

Optimizing the cutoff for the identification of skin sensitizers by the HaCaSens assay: Introducing an ROC-analysis-based cutoff approach

Hyewon Chung^{a,1}, Eunju Kim^{b,1}, Hailian Quan^a, Daun Jung^a, Ravi Gautam^c, Ahrang Cho^c, Chang Yul Kim^c, MiJeong Kang^d, Jeong-Hwan Che^e, Eung-Seok Lee^d, Tae Cheon Jeong^d, Yong Heo^{c,*}, Seung Hyeok Seok^{a,**}

^a Department of Microbiology and Immunology, Institute of Endemic Disease, College of Medicine, Seoul National University, Seoul, 110-799, South Korea

^b College of Pharmacy, Pusan National University, Busan, South Korea

^c Department of Occupational Health, College of Bio-Medical Sciences, Daegu Catholic University, Daegu, 38430, South Korea

^d College of Pharmacy, Yeungnam University, Gyeongsan, 38541, South Korea

^e Biomedical Center for Animal Resource Development, Seoul National University College of Medicine, Seoul 03080, South Korea

ARTICLE INFO

Keywords:

HaCaSens
Skin sensitization
Interleukin-1α
Interleukin-6
Cutoff
ROC analysis

ABSTRACT

The worldwide restricted use of animal testing makes it challenging to identify the skin sensitizing potentials of newly manufactured products. The HaCaSens assay has shown promise as an *in vitro* skin sensitizing assay comparable to existing assays, and is currently under pre-validation. However, there is little agreement on how to assess the results of the assay to discriminate sensitizers from non-sensitizers as the stimulation index (SI) cutoff value was arbitrarily chosen without appropriate statistical methods.

Here, we investigated the SI cutoff values in identifying sensitizers to obtain the optimal value. Sensitivities and specificities were calculated for a set of 30 test substances, and plotted in receiver operator characteristics (ROC) curves. The SI cutoff values with the highest sum of sensitivity and specificity according to LLNA data were 2.2, 1.8 and 3.0 for interleukin 1α (IL-1α), interleukin 6 (IL-6), and the combination of the two cytokines respectively. Also, the same statistical analysis of human data demonstrated optimal SI cutoff values 2.0, 2.0 and 3.2 for the same respective parameters.

When considering the predictive capacity of each possible SI cutoff value determined by ROC curves, the optimal value for HaCaSens is 3.0 for the combination of IL-1α and IL-6 as it had the highest sensitivity (90.9%), specificity (75.0%) and accuracy (86.7%) based on LLNA data. Thus, we recommend the wide use of the SI cutoff value of 3.0 to ensure consistent endpoints.

1. Introduction

Allergic contact dermatitis (ACD), mainly induced by skin sensitization, is so prevalent that it accounts for 15–20% of skin-related diseases (Diepgen, 2003; Diepgen et al., 2016; McDonald et al., 2006; Peiser et al., 2012). In this regard, the skin sensitizing potential of cosmetic ingredients and products must be tested. In the face of the restriction of animal testing (EU, 2006, 2008a, 2009), there is an urgent need for the development of cell-based *in vitro* assays. Hence, a number of *in vitro* skin sensitization tests have been recently developed (Ashikaga et al., 2006; Emter et al., 2010; Gerberick et al., 2004; Piroird

et al., 2015; Ramirez et al., 2014; Sakaguchi et al., 2006).

The newly developed HaCaSens assay, a skin sensitization test using the human keratinocyte HaCaT cell line, was found to be hold promise with a high accuracy of 77%, satisfactory transferability, as well as within-/between-laboratory reproducibility (Chung et al., 2018; Jung et al., 2016). The endpoint for the assay is extracellular production of pro-inflammatory cytokines such that the potency of skin sensitizing is evaluated by a stimulation index (SI) indicating the fold increase in extracellular production of two pro-inflammatory cytokines, interleukin 1α (IL-1α) and interleukin 6 (IL-6), by chemical-treated keratinocytes (Chung et al., 2018). One of the fundamental issues when interpreting

* Corresponding author.

** Corresponding author at: Macrophages Laboratory, Department of Microbiology and Immunology, Seoul National University College of Medicine, Seoul, 110-799, South Korea.

E-mail addresses: yheo@cu.ac.kr (Y. Heo), lamseok@snu.ac.kr (S.H. Seok).

¹ Co-first authors.

and classifying test substances according to results of the assay is defining the SI threshold for distinguishing between sensitizers and non-sensitizers. Previously, we employed a SI value of 3.0 for a positive response to HaCaSens, which is used in the traditional Local Lymph Node Assay (LLNA) (OECD TG 429, 2010), so a chemical that stimulates a threefold or greater increase in extracellular production of cytokine IL-1 α or IL-6 compared with vehicle treated control is considered a sensitizer.

However, it is noteworthy that the SI cutoff value of 3.0 in LLNA (OECD TG 429, 2010) is different from that of its analogous methods LLNA:DA (1.8) (OECD, TG 442A), LLNA:BrDU-ELISA (1.6) (OECD, TG 442B) and LLNA:BrDU-FCM (2.8) (OECD, TG 442B), indicating that the optimal cutoff values for different methods are different. The conventional SI cutoff value 3.0 previously used in the HaCaSens assay was not empirically determined and so a true SI cutoff value remains to be optimized. Defining the optimal cutoff for HaCaSens by an appropriate statistical model is important for further validation of the sensitizing potential of substances.

To establish an optimized prediction model for HaCaSens, two critical factors should be considered: a SI cutoff value and parameter. First, we generated unbiased receiver operating characteristic (ROC) curves previously described (Kim et al., 2016). Potential optimal SI cutoff values were selected based on their predictive capacities such as sensitivity and specificity, and their performance were then reevaluated with all possible SI cutoff values. Second, we assessed the performance of HaCaSens to produce the optimal parameter (IL-1 α and IL-6 alone or in combination) that maximizes accurate discrimination between sensitizers and non-sensitizers.

Here, we aim to (1) define the optimal SI cutoff value for sensitizer-induced production of pro-inflammatory cytokines, and (2) determine the optimal parameter for HaCaSens to classify sensitizers from non-sensitizers with high sensitivity and specificity.

2. Materials and methods

2.1. Selection of test substances

We preferentially selected the substances used in LLNA or other OECD test guidelines (DPRA, ARE-Nrf2-luciferase assay, hCLAT) with concordant results among the previous studies. 30 test substances selected included 22 sensitizers and 8 non-sensitizers and represented a good distribution of 6 strong, 6 moderate, 2 moderate-weak, 8 weak sensitizers along with 8 non-sensitizers. The physical state of the substances was also properly considered (15 liquid and 15 solid substances). Their detailed information is presented in Table 1. The test substances used in this study are as follows: 2,4-dinitrochlorobenzene (DNCB), lactic acid, isopropanol, hexyl cinnamic aldehyde, salicylic acid, 2-mercaptobenzothiazole, formaldehyde, 4-hydroxybenzoic acid, citral, chlorobenzene, 4-phenylenediamine, 2,3-butanedione, cobalt chloride, tetramethylthiuram disulfide, vanillin, benzocaine, nickel sulfate, glycerol, methyl methacrylate, ethylene glycerol dimethacrylate and imidazolidinyl urea were purchased from Sigma-Aldrich (St. Louis, MO, USA). Phenyl benzoate, methyl salicylate, isoeugenol, cinnamic alcohol, eugenol, benzylideneacetone and 4-aminobenzoic acid were purchased from TCI (Tokyo Chemical Industry, Tokyo, Japan). Glutaraldehyde and methylbromo glutaronitrile were purchased from Alfa Aesar (Ward Hill, MA, USA) and CMI/MI (Kathon™ CG Preservative) was from Dow Chemical Group Company (Horgen, Switzerland).

