Toxicology in Vitro 32 (2016) 248–260

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

SENS-IS, a 3D reconstituted epidermis based model for quantifying chemical sensitization potency: Reproducibility and predictivity results from an inter-laboratory study

Françoise Cottrez [a], Elodie Boitel [a], Jean-Claude Ourlin [b], Jean-Luc Peiffer [b], Isabelle Fabre [b], Imène-Sarah Henaoui [c,d], Bernard Mari [c,d], Ambre Vallauri [c,d], Agnes Paquet [c,d], Pascal Barbry [c,d], Claude Auriault [a], Pierre Aeby [e], Hervé Groux [a,*]

[a] ImmunoSearch, Grasse, France
[b] Agence nationale de sécurité du médicament, Vendargues, France
[c] CNRS, Institute of Molecular and Cellular Pharmacology, Sophia Antipolis, France
[d] University of Nice Sophia Antipolis, Nice, France
[e] Independant Consultant, Marly, Switzerland

ARTICLE INFO

Article history:
Received 6 November 2015
Received in revised form 12 January 2016
Accepted 14 January 2016
Available online 18 January 2016

Keywords:
Skin sensitization
Alternatives to animal tests
Toxicogenomics
Reconstituted epidermis
SENS-IS

ABSTRACT

The SENS-IS test protocol for the *in vitro* detection of sensitizers is based on a reconstructed human skin model (Episkin) as the test system and on the analysis of the expression of a large panel of genes. Its excellent performance was initially demonstrated with a limited set of test chemicals. Further studies (described here) were organized to confirm these preliminary results and to obtain a detailed statistical analysis of the predictive capacity of the assay. A ring-study was thus organized and performed within three laboratories, using a test set of 19 blind coded chemicals. Data analysis indicated that the assay is robust, easily transferable and offers high predictivity and excellent within- and between-laboratories reproducibility. To further evaluate the predictivity of the test protocol according to Cooper statistics a comprehensive test set of 150 chemicals was then analyzed. Again, data analysis confirmed the excellent capacity of the SENS-IS assay for predicting both hazard and potency characteristics, confirming that this assay should be considered as a serious alternative to the available *in vivo* sensitization tests.

1. Introduction

The sensitizing potential of new chemicals was and is still (regulation permitting) evaluated using animal tests, such as the murine local lymph node assay (LLNA; OECD TG 429) or the Magnuson–Kligman guinea pig-based tests (OECD TG 406).

In March 2013, the European Commission, through the Cosmetics Regulation finalized and adopted a testing and marketing ban, prohibiting to test finished cosmetic products and cosmetic ingredients on animals or to market them in the EU if tested on animals. In parallel, the REACH regulation (registration, evaluation, authorization and restriction of chemicals) requires the testing of thousands of previously untested or not fully characterized chemicals for their skin sensitization potential, thus increasing the demand for animal testing or suitable *in vitro* alternatives. In this context, methods that can identify potential sensitizers or characterize the skin sensitization potential of chemicals without the use of animals are gaining much importance for the cosmetic and chemical industries as valuable tools to guarantee the safety of new or existing products.

The biological and chemical processes leading to skin sensitization have been analyzed in great details. The key biological events are described in the OECD report on: "The Adverse Outcome Pathway (AOP) for Skin Sensitisation Initiated by Covalent Binding to Proteins" (OECD 2014). Briefly, skin sensitization results from an immune reaction to small reactive chemicals (called haptens) that can penetrate the skin barrier, reach the viable part of the epidermis and bind to endogenous skin proteins (Divkovic et al., 2005). In general, skin sensitizing molecules are small reactive chemicals that have the potential to covalently bind to skin proteins and render them immunogenic. Some molecules need metabolic (pro-haptens) or chemical (pre-haptens) activation to act as haptens (Karlberg et al., 2008). The modified proteins are eventually processed by dendritic cells (DCs) that after proper activation by danger signals migrate to the draining lymph nodes where

* Corresponding author.
E-mail address: hgroux@immunosearch.fr (H. Groux).

http://dx.doi.org/10.1016/j.tiv.2016.01.007
0887-2333/

they complete maturation and present fragments of the modified proteins to the adaptive immune system, instructing the allergen-specific T cells to mount the appropriate immune response by differentiating into a specialized T cell (Th1, Th2, Th17 or Tr1) (Schraml and Reis e Sousa, 2015).

In an attempt to model this complex biological cascade, a range of different *in silico*, *in chemico* (e.g. DPRA (Gerberick et al., 2004)) or *in vitro* cell-based methods (e.g. MUSST (Ade et al., 2006), h-CLAT (Yoshida et al., 2003), KeratinoSens (Emter et al., 2013)) have been developed. *In silico* models uses published skin sensitizers databases to empirically derive rules or structural alerts as predictors of sensitization hazard. *In chemico* approaches measure the chemical reactivity of the test chemicals with nucleophiles, usually model peptides (Gerberick et al., 2004). And cell-based *in vitro* models evaluate certain aspects of the cellular response to skin sensitizers, using either keratinocytes (KeratinoSens, LuSens (Ramirez et al., 2014)) or monocytic cell lines (h-CLAT, MUSST) usually measuring various signals produced by the cell system in response to hapten exposure.

However, most of these approaches reproduce only some particular aspects of the sensitization biological cascade and need to be combined in an integrated test strategy to obtain useful information concerning complex molecules such as cosmetic ingredients. Although some combinations appear to be promising for hazard identification, potency assessment is still difficult. Possible limitations may be due to the inadequate metabolism capacity (compared to native skin) of the test systems, to bioavailability aspects that cannot be reproduced in monolayer cultures, or to the danger signal that may be different in monolayers as compared to a natural tridimensional microenvironment.

We thus decided to take a fresh look on these issues and have developed and proposed an innovative *in vitro* system for the detection of sensitizers. The resulting SENS-IS test protocol (Cottrez et al., 2015) describes a new approach based on a reconstructed human skin model (Episkin) as the test system and on the analysis of the expression of a large panel of genes relevant to the considered biological processes. The modulation of these biomarkers is measured by quantitative reverse transcription-polymerase chain reaction (RT-PCR). Briefly, the test item is applied on the skin model at 4 different concentrations (50%, 10%; 1% and 0.1%) in the appropriate vehicle. After exposure, gene expression of two groups of genes is measured: One group (REDOX group) includes a selection of 17 genes that have an antioxidant responsive element in their promoter and monitor the redox protective signals induced through the interaction of sensitizers binding to cysteine amino acids of the Keap1-NRF2 complex (Uruno and Motohashi, 2011). The second group (SENS-IS group) includes a selection of 21 genes involved in inflammation, danger signals and cell migration to address the complex cascade of events leading to activation of DCs by a sensitizing chemical. Finally, the prediction model (PM), based on the number of over-expressed genes makes a decision according to the following rules: if at a given test concentration (see below) a compound induces at least 7 genes in either the REDOX or SENS-IS group of genes, it is classified as a sensitizer. Potency, using the same classification as in the LLNA, is deduced through the minimal concentration needed to induce the over-expression of ≥7 genes in the REDOX or SENS-IS groups. A compound is classified as an extreme, a strong, a moderate or a weak sensitizer if detected as a sensitizer at a test concentration of 0.1%, 1%, 10% or 50% respectively.

The performance of the SENS-IS protocol was initially evaluated with 50 chemicals (Cottrez et al., 2015). It's hazard (100%) and potency (96%) predictivity were excellent when compared with the corresponding LLNA data.

In order to confirm these promising results, we have organized a ring-study within three laboratories using a test set of 19 reference substances and a comprehensive evaluation of the SENS-IS protocol using a much wider set of 150 chemicals, all tested blind! The results of these evaluations are presented in this paper.

2. Materials and methods

2.1. Chemicals

All test chemicals were obtained from Sigma-Aldrich (Lyon, France). They were selected to include most of the test items proposed by Casati et al. (2009) along with chemicals selected from the list published by the Sens-it-iv consortium (Sens-it-iv, 2009). This selection was completed with many non-sensitizers and other molecules that do not interact with cysteine. Supporting information on these substances and their sensitization potential as determined by LLNA is available in the ICCVAM database (ICCVAM, 2013) and in Basketter et al. (1999).

2.2. The SENS-IS test protocol

The performance of the SENS-IS assay has been described in Cottrez et al. (2015). A brief outline is described below:

A) Cell culture
The reconstructed epidermis, Episkin large model (thirty days old models, 1.07 cm^2) were purchased from Skinethic, Lyon, France. These models are differentiated from human keratinocytes isolated from healthy donors and include a human collagen (Type I) matrix, which stands for the dermis, coated with a thin layer of Type IV human collagen, and a fully stratified and differentiated epidermis covered by a stratum corneum.
Upon receipt, the reconstructed epidermises are maintained in sterile 12-well culture dishes containing 2 ml/well of pre-warmed (37 °C) maintenance medium (DMEM/Ham's F12) supplied with the kit. The medium is replaced (1 ml) each 24 h. The Episkin cultures are shipped onto a nutritive gel for transportation to maintain viability. In accordance with the manufacturer's instructions, shipment quality controls (pH, temperature) were carried out upon receipt and if OK, the epidermis were transferred to sterile 12-well culture dishes containing 2 ml/well of pre warmed maintenance medium.

B) Treatment protocol
At day 14 of culture, the Episkin tissues are exposed to the chemicals. Liquid or viscous chemicals (30 ± 3 μl) are applied with a micropipette onto the stratum corneum of the epidermis. The tissues are then incubated with the chemicals for 15 min at 37 °C/5%CO_2 and washed with 25 ml PBS per epidermis. After washing, the tissues are further incubated at 37 °C, 5% CO_2 for 6 h.

C) Total RNA extraction
Each epidermis model is lifted from the insert using forceps and transferred into a vial for snap freezing in liquid nitrogen. Total RNA is extracted using Qiazol reagent and RNeasy Mini Kit (Qiagen, Courtaboeuf, France). Briefly, the epidermis model is submerged in 1 ml Qiazol reagent and homogenized using two steel beads with a TissueLyser (Qiagen). After centrifugation, the supernatant is collected, completed with 0.2 ml of bromochloropropane (Sigma, Lyon, France) and vortexed. After centrifugation at 12′ 000 g for 15 min at 4 °C, the aqueous phase is collected and added to 600 μl of 70% ethanol, immediately mixed by pipetting and loaded onto a RNeasy spin column placed in a 2 ml collection tube. Total RNA is then extracted according to the manufacturer's instructions.

