because it is not stable in water and falls out of the applicability domain. Due to the attribution of “1” in inconclusive results into the DPRA and the false positive result in h-CLAT, salicylic acid has been misclassified in our prediction model. However, despite the results, this attribution number system allows us to reach a better

positive predictive value of our model than the LLNA, and at the same time to be more conservative.

Moreover, the major limitation of the methods presented is that they no longer permit the use of Quantitative Risk Assessment (QRA) approaches, as was done previously with available LLNA EC3 values. The EC3 value allows the safety assessor to carry out risk quantification, which was essential for the safety evaluation and risk assessment. With the ban on animal testing, we must therefore devise a new strategy to manage the use of skin sensitizing chemicals as effective as the QRA approach to define several potency classes. Unfortunately, unlike the LLNA and associated QRA process, all *in vitro* models currently developed allow only the prediction of hazard without a clear notion of potency.

Concerning our strategy, it should be noticed that this approach seems interesting for simple molecules but difficult to implement for mixtures (e.g. perfumes, essential oils and plant extracts). Indeed, although this seems applicable with the SENS-IS (Cottrez et al., 2016a), this goes beyond the scope of applicability domain for DPRA and h-CLAT. Moreover, preliminary results (data not shown) indicate that this prediction model does not appear to be applicable to mixtures.

Finally, different strategies were proposed in the literature, such as other ITS using the “2 out of 3” prediction model (Urbisch et al., 2015). For example, Bauch et al. (2012) proposed a strategy using DPRA, LuSens and mMUSST. An overall accuracy of 83% and 94% compared to LLNA or human data, respectively, was achieved. Natsch et al. (2013) used the DPRA, Keratino-Sens™ and MUSST in their “2 out of 3” prediction model and reported a similar accuracy for 145 substances (81% compared to LLNA). In the model proposed by Urbisch et al. (2015), the results of the DPRA, KeratinoSens™ and h-CLAT were used for applying this prediction model. All three of these “2 out of 3” prediction models provide very similar accuracy levels as high as those of our prediction model. These “2 out of 3” prediction models revealed a higher accuracy rate when compared to human data for the same set of substances. This indicates that the “2 out of 3” predictions might be more relevant for humans than the LLNA predictions.

However, some specific substances like pre/pro-haptens or molecules with specific physical-properties (e.g. physical state, stability, interference with the detection system) may be misclassified in some *in vitro* models. The major interest of our proposed strategy consists in integrating a 3-D reconstructed epidermal model (SENS-IS) in which solubility problems or technical limitations can be avoided.

5. Conclusion

Our original goal was to develop an ITS based on three *in vitro* methods that cover the main key events of the AOP for skin sensitization. Our prediction model seems to have a high accuracy rate of hazard prediction for human data compared to LLNA.

The interest of this approach is that it integrates a reconstructed human skin model (SENS-IS) avoiding some specific limitations (solubility, substances with high cytotoxicity or some pro/pre haptens) which may be limited in others methods. Finally, it seems interesting to carry out this prediction model sequentially: firstly, DPRA and SENS-IS; and in case of an ambiguous result, a third test, the h-CLAT, must be realized in order to draw a conclusion due to the WoE. Our prediction model could be used as an ITS, allowing the safety assessor to meet the EU Cosmetics Regulation requirements and to overcome the challenges raised by the animal testing ban in terms of skin sensitization.

Transparency document

The <http://dx.doi.org/10.1016/j.tiv.2017.05.014> associated with this article can be found, in online version.

