

Original Article

The dermal sensitization threshold (DST) approach for mixtures evaluated as negative in *in vitro* test methods; mixture DST

Taku Nishijo, Masaaki Miyazawa, Kazutoshi Saito, Yuki Otsubo, Hideyuki Mizumachi
and Hitoshi Sakaguchi

Safety Science Research Laboratories, Kao Corporation, 2606 Akabane, Ichikai, Haga, Tochigi, 321-3497, Japan

(Received October 1, 2018; Accepted November 6, 2018)

ABSTRACT — Cosmetic ingredients often comprise complex mixtures, such as botanical extracts, which may contain skin sensitizing constituents. In our previous study for the sensitivity of the evaluations of skin sensitizing constituents in mixtures using the binary *in vitro* test battery with KeratinoSens™ and h-CLAT, some sensitizers showed higher detection limits in *in vitro* test methods than in murine local lymph node assays (LLNA). Thus, to minimize the uncertainty associated with decreased sensitivity for these sensitizers, a risk assessment strategy was developed for mixtures with negative results from the binary test battery. Assuming that the no expected sensitization induction level of mixtures (mixture NESIL) can be derived for mixtures with negative *in vitro* test results, we assessed 146 sensitizers with *in vitro* and LLNA data according to the assumption of indeterminate constituents in mixtures. Finally, we calculated 95th percentile probabilities of mixture NESILs and derived dermal sensitization thresholds for mixtures (mixture DST) with negative *in vitro* test results of 6010 µg/cm². Feasibility studies indicated that this approach was practical for risk assessments of products in the cosmetic industry. This approach would be a novel risk assessment strategy for incorporating the DST approach and information from *in vitro* test methods.

Key words: Skin sensitization, Mixture, Mixture DST, KeratinoSens™, h-CLAT

INTRODUCTION

Allergic contact dermatitis (ACD) is a delayed-type hypersensitivity reaction of the skin, caused by contact with skin sensitizing substances. Because ACD frequently decreases quality of life for workers and consumers, a standard requirement of cosmetic ingredients is that their potential to induce skin sensitization is assessed under product-use conditions that may lead to ACD in humans. Skin-sensitizing potentials of ingredients are typically identified using animal experiments, such as the guinea pig maximization test (Magnusson and Kligman, 1969) and murine local lymph node assays (LLNA; Basketter *et al.*, 2002). Non-animal test methods, however, are desirable due to regulatory requirements and ethical issues.

In the cosmetic industry, various complex mixtures have been used. Among these, botanical extracts are frequently extracted and solubilized into vehicles, including water, ethanol, and 1,3-butylene glycol, and some contain

skin-sensitizing constituents that may contribute to ACD (Kiken and Cohen, 2002; Lalko and Api, 2006).

Non-animal test methods for skin sensitization have been previously developed for the first three key events (protein binding, keratinocyte activation, and dendritic-cell activation) of the adverse outcome pathway (AOP) of the sensitization induction phase (OECD, 2012). Among non-animal test methods that are accepted in the OECD test guidelines, KeratinoSens™ is addressing keratinocyte activation (Emter *et al.*, 2010), and the human Cell Line Activation Test (h-CLAT) is addressing dendritic-cell activation (Ashikaga *et al.*, 2010). These methods are technically applicable to the testing of multi-constituent substances and mixtures. As a proof-of-concept for the applicability of the binary test battery with KeratinoSens™ and h-CLAT (Otsubo *et al.*, 2017) to mixtures such as botanical extracts, the detection limits (DL) of sensitizing constituents in mixtures have been defined (Nishijo *et al.*, 2019). Positive results from these test methods indicate

that a skin sensitizer may be present in the tested mixture at a concentration that exceeds the respective DL. For example, if a tested mixture gives a positive result in the LLNA after application to mouse ears at 100%, the concentration of the sensitizer in the mixture likely exceeds the EC3 value, which is the estimated concentration of the sensitizer that produces a stimulation index of 3 (Fig. 1). Herein, we developed a feasible skin-sensitization testing strategy using nonanimal test methods. Following positive results in *in vitro* tests of mixtures, we compared DLs of skin sensitizers in mixtures between LLNA and *in vitro* test methods. We identified 146 sensitizers with positive results in KeratinoSens™, h-CLAT, and LLNA, and considered these as potential constituents in mixtures. We showed that 86% of the analyzed sensitizers had DLs in the binary test battery that were lower than those in LLNA, and 14% of these had higher DLs in the binary test battery than in LLNA. When sensitizers had $\geq$ DL in LLNA but $<$ DL in the binary test battery the mixture may give a false negative result on the binary test battery versus LLNA. To evaluate mixtures using KeratinoSens™ and h-CLAT and to formulate into cosmetic products, the presence of these sensitizers should be considered.

The risk of skin sensitization to a particular ingredient in a product under specific usage conditions can be determined based on the evaluation of hazard (i.e., the sensitization potency which is inversely proportional to a no effect level (No Expected Sensitization Induction Level, NESIL)) and the quantification of the actual consumer exposure. For skin sensitization assessment methods without animal testing, the concept of Dermal Sensitization Threshold (DST) approach has been proposed where skin exposure is low (Safford *et al.*, 2008, 2011, 2015). The DST approach is based on the princi-

ples of the Threshold of Toxicological Concern (TTC) concept which is a risk assessment approach intended to identify exposure levels below which there is a low probability of risk to human health (Rulis, 1986), such that a level can be determined for any chemical where there is no appreciable risk of skin sensitization, even when the sensitization hazard of chemical has not been identified. It has been proposed that non-reactive DST value ($900 \mu\text{g}/\text{cm}^2$) or reactive DST value ($64 \mu\text{g}/\text{cm}^2$), which should be selected based on chemical structures and properties, and were values are used as default NESILs for chemicals for which sensitization data had not been determined (Safford *et al.*, 2015). A strategy using non-target analytical screening techniques and the DST approach has been proposed (Koster *et al.*, 2015; Antignac *et al.*, 2011), however, non-target qualitative and qualitative screening analyses of skin-sensitizing constituents in complex mixtures such as ingredient containing indeterminate impurities or botanical extract are often difficult and resource intensive. Hence, the existing DST approach is not applicable to mixtures with constituents that lack chemical structural information. Moreover, skin-sensitization risks of constituents in complex mixtures could not be assessed by exposure assessments alone, even if skin exposures are low.