D) Quantitative Reverse Transcription-Polymerase Chain Reaction (RT-PCR) Analysis
Reverse transcription of total RNA is performed using 1 μg total RNA in 20 μl final volume using Random Primers and SuperScript III Reverse Transcriptase (Invitrogen, Saint Aubin, France), according to the manufacturer's instructions. Quantitative RT-PCR is performed using SYBR Green Real-time PCR Master Mix (ROCHE Diagnostic, Grenoble, France) with 0.4 μM of each oligonucleotide primer in a total volume 25 μl. The reaction is performed in an

LC480 System (ROCHE Diagnostics). The amplification program consists of one cycle at 95 °C with a 1 min hold, followed by 40 amplification cycles (95 °C for 15 s and, annealing at 60 °C for 15 s) and completed by a final 72 °C step for 40 s. The relative amount of each transcript is normalized to the amount of the mean expression levels of 3 house-keeping genes transcripts (Glucuronidase β, β2 microglobuline, and ≪non-POU domain containing, octamer-binding≫ (NONO)).

E) Prediction model
The genes expressed in each sample are measured through an absolute quantification analysis using the second derivative maximum method with the software provided by ROCHE Diagnostics. The level of tissue destruction is evaluated with the Cycle Threshold (CT) value of the HSPAA1 gene. A sample CT value >110% of the corresponding value obtained for the vehicle control indicates an unacceptable level of tissue damage. Moreover, irritation signals are analyzed by measuring the expression level of 23 genes (HAS1, IL-8, IL-1 A, PLAUR, IL-7R, JUN, TNFA, RELT, IER3, IL-4R, ICAM2, CCL20, COX2, MMP10, MKP1, GM-CSF, IL-23, IL-24, ICAM1, CXCL1, MMP3, MMP9, EGR1). If >20 genes are over-expressed (>1.25 × control vehicle value) the sample is rejected and the chemical tested at a lower concentration. A test item is considered as a sensitizer (positive) when at least 7 genes from both the REDOX and the SENS-IS gene lists are over-expressed (>1.25 × control vehicle value)(Cottrez et al., 2015). The lowest test item concentration fulfilling these criteria is then used for potency classification as follow: A chemical is classified as an extreme, a strong, a moderate or a weak sensitizer if found positive at 0.1%, 1%, 10% or 50% respectively. If negative at all tested concentrations, it is considered as a non-sensitizer.

2.3. Transfer of the SENS-IS protocol to participating laboratories

2.3.1. Transferability phase I — training in the lead laboratory

Two laboratories were selected for the transfer phases, e.g. the Agence Nationale de Sécurité des Médicament (ANSM) laboratory in Montpellier France and the CNRS laboratory (Institut de Pharmacologie Moléculaire et Cellulaire (IPMC)) in Sophia Antipolis, France. Both laboratories were well trained in all aspects of quantitative RT-PCR analysis, and both were equipped with the same PCR apparatus (Roche's LC 480) as the one used in the lead laboratory (ImmunoSearch). The assay transfer was performed in 2 steps: First a training in ImmunoSearch's facilities and then a follow-up phase in the participating laboratories (See Transferability phases I–III below),

The training phase consisted in a four days session organized with technicians from the participating laboratories. They performed a full SENS-IS assay, starting with the preparation of the chemicals (weighting, dilutions and selection of vehicle), application on the epidermis, washing followed by tissue collection, RNA and cDNA preparation, quantitative RT-PCR and results analysis. 8 coded chemicals at two concentrations (50% and 10%, see Supplementary data Table 1) were analyzed. Within experiment reproducibility was checked through the use of DMSO as an internal control and as a test chemical. The success criterion (before entering phase II) was a correct classification of the 6 sensitizers and the 2 non-sensitizers (Table 6).

2.3.2. Transferability phase II — transfer of the assay to the participating laboratories

After successful completion of the training phase (see pt. 2.3.1 above), the same 8 chemicals (same batches with a different coding scheme) and the positive and negative controls (i.e. SLS, HCA, TNBS and DMSO, not coded) were distributed to the participating laboratories. The SENS-IS assay was performed according to the provided protocol and the results, obtained after application of the PM, sent back to the lead laboratory for evaluation. The success criterion was the correct classification of the 8 chemicals.

2.3.3. Transferability phase III — results for blind-coded substances

After successful completion of the phase II, laboratories could proceed to phase III, consisting in the testing of 19 blind-coded chemicals (16 sensitizers and 3 non-sensitizers, see Table 1 and the statistical considerations below) independently in the 3 laboratories (ImmunoSearch, ANSM and IPMC). The test items were coded and distributed by the lead laboratory. The SENS-IS assay was performed according to the final protocol and the results obtained after application of the PM sent back to the lead laboratory and evaluated by an independent statistician for within- and between-laboratory reproducibility evaluation (see below). The success criterion for this Transferability phase III was a correct classification of all tested chemicals.

2.4. Within-laboratory reproducibility (WLR) and between-laboratory reproducibility (BLR) evaluation

2.4.1. Statistical considerations

2.4.1.1. Sample size calculation. The sample size for the evaluation of WLR and BLR was calculated based on the expected proportion of concordant classifications obtained in different experiments performed within the same laboratory or in different laboratories, respectively. The methodology applied is the one for comparing two proportions, when one is known, as detailed in Eq. (3.10), from Machin et al. (1997).

Considering that:

π	the expected proportion of concordant classifications among laboratories
π – δ	the lower border of the Confidence Interval for the expected proportion π (the reference proportion)
α	type I error, i.e. the probability to conclude that there is a significant difference in the mean performances of the laboratories when there is not.
1 – β	statistical power, i.e. the probability to detect correctly an existing difference in the performances of the laboratories.
β	type II error, i.e. the probability to not detect a significant difference among the mean performances of the laboratories when there is one.

2.4.1.2. Sample size calculation for the BLR study. The 'expected proportion' of concordant classifications (between laboratories) was set to be 95% on the basis of available pre-validation data generated by ImmunoSearch. The lower border of the Confidence Interval was set at 70%. The power was set to 80% and the α to 0.05.

Using the equation developed by Machin et al.

$$n = \frac{\left[Z_{1-\beta}\sqrt{\pi(1-\pi)} + Z_{1-\alpha/2}\sqrt{(\pi-\delta)(1-\pi+\delta)}\right]^2}{\delta^2}$$

We obtain n = 19 chemicals. We therefore selected a set of 19 chemicals representative of all possible value of the 5 potency classes: 2 extremes, 6 strongs, 4 moderates, 4 weaks and 3 non-sensitizers.

2.4.1.3. Sample size calculation for the WLR study. Similar calculations were performed for the WLR. Based on previous experiences of WLR studies at ImmunoSearch the WLR was set at 95%. The lower border of confidence was set to 80%. The statistical power chosen was 85%, and the α to 0.05. Using the same formula, the minimal sample size was set at n = 45.

Table 1
WLR analysis based on hazard predictions obtained in the three indicated laboratories.

ImmunoSearch						IPMC						ANSM					
Chemical	Concentration	Replicate 1	Replicate 2	Replicate 3	Agreement	Chemical	Concentration	Replicate 1	Replicate 2	Replicate 3	Agreement	Chemical	Concentration	Replicate 1	Replicate 2	Replicate 3	Agreement
PG	100%	NEG	NEG	NEG	Y	PG	100%	NEG	NEG	NEG	Y	PG	100%	NEG	NEG	NEG	Y
PG	10%	NEG	NEG	NEG	Y	PG	10%	NEG	NEG	NEG	Y	PG	10%	NEG	NEG	NEG	Y
DMSO	100%	NEG	NEG	NEG	Y	DMSO	100%	NEG	NEG	NEG	Y	DMSO	100%	NEG	NEG	NEG	Y
DMSO	10%	NEG	NEG	NEG	Y	DMSO	10%	NEG	NEG	NEG	Y	DMSO	10%	NEG	NEG	NEG	Y
DNCB	10%	S	S	S	Y	DNCB	10%	S	S	S	Y	DNCB	10%	S	S	S	Y
DNCB	1%	S	S	S	Y	DNCB	1%	S	S	S	Y	DNCB	1%	S	S	S	Y
DNCB	0.1%	S	S	S	Y	DNCB	0.1%	S	S	S	Y	DNCB	0.1%	S	S	S	Y
Propyl gallate	50%	S	S	S	Y	Propyl gallate	50%	S	S	S	Y	Propyl gallate	50%	S	S	S	Y
Propyl gallate	10%	S	S	S	Y	Propyl gallate	10%	S	S	S	Y	Propyl gallate	10%	NEG	S	S	N
Propyl gallate	1%	S	S	S	Y	Propyl gallate	1%	NEG	NEG	NEG	Y	Propyl gallate	1%	NEG	NEG	NEG	Y
Propyl gallate	0.1%	NEG	NEG	NEG	Y	Geraniol	10%	S	S	S	Y	Geraniol	50%	NEG	S	S	N
Geraniol	50%	S	S	S	Y	Geraniol	1%	NEG	NEG	NEG	Y	Geraniol	10%	S	S	S	Y
Geraniol	10%	S	S	S	Y	Resorcinol	50%	S	S	S	Y	Geraniol	1%	NEG	NEG	NEG	Y
Geraniol	1%	NEG	NEG	NEG	Y	Resorcinol	10%	S	S	S	Y	Resorcinol	50%	S	S	S	Y
Resorcinol	50%	S	S	S	Y	Resorcinol	1%	NEG	NEG	NEG	Y	Resorcinol	10%	S	S	S	Y
Resorcinol	10%	S	S	S	Y	Cinnamal	10%	S	S	S	Y	Resorcinol	1%	NEG	NEG	NEG	Y
Resorcinol	1%	NEG	NEG	NEG	Y	Cinnamal	1%	S	S	S	Y	Cinnamal	10%	S	S	S	Y
Cinnamal	10%	S	S	S	Y	Cinnamal	0.1%	NEG	NEG	NEG	Y	Cinnamal	1%	S	S	S	Y
Cinnamal	1%	S	S	S	Y	O-aminophenol	10%	S	S	S	Y	Cinnamal	0.1%	NEG	NEG	NEG	Y
Cinnamal	0.1%	NEG	NEG	NEG	Y	O-aminophenol	1%	S	S	S	Y	O-aminophenol	10%	S	S	S	Y
O-aminophenol	10%	S	S	S	Y	O-aminophenol	0,1%	NEG	NEG	NEG	Y	O-aminophenol	1%	S	S	S	Y
O-aminophenol	1%	S	S	S	Y	TNBS	1%	S	S	S	Y	O-aminophenol	0.1%	NEG	NEG	NEG	Y
O-aminophenol	0.1%	NEG	NEG	NEG	Y	TNBS	0.1%	S	S	S	Y	TNBS	1%	S	S	S	Y
TNBS	1%	S	S	S	Y	HCA	50%	S	S	S	Y	TNBS	0.1%	S	S	S	Y
TNBS	0.1%	S	S	S	Y	HCA	10%	NEG	NEG	NEG	Y	HCA	50%	S	S	S	Y
HCA	50%	S	S	S	Y	L-Limonene	50%	S	S	S	Y	HCA	10%	NEG	S	S	N
HCA	10%	NEG	NEG	S	N	L-Limonene	10%	NEG	NEG	NEG	Y	HCA	1%	S	S	NEG	N
L-Limonene	50%	S	S	S	Y	MCI (1.5%)	50%	S	S	S	Y	HCA	0.1%	NEG	NEG	NEG	Y
L-Limonene	10%	NEG	NEG	NEG	Y	MCI (1.5%)	10%	NEG	NEG	NEG	Y	L-Limonene	50%	S	S	S	Y
MCI (1,5%)	50%	S	S	S	Y	Trans-anethole	10%	S	S	S	Y	L-Limonene	10%	NEG	NEG	NEG	Y
MCI (1.5%)	10%	NEG	S	NEG	N	Trans-anethole	1%	NEG	NEG	NEG	Y	MCI (1.5%)	50%	S	S	S	Y
Trans-anethole	50%	S	S	S	Y	PDIC	10%	S	S	S	Y	MCI (1.5%)	10%	NEG	NEG	NEG	Y
Trans-anethole	10%	S	S	S	Y	PDIC	1%	S	S	S	Y	Trans-anethole	50%	S	S	S	Y
Trans-anethole	1%	NEG	NEG	NEG	Y	PDIC	0.1%	NEG	NEG	NEG	Y	Trans-anethole	10%	S	S	S	Y
PDIC	10%	S	S	S	Y	BBMG	10%	S	S	S	Y	Trans-anethole	1%	NEG	NEG	NEG	Y
PDIC	1%	S	S	S	Y	BBMG	1%	S	S	S	Y	PDIC	10%	S	S	S	Y
PDIC	0.1%	NEG	NEG	NEG	Y	BBMG	0.1%	NEG	NEG	NEG	Y	PDIC	1%	S	S	S	Y
BBMG	10%	S	S	S	Y	Eugenol	50%	S	NEG	S	N	PDIC	0.1%	NEG	NEG	NEG	Y
BBMG	1%	S	S	S	Y	Eugenol	10%	S	S	S	Y	BBMG	10%	S	S	S	Y
BBMG	0.1%	NEG	NEG	NEG	Y	Eugenol	1%	NEG	NEG	NEG	Y	BBMG	1%	S	S	S	Y
Eugenol	50%	S	S	S	Y	Diethyl sulphate	10%	S	S	S	Y	BBMG	0.1%	NEG	NEG	NEG	Y
Eugenol	10%	S	S	S	Y	diethyl sulphate	1%	NEG	NEG	NEG	Y	Eugenol	50%	S	S	S	Y
Eugenol	1%	NEG	NEG	NEG	Y	SLS	10%	S	S	S	Y	Eugenol	10%	S	S	S	Y
Diethyl sulphate	10%	S	S	S	Y	SLS	1%	NEG	NEG	NEG	Y	Eugenol	1%	NEG	NEG	NEG	Y
Diethyl sulphate	1%	NEG	NEG	NEG	Y	Polysorbate 21	50%	S	S	S	Y	Diethyl sulphate	10%	S	S	S	Y
SLS	10%	S	S	S	Y	Polysorbate 21	10%	NEG	NEG	NEG	Y	Diethyl sulphate	1%	NEG	NEG	NEG	Y
SLS	1%	NEG	NEG	NEG	Y							SLS	10%	S	S	S	Y
Polysorbate 21	50%	S	S	S	Y							SLS	1%	NEG	NEG	NEG	Y
Polysorbate 21	10%	NEG	S	NEG	N							Polysorbate 21	50%	S	S	S	Y
												Polysorbate 21	10%	NEG	NEG	NEG	Y