References

- Ade, N., Martinozzi-Teissier, S., Pallardy, M., Rousset, F., 2006. Activation of U937 cells by contact sensitizers: CD86 expression is independent of apoptosis. *J. Immunotoxicol.* 3, 189–197. <http://dx.doi.org/10.1080/15476910600978038>.
- Adler, S., Basketter, D., Creton, S., Pelkonen, O., Van Benthem, J., Zuang, V., Andersen, K.E., Angers-Loustau, A., Aptula, A., Bal-Price, A., Benfenati, E., Bernauer, U., Bessems, J., Bois, F.Y., Boobis, A., Brandon, E., Bremer, S., Broschard, T., Casati, S., Coecke, S., Corvi, R., Cronin, M., Daston, G., Dekant, W., Felter, S., Grignard, E., Gundert-Remy, U., Heinonen, T., Kimber, I., Kleinjans, J., Komulainen, H., Kreiling, R., Kreysa, J., Leite, S.B., Loizou, G., Maxwell, G., Mazzatorta, P., Munn, S., Pfuhler, S., Phrakonkham, P., Piersma, A., Poth, A., Prieto, P., Repetto, G., Rogiers, V., Schoeters, G., Schwarz, M., Serafimova, R., Tähti, H., Testai, E., Van Delft, J., Van Loveren, H., Vinken, M., Worth, A., Zaldivar, J.M., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Arch. Toxicol.* 85, 367–485. <http://dx.doi.org/10.1007/s00204-011-0693-2>.
- Aeby, P., Ashikaga, T., Bessou-Touya, S., Schepky, A., Gerberick, F., Kern, P., Marrec-Fairley, M., Maxwell, G., Ovigne, J.M., Sakaguchi, H., Reisinger, K., Tailhardat, M., Martinozzi-Teissier, S., Winkler, P., 2010. Identifying and characterizing chemical skin sensitizers without animal testing: Colipa's research and method development program. *Toxicol. in Vitro* 24, 1465–1473. <http://dx.doi.org/10.1016/j.tiv.2010.07.005>.
- Alépée, N., Piroird, C., Aujoulat, M., Dreyfuss, S., Hoffmann, S., Hohenstein, A., Meloni, M., Nardelli, L., Gerbeix, C., Cotovio, J., 2015. Prospective multicentre study of the U-SENS test method for skin sensitization testing. *Toxicol. in Vitro* 30, 373–382. <http://dx.doi.org/10.1016/j.tiv.2015.09.028>.
- Ashikaga, T., Yoshida, Y., Hirota, M., Yoneyama, K., Itagaki, H., Sakaguchi, H., Miyazawa, M., Ito, Y., Suzuki, H., Toyoda, H., 2006. Development of an *in vitro* skin sensitization test using human cell lines: the human cell line activation test (h-CLAT). I. Optimization of the h-CLAT protocol. *Toxicol. in Vitro* 20, 767–773. <http://dx.doi.org/10.1016/j.tiv.2005.10.012>.
- Ashikaga, T., Sakaguchi, H., Sono, S., Kosaka, N., Ishikawa, M., Nukada, Y., Miyazawa, M., Ito, Y., Nishiyama, N., Itagaki, H., 2010. A comparative evaluation of *in vitro* skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Altern. Lab. Anim* 38, 275–284.
- Basketter, D.A., Andersen, K.E., Liden, C., Van Loveren, H., Boman, A., Kimber, I., Alanko, K., Berggren, E., 2005. Evaluation of the skin sensitizing potency of chemicals by using the existing methods and considerations of relevance for elicitation. *Contact Dermatitis* 52, 39–43. <http://dx.doi.org/10.1111/j.0105-1873.2005.00490.x>.
