Fundamental Toxicological Sciences (Fundam. Toxicol. Sci.)
Vol.10, No.3, 113-123, 2023

The Japanese Society of Toxicology

Fundamental Toxicological Sciences
URL : http://www.fundtoxicolsci.org/index_e.html

Original Article

A trial to predict skin irritancy of cosmetic products using cytotoxicity tests

Minori Mezaki[1], Chika Manabe[1], Junichiro Nagumo[1], Tomohisa Yoshimura[1], Tohru Nagahama[1], Hiroshi Itagaki[2] and Akiko Takaoka[1]

[1]R&D Laboratories, Self-medication, Taisho Pharmaceutical Co., Ltd., 1-403, Yoshino-cho, Kita-ku, Saitama-shi, Saitama, Japan
[2]ITACS Consulting, 5-11-19-2504, Minamidai, Minami-ku, Sagamihara-shi, Kanagawa, Japan

(Received April 6, 2023; Accepted April 14, 2023)

ABSTRACT — Alternatives to animal testing are being used to assess the safety of raw ingredients during the process of developing cosmetics. However, since cosmetic products are composed of a variety of raw ingredients, the safety of the product itself should also be tested. In this study, we attempted to evaluate the skin irritation potential of skin lotions by modifying existing alternative test methods. To evaluate the skin irritation potential of commercial skin lotions in the form of an objective indicator, we calculated their Irritation Score (IS) based on the results of searches using keywords related to skin irritation in review statements posted on Japanese cosmetics and beauty websites. We then modified the reconstructed human epidermis (RhE) test method to evaluate irritation by commercial skin lotions. The results showed that exposure to skin lotions with higher ISs tended to result in lower cell viability, and that exposure to many of the lotions with lower ISs resulted in higher cell viability. Next, we tried using an *in vitro* short-term exposure (STE) test method to assess the skin irritation potential of skin lotions. By changing the test-substance exposure concentrations and exposure times in the STE test method, we were able to obtain results that correlated with those obtained by the modified RhE test method. In conclusion, both alternative methods were helpful for assessing the possibility of developing skin irritation of skin lotions. They may also be useful for screening formulations being developed and as means of evaluation before proceeding to human patch tests.

Key words: Skin irritation, Cytotoxicity tests, OECD TG439, OECD TG491, Cosmetic products

INTRODUCTION

The safety of cosmetics has been evaluated by animal testing, but the use of alternatives to animal testing has become necessary from the standpoint of animal welfare. Thus, alternative test methods are being developed. Cosmetic products are generally composed of more than one raw ingredient, each of whose safety has been individually verified by these methods. Nevertheless, cosmetic products sometimes cause adverse events such as irritant dermatitis, allergic contact dermatitis, and depigmentation (Wolf *et al*., 2001; Sasaki *et al*., 2014). Human patch tests are also sometimes performed to confirm the safety of final cosmetic products. On the other hand, safety evaluations of mixtures such as final products by using alternative test methods have not been well established and validated, and few studies have verified the relationship between human adverse events and the results of final product evaluations by alternative test methods. The OECD TG439 or reconstructed human epidermis (RhE)

Correspondence: Minori Mezaki (E-mail: m-mezaki@taisho.co.jp)

test method is a skin irritation test that can be applied to mixtures of ingredients (OECD, 2021). Safety evaluations of cosmetics based on the RhE test have been reported (Kose *et al*., 2018; Ma *et al*., 2021), but issues concerning the conditions for products' dosage forms have remained. Since RhE models are expensive and time-consuming to prepare, other alternative test methods need to be considered.

An *in vitro* short-term exposure (STE) test method (OECD TG491) is an eye irritation test. This test method can be applied to mixtures, and one study has shown that it can be used to evaluate the eye irritation potential of cosmetic products (Abo *et al*., 2018). The STE test method uses a monolayer cell culture, and it is assumed to be a more sensitive predictor of skin irritation than the RhE model, which has a multilayer structure. Based on animal testing, a substance that is negative in an eye irritation test is generally expected to be negative in a skin irritation test performed in the presence of a stratum corneum. However, there has been very little information regarding the use of the STE test method to evaluate skin irritation potential.