The chemicals were tested at the indicated concentration in 3 replicates. Results were ranked either as positive (S) or negative (NEG); see ≪Prediction model≫ section under Materials and methods Section.
PG: propylene glycol; DMSO: dimethyl sulfoxide, DNCB: 2,4-dinitrochlorobenzene, TNBS: 2,4,6-trinitrobenzenesulfonic acid, HCA: hexyl cinnamal, MCI: methylchloroisothiazolinone, PDIC: potassium dichromate, BBMG: 2-bromo-2-(bromomethyl) glutaronitrile.

2.4.2. WLR and BLR measurements

The main determinant of the test method's reliability assessment was on the concordance of classification, as sensitizer or non-sensitizer, for each individual concentration tested, which was determined from the number of over-expressed genes in the groups of genes "REDOX" and "SENS-IS". The same set of 19 test chemicals was used in three independent experiments in each laboratory. Results were defined as negative ("NEG") if the number of genes over-expressed was below 7 in the 2 groups SENS-IS and REDOX, or as sensitizer ("S") if the number of genes over-expressed was ≥7 in at least one of these 2 groups of genes. We considered that assay results were in agreement when, for a given chemical at a given concentration, replicated assays resulted in the same NEG or S results. In the case where one of the replicate failed, only two replicates were used in the agreement assessment. Only concentrations tested at least twice were included in the final concordance analysis.

WLR was evaluated in the three laboratories participating to the ring study with the 19 chemicals analyzed. The main focus of the WLR study was on the concordance of the PM conclusions at a given concentration between the three independent experiments. In addition to the concordance calculation, we used the Fleiss kappa statistic to measure the inter-rater agreement.

The BLR was evaluated based on the results from the ring study in three laboratories testing in triplicates 19 blind coded substances. The main focus of the BLR for the 19 chemicals was on the concordance of the predictions as sensitizers (S) versus non-sensitizers (NS). In addition to the concordance calculation, we used the Fleiss kappa statistic to measure the inter-laboratory agreement.

2.5. Evaluation of a large test set to measure the predictivity of the SENS-IS assay according to Cooper statistics

In order to determine the predictive capacity of the SENS-IS assay, 150 test chemicals were evaluated at ImmunoSearch in at least two independent experiments (Table 5). Cooper statistics values (Sensitivity, specificity and accuracy) (Cooper et al., 1979) were calculated for the SENS-IS assay using human and LLNA data from the literature (Table 6) as references. A 2 × 2 contingency table using sensitizer vs. non-sensitizers data collected from literature (Table 7) was build using data obtained with the SENS-IS assay. Specificity, sensitivity and accuracy were calculated according to Cooper et al. (1979).

3. Results

3.1. Transfer of the SENS-IS protocol to the participating laboratories

3.1.1. Transferability phase I — training in the lead laboratory

The phase I training procedure was organized as described in the Methods section. The results are summarized in the Supplementary data Table 1. During the first experiment performed in ImmunoSearch facilities, both teams correctly detected <7 over-expressed genes in either the "REDOX" and SENS-IS" groups for the non-sensitizers (propylene glycol (PG) and DMSO) and ≥7 over-expressed genes in at least one of the groups of genes for all sensitizers. In some instances, (DNCB or cinnamal tested at 50%), the level of tissue destruction was too high for a proper analysis. The success criterion (correct classification of the test chemicals) was met by both teams and they could proceed to Transferability phase II.

3.1.2. Transferability phase II — transfer of the assay to the participating laboratories

The same 8 chemicals (same batches with a different coding scheme) and the positive and negative controls (i.e. SLS, HCA, TNBS and DMSO, not coded) were distributed to the participating laboratories and tested in their own facilities. The ANSM laboratory obtained a correct classification for 14 out of the 16 tested samples (8 chemicals at 2 different concentrations, see Supplementary data Table 1). For propylene glycol at 50%, the discrepancy was due to poor washing of the skin. The washing procedure was thus adapted, the test item re-analyzed and a correct result obtained. Another discrepancy between the two experiments was observed with propyl gallate at 10%. This chemical was included in the test list because, depending on the batch of Episkin skin model, it can lead to controversial results when tested at 10% (see below). (This skin model variability is taken into account in the SENS-IS test protocol since it is required that results must be confirmed on at least two different batches of Episkin). ANSM was considered to have met the success criterion and could proceed to phase III.

The IPMC laboratory obtained the correct classification for the 16 tested samples. Some minor variability was nevertheless observed. For example geraniol induced heavy skin damages at a 50% concentration in one experiment and not the other. This may be due to variability in the susceptibility of the Episkin batches to corrosion. IPMC was also considered to have met the success criterion and could proceed to phase III.

3.2. WLR analysis in the three participating laboratories using triplicates in separate experiments

We used the data obtained from the 19 blindly tested chemicals of the ring study (see Materials and methods Section, "Transferability phase III" and "Statistical considerations") to calculate the WLR in the three participating laboratories (Table 1). We analyzed the concordance of the PM conclusions at a given concentration between the three independent runs. The WLR for ImmunoSearch, IMPC, and ANSM for the PM conclusion at a given concentration was (46/49) 93.9%, (45/46) 97.8% and (46/50) 92% respectively. The average WLR for the concordance of the PM conclusions at a given concentration from the three laboratories was 94.5%, very close to the target criteria of 95% concordance for the considered sample size (>45).

The WLR was further evaluated using the Fleiss kappa statistic to measure the inter-rater agreement. For ImmunoSearch, the obtained kappa statistic was 0.887, with a p value <0.0001 for IPMC, the kappa value was 0.94 with a p value <0.0001 and for ANSM the kappa statistic obtained was 0.891 with a p value <0.0001. For the three labs the kappa value obtained corresponds to an almost perfect agreement between the results.

When 5 reactivity classes were considered (Extreme, Strong, Moderate, Weak and NS, see Table 2), the WLR was (16/19) 84.2% for ImmunoSearch, 100% for IPMC and (17/19) 89.5% for ANSM with an average WLR of 91.3% for the three laboratories. It is interesting to note that in case of differing classification, the difference was never greater than one class.

3.3. Between-laboratory reproducibility (BLR) evaluation

The main focus of the BLR study for the 19 chemicals was on the concordance of the predictions as sensitizers (S) versus non-sensitizers (NS). As shown in Table 3, the BLR between the three laboratories was 100%, which meets the reproducibility criterion that was set at 95% concordance.

3.4. BLR evaluation of the predictive capacity (PC)

A further analysis of the potency category results (5 classes, see Methods section) was also performed. 14 out of the 16 tested sensitizers were classified in the same category in all laboratories (See Table 4). Discrepant results were observed for propyl gallate that was classified as a strong sensitizer in one laboratory (ImmunoSearch) and as a moderate one in the two other laboratories and for hexyl cinnamal that was classified as a strong sensitizer in one laboratory (ANSM) and as a weak one in the other two laboratories. In summary, the 3 laboratories obtained identical classification for 17 out of 19 chemicals (including the 3 non-sensitizers) resulting in a BLR of 89.5%. The BLR for pair-wise comparison

Table 2
WLR analysis based on the potency category predictions obtained in three laboratories.