To address this problem, we developed a risk assessment strategy for undetermined constituents of complex mixtures by combining exposure assessment in addition to the hazard identification using the binary test battery with KeratinoSens™ and h-CLAT. Initially, we assumed that the concentration of the sensitizer in the mixture was negative in the binary test battery despite the likely presence of sensitizers at low concentrations. When the tested mixture is positive in sensitization tests, it is expected that the concentration of the sensitizer exceeds the DL.

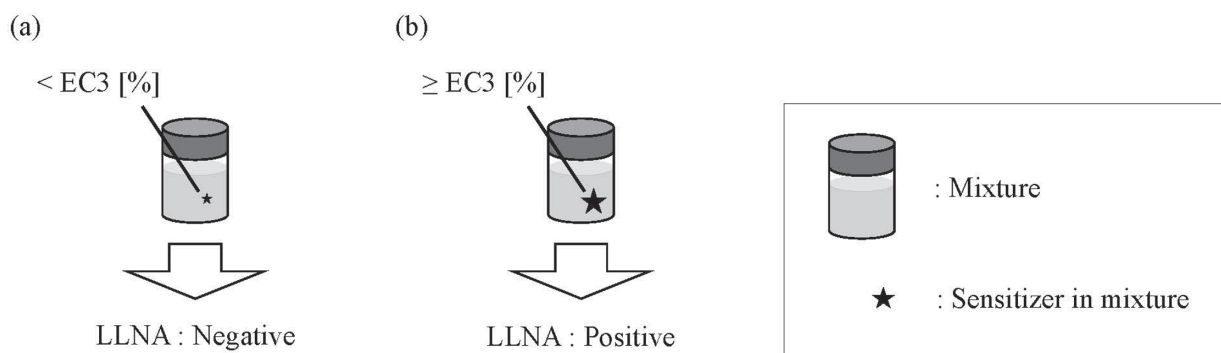


Fig. 1. Expected results from murine local lymph node assays (LLNA) based on the concentration of a sensitizer present in the mixture; (a) If the concentration of a sensitizer in the mixture is below the EC3 value, the mixture is expected to be negative in LLNA. (b) If the concentration of a sensitizer in the mixture exceeds the EC3 value, the mixture is expected to be positive in LLNA.

If the concentration of the sensitizer in the mixture can be estimated, the exposure level of the sensitizer can be calculated based on the exposure level of the mixture. Using thresholds for the induction of skin sensitization (Kimber *et al.*, 2008), we then derived NESILs for each sensitizer. Exposure levels below the NESIL of a sensitizer should fail to induce skin sensitization, even if the sensitizer is present in the mixture. Moreover, if the concentration of the sensitizer in the mixture can be estimated, the NESIL of the mixture can be calculated based on that of the sensitizer. Thus, we defined the mixture NESIL as an exposure level that does not induce skin sensitization according to the NESIL of the sensitizer. If a tested mixture is negative in the binary test battery, the mixture NESIL can be calculated based on the DL from the same analyses. Herein, mixture NESILs were calculated for 146 sensitizers for which *in vitro* and LLNA data were available, and that were assumed as indeterminate constituents in mixtures. A DST value was derived for mixture that was negative in the binary test battery by taking the 95th percentile of calculated mixture NESILs. Finally, we performed feasibility studies of this mixture DST value as an acceptable formulation ratio of mixture evaluated as negative in the binary test battery and performed the mixture DST approach for practical product examples.

MATERIALS AND METHODS

Chemicals

A compilation of currently available data was retrieved from published datasets (Otsubo *et al.*, 2017; Takenouchi *et al.*, 2015; Urbisch *et al.*, 2015; Jaworska *et al.*, 2015). We then selected a total of 147 chemicals with original references that had been evaluated and were positive in LLNA and the binary test battery. Benzo(a)pyrene was excluded from the dataset because its EC3 value came from an extrapolation with a high stimulation index (> 15), suggesting unreliability (Kern *et al.*, 2010). Collectively, a total of 146 chemicals were analyzed (Supplemental Table 1) in this study.

In vitro test methods

KeratinSensTM: OECD test guideline 442D

KeratinSensTM addresses keratinocyte activation according to activation of the keap1–Nrf2–ARE pathway (Emter *et al.*, 2010) and is a reporter cell-based assay using HaCaT cells that exploits luciferase activity as an indicator of ARE activation. *KeratinSensTM* is technically applicable to the testing of multiconstituent substances and mixtures.

HaCaT cells were incubated with test chemicals for

48 hr, and luciferase activity was measured using a luminometer. Test chemicals were classified as positive if luciferase activity was significantly increased by ≥ 1.5 -fold compared with vehicle control for test chemicals with no defined molecular weight. The concentration of sensitizer that leads to a 1.5-fold induction of luciferase (EC1.5) was then calculated using linear interpolation.

h-CLAT: OECD test guideline 442E

h-CLAT addresses the activation of dendritic cells, which are known to up-regulate the expression of the cell surface proteins CD86 and CD54 following exposure to sensitizers. In the present experiments, THP-1 cells were used to represent dendritic cells (Ashikaga *et al.*, 2006, 2010; Sakaguchi *et al.*, 2006) and were treated with test chemicals for 24 hr at concentrations that were based on predetermined concentrations that yield 75% cell viability (CV75). After staining with fluorescence-labeled anti-CD86 and anti-CD54 antibodies, expression levels were determined using flow cytometry. Relative fluorescence intensities (RFI) were compared with those of the vehicle control using flow cytometry. Chemicals were classified as positive when CD86 RFI was $\geq 150\%$ or CD54 RFI was $\geq 200\%$, and the corresponding concentrations (EC150 for CD86 or EC200 for CD54) were calculated as for EC3 values from LLNA. The lower EC value of EC150 or EC200 was chosen as the minimal induction threshold (MIT). *h-CLAT* is technically applicable to the testing of multi-constituent substances and mixtures.

DLs of skin sensitizers in mixtures

DLs of skin sensitizers in mixtures were calculated from exposure concentrations that exceeded the positive criteria for each test method, representing EC3 values from LLNA, EC1.5 values from *KeratinSensTM*, and MIT values from *h-CLAT* (Table 1). DLs of skin sensitizers that were assumed present in the mixtures for LLNA were calculated as EC3 values if the tested mixture gave positive LLNA results following application to ears at 100%. For non-cytotoxic mixtures with unknown molecular weights, desired final concentrations of the test chemicals were identified using serial dilutions from a maximum concentration of 400 $\mu\text{g/mL}$ for *KeratinSensTM*. Finally, mixtures were diluted by a factor of 2500 for testing. Thus, the DLs of skin sensitizers in mixtures were calculated as $\text{EC1.5} \times 2500 \mu\text{g/mL}$. Similarly, serial dilutions from a maximum concentration of 5000 $\mu\text{g/mL}$ were prepared for *h-CLAT*. Mixtures were then diluted by a factor of 200 for testing. DLs of skin sensitizers in the mixture were calculated as $\text{MIT} \times 200 \mu\text{g/mL}$. Lower DL values from *KeratinSensTM*

Table 1. DLs of skin sensitizers that may be present in mixtures were calculated from estimated concentrations that exceeded the positive criteria for each test method (EC3, EC1.5, and MIT).