2.2. HaCaSens procedure

As shown in Fig. 1, this study was performed according to the methods of Chung et al. (2018). To determine the appropriate test concentration of each chemical, lethal concentration causing death in 25% (LC25) was assessed using cell counting kit-8 (CCK-8) (Dojindo,

Kumamoto, Japan). HaCaT cells were plated at 1.2×10^5 cells/ml (100 μ l total volume/well) in a 96-well plate and treated for 24 h with each chemical using various concentrations (1,000, 500, 100, 10, 1 μ g/ml) or vehicle control. Then, 10 μ l of CCK-8 solution was added to each well and cells were incubated for another 4 h at 37°C according to the manufacturer's instructions. LC25 was calculated for each chemical from the O.D. value at 450 nm.

To measure the extracellular production of IL-1 α and IL-6, cells were seeded in 24-well plates at a concentration of 1.5×10^5 cells/ml (500 μ l total volume/well). At 24 h post incubation, the culture medium was removed and cells were treated with 0.01, 0.1, 0.5, and 1 X LC25 concentration of each chemical or vehicle control. After 24 h of incubation, extracellular cytokines produced by cells were measured from supernatants using a DuoSet ELISA kit (R&D Systems, Minneapolis, MN, USA) for IL-6 (DY206, sensitivity: 0.11 pg/ml) and IL-1 α (DY200, sensitivity: 1 pg/ml). The protein content of the cell lysates was measured using the BCA protein assay method (Thermo Scientific, Rockford, IL, USA). SI values were calculated as follows: SI = cytokine produced in chemical-treated cells / cytokine produced in vehicle control.

2.3. Applying the results of HaCaSens to a prediction model

Three independent experiments were performed for 30 test substances for all concentrations (0.01, 0.1, 0.5, and 1 X LC25). The sensitizing potential of each chemical was assessed based on the SI value of IL-1 α and IL-6 at each concentration. The following prediction model was applied: a test substance was considered a sensitizer if the highest SI value among all concentrations was greater than or equal to 3.0 in at least one cytokine. Non-sensitizers produced a SI value under 3.0 in both cytokines at all concentrations. To assess the predictive capacity of the assay, the sensitivity, specificity, accuracy, positive predictivity and negative predictivity were calculated as previously described (Ramirez et al., 2016).

2.4. Determination of optimal SI cutoff values

The receiver operating characteristic (ROC) curve, which illustrates sensitivity against false positive rate, has been frequently used to obtain optimal cutoff values in diagnostic tests. 30 test substances were used to plot a ROC curve. Of those SI values from diverse concentrations (0.01, 0.1, 0.5, 1 X LC25), those with maximum values were used to obtain the ROC curve, since the test substance was defined as positive (sensitizer) if the maximum SI value is greater than the cutoff value, and as negative (non-sensitizer) otherwise.

3. Results

3.1. Dataset in the HaCaSens

A set of 30 test substances including 22 sensitizers and 8 non-sensitizers was tested to characterize the predictive capacity of HaCaSens. Table 1 lists the classification of each test substance according to LLNA, human, and mean values of max SI of three independent HaCaSens experiments with three parameters: IL-1 α , IL-6 and the combination of the two cytokines. Although, overall, the test substances showed similar max SI values between IL-1 α and IL-6, there were also substances with 10-fold differences between the two cytokines such as methyl methacrylate, ethylene glycol dimethacrylate, CMI/MI and benzylideneacetone. Thus, it may be difficult to test skin sensitizing potential of test substances with single parameter.

As shown in Fig. 2, the SI values of sensitizers were significantly higher than non-sensitizers excluding the false positive substances, methyl salicylate. Furthermore, to investigate if the HaCaSens assay could predict skin sensitizing potency of test substances, we classified the 22 sensitizers tested in this study into 3 potency groups (weak, moderate and strong) based on previous publications (Bauch et al.,

Table 1
Maximal SI values for 30 test substances.

No.	Substances	Chemical information				Max SI values (Mean $\pm$ SD)			In vivo prediction	
		CAS No ^a	Supplier	Physical state	Purity (%)	IL-1 α	IL-6	IL-1 α or IL-6	LLNA data ^b	Human data
1	Isopropanol	67-63-0	Sigma-Aldrich	Liquid	99.5	2.0 $\pm$ 0.6	1.9 $\pm$ 0.9	2.4 $\pm$ 0.8	–	–/+ [*]
2	Hexyl cinnamic aldehyde	101-86-0	Sigma-Aldrich	Liquid	85	70.0 $\pm$ 80.3	14.0 $\pm$ 12.9	70.2 $\pm$ 80.1	+	+
3	Salicylic acid	69-72-7	Sigma-Aldrich	Solid	> 99	1.9 $\pm$ 0.6	1.4 $\pm$ 0.6	2.0 $\pm$ 0.6	–	–
4	2-Mercapto benzothiazole	149-30-4	Sigma-Aldrich	Solid	97	9.3 $\pm$ 9.9	10.4 $\pm$ 7.0	13.0 $\pm$ 9.5	+	+
5	Formaldehyde	50-00-0	Sigma-Aldrich	Liquid	36.5-38	5.5 $\pm$ 4.0	6.6 $\pm$ 4.0	8.5 $\pm$ 3.8	+	+
6	4-Hydroxybenzoic acid	99-96-7	Sigma-Aldrich	Solid	$\geq$ 99	1.3 $\pm$ 0.7	1.4 $\pm$ 0.5	1.6 $\pm$ 0.6	–	–
7	Citral	5392-40-5	Sigma-Aldrich	Liquid	95	11.4 $\pm$ 13.4	2.8 $\pm$ 2.0	11.5 $\pm$ 13.3	+	+
8	Phenyl benzoate	93-99-2	TCI	Solid	> 99	9.8 $\pm$ 10.2	6.3 $\pm$ 6.8	11.6 $\pm$ 10.8	+	+
9	Chlorobenzene	108-90-7	Sigma-Aldrich	Liquid	99.9	2.9 $\pm$ 4.8	1.7 $\pm$ 0.7	3.4 $\pm$ 4.7	–	–
10	4-Phenylenediamine	106-50-3	Sigma-Aldrich	Solid	98	40.8 $\pm$ 63.3	5.5 $\pm$ 5.0	41.5 $\pm$ 62.9	+	+
11	Methyl salicylate	119-36-8	TCI	Liquid	> 99	2.1 $\pm$ 1.1	8.4 $\pm$ 3.6	8.4 $\pm$ 3.6	–	–
12	2,3-Butanedione	431-03-8	Sigma-Aldrich	Liquid	97	1.5 $\pm$ 0.7	1.4 $\pm$ 0.6	1.5 $\pm$ 0.7	+	n.a.
13	Isoeugenol	97-54-1	TCI	Liquid	> 97	2.6 $\pm$ 1.1	2.4 $\pm$ 1.0	3.1 $\pm$ 1.0	+	+
14	Glutaraldehyde	111-30-8	Alfa Aesar	Liquid	50	6.7 $\pm$ 10.5	9.8 $\pm$ 6.5	13.5 $\pm$ 12.5	+	+
15	Cobalt chloride	7646-79-9	Sigma-Aldrich	Solid	> 98	3.0 $\pm$ 2.1	2.1 $\pm$ 0.8	3.3 $\pm$ 1.9	+	+
16	Tetramethylthiuram disulfide	137-26-8	Sigma-Aldrich	Solid	97	14.2 $\pm$ 16.1	3.9 $\pm$ 2.5	16.2 $\pm$ 15.3	+	+
17	Cinnamic alcohol	104-54-1	TCI	Solid	> 97	2.5 $\pm$ 0.7	6.6 $\pm$ 0.7	6.6 $\pm$ 0.7	+	+
18	Vanillin	121-33-5	Sigma-Aldrich	Solid	99	1.9 $\pm$ 0.8	1.5 $\pm$ 0.1	2.0 $\pm$ 0.5	–	+ ^{**}
19	Benzocaine	94-09-7	Sigma-Aldrich	Solid	–	1.0 $\pm$ 0.1	1.2 $\pm$ 0.3	1.2 $\pm$ 0.3	+	+
20	Nickel sulfate	10101-97-0	Sigma-Aldrich	Solid	99	2.1 $\pm$ 0.2	5.1 $\pm$ 2.0	5.1 $\pm$ 2.0	+	+
21	Cinnamaldehyde	104-55-2	Sigma-Aldrich	Liquid	> 95	23.8 $\pm$ 33.0	23.8 $\pm$ 33.0	24.5 $\pm$ 32.5	+	+
22	Eugenol	97-53-0	TCI	Liquid	99	1.7 $\pm$ 0.6	4.9 $\pm$ 1.2	4.9 $\pm$ 1.2	+	+
23	Glycerol	56-81-5	Sigma-Aldrich	Liquid	99	1.2 $\pm$ 0.8	1.3 $\pm$ 0.3	1.7 $\pm$ 0.4	–	–
24	Methyl methacrylate	80-62-6	Sigma-Aldrich	Liquid	99	44.0 $\pm$ 34.0	1.2 $\pm$ 0.5	30.3 $\pm$ 33.8	+	+
25	Ethylene glycol dimethacrylate	97-90-5	Sigma-Aldrich	Liquid	98	49.6 $\pm$ 18.8	1.2 $\pm$ 0.1	49.6 $\pm$ 18.8	+	+
26	CMI/MI (Kathon TM CG Preservative)	26172-55-4/ 2682-20-4	Dow Chemical Group Company	Liquid	–	29.5 $\pm$ 25.1	1.3 $\pm$ 0.4	45.0 $\pm$ 52.2	+	+
27	Methyldibromo glutaronitrile	35691-65-7	Alfa Aesar	Solid	99	19.8 $\pm$ 24.5	2.7 $\pm$ 1.3	55.5 $\pm$ 142.0	+	+
28	Benzylideneacetone	122-57-6	TCI	Solid	> 98	120.0 $\pm$ 109.4	2.1 $\pm$ 0.4	82.3 $\pm$ 103.4	+	n.a.
29	Imidazolidinyl urea	39236-46-9	Sigma-Aldrich	Solid	–	22.8 $\pm$ 13.4	8.2 $\pm$ 6.4	22.8 $\pm$ 13.4	+	+
30	4-Aminobenzoic acid	150-13-0	TCI	Solid	> 99	2.3 $\pm$ 1.7	1.1 $\pm$ 0.4	2.4 $\pm$ 1.6	–	n.a.