	ImmunoSearch				IPMC				ANSM			
Chemical	Repro1	Repro2	Repro3	WLR 5class	Repro1	Repro2	Repro3	WLR 5class	Repro1	Repro2	Repro3	WLR 5class
PG	NS	NS	NS	Y	NS	NS	NS	Y	NS	NS	NS	Y
DMSO	NS	NS	NS	Y	NS	NS	NS	Y	NS	NS	NS	Y
DNCB	Extreme	Extreme	Extreme	Y	Extreme	Extreme	Extreme	Y	Extreme	Extreme	Extreme	Y
Propyl gallate	Strong	Strong	Strong	Y	Moderate	Moderate	Moderate	Y	Weak	Moderate	Moderate	N
Geraniol	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y
Resorcinol	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y
Cinnamal	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y
O-aminophenol	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y
TNBS	Extreme	Extreme	Extreme	Y	Extreme	Extreme	Extreme	Y	Extreme	Extreme	Extreme	Y
HCA	Weak	Weak	Moderate	N	Weak	Weak	Weak	Y	Strong	Strong	Moderate	N
L-limonene	Weak	Weak	Weak	Y	Weak	Weak	Weak	Y	Weak	Weak	Weak	Y
MCI	Strong	Extreme	Strong	N	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y
Trans-anethole	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y
PDIC	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y
BBMG	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y	Strong	Strong	Strong	Y
Eugenol	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y
diethyl sulphate	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y	Moderate	Moderate	Moderate	Y
SLS	NS	NS	NS	Y	NS	NS	NS	Y	NS	NS	NS	Y
Polysorbate 21	Weak	Moderate	Weak	N	Weak	Weak	Weak	Y	Weak	Weak	Weak	Y

The chemicals were tested in the indicated laboratories in triplicates (Repros1–3) and classified as extreme, strong, moderate, weak or non sensitizer (NS) as described under Material and methods Section. A "Y" in the WLR column indicates an identical category prediction for all replicates. An "N" indicates differences in the prediction.

was 88.9% (16/18) for ImmunoSearch/ANSM, 94.4% (17/18) for ANSM/IPMC and 94.7% (18/19) for ImmunoSearch/IPMC (Table 4).

When comparing the results obtained with the SENS-IS assay based on majority vote with the results obtained with the LLNA, again only 2 out of 16 tested sensitizers were not rated equally. These were eugenol, (Moderate in the SENS-IS assay vs Weak in the LLNA) and propyl gallate (Moderate in the SENS-IS assay vs Strong in the LLNA, see Table 4). Moreover, when comparing the 5 categories classification obtained by majority vote with the SENS-IS assay to the current LLNA classification, results differed by max. 1 potency category (Table 4).

3.5. Additional analysis: BLR analysis of the number of over-expressed genes in the "SENS-IS" and "REDOX" subgroups

As explained in the Materials and methods Section, the PM of the SENS-IS assay is based on a cut-off number of over-expressed genes. To verify the reproducibility of the number or over-expressed genes in the "SENS-IS" and "REDOX" subgroups, we performed a two-sample Wilcoxon rank test on each concentration of chemicals performed at IS, IPMC and ANSM (Fig. 1). The analysis of the BLR of the measurement of the number of over-expressed REDOX genes shows no significant differences for 33 out of 46 determinations. Similarly, analysis of the BLR measurement of SENS-IS genes over-expressed also shows no significant differences for 33 out of 46 determinations. For 14 out of 26 determinations, (gray stars, Fig. 1) the classification S/N remained unchanged, even though the number of detected genes was different. For the remaining 12 determinations (black stars, Fig. 1) when the determination in SENS-IS or REDOX gene numbers were different, it was always matched by an unambiguous determination in the number of the reciprocal group of genes showing the importance of both gene panel determination (except for propyl gallate (IS04) 10%, HCA (IS10) 10% and eugenol (IS16) 10%). Moreover, an overall Fleiss Kappa value of 0.881, with a p value of 4.26e-13 was obtained.

Table 3
BLR analysis based on hazard predictions obtained in three laboratories.

Name	LLNA classi-fication 2 classes	Lab1 (ImmunoSearch)	Lab2 (ANSM)	Lab3 (IPMC)	Agreement 2 classes
2,4-Dinitrochlorobenzene	S	S	S	S	Y
2,4,6-Trinitrobenzenesulfonic acid	S	S	S	S	Y
Cinnamal	S	S	S	S	Y
Resorsinol	S	S	S	S	Y
2-Bromo-2-(bromomethyl) glutaronitrile	S	S	S	S	Y
Eugenol	S	S	S	S	Y
Hexyl cinnamal	S	S	S	S	Y
Propyl gallate	S	S	S	S	Y
Geraniol	S	S	S	S	Y
o-Aminophenol	S	S	S	S	Y
L-Limonene	S	S	S	S	Y
Methylchloroisothiazolinone solution 1.2%	S	S	S	S	Y
Trans-anethole	S	S	S	S	Y
Potassium dichromate	S	S	S	S	Y
Diethyl sulphate	S	S	S	S	Y
Polysorbate 21	S	S	S	S	Y
Propylene glycol	NS	NS	NS	NS	Y
Dimethyl sulfoxide	NS	NS	NS	NS	Y
Sodium lauryl sulfate*	NS	NS	NS	NS	Y

The chemicals were tested in the indicated laboratories and classified as sensitizers "S" or non-sensitizers "NS" using the SENS-IS assay. The "Labs1–3" columns show the classification obtained in each laboratory. A "Y" in the "Agreement 2 classes" column indicates an identical classification in all laboratories. The LLNA classification of the chemicals ("S" or "NS") is indicated in the "LLNA classification 2 classes" column.

* Sodium lauryl sulfate is ranked as a weak to moderate sensitizer by the LLNA. However since it is a false positive we considered it as a non sensitizer.

Table 4
Evaluation of the concordance in the predicted reactivity class between the laboratories.

Name	LLNA classification 5 classes	Lab1 (IS)	Lab2 (ANSM)	Lab3 (IPMC)	Agreement 5 classes	Majority vote	Agreement with Ref 5 classes
2,4 Dinitrochlorobenzene	EXT	EXT	EXT	EXT	Y	EXT	Y
2,4,6-Trinitrobenzenesulfonic acid	EXT	EXT	EXT	EXT	Y	EXT	Y
Cinnamal	STRONG	STRONG	STRONG	STRONG	Y	STRONG	Y
Resorsinol	MOD	MOD	MOD	MOD	Y	MOD	Y
2-Bromo-2-(bromomethyl) Glutaronitrile	STRONG	STRONG	STRONG	STRONG	Y	STRONG	Y
Eugenol	WEAK	MOD	MOD	MOD	Y	MOD	N
Hexyl cinnamal	WEAK	WEAK	STRONG	WEAK	N	WEAK	Y
Propyl gallate	STRONG	STRONG	MOD	MOD	N	MOD	N
Geraniol	MOD	MOD	MOD	MOD	Y	MOD	Y
O-aminophenol	STRONG	STRONG	STRONG	STRONG	Y	STRONG	Y
L-Limonene	WEAK	WEAK	WEAK	WEAK	Y	WEAK	Y
Methylchloroisothiazolinone solution 1.2%	STRONG	STRONG	STRONG	STRONG	Y	STRONG	Y
Trans-anethole	MOD	MOD	MOD	MOD	Y	MOD	Y
Potassium dichromate	STRONG	STRONG	STRONG	STRONG	Y	STRONG	Y
Diethyl sulphate	MOD	MOD	MOD	MOD	Y	MOD	Y
Polysorbate 21	WEAK	WEAK	WEAK	WEAK	Y	WEAK	Y
Propylene glycol	NS	NS	NS	NS	Y	NS	Y
Dimethyl sulfoxide	NS	NS	NS	NS	Y	NS	Y
Sodium lauryl sulfate*	NS	NS	NS	NS	Y	NS	Y

19 blind-coded chemicals (16 sensitizers and 3 non-sensitizers) were independently evaluated using the SENS-IS assay in 3 laboratories (ImmunoSearch, ANSM and IPMC). The "LLNA classification 5 classes" column shows the LLNA classification of the tested chemical. The "Labs1–3" columns indicate the results obtained with the SENS-IS assay in the indicated laboratory ("EXT" (extreme); "STRONG"; "MOD" (moderate); "WEAK" or "NON" (non-sensitizer)). The "Agreement 5 classes" column indicates an agreement ("Y") or a disagreement ("N") between laboratories in predicting the sensitizer category. The "Majority vote" column indicates the classification obtained by a majority vote approach using the results from the three laboratories. The "Agreement with Ref 5 classes" column indicates the agreement ("Y") or the disagreement ("N") of the majority vote result with the LLNA classification of the corresponding chemical.
* Sodium lauryl sulfate is ranked as a weak to moderate sensitizer by the LLNA. However since it is a false positive we considered it as a non sensitizer.

3.6. Predictivity of the SENS-IS assay according to Cooper statistics

The predictive capacity of the SENS-IS assay was measured using a test set of 150 chemicals that were evaluated at ImmunoSearch in at least two independent experiments. The SENS-IS assay correctly predicted the status (sensitizer versus non-sensitizer), of 145 out of the 150 tested items when compared to the corresponding LLNA data (Table 5). The calculated Cooper statistics values (see Table 6) were: Sensitivity: 97.7%; specificity: 95.2% and overall accuracy: 96.6% (compared to LLNA data). Of the five misclassified molecules, two were false negatives (imidazolinyl urea and isopropyl myristate) and three false positives (dimethylformamide, isopropyl alcohol and benzaldehyde).

When compared to human data that are available for 130 of the 150 tested chemicals (see Table 5), the SENS-IS assay correctly predicted the status (sensitizer versus non-sensitizer) of 125 out of the 130 test items. Cooper statistic values were (Table 6): Sensitivity: 95.8%; specificity: 96.5% and overall accuracy: 96%. Three substances were misclassified as false negatives: isopropyl miristate, benzocaine and hexyl salicylate. However the last two are also rated as negative by the LLNA assay. Two substances were incorrectly rated as sensitizers (false positives): benzaldehyde and isopropyl alcohol.