Sensitization test	Estimated concentration exceeding positive criteria	Maximum tested concentration	Dilution rate	Detection limit of skin sensitizer in water soluble and non-cytotoxic mixture
LLNA	EC3	100 [%]	1	EC3 [%]
KeratinoSens™	EC1.5	400 [µg/mL]	2500	EC1.5 × 2500 [µg/mL]
h-CLAT	MIT	5000 [µg/mL]	200	MIT × 200 [µg/mL]

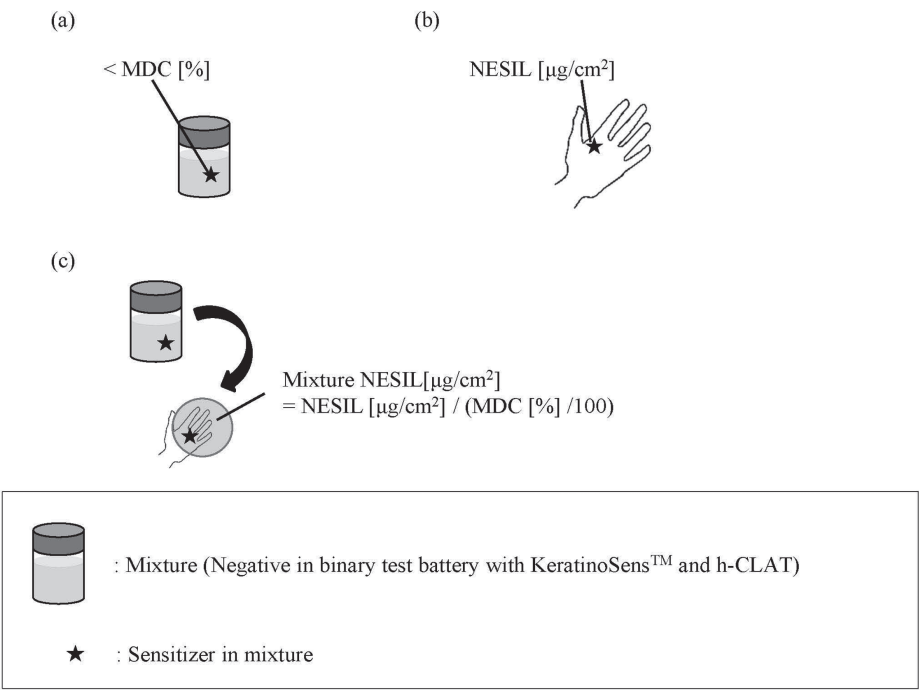


Fig. 2. Derivation of mixture NESIL values; (a) for a mixture evaluated as negative in the binary test battery, it is expected that the concentration of the constituent sensitizer is less than the DL for the binary test battery. (b) NESIL values are derived from intrinsic sensitization potencies of each sensitizer and are estimated from corresponding EC3 values by extrapolating as EC3 × 250 µg/cm². (c) Taking account of (a) and (b), we derived NESILs of mixtures evaluated as negative in the binary test battery (mixture NESIL) is derived.

Sens™ and h-CLAT analyses were defined as the minimal detection concentration (MDC) for the DL of the binary test battery. Subsequently, ratios of MDC/EC3 were calculated for 146 sensitizers, and log(10) ratios were plotted in a probability distribution.

For complex mixtures, it may be difficult to identify constituent skin sensitizers. Nonetheless, we assumed that skin sensitizers that were positive in *in vitro* tests may be present in mixtures, that the mixtures were water soluble and noncytotoxic, and that they did not have a masking or boosting effect on the DLs of their constituent skin sensitizers.

NESIL of mixture evaluated as negative in the binary test battery with KeratinoSens™ and h-CLAT; Mixture NESIL

The present risk-assessment strategy was developed with consideration of exposure assessments and the hazard identification using the binary test battery (Fig. 2). The major elements that were used to derive mixture NESILs are briefly described as follows:

1. For a mixture evaluated as negative in the binary test battery at all tested concentration (≤ 400 µg/mL in KeratinoSens™ and ≤ 5000 µg/mL in h-CLAT), the concentration of sensitizer is expected to be less than the MDC value

(DL for the binary test battery; Fig. 2a).

2. There exist thresholds for the induction of skin sensitization (Kimber *et al.*, 2008), and NESILs were derived from intrinsic sensitization potencies of each sensitizer. NESILs were then estimated from EC3 values by extrapolating $EC3 \times 250 \mu\text{g}/\text{cm}^2$ based on the application of 25- μL test solutions to 1 cm^2 areas of mouse ears (Fig. 2b). Because sensitization potencies in mouse LLNA correlated well with those determined in human HRIPT (Basketter *et al.*, 2005, 2018), no additional extrapolation was made to accommodate differences between mice and humans, as in a previous DST study (Safford *et al.*, 2008, 2011, 2015).

3. Considering the above elements, NESILs of mixtures evaluated as negative in the binary test battery (mixture NESIL) were derived as follows (Fig. 2c):

$$\text{Mixture NESIL } [\mu\text{g}/\text{cm}^2] = \text{NESIL } [\mu\text{g}/\text{cm}^2 \text{ (of putative sensitizer in mixture)}] / (\text{MDC} / 100)$$

Mixture NESILs are those of the whole mixture and are defined as the exposure level that produces no skin sensitization mixture (i.e., not to exceed NESIL of the putative sensitizer in mixture). Thus the mixture NESIL is not applicable to the sensitizing constituent.

In addition, the mixture NESIL is calculated from MDC values for each sensitizer (DL from the binary test battery) and from corresponding NESILs, which were estimated using EC3 values from LLNA. Accordingly, the mixture NESIL is inversely proportional to the MDC/EC3 ratio, as indicated by the following equations:

$$\text{NESIL} = EC3 \times 250 \mu\text{g}/\text{cm}^2 \quad (1)$$

$$\text{Mixture NESIL } [\mu\text{g}/\text{cm}^2] = \text{NESIL } [\mu\text{g}/\text{cm}^2 \text{ (of putative sensitizer in mixture)}] / (\text{MDC}/100) \quad (2)$$

Based on (1) and (2), the mixture NESIL $[\mu\text{g}/\text{cm}^2] = EC3 \times 250 \mu\text{g}/\text{cm}^2 / (\text{MDC}/100) = EC3/\text{MDC} \times 25000$

Thus, $\log(10)$ MDC/EC3 and negative $\log(10)$ mixture NESIL values have similar distributions.