In this study, we tried to establish test methods that can predict the skin irritation potential of cosmetic products in humans by modifying the RhE test method and the STE test method. Since cosmetics have multiple formulations that differ in terms of base materials, skin permeability, and irritation, each formulation must be evaluated. We began by evaluating water-soluble, easy-to-evaluate lotions.

MATERIALS AND METHODS

Posted review analysis

We analyzed reviews of commercial skin lotions posted by users on the comprehensive cosmetic information site (@cosme). We then selected 148 commercial skin lotions and used their top 50 posted review sentences as the subject of our analysis. As shown in Table 1, words related to skin trouble, implying skin irritation, were extracted, and scored, and the Irritation Scores (ISs) of the skin lotions were calculated.

The survey period was May 1-December 25, 2015. Whenever there were fewer than 50 posted reviews, we calculated the IS per 50 posted reviews by using the formula:

IS = Total score/Number of posted reviews × 50.

Table 1. The Irritation Score (IS) for each type of skin trouble in the posted reviews

Score	Skin irritation and related symptoms
1	burning/stinging/smarting pain/sensory irritation
3	redness/eczema/inflammation/swelling

Materials

We obtained 79 commercially available cosmetic skin lotions. The LabCyte EPI-MODEL 24 reconstructed human cultured epidermal model and assay medium were purchased from Japan Tissue Engineering Co., Ltd. (Aichi, Japan). SIRC cells (CCL-60) were obtained from DS-Pharma Biomedical Co., Ltd. (Osaka, Japan). Distilled water and saline were purchased from Otsuka Pharmaceutical Factory, Inc. (Tokushima, Japan). Sodium lauryl sulfate (SLS), 2-propyl alcohol, and Dulbecco's phosphate buffered saline (DPBS) were purchased from Wako Pure Chemical Co., Ltd. (Osaka, Japan). 2-Propyl alcohol and hydrogen chloride (HCl) were purchased from Kanto Chemical Industry Co., Ltd. (Tokyo, Japan). 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was obtained from Sigma-Aldrich (St. Louis, MO, USA). DPBS was also purchased from Thermo Fisher Scientific, Inc. (Waltham, MA, USA).

Cytotoxicity assay in the RhE model

The experiments were performed according to OECD TG439 (OECD, 2020), but some procedures were modified. LabCyte EPI-MODEL 24 was pre-cultured overnight in an incubator (37°C, 5% CO_2). The epidermal surface was exposed to 50 µL of the test substance. The RhE model was incubated for 3-42 hr and the test substance was then removed by washing with DPBS. No post-culture treatment was performed.

Tissue samples were placed in medium containing a 0.5 mg/mL concentration of MTT for 3 hr, and each tissue sample was then immersed in 2-propyl alcohol for at least 15 hr to extract MTT formazan. The absorbances of extracts at 570 nm and 650 nm were measured in a microplate reader (EnSpire™2300, Perkin Elmer Inc., Waltham, MA, USA).

The value obtained by subtracting the absorbance at 650 nm from the absorbance at 570 nm was used as the measurement value, and the absorbance of 2-propyl alcohol was measured as a blank. Cell viability was calculated as the ratio of the MTT formazan absorbance of each test substance to the absorbance of the negative control and expressed as a percentage.