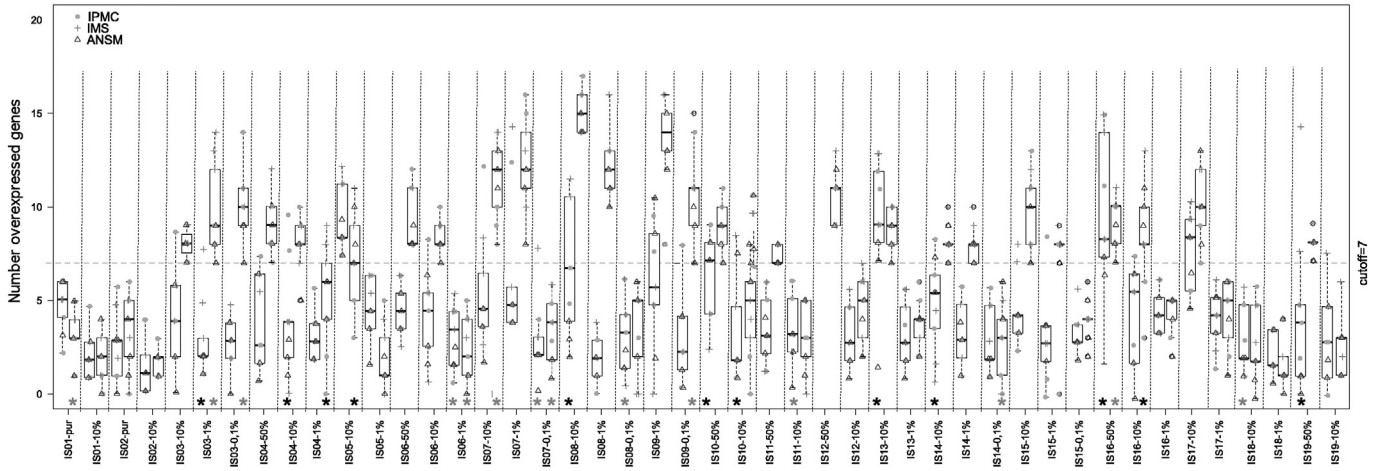

Fig. 1. Statistical analysis of the variability in the measured number of over-expressed genes in the "SENS-IS" and "REDOX" groups of genes. The number of over-expressed gens (>1.25 × control vehicle value) measured in each laboratories (dot: IMPC, cross: ImmunoSearch, triangle: ANSM) plotted as box plots. Median is represented by the horizontal bar, upper and lower quartile by the upper and lower values of the boxes and values outside the interquartile as individual values. Each chemical (coded IS01 to IS19 see below) was tested in each laboratory at the indicated concentration in 3 distinct experiments on different batches of EPISKIN. For each concentration the number of over-expressed genes in the "SENS-IS" group (left box) or "REDOX" group (right box) was analyzed. A two sample Wilcoxon rank test was performed on each concentration tested. Significant differences within each tested concentration (p value <0.05) are indicated by a star. A gray star indicates a significant difference observed that did not change the PM conclusion (number of overexpressed genes above or below 7). A black star indicates a significant difference that could modify the PM conclusion. The coded chemicals were: IS01: propylene glycol, IS02: dimethyl sulfoxide (DMSO), IS03: 2,4-dinitrochlorobenzene (DNCB), IS04: propyl gallate, IS05: geraniol, IS06: resorcinol, IS07: cinnamal, IS08: o-aminophenol, IS09: 2,4,6-trinitrobenzenesulfonic acid (TNBS), IS10: hexyl cinnamal, IS11: l-limonene, IS12: 5-chloro-2-methylisothiazol-3-one (MCI) at 1.5%, IS13: trans-anethole, IS14: potassium dichromate, IS15: 2-bromo-2-(bromomethyl) glutaronitrile, IS16: eugenol, IS17: diethyl sulfate, IS18: sodium lauryl sulfate (SLS), IS19: polysorbate 21.

As the SENS-IS PM classifies sensitizers in 5 categories based on the lowest concentration able to induce a positive response (See "PM evaluation" in the Materials and methods Section), we analyzed the concordance of prediction according to the 5 classes of the LLNA assay or the 6 human classes recently proposed by Basketter et al. (2014) (Table 6). When compared to the LLNA data (150 test substances) the Cooper statistic values for overall accuracy was 92.6%, and 90.6% when compared to human data (64 substances, see Table 7).

Table 5
Analysis of the predictive capacity of the SENS-IS assay using 150 chemicals.

Name	CAS	Vehicle used	Physical form	LLNA classification	Human classification	SENS-IS classification
2,4 Dinitrochlorobenzene*	97-00-7	DMSO	SOL	EXT	1[a]	EXT
2,4,6-Trinitrobenzenesulfonic acid	2508-19-2	water	SOL	EXT	No data	EXT
4-Nitrobenzylbromide*	100-11-8	DMSO	SOL	EXT	No data	EXT
Oxazolone*	15646-46-5	OO	SOL	EXT	+[b]	EXT
Ammonium tetrachloroplatinate	13820-41-2	OO	SOL	EXT	No data	EXT
Diphenylcyclopropenone	886-38-4	OO	SOL	EXT	1[a]	EXT
Benzoylacetone	93-91-4	OO	LIQ	EXT	No data	EXT
p-Benzoquinone+	106-51-4	water	SOL	EXT	+[b]	EXT
1-Fluoro-2,4-dinitrobenzene	70-34-8	DMSO	SOL	EXT	No data	EXT
Propyl gallate	121-79-9	DMSO	SOL	STRONG	2[a]	STRONG
O-aminophenol	95-55-6	OO	SOL	STRONG	2[a]	STRONG
P-aminophenol	123-30-8	DMSO	SOL	STRONG	+	STRONG
Methylchloroisothiazolinone+	26172-55-4	Water	LIQ	STRONG	1[a]	STRONG
Potassium dichromate	7778-50-9	Water	SOL	STRONG	1[a]	STRONG
Benzoyl peroxide	94-36-0	DMSO	SOL	STRONG	3[a]	STRONG
2-Bromo-2-(bromomethyl) glutaronitrile*	35691-65-7	DMSO	SOL	STRONG	2[a]	STRONG
Chloroatranol	57074-21-2	OO	SOL	STRONG	1[a]	STRONG
Formaldehyde+	50-00-0	Water	LIQ	STRONG	2[a]	STRONG
Glutaral*	111-30-8	Water	LIQ	STRONG	2[a]	STRONG
Hydroquinone	123-31-9	Water	SOL	STRONG	3[a]	STRONG
Dodecyl gallate	1166-52-5	OO	SOL	STRONG	2[a]	MOD
Phthalic anhydride	85-44-9	Water	SOL	STRONG	+	STRONG
P-phenylenediamine*	106-50-3	water	SOL	STRONG	1[a]	STRONG
Methyl-2-octynoate+	111-12-6	DMSO	LIQ	STRONG	2[a]	STRONG
4-Methylaminophenol+	55-55-0	DMSO	SOL	STRONG	3[a]	STRONG
Cinnamal*,+	104-55-2	OO	LIQ	STRONG	2[a]	STRONG
Thioglycerin+	96-27-5	PBS	LIQ	MOD	2[a]	MOD
2-Methoxy-p-creosol	93-51-6	OO	LIQ	MOD	No data	MOD
1,2-Dichloro-4-nitro-benzene+	611-06-3	DMSO	SOL	MOD	No data	STRONG
Dimethylaminopropylamine	109-55-7	DMSO	LIQ	MOD	2[a]	MOD
Geraniol	106-24-1	OO	LIQ	MOD	4[a]	MOD
Trans-anethole	4180-23-8	DMSO	LIQ	MOD	+	MOD
Diethyl sulphate	64-67-5	OO	LIQ	MOD	+	MOD
Disulfiram*	137-26-8	DMSO	SOL	MOD	3[a]	MOD
Diphenylmethane-4,4′-diisocyanate*,+	101-68-8	water	SOL	MOD	+	MOD
Glyoxal*	107-22-2	water	LIQ	MOD	2[a]	MOD
2-Mercaptobenzothiazole*	149-30-4	DMSO	SOL	MOD	3[a]	MOD
2,4-Hexadienal	142-83-6	OO	LIQ	MOD	+	MOD
Benzisothiazolinone+	2634-33-5	water	SOL	MOD	+	MOD
Isoeugenol*	97-54-1	OO	LIQ	MOD	2[a]	MOD
Sodium metabisulfite	7681-57-4	Water	SOL	MOD	+	MOD
2-Hydroxyethyl acrylate	818-61-1	Water	LIQ	MOD	+	MOD
Citral+	5392-40-5	OO	LIQ	MOD	3[a]	MOD
3,4-Dihydrocoumarin	119–84-6	Water	LIQ	MOD	+[c]	MOD
Ethylenediamine+	107-15-3	Water	LIQ	MOD	3[a]	MOD
Maleic anhydride*+	108-31-6	water	SOL	MOD	+	MOD
Nickel sulphate+	7786-81-4	OO	SOL	MOD	+	MOD
Ethyl-3-hydroxy-3-phenylpropionate	764-85-2	OO	LIQ	MOD	No data	MOD
Phenylacetaldehyde+	122–78-1	OO	LIQ	MOD	+[c]	MOD
Trimellitic anhydride*+	552-30-7	water	SOL	MOD	+	WEAK
Eugenol*	97-53-0	OO	LIQ	WEAK	3[a]	MOD
Imidazolidinyl urea+	39236-46-9	DMSO	SOL	WEAK	3[a]	NON
Glycol dimethacrylate+	97-90-5	OO	LIQ	WEAK	4[a]	WEAK
Butyl glycidyl ether+	2426-08-6	Water	LIQ	WEAK	+[c]	STRONG
Cinnamyl alcohol*	104-54-1	OO	**SOL**	WEAK	3[a]	WEAK
Polysorbate 21	9005-64-5	DMSO	LIQ	WEAK	No data	WEAK
Resorsinol*	108-46-3	water	SOL	WEAK	4[a]	MOD
L-Limonene	5989-54-8	DMSO	LIQ	WEAK	5[a]	WEAK
Hexyl cinnamal+	101-86-0	OO	LIQ	WEAK	5[a]	WEAK
2,2,6,6-Tetramethyl-3,5-heptanedione	1118-71-4	OO	LIQ	WEAK	No data	WEAK
Abietic acid+	514-10-3	DMSO	SOL	WEAK	3[a]	WEAK
Amyl cinnamal	122-40-7	OO	LIQ	WEAK	4[a]	WEAK
Benzyl benzoate	120-51-4	OO	LIQ	WEAK	5[a]	WEAK
Diethyl maleate	141-05-9	DMSO	LIQ	WEAK	2[a]	WEAK
Farnesal	19317-11-4	DMSO	LIQ	WEAK	No data	WEAK
Hydroxycitronellal+	107-75-5	OO	LIQ	WEAK	4[a]	WEAK

(continued on next page)

Table 5 (*continued*)