^a CAS No (Chemical Abstracts Service Number).

^b Testing results using the LLNA (local lymph node assay) and data taken from OECD, test guideline (TG429). +; sensitizer, –; non-sensitizer.

* presumed to be a sensitizer in OECD, TG429, a non-sensitizer based on literature from Bauch C et al., 2011 and Ramirez T., 2016 and classified into human class 5, a weak sensitizer based on the literature from Basketter et al., 2014 and Cottrez et al., 2016.

** classified into human class 5 based on the literature from Basketter et al., 2014 and Cottrez et al., 2016.

2011; Natsch et al., 2013; Nukada et al., 2013; Tsujita-Inoue et al., 2014). When comparing the max SI value of each sensitizer according to sensitizing potency, we found that the sensitizers with higher sensitizing potency exhibit the higher SI values, suggesting a positive correlation between the potency and SI values, although not significant (Table S1).

3.2. Identification of optimal SI cutoff values by ROC analysis

To find an optimal SI cutoff value for the discrimination of sensitizers from non-sensitizers by HaCaSens, we plotted ROC curves using the data of the set of 30 substances. Based on ROC analysis, the area under the ROC curves (AUCs) for this assay were 0.799 (95% CI 0.757–0.96), 0.707 (95% CI 0.652–0.763) and 0.806 (95% CI 0.763–0.850) and the optimal SI cutoff values were also plotted on the curve (Fig. 3). For single parameters IL-1 α and IL-6, the optimal SI cutoff values yielding maximal sensitivity plus specificity of over 70% were 2.2 and 1.8 respectively. For the combination of IL-1 α and IL-6, the optimal SI cutoff value was 3.2 with superior sensitivity (78.1%) and specificity (78.6%) for predicting skin sensitization compared to other parameters but with no significant increase in AUC.

Considering the discrepancy of sensitizing potential between LLNA and human data, we also defined the optimal SI cutoff values based on human data using the same approach. Using the same data from 26 of the test substances excluding not applicable substances

(benzylideneacetone, 2,3-butanedione, 4-aminobenzoic acid) and equivocally predicted chemical (isopropanol), the AUCs for this assay were 0.811 (95% CI 0.766–0.857), 0.658 (95% CI 0.591–0.724) and 0.784 (95% CI 0.32–0.836) and optimal SI cutoff values were 2.0, 2.0, 3.2 of IL-1 α , IL-6 and the combination of the two cytokines respectively (Fig. 4). These optimal SI cutoff values for each parameter were similar to those based on LLNA data. The overall statistical analysis results including AUC showed no significant differences as well.

3.3. Validation of the SI cutoff values

We selected the ROC-analysis based optimal SI cutoff values 1.8, 2.0, 2.2, 3.0 and 3.2 and applied them to 30 test substances to classify them as sensitizers (+) or non-sensitizer (–) to further confirm these acquired cutoff values. As shown in Table 2, we classified the test substances following the prediction model same as above: if the highest SI value between IL-1 α and IL-6 was greater than or equal to SI cutoff values, the substance was considered a sensitizer (Chung et al., 2018).

Next, we calculated the sensitivity, specificity and accuracy with each SI cutoff value to determine the optimal SI value for prediction (Table 3 and 4). All SI cutoff values showed the same sensitivity of 90.9%, except SI cutoff value 3.2 which resulted in a lower sensitivity of 86.3%. There was a greater variation in specificity: 25.0, 25.0, 50.0, 75.0 and 75.0% for SI cutoff value 1.8, 2.0, 2.2, 3.0 and 3.2 respectively. For accuracy, HaCaSens produced 73.3, 73.3, 80.0, 86.7 and