Cell culture

SIRC cells were cultured in Eagle's minimum essential medium (Sigma-Aldrich Co.) containing 10% fetal bovine serum (FBS), 100 units/mL penicillin, 100 µg/mL

Table 2. Number of commercial skin lotions rated according to Irritation Score (IS)

IS	Number	IS	Number	IS	Number	IS	Number
0	64 (6)	7	4 (4)	14	2 (1)	27	2 (2)
1	15 (14)	8	5 (4)	16	1 (1)	30	2 (1)
2	12 (9)	9	1 (1)	19	1 (1)	40	1 (1)
3	7 (6)	10	1 (1)	20	1 (1)	46	1 (1)
4	6 (6)	11	2 (2)	23	1 (0)	-	-
5	8 (8)	12	1 (1)	24	1 (1)	-	-
6	4 (4)	13	3 (3)	25	2 (0)	-	-

The numbers in the parenthesis are the numbers of products selected for *in vitro* evaluation.

streptomycin, and 2 mM L-glutamine (Thermo Fisher Scientific). Cells seeded at a concentration of 6.0×10^3 cells/well or 3.0×10^3 cells/well in 96-well plate (MS-8096F, Sumitomo Bakelite Co., Ltd.) and then cultured for 4 days or 5 days (37°C, 5% CO_2) until confluence was achieved.

Cytotoxicity assay in SIRC cells

The cytotoxicity assays in SIRC cells were performed according to the OECD TG491 (OECD, 2020) except for the following steps. The test substances were dissolved in saline to form 10% to 50% (w/w) solutions. Cells were exposed to 200 µL/well of the test substances, saline, and medium for 5 min to 20 min. The test substances were removed by washing the wells with DPBS, and 200 µL of medium containing a 0.5 mg/mL concentration of MTT was added. After culturing for 2 hr, MTT formazan was extracted by adding 200 µL of 2-propyl alcohol containing 0.04 N HCl and allowing the solution to stand for at least 1 hr. Absorbances at 570 nm and 650 nm were measured with a microplate reader. Relative cell viability was calculated as the ratio of the MTT formazan absorbance of each test substance to the MTT formazan absorbance of the test solvent control and expressed as a percentage. The mean viability of the cells in three wells was calculated for each concentration of each test substance.

Calculation of predictive performance

The predictive performance of the test was calculated. Sensitivity was calculated as the percentage of substances tested that were correctly predicted to be irritant substances by the test. Specificity was calculated as the percentage of substances tested that were correctly predicted to be non-irritant substances by the test.

Accuracy was calculated as the percentage of substances tested that were correctly predicted to be in the correct class of substances by the test.

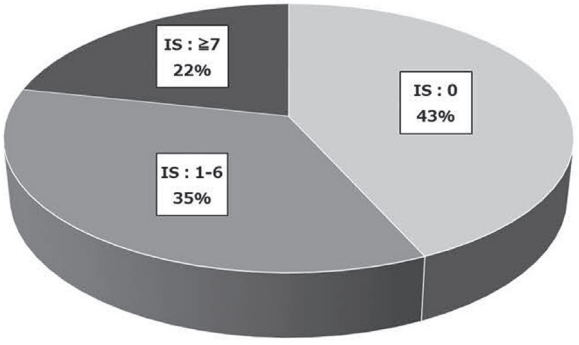

Fig. 1. Percentages of commercial skin lotions according to their Irritation Score (IS). The percentages of 148 commercial skin lotions according to their IS was calculated: There were 64 products with an IS of 0, accounting for 43% of the total; 52 products with an IS of 1-6, accounting for 35% of the total, and 32 products with an IS of 7 or more, accounting for 22% of the total.

RESULTS

Scoring skin irritation by commercial products

The 148 commercially available skin lotions were rated according to their IS, and their ISs are shown in Table 2. The percentages skin lotions according to their ISs are shown in Fig. 1, and 64 products had an IS of 0. Expressions related to skin irritation were found in the review sentences regarding the other 84 products, and after calculating their ISs, we selected the 79 of them for which there were at least 30 posted reviews and tested them in the subsequent *in vitro* evaluations. The numbers of products selected are shown in parenthesis in Table 2.