- Basketter, D.A., Alepee, N., Ashikaga, T., Barroso, J., Gilmour, N., Goebel, C., Hibatallah, J., Hoffmann, S., Kern, P., Martinozzi-Teissier, S., Maxwell, G., Reisinger, K., Sakaguchi, H., Schepky, A., Tailhardat, M., Templier, M., 2014. Categorization of chemicals according to their relative human skin sensitizing potency. *Dermat.* 25, 11–21. (contact, atopic, Occup. drug). <http://dx.doi.org/10.1097/DER.0000000000000003>.
- Bauch, C., Kolle, S.N., Fabian, E., Pachel, C., Ramirez, T., Wiench, B., Wruck, C.J., van Ravenzwaay, B., Landsiedel, R., 2011. Intralaboratory validation of four *in vitro* assays for the prediction of the skin sensitizing potential of chemicals. *Toxicol. in Vitro* 25, 1162–1168. <http://dx.doi.org/10.1016/j.tiv.2011.05.030>.
- Bauch, C., Kolle, S.N., 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, 489–504. <http://dx.doi.org/10.1016/j.yrtph.2012.05.013>.
- Bruckner, A.L., Weston, W.L., Morelli, J.G., 2000. Does sensitization to contact allergens begin in infancy? *Pediatrics* 105, e3.
- Cooper 2nd, J.A., Saracci, R., Cole, P., 1979. Describing the validity of carcinogen screening tests. *Br. J. Cancer* 39, 87–89.
- Cottrez, F., Boitel, E., Auriault, C., Aeby, P., Groux, H., 2015. Genes specifically modulated in sensitized skins allow the detection of sensitizers in a reconstructed human skin model. Development of the SENS-IS assay. *Toxicol. in Vitro* 29, 787–802. <http://dx.doi.org/10.1016/j.tiv.2015.02.012>.
- Cottrez, F., Boitel, E., Dalhuchyts, H., Bidan, C., Rattier, S., Ferret, P., Groux, H., 2016a. *In Vitro* Skin Sensitization Analysis of Mixtures: Results Using the SENS-IS Assay.
- Cottrez, F., Boitel, E., Ourlin, J.C., Peiffer, J.L., Fabre, I., Henaoui, I.S., Mari, B., Vallauri, A., Paquet, A., Barbry, P., Auriault, C., Aeby, P., Groux, H., 2016b. SENS-IS, a 3D reconstituted epidermis based model for quantifying chemical sensitization potency: reproducibility and predictivity results from an inter-laboratory study. *Toxicol. in Vitro* 32, 248–260. <http://dx.doi.org/10.1016/j.tiv.2016.01.007>.
- Dearman, R.J., Basketter, D.A., Kimber, I., 2013. Inter-relationships between different classes of chemical allergens. *J. Appl. Toxicol.* 33, 558–565. <http://dx.doi.org/10.1002/jat.1758>.
- EURL-ECVAM, 2016. EURL ECVAM status report on the development. In: Validation and Regulatory Acceptance of Alternative Methods and Approaches, <http://dx.doi.org/10.2788/86047>.
- European Chemicals Agency, 2016. Guidance on information requirements and chemical safety assessment. In: Chapter R7a: Endpoint Specific Guidance. European Chemicals Agency. <http://dx.doi.org/10.2823/2611>.
- European Commission, 2015. EURL ECVAM Recommendation on the Human Cell Line Activation Test (h-CLAT) for Skin Sensitisation Testing. <http://dx.doi.org/10.2788/29986>.
- European Parliament, 2016. Consolidated version of the REACH Regulation. Commission Regulation (EU) 2016/1688 of 20 September 2016 Amending Annex VII to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Res.
- European Union, 2010. Directive 2010/63/EU of the European Parliament and of the