For example, mixture containing sensitizer A is summarized as follows:

Sensitizer A,

$$EC3 = 0.1\% \quad (25 \mu\text{g}/\text{cm}^2 = 0.1 \times 250)$$

$$\text{MDC (DL for the binary test battery)} = 0.5\%$$

The mixture NESIL could be calculated as $5000 \mu\text{g}/\text{cm}^2 = 25 \mu\text{g}/\text{cm}^2 / (0.5 / 100)$. Accordingly, the exposure level of sensitizer A is expected to be less than $25 \mu\text{g}/\text{cm}^2$ if the exposure level of the mixture is less than $5000 \mu\text{g}/\text{cm}^2$. Mixture NESILs were calculated for 146

analyzed sensitizers, and negative $\log(10)$ values were plotted in a probability distribution.

Probability distribution analyses

Normal and gamma probability distributions were fitted for $\log(10)$ values of MDC/EC3 ratios and negative $\log(10)$ values of mixture NESILs. Mixture DST values are those of mixtures evaluated as negative values in the binary test battery and were derived by determining 95th percentiles of mixture NESILs.

Quantitative risk assessment (QRA)

In the QRA approach, an acceptable exposure level (AEL) in $\mu\text{g}/\text{cm}^2/\text{day}$ is estimated to protect consumers from skin sensitization (Api *et al.*, 2008). To derive AELs, NESILs were modified using various sensitization assessment factors (SAFs; Felter *et al.*, 2002, 2003; Basketter and Safford, 2016). Generally, NESILs are derived from EC3 values in LLNA or from no observed effect levels (NOEL) in human sensitization tests, such as the human repeated insult patch test (HRIPT) and human maximization test (HMT). SAFs were applied to extrapolate NESILs from controlled experimental conditions to levels that were safe during exposure of consumers to products. The consumer exposure level (CEL) is an essential element of QRA and was determined in exposure assessments of frequencies of use, habits, practices, durations of use, and amounts of product used per application. If $AEL > CEL$ for a particular ingredient, the proposed exposure to the ingredient is considered acceptable in terms of the risk of skin sensitization.

Case studies for practical examples in the cosmetic industry

To assess the practical feasibility of the mixture DST approach, the mixture DST needs to be converted to a product formulation ratio by considering specific products and exposure situations. Shower gel, rinse-off hair conditioner, body lotion, hand cream and deodorant non-spray were selected as practical examples of rinse-off and leave-on products in the cosmetic industry. CELs of these products were derived from SCCS (2016). For each product example, acceptable formulation ratios of mixtures evaluated as negative in the binary test battery were calculated using the mixture DST approach.

RESULTS

Ratios of MDC/EC3

In the histogram of $\log(10)$ MDC/EC3 ratios for 146 analyzed sensitizers (Fig. 3), 14% (21/146) of the ana-

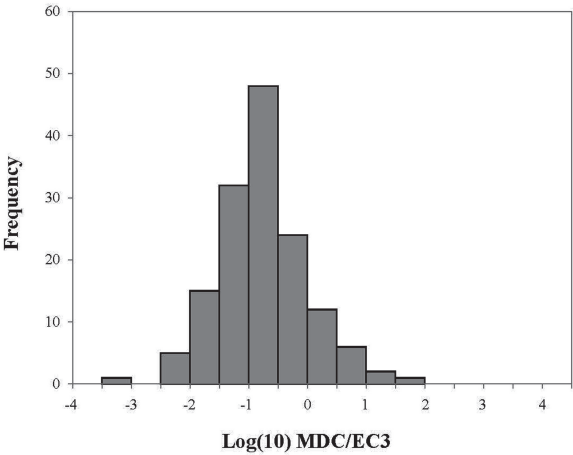


Fig. 3. Histogram of log(10) MDC/EC3 values for 146 analyzed sensitizers.

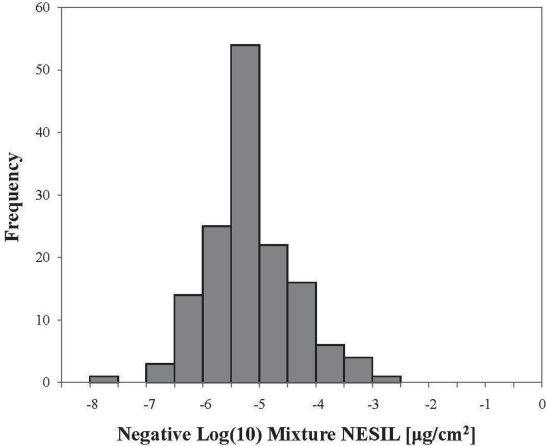


Fig. 4. Histogram of negative log(10) mixture NESIL for 146 analyzed sensitizers.

Table 2. 95th percentiles of mixture NESILs from the distributions shown in Fig. 5.

	Normal Distribution		Gamma Distribution	
	Negative Log(10) mixture NESIL [μg/cm²]	Mixture NESIL [μg/cm²]	Negative Log(10) mixture NESIL [μg/cm²]	Mixture NESIL [μg/cm²]
95th Percentile	-3.90	7894	-3.78	6010

lyzed sensitizers with positive log(10) MDC/EC3 ratios had higher DLs in the binary test battery than in LLNA (Nishijo *et al.*, 2019). Of the 146 analyzed sensitizers, that with the highest MDC/EC3 ratio (37.3) was maleic anhydride.

DST of mixtures evaluated as negative in the binary test battery with KeratinoSens™ and h-CLAT; Mixture DST

The histogram of negative log(10) mixture NESIL values are shown in Fig. 4. As described above, the mixture NESIL is inversely proportional to the MDC/EC3 ratio. Thus, log(10) MDC/EC3 and negative log(10) mixture NESIL values had similar distributions (Fig. 3 and 4). And cumulative probability distributions of negative log(10) mixture NESIL values are shown in Fig. 5. Normal (Fig. 4a) and gamma (Fig. 4b) distributions fitted well with respective experimental data. The mixture NESIL values corresponding to the percentiles from the distributions depicted in Fig. 4 are listed (Table 2).