Name	CAS	Vehicle used	Physical form	LLNA classification	Human classification	SENS-IS classification
Butylphenyl methylpropional	80-54-6	OO	LIQ	WEAK	4[a]	WEAK
Oxalic acid	144-62-7	water	SOL	WEAK	No data	WEAK
Phenyl benzoate[+]	93-99-2	DMSO	SOL	WEAK	3[a]	WEAK
Pyridine	110-86-1	water	LIQ	WEAK	5[a]	WEAK
Salicylaldehyde	90-02-8	OO	LIQ	WEAK	+	WEAK
Alpha-terpineol	98-55-5	OO	LIQ	WEAK	+	WEAK
2-Hydroxypropyl methacrylate	923-26-2	OO	LIQ /	WEAK	No data	WEAK
Cyclamen aldehyde	103-95-7	OO	LIQ	WEAK	No data	WEAK
Benzyl salicylate[+]	118-58-1	OO	LIQ	WEAK	5[a]	WEAK
Linalool	78-70-6	OO	LIQ	WEAK	4[a]	WEAK
Methylparaben	99-76-3	OO	SOL	WEAK	+[d]	WEAK
Eucalyptol	470-82-6	DMSO	LIQ	WEAK	No data	WEAK
6-Methylcoumarin	92-48-8	DMSO	SOL	WEAK	+[b]	WEAK
Chlorhexidine	55-56-1	DMSO	SOL	WEAK	4[a]	WEAK
Diacetyl	431-03-8	water	LIQ	WEAK	No data	WEAK
9-Decen-2-one	35194-30-0	OO	LIQ	WEAK	No data	WEAK
Ammonium hexachloroplatinate[*,+]	16919-58-7	water	SOL	WEAK	No data	WEAK
Hexamethylenediisocyanate[*,+]	822-06-0	OO	LIQ	WEAK	+	WEAK
Isopropyl myristate	110-27-0	OO	LIQ	WEAK	5[a]	NON
Citronellol	106-22-9	DMSO	LIQ	WEAK	5[a]	WEAK
Methyl salicylate[+]	119-36-8	OO	LIQ	WEAK	5[a]	WEAK
Isopropyl alcohol[+]	67-63-0	Water	LIQ	NON	5[a]	WEAK
Dimethylformamide	68-12-2	DMSO	LIQ	NON	+[e]	WEAK
Propylparaben	94-13-3	DMSO	SOL	NON	5[a]	NON
Sodium lauryl sulfate[*]	151-21-3	water	SOL	WEAK (FP)	6[a]	NON
Salicylic acid[*]	69-72-7	DMSO	SOL	NON	6[a]	NON
Phenol[*]	108-95-2	water	SOL	NON	6[a]	NON
Sodium benzoate	532-32-1	DMSO	SOL	NON	−[f]	NON
Glycerin[*]	56-81-5	water	LIQ	NON	6[a]	NON
Lactic acid[*]	50-21-5	water	LIQ	NON	6[a]	NON
Triethanolamine	102-71-6	DMSO	LIQ	NON	5[a]	NON
Hexane	110-54-3	OO	LIQ	NON	6[a]	NON
Chlorobenzene[*]	108-90-7	DMSO	LIQ	NON	+	NON
Propylene glycol[+]	57-55-6	Water	LIQ	NON	5[a]	NON
Dimethyl sulfoxide	67-68-5	Water	LIQ	NON	6[a]	NON
Benzalkonium chloride	8001-54-5	Water	SOL	NON	−	NON
Xylene	1330-20-7	OO	LIQ	NON	−	NON
3-Chloronitrobenzene	121-73-3	DMSO	SOL	NON	−	NON
Dipropyleneglycol monobutylether	29911-28-2	OO	LIQ	NON	−	NON
Erucamide	112-84-5	DMSO	SOL	NON	−	NON
Heptyl butyrate	5870-93-9	OO	LIQ	NON	−	NON
Hexyl salicylate	6259-76-3	OO	LIQ	NON	4[a]	NON
Lauric acid	143-07-7	DMSO	SOL	NON	−	NON
Methyl palmitate	112-39-0	OO	SOL	NON	−	NON
Sodium bicarbonate	144-55-8	OO	SOL	NON	−	NON
Sodium bisulfite	7631-90-5	Water	SOL	NON	−	NON
Methoxyaminophenone	100-06-1	Water	SOL	NON	−[g]	NON
Sulfanilic acid[+]	121-57-3	DMSO	LIQ	NON	−	NON
Vanillin[+]	121-33-5	Water	SOL	NON	5[a]	NON
N-butyl-alcohol[+]	71-36-3	Water	LIQ	NON	6[a]	NON
Aluminum hydroxide	21645-51-2	DMSO	SOL	NON	−	NON
Benzocaine[+]	94-09-7	DMSO	LIQ	NON	4[a]	NON
Cetrimonium bromide	57-09-0	Water	SOL	NON	5[a]	NON
1-Bromo-4-chlorobutane	6940-78-9	OO	LIQ	NON	−	NON
Glyceraldehyde	56-82-6	Water	SOL	NON	−	NON
1-Bromopentane	110-53-2	OO	SOL	NON	−	NON
1-Methyl-3-phenylpiperazine	5271-27-2	Water	LIQ	NON	−	NON
4-Methylthiobenzaldehyde	3446-89-7	DMSO	LIQ	NON	−	NON
Di-n-propyl disulfide	629-19-6	OO	LIQ	NON	−	NON
Heptanal	111-71-7	OO	SOL	NON	−	NON
Methyl laurate	111-82-0	OO	LIQ	NON	−	NON
Potassium hydroxide	1310-58-3	Water	SOL	NON	−	NON
A-terpinyl acetate	80-26-2	OO	LIQ	NON	−	NON
Tetrachloroethylene	127-18-4	OO	LIQ	NON	−	NON
Decyl alcohol	112-30-1	Water	LIQ	NON	−	NON
Benzenethiol, 5-(1,1-Dimethylethyl)-2-methyl	7340-90-1	OO	LIQ	NON	−	NON
Butyl methacrylate	97-88-1	OO	LIQ	NON	−	NON
Tri-isobutyl phosphate	126-71-6	OO	LIQ	NON	−	NON
2-Chloromethyl-3,5-dimethyl-4-methoxypyridine hydrochloride	86604-75-3	Water	SOL	NON	−	NON
2-Ethoxyethyl methacrylate	2370-63-0	OO	LIQ	NON	−	NON
Allyl heptanoate	142-19-8	OO	LIQ	NON	−	NON
Ascorbic acid	50-81-7	Water	SOL	NON	6[a]	NON
Bromobutane	109-65-9	DMSO	SOL	NON	−	NON
Cyclopentasiloxane	541-02-6	OO	LIQ	NON	−	NON
Benzoic acid[+]	65-85-0	DMSO	LIQ	NON	−	NON
Isoleucine	73-32-5	DMSO	SOL	NON	−	NON

Table 5 (*continued*)

Name	CAS	Vehicle used	Physical form	LLNA classification	Human classification	SENS-IS classification
Polyethylene glycol	2532-68-3	Water	LIQ	NON	—	NON
4-Aminobenzoic acid	150-13-0	DMSO	SOL	NON	—[b]	NON
Fumaric acid	110-17-8	water	SOL	NON	—[h]	NON
Benzaldehyde*	100-52-7	water	LIQ	NON	5[a]	MOD
Diethyl phthalate*	84-66-2	DMSO	LIQ	NON	6[a]	NON
Caprylic acid*	124-07-2	OO	LIQ	NON	6[a]	NON
4-Hydroxybenzoic acid*	99-96-7	DMSO	SOL	NON	—	NON
Tartaric acid	87-69-4	DMSO	SOL	NON	—	NON

150 chemicals covering the whole range of sensitizing potency were evaluated at ImmunoSearch in at least two independent experiments. The name, CAS Nr., vehicle used for the test, their physical form, LLNA and human classification are indicated in the first six columns. The last column (SENS-IS classification) shows the classification predicted by the SENS-IS assay ("EXT" (extreme); "STRONG"; "MOD" (moderate); "WEAK" or "NON" (non-sensitizer)).

* Chemicals present in the Sens-it-iv list (Sens-it-iv, 2009).

\+ Chemicals present in the silver list (Emter et al., 2010).

[a] Basketter et al. (2014).

[b] Basketter et al. (1999).

[c] Schneider and Akkan (2004).

[d] Andersen et al. (2008).

[e] Lepoittevin and Coz (2007).

[f] SCCP (2005).

[g] Gerberick et al. (2001).

[h] Hansson and Thörneby-Andersson (2003).

3.7. Repeatability analysis of the SENS-IS assay

In order to measure the robustness of the assay over experiments, repeatability was measured using the results obtained with the four internal controls over 30 experiments. It was considered as excellent with an interquartile range (Q3 – Q1) of less than 5 between experiments (Fig. 2). Moreover, the PM always gave the correct sensitizer versus non-sensitizer status, i.e. the number of over-expressed genes in the SENS-IS or ARE groups were <7 for the two non-sensitizers (SLS and DMSO) whereas it was always >7 (in the REDOX group) for the two sensitizers (HCA and TNBS) (Fig. 2).

4. Discussion

The SENS-IS test protocol makes use of a 3D reconstructed human skin model (Episkin) as the test system and measures the expression of a large panel of genes relevant to the sensitization process as endpoints. This combination attempts to reproduce the human skin situation and aims to deliver a detailed analysis of the skin response to the stress induced by the exposure to a test chemical or a mixture of ingredients. To evaluate both hazard and potency, 4 dilutions (50%, 10%; 1% and 0.1%) of the test chemical are applied on the skin model in an appropriate vehicle. The lowest concentration that eventually induces a positive response is used to classify the chemical as a weak, moderate, strong or extreme sensitizer. In a stepwise approach, the highest concentration is first analyzed and if found negative, the next lower concentration is evaluated. Positive or negative responses are always analyzed on two separate epidermis batches and in case of discrepant results, on a third batch (see Methods section). The sensitizing potency of a chemical is thus predicted according to a threshold value (the minimum test concentration necessary to induce the over-expression of a given number of genes in two groups of genes). This is very similar to the threshold approaches used for standard animal models such as the LLNA.

The SENS-IS assay requires the measurement of multiple gene expression pathways in order to declare a test item as a sensitizer or a non sensitizer. This approach, although complex, yields more precise and detailed results than single biomarker approaches. Albrekt et al. (2014) made similar observations when investigating the mechanisms of action of skin sensitizers. However, their approach for predicting

Table 6
Predictivity of SENS-IS assay according to Cooper statistics.