Effect of exposure time on the evaluations of skin irritation in the RhE models

We attempted to increase the sensitivity of the RhE test

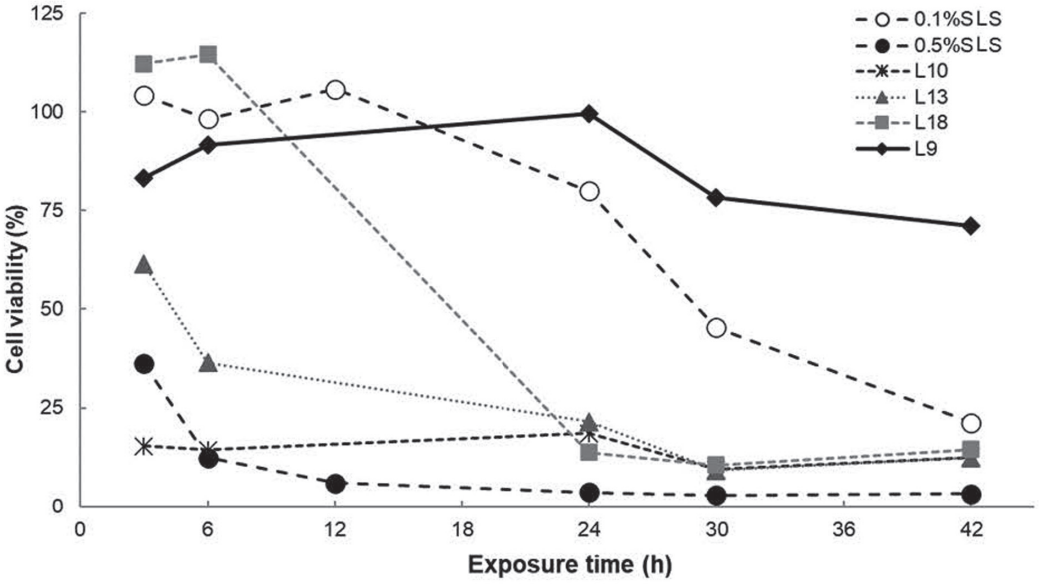

Fig. 2. Effect of exposure time on cell viability in the RhE model. The positive controls were 0.5% sodium lauryl sulfate (SLS) and 0.1% SLS solutions. Test substances were Lotion 10 (IS: 40), Lotion 13 (IS: 27), Lotion 18 (IS:14), and Lotion 9 (IS:0). The RhE model was exposed to the positive controls and test substances for 3-42 hr, and cell viability was measured.

method and evaluate the cytotoxicity of the products. We therefore extended the exposure time from 3 hr to 42 hr and did not perform any post-culture.

Because 0.5% SLS has been shown to be a mild irritant in the human patch test, and 0.1% SLS may cause irritation in some subjects, both were selected as positive controls (Kanto *et al*., 2013; Basketter *et al*., 1996). Lotion 10 (IS: 40), Lotion 13 (IS: 27), and Lotion 18 (IS:14), which had higher ISs, and Lotion 9, which had an IS of 0, were also selected as test substances. Cell viability was evaluated by exposing RhE models to each of them. The results showed that cell viability after exposure to 0.5% SLS and 0.1% SLS was reduced by extending the exposure time (Fig. 2). Cell viability after exposure to Lotion 10, Lotion 13, and Lotion 18 also decreased as the exposure time was extended and fell to less than 50% after more than 24 hr. Cell viability after exposure to Lotion 9, however, was never less than 50% even when the exposure time was extended.

Based on these results, we set the exposure time for product evaluation in the RhE models at 30 hr. The criterion for predicting skin irritation by products in this test model was tentatively set at 50% cell viability in accordance with OECD TG439. We defined this test as the modified RhE test method in this study.