- Council of 22 September 2010 on the protection of animals used for scientific purposes. Off. J. Eur. Union 33–79 (doi: 32010L0063).
- Ezendam, J., Braakhuis, H.M., Vandebriel, R.J., 2016. State of the art in non-animal approaches for skin sensitization testing: from individual test methods towards testing strategies. Arch. Toxicol. 90, 2861–2883. <http://dx.doi.org/10.1007/s00204-016-1842-4>.
- Fujita, M., Yamamoto, Y., Tahara, H., Kasahara, T., Jimbo, Y., Hioki, T., 2014. Development of a prediction method for skin sensitization using novel cysteine and lysine derivatives. J. Pharmacol. Toxicol. Methods 70, 94–105. <http://dx.doi.org/10.1016/j.vascn.2014.06.001>.
- Gerberick, G.F., Vassallo, J.D., Bailey, R.E., Chaney, J.G., Morrall, S.W., Lepoittevin, J.-P., 2004. Development of a peptide reactivity assay for screening contact allergens. Toxicol. Sci. 81, 332–343. <http://dx.doi.org/10.1093/toxsci/kfh213>.
- Gerberick, G.F., Vassallo, J.D., Foertsch, L.M., Price, B.B., Chaney, J.G., Lepoittevin, J.-P., 2007. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol. Sci. 97, 417–427. <http://dx.doi.org/10.1093/toxsci/kfm064>.
- Gerberick, F., Aleksic, M., Basketter, D., Casati, S., Karlberg, A.-T., Kern, P., Kimber, I., Lepoittevin, J.P., Natsch, A., Ovigne, J.M., Rovida, C., Sakaguchi, H., Schultz, T., 2008. Chemical reactivity measurement and the predictive identification of skin sensitizers. The report and recommendations of ECVAM Workshop 64. Altern. Lab. Anim 36, 215–242.
- Gerberick, G.F., Troutman, J.A., Foertsch, L.M., Vassallo, J.D., Quijano, M., Dobson, R.L.M., Goebel, C., Lepoittevin, J.-P., 2009. Investigation of peptide reactivity of pro-hapten skin sensitizers using a peroxidase-peroxide oxidation system. Toxicol. Sci. 112, 164–174. <http://dx.doi.org/10.1093/toxsci/kfp192>.
- Griem, P., Goebel, C., Scheffler, H., 2003. Proposal for a risk assessment methodology for skin sensitization based on sensitization potency data. Regul. Toxicol. Pharmacol. 38, 269–290.
- ICCVAM, 1999. The murine Local Lymph Node Assay: a test method for assessing the allergic contact dermatitis for potential of chemicals/compounds. In: 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 Center for the Evaluation of Alternative Toxicological Methods (NICEATM), NIH Publication No. 99-4494. (February).
- Jaworska, J., Dancik, Y., Kern, P., Gerberick, F., Natsch, A., 2013. Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice. J. Appl. Toxicol. <http://dx.doi.org/10.1002/jat.2869>.
- Jaworska, J.S., Natsch, A., Ryan, C., Strickland, J., Ashikaga, T., Miyazawa, M., 2015. Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative Weight of Evidence and adaptive testing strategy. Arch. Toxicol. 89, 2355–2383. <http://dx.doi.org/10.1007/s00204-015-1634-2>.
- Jordan, W.P., 1983. Allergic contact dermatitis to simple chemicals: a molecular approach. J. Am. Acad. Dermatol. 8, 575. [http://dx.doi.org/10.1016/S0190-9622\(83\)80070-X](http://dx.doi.org/10.1016/S0190-9622(83)80070-X).
- Karlberg, A.-T., Bergstrom, M.A., Borje, A., Luthman, K., Nilsson, J.L.G., 2008. Allergic contact dermatitis—formation, structural requirements, and reactivity of skin sensitizers. Chem. Res. Toxicol. 21, 53–69. <http://dx.doi.org/10.1021/tx7002239>.
- Kimber, I., 1995. Contact and respiratory sensitization by chemical allergens: uneasy relationships. Am. J. Contact Dermat. 6, 34–39. [http://dx.doi.org/10.1016/1046-199X\(95\)90068-3](http://dx.doi.org/10.1016/1046-199X(95)90068-3).
- Kimber, I., Basketter, D.A., 1992. The murine local lymph node assay: a commentary on collaborative studies and new directions. Food Chem. Toxicol. 30, 165–169. [http://dx.doi.org/10.1016/0278-6915\(92\)90153-C](http://dx.doi.org/10.1016/0278-6915(92)90153-C).
- Kimber, I., Basketter, D.A., Gerberick, G.F., Ryan, C.A., Dearman, R.J., 2011. Chemical allergy: translating biology into hazard characterization. Toxicol. Sci. 120 (Suppl), S238–S268. <http://dx.doi.org/10.1093/toxsci/kfq346>.