Comparing the actual data points to the fitted theoretical data points in a probability plot (a graphical technique to determine the goodness of fit of the data set to a spe-

cific distribution) showed that the gamma distribution fits better at the top end (i.e. the lower mixture NESIL) (data not shown). Thus, a gamma distribution was used to derive the mixture DST. Because the mixture DST value of 6010 μg/cm² was calculated as the 95th percentile of mixture NESIL values from 146 analyzed sensitizers, it is expected that 5% of sensitizers have lower mixture NESILs than this value. Among the 146 analyzed sensitizers, nine had lower mixture NESILs than 6010 μg/cm² (Table 3). A workflow for the mixture DST approach for mixtures evaluated as negative in the binary test battery is proposed in Fig. 6.

Case studies

Shower gel

The CEL value of 0.011 μg/cm²/day for shower gel from SCCS (2016) was used in this study. The acceptable exposure level of mixture evaluated as negative in the binary test battery (mixture AEL, 60.1 μg/cm²) was determined by dividing the mixture DST of 6010 μg/cm² by SAF = 100 (Api *et al.*, 2008). Therefore, the acceptable formulation ratio of the mixture was calculated as > 100% (60.1 / (0.011 × 1000)); Table 4).

Dermal sensitization threshold for mixtures

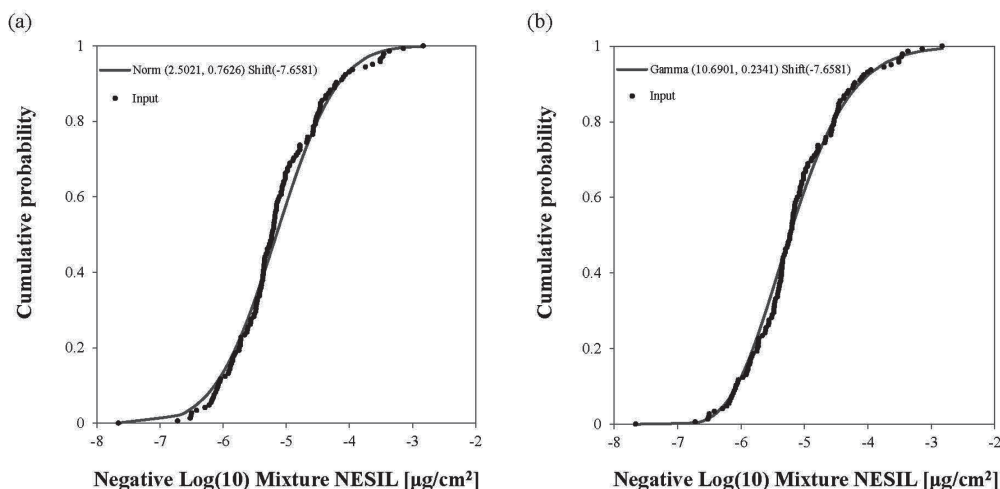


Fig. 5. Cumulative probability distribution of negative log(10) mixture NESIL values from the 146 analyzed sensitizers.

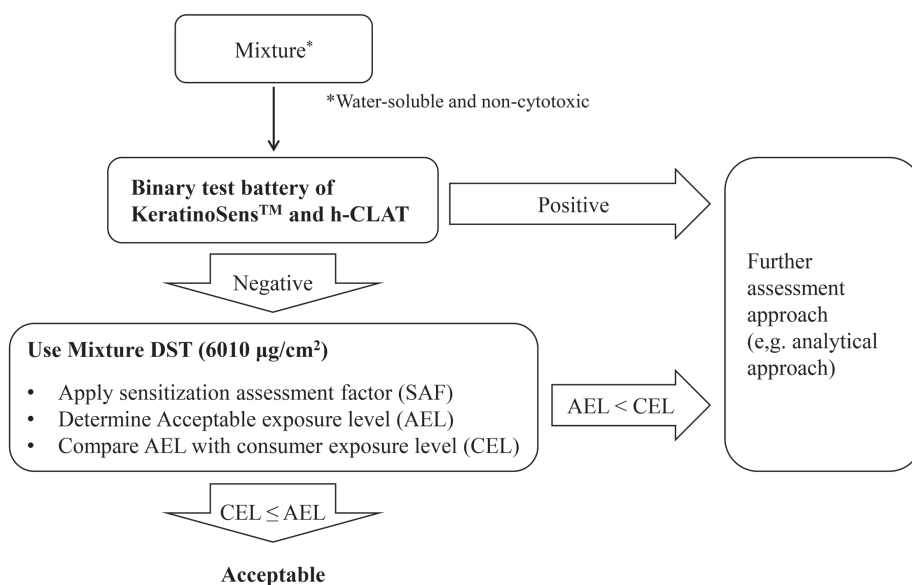


Fig. 6. Proposed workflow for the Dermal Sensitization Threshold (mixture DST) approach for mixtures evaluated as negative in the binary test battery with KeratinoSens™ and h-CLAT. Mixture is considered as negative in the binary test battery if negative results are obtained at all tested concentration ($\leq 400 \mu\text{g/mL}$ in KeratinoSens™ and $\leq 5000 \mu\text{g/mL}$ in h-CLAT).

Rinse-off hair conditioner

The CEL of $0.027 \mu\text{g/cm}^2/\text{day}$ for rinse-off hair conditioner was retrieved from SCCS (2016). A mixture AEL of $60.1 \mu\text{g/cm}^2$ was determined by dividing the mixture DST of $6010 \mu\text{g/cm}^2$ by $\text{SAF} = 100$ (Api *et al.*, 2008). Therefore, the acceptable formulation ratio of the mixture was calculated as $> 100\%$ ($60.1 / (0.027 \times 1000)$; Table 4).

Body lotion

The CEL of $0.50 \mu\text{g/cm}^2/\text{day}$ for body lotion was retrieved from SCCS (2016). A mixture AEL of $60.1 \mu\text{g/cm}^2$ was determined by dividing the mixture DST of $6010 \mu\text{g/cm}^2$ by $\text{SAF} = 100$ (Api *et al.*, 2008). Therefore, the acceptable formulation ratio of the mixture was calculated as $> 12.0\%$ ($60.1 / (0.50 \times 1000)$; Table 4).

Table 3. Nine sensitizers had lower mixture NESILs than the mixture DST of 6010 $\mu\text{g}/\text{cm}^2$ based on 95th percentiles in the gamma distributions shown in Fig. 5.