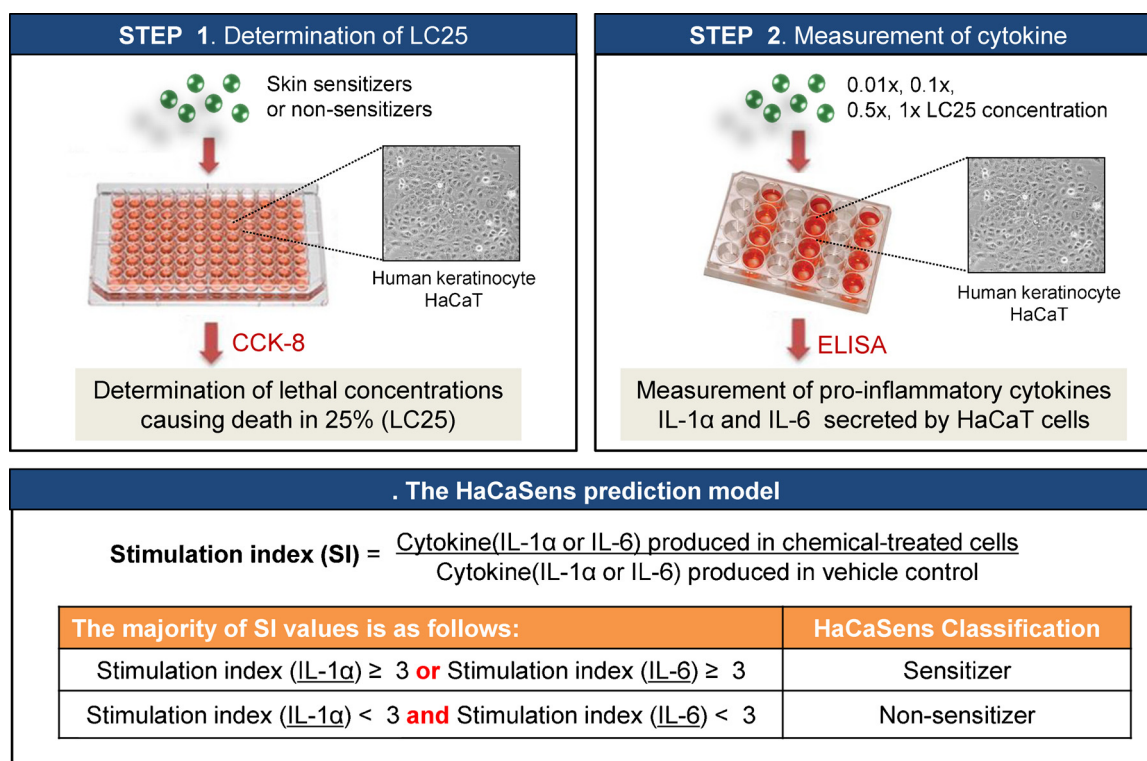


Fig. 1. The schematic image of the overall HaCaSens assay.

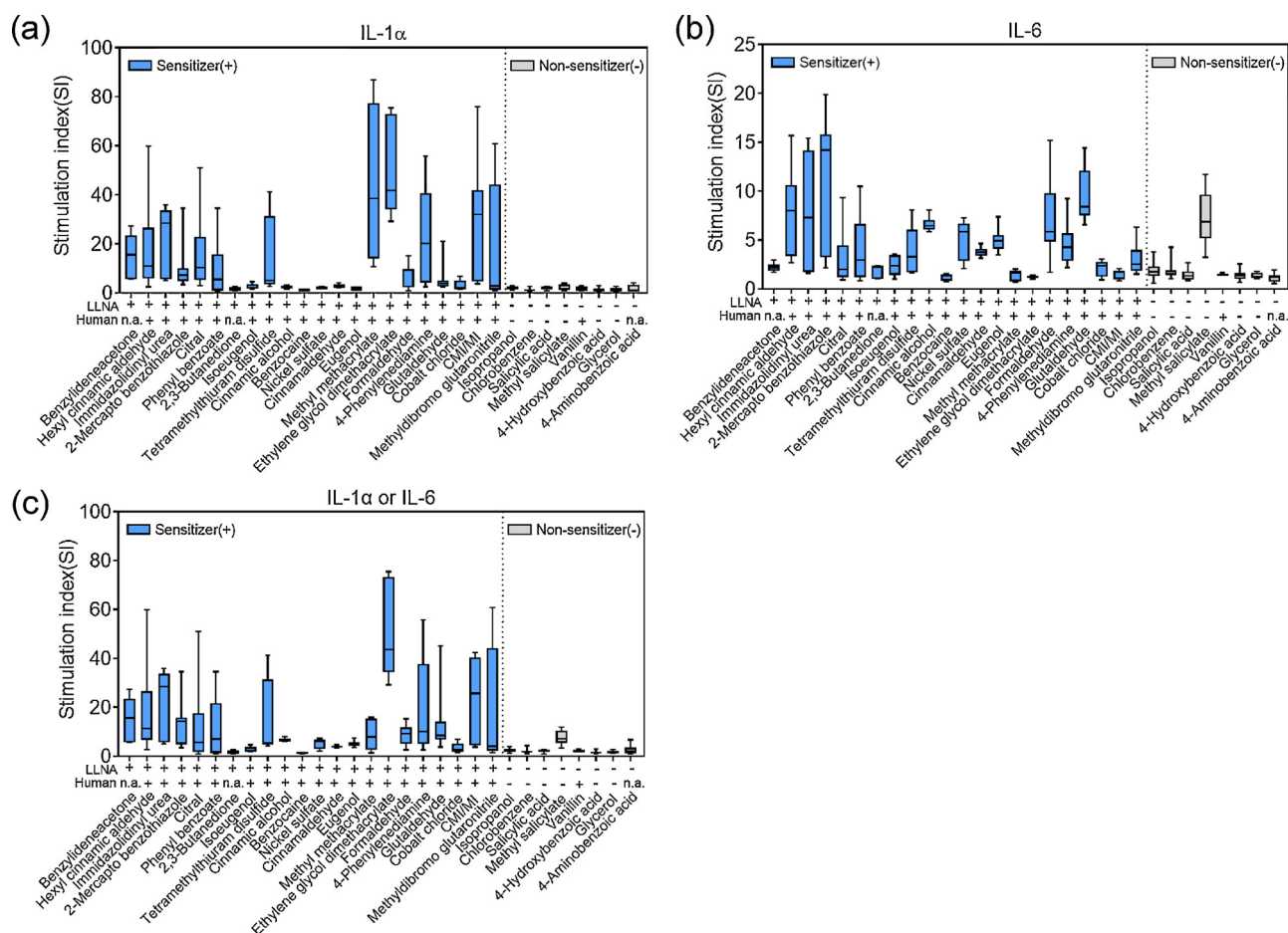


Fig. 2. Max SI values of 30 test substances with 22 sensitizers and 8 non-sensitizer. (a) IL-1 α , (b) IL-6, (c) IL-1 α or IL-6 and (d) overall predictive capacity of the HaCaSens with a conventional cutoff SI 3.0.

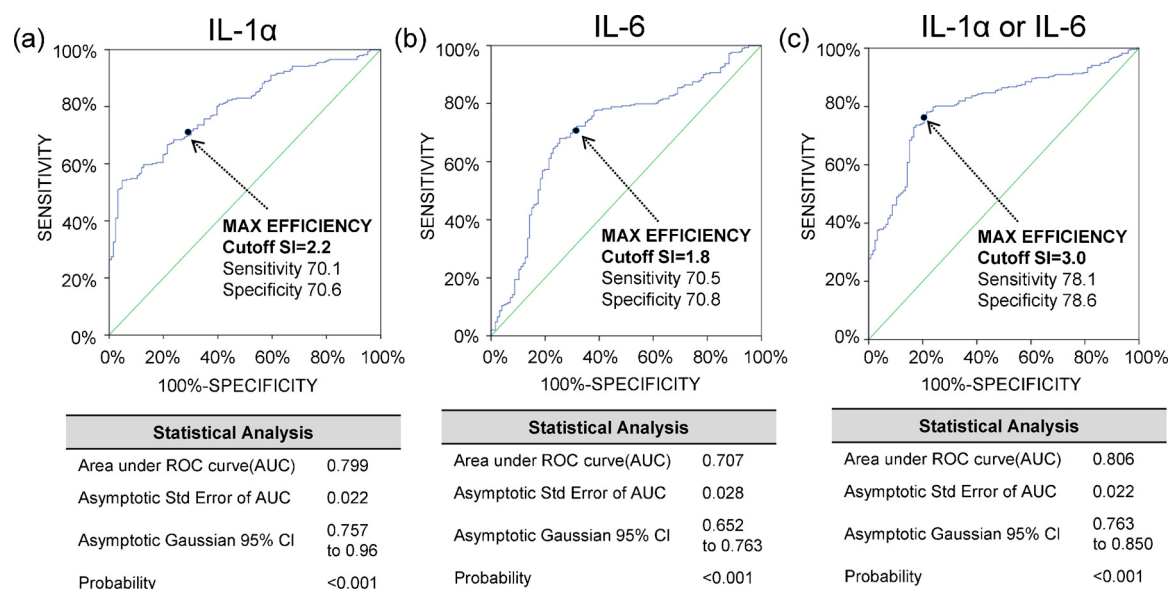


Fig. 3. ROC curves for maximum stimulation index (max SI) to discriminate sensitizers from non-sensitizers with pro-inflammatory cytokine IL-1α (a), IL-6 (b) and combination of two cytokines (IL-1α or IL-6) (c) for 30 test substances based on LLNA data, cutoff SI yielding the maximal sensitivity plus specificity for predicting skin sensitizing potential of the test substances.