In comparison to		
	Human	LLNA
n	130	150
Sensitivity	95.8%	97.7%
Specificity	96.5%	95.2%
PPV	97%	96.6%
NPV	95%	96.7%
Accuracy	96%	96.6%

150 test chemicals were evaluated at ImmunoSearch in at least two independent experiments. Cooper statistics values (Sensitivity, Specificity and Accuracy) were calculated for the SENS-IS assay using human (see human column) and LLNA (see LLNA column) data from the literature as references.
"n": Number of results included in the calculation (depending on available reference data); "PPV »: Positive Prediction Value; ≪NPV≫: Negative Prediction Value.

Table 7
Concordance of prediction according to the 5 classes of the LLNA assay or the 6 classes for human data.

LLNA Class	Extreme	Strong	Moderate	Weak	NS	Total
Extreme	9					9
Strong		15	1	1		17
Moderate		1	23	2	1	27
Weak			1	33	2	36
NS				2	59	61
Total	9	16	25	38	62	150
	Concordance	92.66%				

Human class	#1	#2	#3	#4	#5	#6	Total
Extreme	2						2
Strong	4	7	2				13
Moderate		5	4	2	1		12
Weak		1	4	5	7		17
NS				2	7	11	20
Total	6	13	10	9	15	11	64
	concordance	90,62					

Comparison between data obtained with the SENS-IS assay and LLNA was performed based on exact concordance in potency class prediction. The EC3 value obtained with the LLNA and the MID (Minimal Inducing Dose) value obtained with the SENS-IS assay were indentical to classify chemicals' potency i.e.: 0.1 > EC3 or MID: Extreme; 0.1 < EC3 or MID ≤ 1: Strong; 1 < EC3 or MID ≤ 10: Moderate; 10 < EC3 or MID: Weak.
Since the human data obtained from Basketter et al. (2014) are split into 6 classes vs 5 for the SENS-IS assay comparison was based on the assumption that human cat 1 can be classified as Extreme or Strong, human cat 2 can be classified as Strong or Moderate, human cat 3 and 4 can be classified as Moderate or Weak, human cat 5 can be classified as Weak or NS.

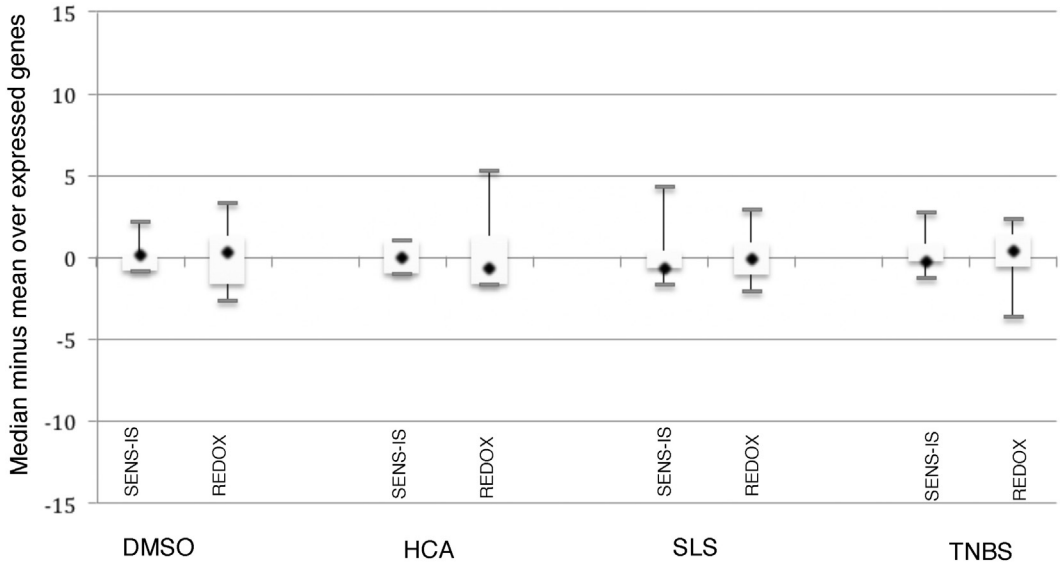

Fig. 2. Repeatability analysis of the SENS-IS assay. Box-plot representation of the number of over-expressed genes in the "SENS-IS" and "REDOX" group of genes. The number of over-expressed genes induced by the four indicated chemicals (DMSO, (100%), HCA (50% *v*/*v* in DMSO), SLS (5% *w*/*v* in PBS) and TNBS (1% *w*/*v* in PBS) was measured in 30 consecutive independent experiments. Black dots indicate the median of the 30 measurements. Boxes upper and lower limits indicate the upper and lower quartiles and the whiskers show the highest and lowest measured values.

the sensitizing potency of chemicals is based on the activation of key cellular events. This is different from the threshold approach (see above) used in the SENS-IS protocol.

The SENS-IS assay has already been shown to deliver high quality data on the hazard and potency characteristics of a wide variety of sensitizers (Cottrez et al., 2015).

In the present paper, we present the results of a comprehensive evaluation of the SENS-IS protocol using a large set of 150 chemicals. Furthermore, the transferability, robustness and reproducibility of the assay were evaluated through a ring-study within three laboratories using a test set of 19 reference substances.

Due to the multi-parametric level of complexity intrinsic to the SENS-IS assay, a very careful analysis of the possible sources of variability was performed before transferring the assay to other laboratories. This included the inherent variability of the reconstructed epidermis and the relative complexity of the RT-PCR procedure. Multiple internal standards were used (3 non-sensitizers also used as solvent, one irritant and 2 sensitizers) and further control procedures for tissue damage, cell lysis and irritation level ensured that the intrinsic variability of the experimental procedure could be well controlled. Moreover, we developed a step-wise approach for the training of the recipient laboratories in order to monitor all the possible mistakes.

The assay was successfully transferred to two laboratories. The expected test results were obtained for most test substances (see Supplementary data Table 1). The discrepancies observed between the obtained and expected results could be easily explained (washing procedure for propylene glycol in one of the recipient lab or skin model batch variability for propyl gallate) and resolved. The two laboratories were then considered to have met the transferability success criterion as described in the "Transferability phase I–II" of the Materials and methods Section.

Before proceeding to the blind testing phase (Transferability phase III), an independent biostatistician calculated the number of chemicals that would be required to properly evaluate the performance of the test in light of the primary study goals, i.e. to perform a statistically sound evaluation of the WLR and of the BLR. As explained in the Method section, it was calculated that 19 chemicals are required, based on evaluation of historical data and a 95% target reproducibility value. In order to cover the full range of sensitization potencies and to allow the evaluation of the reproducibility of hazard and also of potency predictions, we designed a chemical test set using evenly distributed chemicals covering the 5 LLNA classes (2 extremes, 6 strongs, 4 moderates, 4 weaks and 3 non-sensitizers).

The WLR was analyzed using the dataset generated with the 19 chemicals selected for the blind ring study. Each chemical was tested in triplicate in all three laboratories (see Table 1). The resulting sample size thus corresponded to 45 determinations at different concentrations (Table 1), allowing reaching the desired statistical confidence (as described in the Method section). The ratio of similar results obtained for a given concentration was 93.9% for ImmunoSearch, 97.8% for IPMC and 92% for ANSM. The discrepancies were generally observed for the same chemical in all three laboratories. Propyl gallate ends to produce different results with different batches of the Episkin model. This may be due to a high sensitivity toward the quality of the skin barrier resulting in an inconsistent bioavailability (Cosmetic ingredient expert panel, 2007). Similarly, HCA, although used as a positive control at 50%, produced highly variable results at concentrations ≤10% in two out of three laboratories. Variability in the determination of the sensitizing potential of HCA has already been discussed by Basketter et al. (2015). An explanation could be the differing proportion of cis-trans isomers and their respective capacity to form adduct with lysine and cysteine residues or the formation of oxidative products. The other discrepancies were due to chemicals evaluated at their lowest test concentration e.g. polysorbate 21 and MCI in the second experiment at ImmunoSearch. This may be due to the inherent variability of the Episkin model. This variability is taken into account by the PM, which requires two identical determinations at a given concentration. Finally, highly irritant and volatile products, e.g. geraniol and eugenol for IPMC and ANSM respectively were rated negative in one out of three experiments. This could be due to damage to the skin model. Again, skin model integrity is a parameter checked by the PM and the issue was solved through testing at a lower concentration.

We also analyzed the WLR for the determination of the potency prediction based on the five potency classes defined by ECETOC for the LLNA assay (Loveless et al., 2010). As expected (See Table 2), discrepancies in potency determination appeared with the same chemicals, i.e. HCA, propyl gallate, MCI and polysorbate 21. However, the potency classes (SENS-IS prediction versus published data) never differed by more than one class. A mean WLR of 91.3% was calculated.

As confirmed by an evaluation using the Fleiss kappa statistic to measure inter-rater agreement for the 3 laboratories, these results correspond to an almost perfect agreement between the results obtained in the three laboratories, thus confirming the excellent transferability and robustness of the assay.

As the next step, the BLR was first evaluated on hazard determination (sensitizer versus non-sensitizer) using the selected set of 19 chemicals. A perfect BLR of 100% was obtained (Table 3). Moreover, as mentioned above, the assay was designed to cover the 5 LLNA classes thus allowing a BLR evaluation on potency prediction. As shown in Table 4, 17 out of the 19 chemicals were assigned in the same reactivity class by the 3 laboratories. Discrepancies were again (see above) observed for HCA and propyl gallate. These results confirm the good reproducibility and robustness of the SENS-IS assay and showed that it is easily transferable. To further analyze these results, we then evaluated the level of discrepancies in the determination of the number of over-expressed genes in the "SENS-IS" and "REDOX" groups of genes used by the PM. As expected for a multi-step assay, small variations were observed between laboratories (See Fig. 1). However, no significant difference (see Fig. 1) between the three laboratories was observed for the majority (71% in both group of genes) of measurements. Among the 29% measurement showing a significant difference, more than half (34 out of 46, indicated by a gray star in Fig. 1) did not change the PM conclusion. The other 12 measurements with a significant difference in the number of over-expressed genes in either the "SENS-IS" or "REDOX" (indicated by a black star in Fig. 1) could indeed modify the PM conclusion. But for 9 of them, an unambiguous and reproducible determination of the number of over-expressed genes was obtained in the complementary group of genes, emphasizing the importance of using these two gene panels. As expected, the only 3 chemicals inducing significant differences were propyl gallate that was described as difficult to assess when applied topically (Cosmetic ingredient expert panel, 2007), HCA and eugenol for which EC3 determination by LLNA varies from 1.3 to 33 and from 5.4 to 40 respectively (Iccvam documentation, 2013). In conclusion, even though small differences in the number of genes detected are observed between laboratories, the overall results of the assays (given by the number of genes above the specified threshold) are very consistent, as demonstrated by the Fleiss Kappa value of 0.881, with a p value of 4.26^{e-13}!