Chemicals	CAS	LLNA EC3 [%]	NESIL [$\mu\text{g}/\text{cm}^2$]	Binary test battery MDC [%]	Mixture NESIL [$\mu\text{g}/\text{cm}^2$]
Maleic anhydride	108-31-6	0.16	40	5.97	670
Oxazolone	15646-46-5	0.003	0.75	0.05	1384
Allyl phenoxyacetate	7493-74-5	3.1	775	33.51	2313
MCI/MI	26172-55-4 & 2682-20-4	0.005	1.25	0.04	2828
Diphenylcyclopropenone	886-38-4	0.003	0.75	0.03	2909
chloramine T	127-65-1	0.6	150	4.68	3203
2-Methyl-4H,3,1-benzoxazin-4-one	525-76-8	0.7	175	5.44	3217
Hexyl salicylate	6259-76-3	0.18	45	1.05	4267
1,4-Benzoquinone	106-51-4	0.0099	2.48	0.04	5525

Hand cream

The CEL of 2.5 $\mu\text{g}/\text{cm}^2/\text{day}$ for hand cream was retrieved from SCCS (2016). A mixture AEL of 60.1 $\mu\text{g}/\text{cm}^2$ was determined by dividing the mixture DST of 6010 $\mu\text{g}/\text{cm}^2$ by $\text{SAF} = 100$ (Api *et al.*, 2008). Therefore, the acceptable formulation ratio of the mixture was calculated as $> 2.4\%$ ($60.1 / (2.5 \times 1000)$; Table 4).

Deodorant non-spray

The CEL of 7.5 $\mu\text{g}/\text{cm}^2/\text{day}$ for deodorant non-spray was retrieved from SCCS (2016). A mixture AEL of 20.0 $\mu\text{g}/\text{cm}^2$ was determined by dividing the mixture DST of 6010 $\mu\text{g}/\text{cm}^2$ by $\text{SAF} = 300$ (Api *et al.*, 2008). Therefore, the acceptable formulation ratio of the mixture was calculated as $> 0.3\%$ ($20.0 / (7.5 \times 1000)$; Table 4).

DISCUSSION

For skin sensitization assessment methods of complex mixtures such as botanical extracts without animal testing, a strategy using analytical screening techniques has been proposed (Koster *et al.*, 2015; Antignac *et al.*, 2011). However, non-target qualitative and qualitative screening analyses of skin-sensitizing constituents in complex mixture are often difficult and resource intensive. To address this problem, we developed a risk assessment strategy using *in vitro* test methods. In our previous sensitivity study of the binary test battery with KeratinoSensTM and h-CLAT for detecting sensitizing constituents in mixtures, some sensitizers showed higher DLs in the binary test battery than in LLNA (Nishijo *et al.*, 2019). We developed a risk assessment strategy for mixtures that had been evaluated using the binary test battery. To this end, the risk that the tested mixture contains sensitizers at concentrations below the DLs of the binary test

battery should be considered, even if the tested mixtures have negative values in the binary test battery. There exist thresholds for the induction of skin sensitization and each sensitizer has NESIL. Thus we focused that the NESIL of mixture (Mixture NESIL) can be calculated based on the NESIL of sensitizer in mixture if the concentration of the sensitizer in the mixture could be estimated. If the tested mixture was negative in the binary test battery, the mixture NESIL can be calculated based on the corresponding DL. In this way, mixture NESILs were calculated for sensitizers that were assumed indeterminate constituents of mixture. We also assumed that the mixtures were water soluble and noncytotoxic and that the DLs of skin sensitizers were not masked in the mixtures. Subsequently, using 95th percentiles of calculated mixture NESILs, we calculated a corresponding DST of 6010 $\mu\text{g}/\text{cm}^2$ for mixtures evaluated as negative in the binary test battery.

The mixture DST is defined as a threshold at which the whole mixture does not induce skin sensitization. However, this value is not applicable to each constituent, thus distinguishing this approach from TTC and DST. In addition, the mixture NESIL needs to be applied for just mixture evaluated as negative in the binary test battery. This is an essential difference from existing DST approaches, which give nonreactive and reactive DST values of 900 and 64 $\mu\text{g}/\text{cm}^2$, respectively, following selection based on chemical structures and chemical property information. On the other hand, the mixture DST approach can be applied to the complex mixture such as botanical extract containing indeterminate constituent or impurity without chemical information as long as the mixture is negative in the binary test battery. It was suggested that the mixture DST approach could be a novel risk assessment strategy using the DST approach incorporating the information from *in vitro* test methods.

Dermal sensitization threshold for mixtures

Table 4. Acceptable formulation ratios of mixture evaluated as negative in the binary test battery based on the mixture DST approach for practical product examples.

Product type	Estimated daily amount applied* [g]	Retention factor*	Calculated daily exposure* [g/day]	Skin surface area involved*		CEL* [mg/cm ² /day]	SAF**	Mixture AEL	Acceptable formulation ratio in product [%]
				Surface area [cm ²]	Parameters (if specified)				
Shower gel	18.67	0.01	0.19	17500	total body area	0.011	100	60.1	>100
Hair conditioner	3.92	0.01	0.04	1440	area hands + 1/2 area head	0.027	100	60.1	>100
Body lotion	7.82	1.0	7.82	15670	area body - area hand female	0.50	100	60.1	12.0
Hand cream	2.16	1.0	2.16	860	area hands	2.5	100	60.1	2.4
Deodorant non-spray	1.50	1.0	1.50	200	both axillae	7.5	300	20.0	0.3

* The data was retrieved from SCCS, 2016.

**SAF was retrieved from Api *et al.*, 2008

In the present case study, using the mixture DST approach for practical product examples, acceptable formulation ratios of mixture evaluated as negative in the binary test battery were calculated as $> 100\%$ for rinse-off products and 0.3 - 12 percent for leave-on products, with the exception of circumstances such as aggregate exposure. It was suggested that the mixture DST is a practical approach for risk assessments of many products in the cosmetic industry. If for assessed mixture the relationship was $CEL > AEL$ from mixture DST, further assessment approach (e.g. reduce indeterminate constituents by performing analytical approach) is needed. Qualitative and qualitative analyses of sensitizing constituents in mixtures would permit to perform the existing DST approach (Safford *et al.*, 2015) if skin exposure to sensitizing constituent was low.