83.3% identification of sensitizers for each respective SI cutoff value (Table 3). When compared with human data, sensitivity was 95.2, 95.2, 90.5, 90.5, 85.7% and specificity was 40.0, 40.0, 60.0, 60.0, 60.0% for SI cutoff value 1.8, 2.0, 2.2, 3.0 and 3.2 respectively (Table 4). Although the variation of the sensitivity and specificity among the SI cutoff values was more significant than that of LLNA data, all SI cutoff values had the accuracy of 84.6%. Compared with human classification, HaCaSens repeatedly showed the lowest predictivity with SI cutoff value 3.2 in terms of sensitivity and accuracy.

3.4. Determination of the optimal parameter

Comparing the five possible SI cutoff values in terms of sensitivity, specificity and accuracy, we propose the SI cutoff value 3.0 as the optimized cutoff for HaCaSens. With this SI cutoff value 3.0, HaCaSens

demonstrated the highest accuracy based on both LLNA (86.7%) and human data (84.6%) (Tables 3 and 4). Actually the SI cutoff value 3.0 outperformed any other value in identifying sensitizers in all aspects of predictive capacity including positive predictive value (90.9%) and negative predictive value (75.0%) based on LLNA. With respect to human data, the cutoff value 3.0 acquired the highest specificity (60.0%), positive predictive value (90.5%) and accuracy (84.6%) despite slightly lower sensitivity and negative predictive value. Although the sensitivity of SI cutoff values 1.8 and 2.0 was 95.2%, which is slightly higher than that of cutoff value 3.0 (90.5%), the specificity of cutoff values 1.8 and 2.0 (40.0%) was far lower than 3.0 (60.0%).

We further investigated the optimal parameters for SI cutoff value 3.0. HaCaSens uses the pro-inflammatory cytokines IL-1α and IL-6 produced by keratinocytes as parameters for discriminating sensitizers from non-sensitizers. With the same set of test substances, we compared

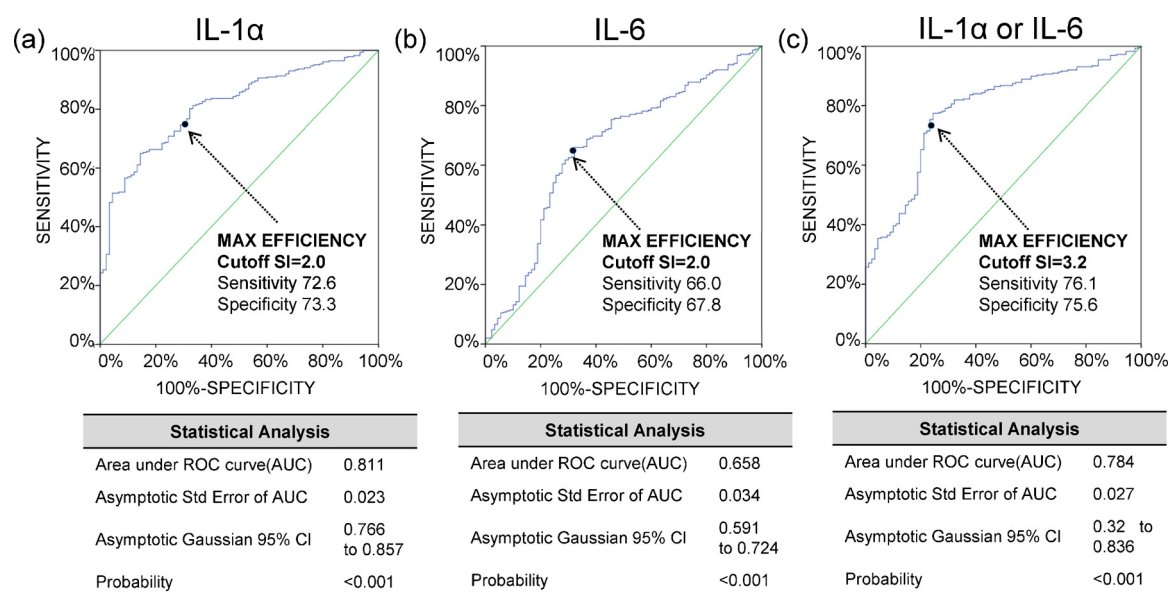


Fig. 4. ROC curves for maximum stimulation index (max SI) to discriminate sensitizers from non-sensitizers with pro-inflammatory cytokine IL-1α (a), IL-6 (b) and combination of two cytokines (IL-1α or IL-6) (c) for 26 test substances based on human data, cutoff SI yielding the maximal sensitivity plus specificity for predicting skin sensitizing potential of the test substances.

Table 2

Decision based on conventional SI cutoff value 3.0 and SI cutoff values 1.8, 2.0, 2.2 and 3.2.

No.	Substances	LLNA data	Human data	HaCaSens						
				Group	Max SI	Cutoff 1.8	Cutoff 2.0	Cutoff 2.2	Cutoff 3.0	Cutoff 3.2
1	Isopropanol	–	– / + *	0.1x	2.4 ± 0.8	+	+	+	–	–
2	Hexyl cinnamic aldehyde	+	+	1x	70.2 ± 80.1	+	+	+	+	+
3	Salicylic acid	–	–	1x	2.0 ± 0.6	+	+	–	–	–
4	2-Mercapto benzothiazole	+	+	1x	13.0 ± 9.5	+	+	+	+	+
5	Formaldehyde	+	+	1x	8.5 ± 3.8	+	+	+	+	+
6	4-Hydroxybenzoic acid	–	–	1x	1.6 ± 0.6	–	–	–	–	–
7	Citral	+	+	1x	11.5 ± 13.3	+	+	+	+	+
8	Phenyl benzoate	+	+	1x	11.6 ± 10.8	+	+	+	+	+
9	Chlorobenzene	–	–	1x	3.4 ± 4.7	+	+	+	+	+
10	4-Phenylenediamine	+	+	0.5x	41.5 ± 62.9	+	+	+	+	+
11	Methyl salicylate	–	–	1x	8.4 ± 3.6	+	+	+	+	+
12	2,3-Butanedione	+	n.a.	1x	1.5 ± 0.7	–	–	–	–	–
13	Isoeugenol	+	+	1x	3.1 ± 1.0	+	+	+	+	–
14	Glutaraldehyde	+	+	1x	13.5 ± 12.5	+	+	+	+	+
15	Cobalt chloride	+	+	1x	3.3 ± 1.9	+	+	+	+	+
16	Tetramethylthiuram disulfide	+	+	1x	16.2 ± 15.3	+	+	+	+	+
17	Cinnamic alcohol	+	+	1x	6.6 ± 0.7	+	+	+	+	+
18	Vanillin	–	+	1x	2.0 ± 0.5	+	+	–	–	–
19	Benzocaine	+	+	1x	1.2 ± 0.3	–	–	–	–	–
20	Nickel sulfate	+	+	1x	5.1 ± 2.0	+	+	+	+	+
21	Cinnamaldehyde	+	+	1x	24.5 ± 32.5	+	+	+	+	+
22	Eugenol	+	+	1x	4.9 ± 1.2	+	+	+	+	+
23	Glycerol	–	–	0.01x	1.7 ± 0.4	–	–	–	–	–
24	Methyl methacrylate	+	+	0.5x	30.3 ± 33.8	+	+	+	+	+
25	Ethylene glycol dimethacrylate	+	+	0.5x	49.6 ± 18.8	+	+	+	+	+
26	CMI/MI (Kathon™ CG Preservative)	+	+	1x	45.0 ± 52.2	+	+	+	+	+
27	Methyldibromo glutaronitrile	+	+	1x	55.5 ± 142.0	+	+	+	+	+
28	Benzylideneacetone	+	n.a.	0.5x	82.3 ± 103.4	+	+	+	+	+
29	Imidazolidinyl urea	+	+	0.5x	22.8 ± 13.4	+	+	+	+	+
30	4-Aminobenzoic acid	–	n.a.	0.01x	2.4 ± 1.6	+	+	+	–	–

Table 3

Predictive capacity of the HaCaSens for 30 test substances with possible SI cutoff values based on LLNA data.