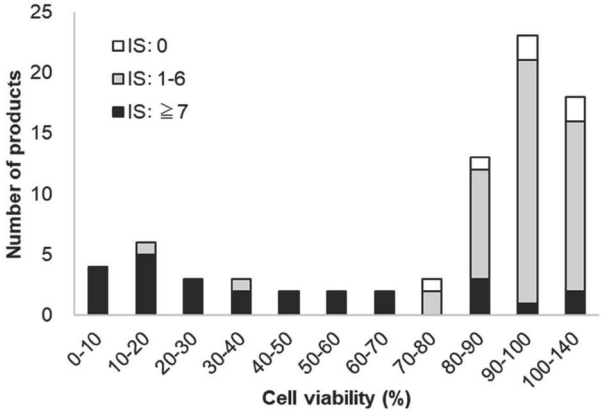

Fig. 3. Cell viability after exposure to 79 commercial skin lotions by the modified RhE test method. The RhE model was exposed to the test products for 30 hr, and cell viability was measured. The bars represent the number of products according to cell viability range. (Products with an IS of 0: white bars; products with an IS between 1 and 6: gray bars; products with an IS of 7 or more: black bars.)

Table 3. List of test lotions and the results of their IS and cell viability evaluations

No.	IS	CV (%) *1	CV (%) *2	No.	IS	CV (%) *1	CV (%) *2
L1	46	32.9	13.9	L41	5	35.5	3.0
L2	13	64.2	70.1	L42	5	91.7	105.9
L3	7	103.3	110.5	L43	4	94.8	104.9
L4	5	87.2	81.0	L44	4	87.2	114.2
L5	3	104.4	83.0	L45	4	107.6	99.6
L6	2	95.7	98.0	L46	6	107.0	110.4
L7	1	93.0	103.3	L47	4	102.1	97.8
L8	0	81.4	94.5	L48	4	93.2	103.8
L9	0	78.3	10.1	L49	4	79.0	10.4
L10	40	9.4	1.7	L50	3	105.1	101.5
L11	30	12.6	2.4	L51	3	97.7	99.0
L12	27	6.7	2.3	L52	3	96.2	115.0
L13	27	9.1	0.0	L53	3	93.5	107.9
L14	24	12.1	6.3	L54	1	93.9	72.5
L15	20	6.6	35.9	L55	3	88.6	96.3
L16	19	42.4	30.6	L56	2	101.9	93.1
L17	16	81.8	60.3	L57	2	112.5	103.9
L18	14	10.5	0.0	L58	2	97.4	107.6
L19	13	20.6	5.9	L59	2	97.0	104.1
L20	12	84.0	93.2	L60	2	92.8	35.5
L21	11	110.0	106.1	L61	2	89.4	7.3
L22	10	83.3	94.4	L62	2	70.1	103.3
L23	9	47.5	98.2	L63	2	18.5	0.0
L24	11	36.5	4.9	L64	1	108.1	51.6
L25	8	94.8	105.3	L65	1	102.8	84.6
L26	8	68.9	25.6	L66	1	100.7	79.4
L27	13	51.0	57.7	L67	1	98.2	16.8
L28	8	26.0	17.9	L68	1	96.9	38.6
L29	8	18.2	21.0	L69	1	96.7	97.4
L30	7	58.9	14.2	L70	1	92.8	82.2
L31	7	20.5	2.7	L71	1	90.8	80.7
L32	7	15.6	0.2	L72	1	90.2	95.0
L33	6	104.7	80.6	L73	1	88.0	92.5
L34	6	102.0	106.6	L74	1	83.8	99.3
L35	6	99.0	100.9	L75	1	81.6	117.2
L36	5	95.8	0.2	L76	0	112.2	91.7
L37	5	84.8	6.6	L77	0	105.1	74.5
L38	5	137.0	43.2	L78	0	98.1	110.8
L39	5	100.9	80.3	L79	0	92.4	104.9
L40	5	85.8	44.3				

IS: Irritation score
CV: Cell viability
*1 Modified RhE test method
*2 Modified STE test method