It should be considered that there are uncertainties in this study due to limited analyzed sensitizers against indeterminate constituents with illimitable variety in complex mixture. It should be remembered that the mixture DST approach, however, is probabilistic and may not be protective against all sensitizing constituents, as is also the case for TTC and DST approaches. Because the mixture DST was set at the 95th percentile, it follows that there is a 5% probability that a sensitizing constituent causes skin sensitization. Based on an examination of the European List of Notified Chemical Substances database, it was previously suggested that sensitizers comprise approximately 20% of all chemicals (Safford, 2008). This figure was also validated in an examination of the database of Annex I of Directive 67/548 EEC, the European Flavour and Fragrance Association database, and the IFRA/RIFM dataset of fragrance ingredients (Keller *et al.*, 2009). Given this assumption, to achieve a 95% chance of not falling below the threshold, the DST was devised around 75th percentiles, corresponding with 25% probabilities that an untested chemical might be a sensitizer among analyzed sensitizers ($1 - (0.20 \times 0.25) = 0.95$; Safford, 2008). In contrast, our probabilistic assessment uses the 95th percentile of mixture NESIL values from analyzed sensitizers. This is more conservative than taking the 95th percentile of all substances and is equivalent to a 1% ($20\% \times (1 - 0.95)$) probability that indeterminate constituents in mixtures will cause skin sensitization. Of 146 analyzed sensitizers, nine had mixture NESILs of less than $6010 \mu\text{g}/\text{cm}^2$ (Table 3). Hexyl salicylate was falsely positive in LLNA, as confirmed in HRIPT and HMT analyses, which showed no induction at 35,433 and $20,654 \mu\text{g}/\text{cm}^2$, respectively, in the database of the Research Institute for Fragrance Materials (Urbisch *et al.*, 2015). Maleic anhydride is susceptible to hydrolysis, and might be hydro-

lyzed and changed to non-sensitizing hydrolyzates during formulation into water-based final cosmetic products. Remaining seven sensitizers in analyzed sensitizers might give the risk of skin sensitization. However, it should be noted that these chemicals are unlikely to be present in complex mixtures.

To validate LLNA, Hoffmann (2015) analyzed data variability for a set of chemicals that have been described in multiple studies. They suggested the importance of considering variability in reference data. On the other hand, this study provides a proof of principle using a probabilistic approach to risk assessments of minor sensitizers in mixtures without considering the variability of NESIL, EC3, EC1.5, and MIT values. Thus, further investigation on the variability of EC3, EC1.5, and MIT values may be needed. Moreover, to derive NESILs for each analyzed sensitizer, we did not extrapolate from mouse to human data, because LLNA have been shown to correlate well with human skin-sensitization threshold data (Basketter *et al.*, 2005, 2018). This is the same approach as in previous DST studies (Safford *et al.*, 2008, 2011, 2015). However, an interspecies assessment factor of 15 was recently suggested (Bil *et al.*, 2017) and this is controversial discussion (Basketter *et al.*, 2018; SCCS, 2018). In addition, one thing we can say for sure is that a case-by-case evaluation taking account of specific circumstances and chemical properties that can lead to over/under-prediction (e.g. Roberts and Api, 2018) and further investigation is thus needed for more robust threshold values. To minimize the risk of these uncertainties, clinical and epidemiological studies, post marketing surveillance data and case reports also might need to be considered carefully. One final note is that the mixture DST would be protective for the induction of sensitization, and not for the elicitation of sensitization because it is based on AELs, which are exposure levels that do not induce sensitization. It is generally accepted that, in subjects already sensitized to chemicals, elicitation reactions occur at lower levels of exposure than those required to induce the sensitization.

In this study, we developed a novel DST approach for mixtures evaluated as negative in the binary test battery with KeratinoSensTM and h-CLAT. A probabilistic assessment using 95th percentiles of calculated NESILs for mixtures (mixture NESILs) was performed, and a DST of $6010 \mu\text{g}/\text{cm}^2$ was derived for mixtures that were negative in the binary test battery. However, a 5% probability that a sensitizing constituent (1% probability that all indeterminate constituents) will cause skin sensitization remains. In addition, the mixture DST would be protective not for the elicitation of sensitization. Further evalu-

ations that incorporate human skin sensitization threshold data are warranted to minimize uncertainty and derive a more robust threshold value.

Conflict of interest---- The authors declare that there is no conflict of interest.