- Landsteiner, K., Jacobs, J., 1936. Studies on the sensitization of animals with simple chemical compounds. II. J. Exp. Med. 64, 625–639.
- Loveless, S.E., Ladies, G.S., Gerberick, G.F., Ryan, C.A., Basketter, D.A., Scholes, E.W., House, R.V., Hilton, J., Dearman, R.J., Kimber, I., 1996. Further evaluation of the local lymph node assay in the final phase of an international collaborative trial. Toxicology 108, 141–152. [http://dx.doi.org/10.1016/0300-483X\(95\)03279-0](http://dx.doi.org/10.1016/0300-483X(95)03279-0).
- MacKay, C., Davies, M., Summerfield, V., Maxwell, G., 2013. From pathways to people: applying the adverse outcome pathway (AOP) for skin sensitization to risk assessment. ALTEX 30, 473–486.
- Martin, S.F., Esser, P.R., Weber, F.C., Jakob, T., Freudenberger, M.A., Schmidt, M., Goebeler, M., 2011. Mechanisms of chemical-induced innate immunity in allergic contact dermatitis. Allergy 66, 1152–1163. <http://dx.doi.org/10.1111/j.1398-9995.2011.02652.x>.
- McKim, J.M.J., Keller 3rd, D.J., Gorski, J.R., 2010. A new in vitro method for identifying chemical sensitizers combining peptide binding with ARE/EpRE-mediated gene expression in human skin cells. Cutan. Ocul. Toxicol. 29, 171–192. <http://dx.doi.org/10.3109/15569527.2010.483869>.
- McKim, J.M.J., Keller 3rd, D.J., Gorski, J.R., 2012. An in vitro method for detecting chemical sensitization using human reconstructed skin models and its applicability to cosmetic, pharmaceutical, and medical device safety testing. Cutan. Ocul. Toxicol. 31, 292–305. <http://dx.doi.org/10.3109/15569527.2012.667031>.
- Natsch, A., Emter, R., 2008. Skin sensitizers induce antioxidant response element dependent genes: application to the in vitro testing of the sensitization potential of chemicals. Toxicol. Sci. 102, 110–119. <http://dx.doi.org/10.1093/toxsci/kfm259>.
- Natsch, A., Ryan, C.A., Foertsch, L., Emter, R., Jaworska, J., Gerberick, F., Kern, P., 2013. A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. J. Appl. Toxicol. 33, 1337–1352. <http://dx.doi.org/10.1002/jat.2868>.
- Natsch, A., Emter, R., Gfeller, H., Haupt, T., Ellis, G., 2015. Predicting skin sensitizer potency based on in vitro data from KeratinoSens and kinetic peptide binding: global versus domain-based assessment. Toxicol. Sci. 143, 319–332. <http://dx.doi.org/10.1093/toxsci/kfu229>.
- Newcombe, D.S., 1992. Clinical Immunotoxicology.
- Nguyen, T., Nioi, P., Pickett, C.B., 2009. The Nrf2-antioxidant response element signaling pathway and its activation by oxidative stress. J. Biol. Chem. 284, 13291–13295. <http://dx.doi.org/10.1074/jbc.R900012000>.
- Nukada, Y., Miyazawa, M., Kazutoshi, S., Sakaguchi, H., Nishiyama, N., 2013. Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. Toxicol. in Vitro 27, 609–618. <http://dx.doi.org/10.1016/j.tiv.2012.11.006>.
- OECD, 1992. Test No. 406: Skin Sensitisation, Guideline for Testing of Chemicals. <http://dx.doi.org/10.1787/9789264070684-en>.
- OECD, 2010a. Test No. 429: Skin Sensitisation: Local Lymph Node Assay, Guideline for Testing of Chemicals. <http://dx.doi.org/10.1787/9789264071100-en>.
- OECD, 2010b. Test No. 442A: Skin Sensitization: Local Lymph Node Assay: DA. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264090972-en>.
- OECD, 2010c. Test No. 442B: Skin Sensitization: Local Lymph Node Assay: BrdU-ELISA.
- OECD, 2012. The adverse outcome pathway for skin sensitization initiated by covalent binding for proteins. Part 1. In: Scientific Evidence OECD Environment, Health and Safety Publications Series on Testing and Assessment.
- OECD, 2015a. Test No. 442C: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA).
- OECD, 2015b. Test No. 439: In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method. <http://dx.doi.org/10.1787/9789264242845-en>.
- OECD, 2016a. Guidance document on the reporting of defined approaches to be used within integrated approaches to testing and assessment. In: ENV/JM/MONO(2016) 28. OECD Publishing, Paris.