The predictivity and specificity of the SENS-IS assay were evaluated in at least two independent experiments using a test set comprising 150 chemicals. Cooper statistics (sensitivity, specificity and accuracy) were computed against human and LLNA data from the literature. Against LLNA data (available for all 150 test items), the Cooper values (see Table 2) were: sensitivity: 97.7%; specificity: 95.2% and overall accuracy: 96.6%. Two molecules, imidazolinyl urea and isopropyl myristate were tested false negative. Isopropyl myristate, although weakly positive in the LLNA (EC3 = 44 – Iccvam documentation, 2013), has been classified as a non-sensitizer in guinea pig and human studies (Andersen et al., 2008). Similarly, imidazolinyl urea (EC3 = 24, Gerberick et al. (2005)) was negative in a Magnusson-Kligman assay and is a rare sensitizer in human. Besides, imidazolinyl urea acts through the release of formaldehyde, which in small quantities could be trapped in the corneal layer. Regarding the three false positive molecules (as compared to the LLNA results), dimethylformamide has been described as a sensitizer (Lepoittevin and Coz, 2007) in a compilation of human sensitizers, whereas allergic reaction, although rare, have been described in human for benzaldehyde (Andersen, 2006) and isopropinyl alcohol (García-Gavín et al., 2011).

When compared to the available human data (130 chemicals), the sensitivity, specificity and accuracy values were respectively 95.8%, 96.5% and 96%. The three chemicals misclassified as negatives were isopropyl myristate, benzocaine and hexyl salicylate. Benzocaine, although a proven weak human sensitizer, is an occasional positive in animal tests (Basketter et al., 1995). Further studies on similar "-caine" products should be conducted to better understand the reaction of the SENS-IS assay to similar chemicals. Hexyl salicylate is a very rare sensitizer in both human and animal models (RIFM Expert Panel et al., 2007).

In line with the LLNA data comparison, the same two chemicals (benzaldehyde and isopropinyl alcohol) were misclassified as sensitizers (false positives).

These results together with the corresponding detailed statistical analysis indicate that the SENS-IS test protocol delivers high quality results that support reliable predictions concerning the sensitizing hazard and potency of the tested items. Moreover, numerous tests on complex products and mixtures (manuscript in preparation) have demonstrated the capacity of the SENS-IS assay to correctly assess and classify sensitizers even in complex mixtures. The applicability domain of the SENS-IS assay is thus extremely wide compared to cell suspension-based assays. This is mainly due to the use of a reconstructed human skin model (Episkin) as the test system and to the analysis of a large panel of genes covering the many biological processes involved in skin sensitization. Indeed, a large set of chemicals covering a wide spectrum of chemical properties and various mixtures could be successfully evaluated.

In conclusion, the SENS-IS assay has been shown to be robust and easily transferable. Its robustness (Fig. 2) is also supported by excellent WLR and BLR. It's capacity to predict hazard is excellent as demonstrated by Cooper statistics values over 95% on a large panel of chemicals. In addition, it could correctly predict the potency of most tested items (Cooper statistic values of 92.6%, and 90.6% when compared to LLNA or human data respectively (see Table 7). The SENS-IS assay being based on a reconstructed human skin model as the test system, it is providing a wide applicability domain very similar to the expected "typical use" of the tested products. Furthermore, due to the very large amount of data that are produced for each test item through the measurement of the expression of the 62 selected biomarkers on different 3D skin batches and at various concentrations, the assay could benefit from big data approaches such as supervised computer learning, that may further optimize the scores obtained for all chemical categories (LeCun et al., 2015). The SENS-IS assay thus represents a serious alternative to the available *in vivo* sensitization tests.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.tiv.2016.01.007.

Transparency document

The Transparency document associated with this article can be found, in the 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.

Albrekt, A.N., Johansson, H., Börge, A., Borrebaeck, C., Lindstedt, M., 2014. Skin sensitizers differentially regulate signaling pathways in MUTZ-3 cells in relation to their individual potency. BMC Pharmacol. Toxicol. 15, 5.

Andersen, A., 2006. Final report on the safety assessment of benzaldehyde. Int. J. Toxicol. 25 (Suppl. 1), 11–27. http://dx.doi.org/10.1080/10915810600716612.

Andersen, et al., 2008. Final amended report on the safety assessment of methylparaben, ethylparaben, propylparaben, isopropylparaben, butylparaben, isobutylparaben, and benzylparaben as used in cosmetic products. Int. J. Toxicol. 27 (Suppl. 4), 1–82. http://dx.doi.org/10.1080/10915810802548359.

Basketter, D.A., Scholes, E.W., Wahlkvist, H., Montelius, J., 1995. An evaluation of the suitability of benzocaine as a positive control skin sensitizer. Contact Dermatitis 33, 28–32.

Basketter, D.A., Lea, L.J., Cooper, K., Stocks, J., Dickens, A., Pate, I., Dearman, R.J., Kimber, I., 1999. Threshold for classification as a skin sensitizer in the local lymph node assay: a statistical evaluation. Food Chem. Toxicol. 37, 1167–1174.

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

Basketter, D., White, I.R., McFadden, J.P., Kimber, I., 2015. Hexyl cinnamal: consideration of skin-sensitizing properties and suitability as a positive control. Cutan. Ocul. Toxicol. 34, 227–231. http://dx.doi.org/10.3109/15569527.2014.933973.

Casati, S., Aeby, P., Kimber, I., Maxwell, G., Ovigne, J.M., Roggen, E., Rovida, C., Tosti, L., Basketter, D., 2009. Selection of chemicals for the development and evaluation of in vitro methods for skin sensitisation testing. Altern. Lab. Anim. 37, 305–312.

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

Cosmetic ingredient expert panel, 2007. Final report on the amended safety assessment of propyl gallate. Int. J. Toxicol. 26 (Suppl. 3), 89–118. http://dx.doi.org/10.1080/10915810701663176.

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.

Divkovic, M., Pease, C.K., Gerberick, G.F., Basketter, D.A., 2005. Hapten-protein binding: from theory to practical application in the in vitro prediction of skin sensitization. Contact Dermatitis 53, 189–200. http://dx.doi.org/10.1111/j.0105-1873.2005.00683.x.

Emter, R., Ellis, G., Natsch, A., 2010. Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. Toxicol. Appl. Pharmacol. 245, 281–290. http://dx.doi.org/10.1016/j.taap.2010.03.009.

Emter, R., van der Veen, J.W., Adamson, G., Ezendam, J., van Loveren, H., Natsch, A., 2013. Gene expression changes induced by skin sensitizers in the KeratinoSens™ cell line: discriminating Nrf2-dependent and Nrf2-independent events. Toxicol. in Vitro 27, 2225–2232. http://dx.doi.org/10.1016/j.tiv.2013.09.009.

García-Gavín, J., Lissens, R., Timmermans, A., Goossens, A., 2011. Allergic contact dermatitis caused by isopropyl alcohol: a missed allergen? Contact Dermatitis 65, 101–106. http://dx.doi.org/10.1111/j.1600-0536.2011.01936.x.

Gerberick, G.F., Robinson, M.K., Ryan, C.A., Dearman, R.J., Kimber, I., Basketter, D.A., Wright, Z., Marks, J.G., 2001. Contact allergenic potency: correlation of human and local lymph node assay data. Am. J. Contact Dermat. 12, 156–161. http://dx.doi.org/10.1053/ajcd.2001.23926.

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., Ryan, C.A., Kern, P.S., Schlatter, H., Dearman, R.J., Kimber, I., Patlewicz, G.Y., Basketter, D.A., 2005. Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis 16, 157–202.

Hansson, C., Thörneby-Andersson, K., 2003. Stereochemical considerations on concomitant allergic contact dermatitis to ester of the cis-trans isomeric compounds maleic acid and fumaric acid. Skin Pharmacol. Appl. Ski. Physiol. 16, 117–122.

Iccvam documentation, 2013. http://ntp.niehs.nih.gov/iccvam/methods/immunotox/niceatm-llnadatabase-23dec2013.xls.

Karlberg, A.-T., Bergström, M.A., Börje, 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.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. http://dx.doi.org/10.1038/nature14539.

Lepoittevin, J.-P., Coz, C.J., 2007. Dictionary of contact allergens. Springer Science & Business Media, Berlin, Heidelberg http://dx.doi.org/10.1007/978-3-540-74165-7.

Loveless, S.E., Api, A.-M., Crevel, R.W.R., Debruyne, E., Gamer, A., Jowsey, I.R., Kern, P., Kimber, I., Lea, L., Lloyd, P., Mehmood, Z., Steiling, W., Veenstra, G., Woolhiser, M., Hennes, C., 2010. Potency values from the local lymph node assay: application to classification, labelling and risk assessment. Regul. Toxicol. Pharmacol. 56, 54–66. http://dx.doi.org/10.1016/j.yrtph.2009.08.016.

Machin, D., Campbell, M., Fayers, P., Pinol, A., 1997. Sample Size Tables for Clinical Studies. second ed. Blackwell Science, Oxford.

OECD, 2014. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. OECD Series on Testing and Assessment. No. 168, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264221444-en.

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.

RIFM Expert Panel, Belsito, D., Bickers, D., Bruze, M., Calow, P., Greim, H., Hanifin, J.M., Rogers, A.E., Saurat, J.H., Sipes, I.G., Tagami, H., 2007. A toxicologic and dermatologic assessment of salicylates when used as fragrance ingredients. Food Chem. Toxicol. 45 (Suppl. 1), S318–S361. http://dx.doi.org/10.1016/j.fct.2007.09.066.

SCCP, 2005. Scientific committee on consumer products (SCCP) opinion on benzoic acid and sodium benzoate. p. 30.

Sens-it-iv list, 2009. http://chelist.jrc.ec.europa.eu/index.jsp?id=2&id_list=26.

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.

Schraml, B.U., Reis e Sousa, C., 2015. Defining dendritic cells. Curr. Opin. Immunol. 32, 13–20. http://dx.doi.org/10.1016/j.coi.2014.11.001.

Uruno, A., Motohashi, H., 2011. The Keap1-Nrf2 system as an in vivo sensor for electrophiles. Nitric Oxide 25, 153–160. http://dx.doi.org/10.1016/j.niox.2011.02.007.

Yoshida, Y., Sakaguchi, H., Ito, Y., Okuda, M., Suzuki, H., 2003. Evaluation of the skin sensitization potential of chemicals using expression of co-stimulatory molecules, CD54 and CD86, on the naive THP-1 cell line. Toxicol. in Vitro 17, 221–228.