	HaCaSens				
	Cutoff 1.8	Cutoff 2.0	Cutoff 2.2	Cutoff 3.0	Cutoff 3.2
Total number (N)	30	30	30	30	30
Sensitivity (%)	90.9%	90.9%	90.9%	90.9%	86.3%
Specificity (%)	25.0%	25.0%	50.0%	75.0%	75.0%
Positive predictive value (%)	76.9%	76.9%	83.3%	90.9%	90.5%
Negative predictive value (%)	50.0%	50.0%	66.7%	75.0%	66.7%
Accuracy (%)	73.3%	73.3%	80.0%	86.7%	83.3%

Table 4

Predictive capacity of the HaCaSens for 30 test substances with possible SI cutoff values based on human data.

	HaCaSens				
	Cutoff 1.8	Cutoff 2.0	Cutoff 2.2	Cutoff 3.0	Cutoff 3.2
Total number (N)	26	26	26	26	26
Sensitivity (%)	95.2%	95.2%	90.5%	90.5%	85.7%
Specificity (%)	40.0%	40.0%	60.0%	60.0%	60.0%
Positive predictive value (%)	87.0%	87.0%	90.5%	90.5%	90.0%
Negative predictive value (%)	66.7%	66.7%	60.0%	60.0%	50.0%
Accuracy (%)	84.6%	84.6%	84.6%	84.6%	80.8%

its predictive capacity according to all possible parameters, single respective parameters IL-1 α , IL-6 and the combination of the two cytokines to identify the optimal parameter with maximal predictive capacity (Table 5). HaCaSens produced 72.7% sensitivity, 100.0%

Table 5

Predictive capacity of the HaCaSens for 30 test substances using possible parameters based on LLNA and human data respectively.

	HaCaSens with Cutoff 3.0	LLNA data			Human data		
		IL-1 α	IL-6	IL-1 α or IL-6	IL-1 α	IL-6	IL-1 α or IL-6
Total number (N)	30	30	30	30	26	26	26
Sensitivity (%)	72.7%	54.5%	90.9%	90.9%	71.4%	71.4%	90.5%
Specificity (%)	100.0%	87.5%	75.0%	75.0%	100.0%	80.0%	60.0%
Accuracy (%)	80.0%	63.3%	86.7%	86.7%	76.9%	73.1%	84.6%

specificity and 80.0% accuracy for IL-1 α , 54.5% sensitivity, 87.5% specificity and 63.3% accuracy for IL-6 and 90.9% sensitivity, 75.0% specificity and 86.7% accuracy for the combination based on LLNA classification. The same analysis with human classification also showed the highest predictive capacity for the combination of the two cytokines IL-1 α and IL-6 (90.5% sensitivity, 60.0% specificity and 84.6% accuracy).

Next, we compared predictive capacity of the HaCaSens with that of KeratinoSens™, DPRA or hCLAT with the goal of assessing any advantages over the currently accepted OECD TG guidelines (Table 6). To this end, two criteria were used for the same set of 30 test substances; 1) test substances not applicable should be excluded, 2) In case of test substances with equivocal results among studies, they are considered as false prediction, so that we could thoroughly compare the predictive capacity with high reproducibility. Based on these criteria, the HaCaSens was shown to have better sensitivity of 90.9% and lower specificity of 75.0% compared to KeratinoSens™ (81.0% and 100%) and DPRA (66.7% and 85.7%) while the hCLAT acquired similar predictive capacity (90.5% and 75.0%) with 29 test substances without benzocaine (not applicable). With respect to accuracy, HaCaSens (86.7%),

Table 6

Comparison of predictive capacity among HaCaSens, KeratinoSens™, DPRA and hCLAT (+: positive/sensitizer prediction; -: negative/non-sensitizer prediction; +/–: equivocal prediction; n.a.: not applicable).

No.	Substances	LLNA	HaCaSens	KeratinoSens™ ^a	DPRA ^b	hCLAT ^c
1	Isopropanol	–	–	–	–	–
2	Hexyl cinnamic aldehyde	+	+	+	–	+ / –
3	Salicylic acid	–	–	–	–	+
4	2-Mercapto benzothiazole	+	+	+	+	+
5	Formaldehyde	+	+	+	+	+
6	4-Hydroxybenzoic acid	–	–	–	–	–
7	Citral	+	+	+	+	+
8	Phenyl benzoate	+	+	–	+ / –	+
9	Chlorobenzene	–	+	–	–	+
10	4-Phenylenediamine	+	+	+	+	+
11	Methyl salicylate	–	+	–	–	–
12	2,3-Butanedione	+	–	+	+	+
13	Isoeugenol	+	+	+	+	–
14	Glutaraldehyde	+	+	+	+	+
15	Cobalt chloride	+	+	+	n.a.	+
16	Tetramethylthiuram disulfide	+	+	+	+	+
17	Cinnamic alcohol	+	+	+	+ / –	+
18	Vanillin	–	–	–	+ / –	–
19	Benzocaine	+	–	n.a.	–	n.a.
20	Nickel sulfate	+	+	+	–	+
21	Cinnamaldehyde	+	+	+	+	+
22	Eugenol	+	+	+ / –	+	+
23	Glycerol	–	–	–	–	–
24	Methyl methacrylate	+	+	–	+ / –	+
25	Ethylene glycol dimethacrylate	+	+	+	+	+
26	CMI/MI (Kathon™ CG Preservative)	+	+	+	+	+
27	Methyldibromo glutaronitrile	+	+	+	+	+
28	Benzylideneacetone	+	+	+	+	+
29	Imidazolidinyl urea	+	+	+ / –	+ / –	+
30	4-Aminobenzoic acid	–	–	n.a.	n.a.	–
Total number			30	28	28	29
Sensitivity (%)			90.9 (20/22)	81.0 (17/21)	66.7 (14/21)	90.5 (19/21)
Specificity (%)			75.0 (6/8)	100.0 (7/7)	85.7 (6/7)	75.0 (6/8)
Accuracy (%)			86.7 (26/30)	85.7 (24/28)	71.4 (20/28)	86.2 (25/29)

^a Literature data from Natsch (2010); Ashikaga et al., (2010); Bauch et al., (2011); Nukada et al., (2013); Natsch et al., (2013); Ramirez et al., (2014); Tsujita-Inoue et al., (2014) and Reisinger et al. (2015).

^b Literature data from Natsch (2010); Bauch et al., (2011); Cho et al. (2013); Nukada et al., (2013); Natsch et al., (2013) and Reisinger et al. (2015).

^c Literature data from Natsch (2010); Ashikaga et al., (2010); Bauch et al., (2011); Cho et al. (2013); Nukada et al., (2013); Natsch et al., (2013); Tsujita-Inoue et al., (2014), and Reisinger et al. (2015).

KeratinoSens™ (85.7%), and hCLAT (86.2%) had similar results except DPRA (71.4%). The satisfactory predictive capacity of the HaCaSens shows great promise as an *in vitro* assay comparable with other assays.