- OECD, 2016b. 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. In: ENV/JM/MONO(2016)29. OECD Publishing, Paris.
- OECD, 2016c. Test No. 442E: In Vitro Skin Sensitisation: Human Cell Line Activation Test (h-CLAT). <http://dx.doi.org/10.1787/9789264264359-en>.
- Patlewicz, G., Kuseva, C., Kesova, A., Popova, I., Zhechev, T., Pavlov, T., Roberts, D.W., Mekenyian, O., 2014. Towards AOP application—implementation of an integrated approach to testing and assessment (IATA) into a pipeline tool for skin sensitization. Regul. Toxicol. Pharmacol. 69, 529–545. <http://dx.doi.org/10.1016/j.yrtph.2014.06.001>.
- Piroid, C., Ovigne, J.-M., Rousset, F., Martinuzzi-Teissier, S., Gomes, C., Cotovio, J., Alépée, N., 2015. The myeloid U937 skin sensitization test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. Toxicol. in Vitro 29, 901–916. <http://dx.doi.org/10.1016/j.tiv.2015.03.009>.
- Ramirez, T., Mehling, A., Kolle, S.N., Wruck, C.J., Teubner, W., Eltze, T., Aumann, A., Urbisch, D., van Ravenzwaay, B., Landsiedel, R., 2014. LuSens: a keratinocyte based ARE reporter gene assay for use in integrated testing strategies for skin sensitization hazard identification. Toxicol. in Vitro 28, 1482–1497. <http://dx.doi.org/10.1016/j.tiv.2014.08.002>.
- Ramirez, T., Stein, N., Aumann, A., Remus, T., Edwards, A., Norman, K.G., Ryan, C., Bader, J.E., Fehr, M., Burleson, F., Foertsch, L., Wang, X., Gerberick, F., Beilstein, P., Hoffmann, S., Mehling, A., van Ravenzwaay, B., Landsiedel, R., 2016. Intra- and inter-laboratory reproducibility and accuracy of the LuSens assay: a reporter gene cell line to detect keratinocyte activation by skin sensitizers. Toxicol. in Vitro 32, 278–286. <http://dx.doi.org/10.1016/j.tiv.2016.01.004>.
- Rastogi, S.C., Lepoittevin, J.P., Johansen, J.D., Frosch, P.J., Menne, T., Bruze, M., Dreier, B., Andersen, K.E., White, I.R., 1998. Fragrances and other materials in deodorants: search for potentially sensitizing molecules using combined GC-MS and structure activity relationship (SAR) analysis. Contact Dermatitis 39, 293–303.
- Reisinger, K., Hoffmann, S., Alépée, N., Ashikaga, T., Barroso, J., Elcombe, C., Gellatly, N., Galbiati, V., Gibbs, S., Groux, H., Hibatallah, J., Keller, D., Kern, P., Klaric, M., Kolle, S., Kuehn, J., Lambrechts, N., Lindstedt, M., Millet, M., Martinuzzi-Teissier, S., Natsch, A., Petersohn, D., Pike, I., Sakaguchi, H., Schepky, A., Tailhardat, M., Templier, M., Van Vliet, E., Maxwell, G., 2015. Systematic evaluation of non-animal test methods for skin sensitisation safety assessment. Toxicol. in Vitro 29, 259–272. <http://dx.doi.org/10.1016/j.tiv.2014.10.018>.
- Russell, Burch, R.L., 1959. The Principles of Humane Experimental Technique. Methuen, London UK.
- Ryan, C. a, Kimber, I., Basketter, D. a, Pallardy, M., Gildea, L. a, Gerberick, G.F., 2007. Dendritic cells and skin sensitization: biological roles and uses in hazard identification. Toxicol. Appl. Pharmacol. 221, 384–394. <http://dx.doi.org/10.1016/j.taap.2007.03.006>.
- Sakaguchi, H., Ashikaga, T., Miyazawa, M., Yoshida, Y., Ito, Y., Yoneyama, K., Hirota, M., Itagaki, H., Toyoda, H., Suzuki, H., 2006. Development of an in vitro skin sensitization test using human cell lines; human cell line activation test (h-CLAT). II. An inter-laboratory study of the h-CLAT. Toxicol. in Vitro 20, 774–784. <http://dx.doi.org/10.1016/j.tiv.2005.10.014>.
- Sakaguchi, H., Ashikaga, T., Miyazawa, M., Kosaka, N., Ito, Y., Yoneyama, K., Sono, S., Itagaki, H., Toyoda, H., Suzuki, H., 2009. The relationship between CD86/CD54 expression and THP-1 cell viability in an in vitro skin sensitization test - human cell line activation test (h-CLAT). Cell Biol. Toxicol. 25, 109–126. <http://dx.doi.org/10.1007/s10565-008-9059-9>.
- Sakaguchi, H., Ryan, C., Ovigne, J.-M., Schroeder, K.R., Ashikaga, T., 2010. Predicting skin sensitization potential and inter-laboratory reproducibility of a human cell line activation test (h-CLAT) in the European Cosmetics Association (COLIPA) ring trials.