REFERENCES

- Antignac, E., Nohynek, G.J., Re, T., Clouzeau, J. and Toutain, H. (2011): Safety of botanical ingredients in personal care products/cosmetics. *Food Chem. Toxicol.*, **49**, 324-341.
- Api, A.M., Basketter, D.A., Cadby, P.A., Cano, M.-F., Ellis, G., Gerberick, G.F., Griem, P., McNamee, P.M., Ryan, C.A. and Safford, R. (2008): Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. *Regul. Toxicol. Pharmacol.*, **52**, 3-23.
- Ashikaga, T., Sakaguchi, H., Sono, S., Kosaka, N., Ishikawa, M., Nukada, Y., Miyazawa, M., Ito, Y., Nishiyama, N. and Itagaki, H. (2010): A comparative evaluation of *in vitro* skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). *Altern. Lab. Anim.*, **38**, 275-284.
- Ashikaga, T., Yoshida, Y., Hirota, M., Yoneyama, K., Itagaki, H., Sakaguchi, H., Miyazawa, M., Ito, Y., Suzuki, H. and 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.
- Basketter, D. and Safford, B. (2016): Skin sensitization quantitative risk assessment: A review of underlying assumptions. *Regul. Toxicol. Pharmacol.*, **74**, 105-116.
- Basketter, D.A., Clapp, C., Jefferies, D., Safford, B., Ryan, C.A., Gerberick, F., Dearman, R.J. and Kimber, I. (2005): Predictive identification of human skin sensitization thresholds. *Contact Dermat.*, **53**, 260-267.
- Basketter, D.A., Natsch, A., Ellis, G., Api, A.M., Irizar, A., Safford, B., Ryan, C. and Kern, P. (2018): Interspecies assessment factors and skin sensitization risk assessment. *Regul. Toxicol. Pharmacol.*, **97**, 186-188.
- Basketter, D.A., Evans, P., Fielder, R.J., Gerberick, G.F., Dearman, R.J. and Kimber, I. (2002): Local lymph node assay - validation, conduct and use in practice. *Food Chem. Toxicol.*, **40**, 593-598.
- Bil, W., Schuur, A.G., Ezendam, J. and Bokkers, B.G. (2017): Probabilistic derivation of the interspecies assessment factor for skin sensitization. *Regul. Toxicol. Pharmacol.*, **88**, 34-44.
- Dumont, C., Barroso, J., Matys, I., Worth, A. and Casati, S. (2016): Analysis of the Local Lymph Node Assay (LLNA) variability for assessing the prediction of skin sensitisation potential and potency of chemicals with non-animal approaches. *Toxicol. In Vitro*, **34**, 220-228.
- Emter, R., Ellis, G. and Natsch, A. (2010): Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers *in vitro*. *Toxicol. Appl. Pharmacol.*, **245**, 281-290.
- Felter, S.P., Robinson, M.K., Basketter, D.A. and Gerberick, G.F. (2002): A review of the scientific basis for uncertainty factors for use in quantitative risk assessment for the induction of allergic contact dermatitis. *Contact Dermat.*, **47**, 257-266.
- Felter, S.P., Ryan, C.A., Basketter, D.A., Gilmour, N.J. and Gerberick, G.F. (2003): Application of the risk assessment paradigm to the induction of allergic contact dermatitis. *Regul. Toxicol. Pharmacol.*, **37**, 1-10.
- Hoffmann, S. (2015): LLNA variability: an essential ingredient for a comprehensive assessment of non-animal skin sensitization test methods and strategies. *ALTEX*, **32**, 379-383.
- Jaworska, J.S., Natsch, A., Ryan, C., Strickland, J., Ashikaga, T. and Miyazawa, M. (2015): Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. *Arch. Toxicol.*, **89**, 2355-2383.
- Keller, D., Krauledat, M. and Scheel, J. (2009): Feasibility study to support a threshold of sensitization concern concept in risk assessment based on human data. *Arch. Toxicol.*, **83**, 1049-1060.
- Kern, P.S., Gerberick, G.F., Ryan, C.A., Kimber, I., Aptula, A. and Basketter, D.A. (2010): Local lymph node data for the evaluation of skin sensitization alternatives: a second compilation. *Dermatitis*, **21**, 8-32.
- Kiken, D.A. and Cohen, D.E. (2002): Contact dermatitis to botanical extracts. *Am. J. Contact Dermat.*, **13**, 148-152.
- Kimber, I., Dearman, R.J., Basketter, D.A., Ryan, C.A., Gerberick, G.F., McNamee, P.M., Lalko, J. and Api, A.M. (2008): Dose metrics in the acquisition of skin sensitization: thresholds and importance of dose per unit area. *Regul. Toxicol. Pharmacol.*, **52**, 39-45.
- Kimber, I., Gerberick, G.F. and Basketter, D.A. (2017): Quantitative risk assessment for skin sensitization: success or failure? *Regul. Toxicol. Pharmacol.*, **83**, 104-108.
- Koster, S., Leeman, W., Verheij, E., Dutman, E., van Stee, L., Nielsen, L.M., Ronsmans, S., Noteborn, H. and Krul, L. (2015): A novel safety assessment strategy applied to non-selective extracts. *Food Chem. Toxicol.*, **80**, 163-181.
- Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J.G. and Würtzen, G.; European branch of the International Life Sciences Institute. (2004): Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food Chem. Toxicol.*, **42**, 65-83.
- Lalko, J. and Api, A.M. (2006): Investigation of the dermal sensitization potential of various essential oils in the local lymph node assay. *Food Chem. Toxicol.*, **44**, 739-746.
- Magnusson, B. and Kligman, A.M. (1969): The identification of contact allergens by animal assay. The guinea pig maximization test. *J. Invest. Dermatol.*, **52**, 268-276.
- Nishijo, T., Miyazawa, M., Saito, K., Otsubo, Y., Mizumachi, H. and Sakaguchi, H. (2019): Sensitivity of KeratinoSens™ and h-CLAT for detecting minute amounts of sensitizers to evaluate botanical extracts. *J. Toxicol. Sci.*, **44**, 13-21.
- OECD (2012): The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins Part 1: Scientific Evidence. Series on Testing and Assessment No. 168. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono\(2012\)10/part1&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2012)10/part1&doclanguage=en)
- OECD (2015): OECD Guidelines for the Testing of Chemicals. Test No. 442D: In Vitro Skin Sensitization: ARE-Nrf2 Luciferase Test Method. http://www.oecd-ilibrary.org/environment/test-no-442d-in-vitro-skin-sensitisation_9789264229822-en
- OECD (2016): OECD Guidelines for the Testing of Chemicals. Test No. 442E: In Vitro Skin Sensitization: human Cell Line Activation Test (h-CLAT). http://www.oecd-ilibrary.org/environment/test-no-442e-in-vitro-skin-sensitisation_9789264264359-en
- Otsubo, Y., Nishijo, T., Miyazawa, M., Saito, K., Mizumachi, H. and

- Sakaguchi, H. (2017): Binary test battery with KeratinoSens™ and h-CLAT as part of a bottom-up approach for skin sensitization hazard prediction. *Regul. Toxicol. Pharmacol.*, **88**, 118-124.
- Roberts, D.W. and Api, A.M. (2018): Chemical applicability domain of the local lymph node assay (LLNA) for skin sensitisation potency. Part 4. Quantitative correlation of LLNA potency with human potency. *Regul. Toxicol. Pharmacol.*, **96**, 76-84.
- Rulis, A.M. (1986): De minimis and the threshold of regulation. In: Food protection technology (Felix, C.W., ed.), pp.29-37, Lewis Publishers, Chelsea.
- Safford, R.J. (2008): The Dermal Sensitisation Threshold- a TTC approach for allergic contact dermatitis. *Regul. Toxicol. Pharmacol.*, **51**, 195-200.
- Safford, R.J., Api, A.M., Roberts, D.W. and Lalko, J.F. (2015): Extension of the Dermal Sensitisation Threshold (DST) approach to incorporate chemicals classified as reactive. *Regul. Toxicol. Pharmacol.*, **72**, 694-701.
- Safford, R.J., Aptula, A.O. and Gilmour, N. (2011): Refinement of the Dermal Sensitisation Threshold (DST) approach using a larger dataset and incorporating mechanistic chemistry domains. *Regul. Toxicol. Pharmacol.*, **60**, 218-224.
- SCCS (2016): Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation, 9th revision, SCCS/1564/15, revision of 25 April 2016, http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_190.pdf
- SCCS (2018): Opinion on Skin Sensitisation Quantitative Risk Assessment for Fragrance Ingredients (QRA2) Submission I (SCCS/1589/17) https://ec.europa.eu/health/sites/health/files/scientific_committees/consumer_safety/docs/sccs_o_211.pdf
- 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. and 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.
- Urbisch, D., Mehling, A., Guth, K., Ramirez, T., Honarvar, N., Kolle, S., Landsiedel, R., Jaworska, J., Kern, P.S., Gerberick, F., Natsch, A., Emter, R., Ashikaga, T., Miyazawa, M. and Sakaguchi, H. (2015): Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul. Toxicol. Pharmacol.*, **71**, 337-351.