4. Discussion

The present study, which was based on statistical ROC-analysis, examined various SI cutoff values and parameters to find the optimal prediction model for HaCaSens that would maximize predictive capacity (sensitivity, specificity and accuracy). The production of pro-inflammatory cytokines IL-1 α and IL-6 in keratinocytes, the endpoint of the HaCaSens assay, was shown to discriminate sensitizers from non-sensitizers.

The fact that the assay uses a non-genetically transformed cell line similar to normal keratinocytes and does not target a specific signaling pathway such as keap-1/nrf-2 to identify sensitizers also enhances accuracy over existing similar assays (Chung et al., 2018; Jung et al., 2016). However, the arbitrary selection of the SI cutoff value 3.0 for positive prediction in HaCaSens following LLNA has been debated as it was selected without statistical consideration. Given the advantages and high predictive capacity of HaCaSens, it should undergo further validation to support its candidacy as a new OECD test guideline. Defining the optimal cutoff and parameters for the greatest predictive capacity should be determined based on established statistical models prior to further validation study.

We investigated the optimal SI cutoff values for identifying

sensitizers using HaCaSens and a set of 30 test substances. ROC curves plotted with the data showed several possible SI cutoff values with three respective parameters: IL-1 α , IL-6 and the combination of IL-1 α and IL-6 (Fig. 3). The SI cutoff values for IL-1 α (2.2) and IL-6 (1.8) were similar and lower than the SI cutoff value for the combination (3.2) which could be due to the fact the highest SI values between the two cytokines were additive in the combination parameter. When considering predictive capacity such as sensitivity and specificity, the combination of the two cytokines significantly outperformed the individual parameters with the SI cutoff value 3.0. Of note, the SI cutoff value 3.0 prediction model with the highest sensitivity and specificity was also used in a previous study using HaCaSens (Chung et al., 2018).

Some substances show different skin sensitizing abilities in human compared to LLNA, so there may be different optimal SI cutoff values in human classifications data. In particular, applying the same analysis with human data is meaningful since HaCaSens uses HaCaT cells, human keratinocytes. The optimal SI cutoff values as well as the overall statistical analysis results were almost same for LLNA and human data (Fig. 4). The combination of the two cytokines with SI cutoff value 3.2 showed the highest predictive capacity with 76.1% sensitivity and 75.6% specificity. Among the 26 test substances following exclusion of equivocal substances, vanillin, a non-sensitizer according to LLNA data, has been reported to have very low skin sensitization ability in humans (Basketter et al., 2014, 2001; Cottrez et al., 2016). As shown in Table 1, the mean SI value of vanillin is below the cutoff to incorrectly categorize it as a non-sensitizer compared to human data, which is why the

predictive capacity based on human had lower sensitivity and specificity than LLNA.

Next, we identified five possible SI cutoff values 1.8, 2.0, 2.2, 3.0 and 3.2 by ROC analysis and validated these values by evaluating their predictive capacities of HaCaSens with same set of test substances as in the LLNA and human data (Table 2). We found that relatively lower SI cutoff (1.8 and 2.0) had lower specificity and negative predictive value, mainly because of the lower threshold for positive prediction, but there was no significant difference among all SI cutoff values in terms of sensitivity and positive predictive value. Also, the cutoff 3.0 produced the highest accuracy of 86.7% based on LLNA data, but all except cutoff 3.2 (80.8%) had the same 84.6% accuracy based on human data (Tables 3 and 4). Taken together, we propose 3.0 as the optimal SI cutoff value for the identification of sensitizers as it yielded the maximal sensitivity plus accuracy, 90.9%, 86.7% and 90.5%, 84.7% based on LLNA and human data respectively.

HaCaSens parameters were also further evaluated with ROC-analysis using the optimal SI cutoff value 3.0. Among the cytokines secreted by the keratinocytes, the pro-inflammatory cytokines IL-1 α and IL-6 increased upon exposure to sensitizers. Thus the production of these two cytokines corresponded to *in vivo* classification data from LLNA (Jung et al., 2016). When we consider the importance and indeed necessity of optimizing the assay to make it simple to perform and to ensure higher predictive capacity as well as intra/inter-laboratory reproducibility, it is vital to thoroughly investigate whether single cytokine or the combination of IL-1 α and IL-6, is sufficient to discriminate sensitizers from non-sensitizers. Herein, we compared the predictive capacity of the assay applying all possible parameters, IL-1 α , IL-6 and the combination of the two cytokines with the same set of test substances. Consistent with the previous results, use of the two cytokines as a parameter led to the highest accuracy 86.7% and 84.6% based on LLNA and human classification, respectively, with corresponding sensitivities 90.9% and 90.5% (Table 5). Considering that HaCaSens is used to initially screen test substances for skin sensitization, identifying a cutoff value with high sensitivity as well as accuracy is a high priority. Thus, the combination of the two cytokines is suggested to be the appropriate parameter despite having the lowest specificity (75.0% and 60.0%) among the parameters. Using two cytokines as a final parameter, it is also possible to implement sequential testing strategy (STS) developed by Kao for prediction of the skin sensitizing potential (Kleinstreuer et al., 2018; Nukada et al., 2013; Takenouchi et al., 2015). When the STS approach is applied to the HaCaSens, chemicals being positive at first cytokine (SI $\geq$ 3.0) are classified as having sensitizing potential with no need to be further tested with another cytokine. If the chemical is negative at first cytokine, then analysis of another cytokine should be performed as a second. This STS-based prediction model is expected to enable the sophistication of the HaCaSens assay in terms of reduced cost as well as labor force.

In conclusion, the present study is the first to use statistical methods to define a SI cutoff value for HaCaSens. We used unbiased ROC curves to evaluate the performance of HaCaSens and showed that SI cutoff value 3.0 with parameter of the combination of IL-1 α and IL-6 had the highest accuracy 86.7% with 90.9% sensitivity and 75.0% specificity compared with other ROC-analysis-based SI cutoff values. Thus the results found in this study confirm the SI value 3.0 as the optimal cutoff for HaCaSens to discriminate sensitizers from non-sensitizers and validated this value with a set of 30 test substances. Given the unmet clinical need for risk assessment, the application of the HaCaSens with this optimized prediction model paves the way for successful identification of sensitizing potential. Further studies must be performed to validate this SI cutoff value and parameter with additional test substances.

5. Transparency document

The Transparency document associated with this article can be

found in the online version.

Acknowledgements

This research was supported by grants (16182MFD523) from the Ministry of Food Drug Safety.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.toxlet.2018.09.004>.