- Toxicol. in Vitro 24, 1810–1820. <http://dx.doi.org/10.1016/j.tiv.2010.05.012>.
- Schneider, K., Akkan, Z., 2004. Quantitative relationship between the local lymph node assay and human skin sensitization assays. *Regul. Toxicol. Pharmacol.* 39, 245–255. <http://dx.doi.org/10.1016/j.yrtph.2004.02.002>.
- Strickland, J., Zang, Q., Kleinstreuer, N., Paris, M., Lehmann, D.M., Choksi, N., Matheson, J., Jacobs, A., Lowit, A., Allen, D., Casey, W., 2016. Integrated decision strategies for skin sensitization hazard. *J. Appl. Toxicol.* 36, 1150–1162. <http://dx.doi.org/10.1002/jat.3281>.
- Thyssen, J.P., Linneberg, A., Menné, T., Johansen, J.D., 2007. The epidemiology of contact allergy in the general population—prevalence and main findings. *Contact Dermatitis* 57, 287–299. <http://dx.doi.org/10.1111/j.1600-0536.2007.01220.x>.
- Tollefsen, K.E., Scholz, S., Cronin, M.T., Edwards, S.W., de Knecht, J., Crofton, K., Garcia-Reyero, N., Hartung, T., Worth, A., Patlewicz, G., 2014. Applying adverse outcome pathways (AOPs) to support integrated approaches to testing and assessment (IATA). *Regul. Toxicol. Pharmacol.* 70, 629–640. <http://dx.doi.org/10.1016/j.yrtph.2014.09.009>.
- Tsujita-Inoue, K., Hirota, M., Ashikaga, T., Atobe, T., Kouzuki, H., Aiba, S., 2014. Skin sensitization risk assessment model using artificial neural network analysis of data from multiple in vitro assays. *Toxicol. in Vitro* 28, 626–639. <http://dx.doi.org/10.1016/j.tiv.2014.01.003>.
- Urbisch, D., Mehling, A., Guth, K., Ramirez, T., Honarvar, N., Kolle, S., Landsiedel, R., Jaworska, J., Kern, P.S., Gerberick, F., Natsch, A., Emter, R., Ashikaga, T., Miyazawa, M., Sakaguchi, H., 2015. Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul. Toxicol. Pharmacol.* 71, 337–351. <http://dx.doi.org/10.1016/j.yrtph.2014.12.008>.
- van der Veen, J., Vandebriel, R., van Loveren, H., Ezendam, J., 2011. Keratinocytes, Innate Immunity and Allergic Contact Dermatitis-Opportunities for the Development of In Vitro Assays to Predict the Sensitizing Potential of Chemicals. *Ro YS, (Contact de. ed)*.
- van der Veen, J.W., Pronk, T.E., van Loveren, H., Ezendam, J., 2013. Applicability of a keratinocyte gene signature to predict skin sensitizing potential. *Toxicol. in Vitro* 27, 314–322. <http://dx.doi.org/10.1016/j.tiv.2012.08.023>.
- van der Veen, J.W., Rorije, E., Emter, R., Natsch, A., Van, L.H., Ezendam, J., 2014a. Evaluating the performance of integrated approaches for hazard identification of skin sensitizing chemicals. *Regul. Toxicol. Pharmacol.* 69, 371–379. <http://dx.doi.org/10.1016/j.yrtph.2014.04.018>.
- van der Veen, J.W., Soeteman-Hernandez, L.G., Ezendam, J., Stierum, R., Kuper, F.C., van Loveren, H., 2014b. Anchoring molecular mechanisms to the adverse outcome pathway for skin sensitization: analysis of existing data. *Crit. Rev. Toxicol.* 44, 590–599. <http://dx.doi.org/10.3109/10408444.2014.925425>.
- Weltzien, H.U., Corsini, E., Gibbs, S., Lindstedt, M., Borrebaeck, C., Budde, P., Schulz-Knappe, P., Thierse, H.-J., Martin, S.F., Roggen, E.L., 2009. Safe cosmetics without animal testing? Contributions of the EU Project Sens-it-iv. *J. Verbr. Lebensm.* 4, 41–48. <http://dx.doi.org/10.1007/s00003-009-0510-5>.