References

- 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.
- 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 sensitization 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., Wright, Z.M., Warbrick, E.V., Dearman, R.J., Kimber, I., Ryan, C.A., Gerberick, G.F., White, I.R., 2001. Human potency predictions for aldehydes using the local lymph node assay. *Contact Derm.* 45, 89–94.
- 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. *Dermatitis* 25, 11–21.
- 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.
- Chung, H., Quan, H., Jung, D., Ravi, G., Cho, A., Kang, M.J., Kim, E., Che, J.H., Lee, E.S., Jeong, T.C., Heo, Y., Seok, S.H., 2018. Intra- and inter-laboratory reproducibility and predictivity of the HaCaSens assay: a skin sensitization test using human keratinocytes, HaCaT. *Toxicol. In Vitro* 46, 304–312.
- Diepgen, T.L., 2003. Occupational skin-disease data in Europe. *Int. Arch. Occup. Environ. Health* 76, 331–338.
- Diepgen, T.L., Ofenloch, R.F., Bruze, M., Bertuccio, P., Cazzaniga, S., Coenraads, P.J., Elsner, P., Goncalo, M., Svensson, A., Naldi, L., 2016. Prevalence of contact allergy in the general population in different European regions. *Br. J. Dermatol.* 174, 319–329.
- 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.
- EU, 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official J. No L 396*, 1–1355.
- EU, 2008a. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending 15 and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. *Off. J. No L 353*, 1–1355.
- EU, 2009. Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products. *Off. J. No L 342*, 1–59.
- 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.
- Jung, D., Che, J.H., Lim, K.M., Chun, Y.J., Heo, Y., Seok, S.H., 2016. Discrimination of skin sensitizers from non-sensitizers by interleukin-1 α and interleukin-6 production on cultured human keratinocytes. *J. Appl. Toxicol.* 36, 1129–1136.
- Kim, D.E., Yang, H., Jang, W.H., Jung, K.M., Park, M., Choi, J.K., Jung, M.S., Jeon, E.Y., Heo, Y., Yeo, K.W., Jo, J.H., Park, J.E., Sohn, S.J., Kim, T.S., Ahn, I.Y., Jeong, T.C., Lim, K.M., Bae, S., 2016. Predictive capacity of a non-radioisotopic local lymph node assay using flow cytometry, LLNA:BrU-FCM: comparison of a cutoff approach and inferential statistics. *J. Pharmacol. Toxicol. Methods* 78, 76–84.
- Kleinstreuer, N.C., Hoffmann, S., Alepee, N., Allen, D., Ashikaga, T., Casey, W., Clouet, E., Cluzel, M., Desprez, B., Gellatly, N., Gobel, C., Kern, P.S., Klaric, M., Kuhn, J., Martinozzi-Teissier, S., Mewes, K., Miyazawa, M., Strickland, J., van Vliet, E., Zang, Q., Petersohn, D., 2018. Non-animal methods to predict skin sensitization (II): an assessment of defined approaches (*). *Crit. Rev. Toxicol.* 48, 359–374.
- McDonald, J.C., Beck, M.H., Chen, Y., Cherry, N.M., 2006. Incidence by occupation and industry of work-related skin diseases in the United Kingdom, 1996–2001. *Occup. Med. (Lond.)* 56, 398–405.
- Natsch, A., 2010. The Nrf2-Keap1-ARE toxicity pathway as a cellular sensor for skin sensitizers—functional relevance and a hypothesis on innate reactions to skin sensitizers. *Toxicol. Sci.* 113, 284–292.
- 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.
- 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.
- OECD test guidelines (TG) 442A. Skin Sensitization: Local Lymph Node Assay: DA. Adopted: 22 July 2010.
- OECD test guidelines (TG) 442B. Skin Sensitization: Local Lymph Node Assay: BrdU-ELISA or -FCM. Adopted: 25 June 2018.
- OECD TG 429. Skin sensitization, local lymph node assay (adopted 22.7. 2010).
- Peiser, M., Tralau, T., Heidler, J., Api, A.M., Arts, J.H., Basketter, D.A., English, J., Diepgen, T.L., Fuhlbrigge, R.C., Gaspari, A.A., Johansen, J.D., Karlberg, A.T., Kimber, I., Lepoittevin, J.P., Liebsch, M., Maibach, H.I., Martin, S.F., Merk, H.F., Platzeck, T., Rustemeyer, T., Schnuch, A., Vandebriel, R.J., White, I.R., Luch, A., 2012. Allergic contact dermatitis: epidemiology, molecular mechanisms, in vitro methods and regulatory aspects. Current knowledge assembled at an international workshop at BfR, Germany. *Cell Mol. Life Sci.* 69, 763–781.
- Piroird, C., Ovigne, J.M., Rousset, F., Martinozzi-Teissier, S., Gomes, C., Cotovio, J., Alepee, 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.
- 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.
- 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.
- Reisinger, K., Hoffmann, S., Alepee, 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., Martinozzi-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, 259–270.
- 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.
- Takenouchi, O., Fukui, S., Okamoto, K., Kurotani, S., Imai, N., Fujishiro, M., Kyotani, D., Kato, Y., Kasahara, T., Fujita, M., Toyoda, A., Sekiya, D., Watanabe, S., Seto, H., Hirota, M., Ashikaga, T., Miyazawa, M., 2015. Test battery with the human cell line activation test, direct peptide reactivity assay and DEREK based on a 139 chemical data set for predicting skin sensitizing potential and potency of chemicals. *J. Appl. Toxicol.* 35, 1318–1332.
- 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.

Supplementary Table 1. Comparison of SI cutoff values of the HaCaSens according to sensitizing potency

No	Substances	<i>In vivo</i> potency	Reference	HaCaSens	Max SI values (Mean±SD)
1	Isopropanol	Non-sensitizer (8 substances)	OECD TG 429 (2010), TG 442D (2015), TG 442E (2016)	-	2.4±0.8
2	Chlorobenzene		OECD TG 429 (2010)	+	3.4±4.7
3	Salicylic acid		OECD TG 429 (2010), TG 442D (2015)	-	2.0±0.6
4	4-Hydroxybenzoic acid		Nukada Y. et al. (2012) Tsujita-Inoue K. et al. (2014)	-	1.6±0.6
5	Methyl salicylate		OECD TG 429 (2010)	+	8.4±3.6
6	4-Hydroxybenzoic acid		Nukada Y. et al. (2012) Tsujita-Inoue K. et al. (2014)	-	1.6±0.6
7	Vanillin		Nukada Y. et al. (2012) Tsujita-Inoue K. et al. (2014)	-	2.0±0.5
8	Glycerol		OECD TG 442D (2015), TG 442E (2016),	-	1.7±0.4
9	2,3-Butanedione	Weak (8 substances)	Natsch A. et al.(2009) Nukada Y. et al. (2012)	-	1.5±0.7
10	Cinnamic alcohol		OECD TG 429 (2010), TG 442D (2015)	+	6.6±0.7
11	Benzocaine		Arkusz J. et al.(2010) Cho S. et al.(2013)	-	1.2±0.3
12	Methyl methacrylate		OECD TG 429 (2010)	+	30.3±33.8
13	Ethylene glycol dimethacrylate		OECD TG 429 (2010), TG 442D (2015)	+	49.6±18.8
14	Imidazolidinyl urea		OECD TG 429 (2010), TG 442E (2016)	+	22.8±13.4
15	Phenyl benzoate		OECD TG 429 (2010)	+	11.6±10.8

16	Eugenol			OECD TG 429 (2010)	+	4.9±1.2
17	Hexyl aldehyde	cinnamic	Moderate/Weak (2 substances)	OECD TG 429 (2010)	+	70.2±80.1
18	Citral			OECD TG 429 (2010)	+	11.5±13.3
19	Tetramethylthiuram disulfide		Moderate (6 substances)	Ashikaga et al., 2010 Natsch A. et al.(2013)	+	16.2±15.3
20	Nickel sulfate			OECD TG 442E (2016)	+	5.1±2.0
21	Cinnamaldehyde			Natsch A. et al.(2013) Ramirez T. et al. (2014)	+	24.5±32.5
22	Benzylideneacetone			OECD TG 442C (2015)	+	82.3±103.4
23	Isoeugenol			OECD TG 429 (2010)	+	3.1±1.0
24	2-Mercapto benzothiazole			OECD TG 429 (2010), TG 442D (2015), TG 442E (2016),	+	13.0±9.5
25	Formaldehyde		Strong (6 substances)	OECD TG 442C (2015),	+	8.5±3.8
26	4-Phenylenediamine			OECD TG 429 (2010), TG 442E (2016)	+	41.5±62.9
27	Glutaraldehyde			Basketter D. et al. (2008) Natsch A. et al.(2013)	+	13.5±12.5
28	Cobalt chloride			OECD TG 429 (2010)	+	3.3±1.9
29	Methyldibromo glutaronitrile			OECD TG 442D (2015)	+	55.5±142.0
30	CMI/MI (Kathon™ CG Preservative)			OECD TG 429 (2010)	+	45.0±52.2