Evaluation of the skin irritation potential of commercial skin lotions

Next, we evaluated cell viability after exposure to 79 commercial skin lotions according to the modified RhE test method. The results for cell viability and the ISs of the test products are shown in Fig. 3 and Table 3. Many products that resulted in a cell viability of 70% or more had an IS below 6, although exposure to some products with an IS of 7 or more resulted in low cell viability. These findings suggested the possibility of using the modified RhE test method to predict skin irritation by cosmetic products in humans. We assumed that products with an IS of 7 or more were irritant and calculated the correlation between IS and cell viability in the RhE model. When the positivity criterion of the modified RhE test method was set at 50%, sensitivity was 61.5% (16/26), specificity 96.2% (51/53), and accuracy 84.8% (67/79). On the other hand, when the criterion was set at 70% based on the results obtained with the products that had an IS of 7 or more, sensitivity was 76.9% (20/26), specificity 96.2% (51/53), and accuracy 89.9% (71/79).

Better accuracy was obtained when the criterion was set at 70%.

Evaluation of the skin irritation potential of products by modification of the STE test method

We turned our attention to the STE test method, which provides a relatively high evaluation throughput, and assessed the conditions under which it might be able to predict the skin irritation potential of cosmetic products. First, we modified the exposure concentrations and exposure time. The cell viability was not reduced when the test substances were diluted to 10%-50% and the exposure time was 5 min (data not shown).

In addition, since a decrease in cell viability was observed when exposed to saline, the solvent control, for 20 min (data not shown), we set the exposure time for the evaluations at 10 min.

We evaluated the 6 products that resulted in cell viability of 70% or more (L54, L65, L72, L76 L78, L79; Group A) in the modified RhE test method, the 3 products that resulted in cell viability of 50%-70% (L26, L27, L30; Group B), and the 6 products that resulted in cell viability of under 50% (L11, L12, L13, L14, L28, L29; Group C). The results are shown in Fig. 4. Exposure to the products in Group A resulted in high cell viability at concentrations of 40%. Exposure to Group B resulted in decreased cell viability in a concentration-dependent manner, and exposure to Group C resulted in a trend toward lower cell viability. At the 40% exposure concentration, the differ-

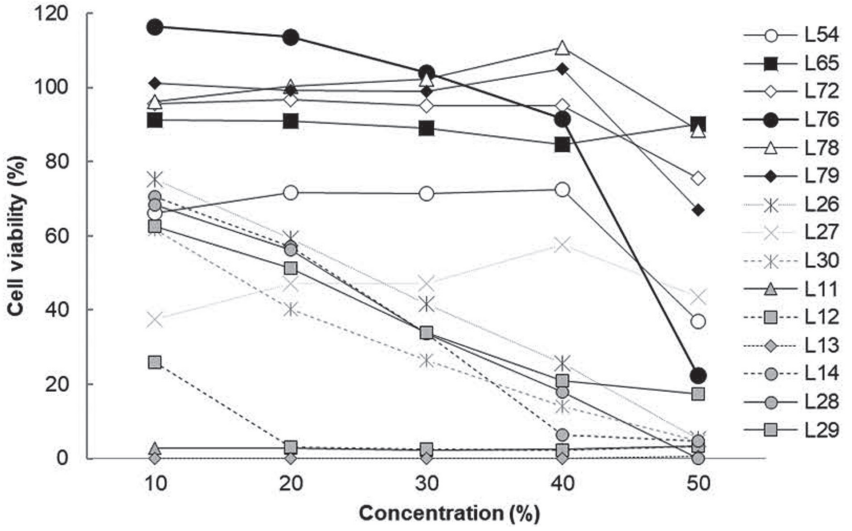

Fig. 4. Relationship between exposure concentrations of the products and cell viability measured by the modified STE test method. SIRC cells were exposed for 10 min to the test products diluted to 10%-50% concentrations with saline, and cell viability was measured. The test products selected were L54, L65, L72, L76 L78, and L79, exposure to which resulted in cell viability values greater than 70% by the modified RhE test method, L26, L27, and L30, exposure to which resulted in cell viability values of 50% to 70% by the modified RhE test method, and L11, L12, L13, L14, L28, and L29, exposure to which resulted in cell viability values of 50% or less by the modified RhE test method.

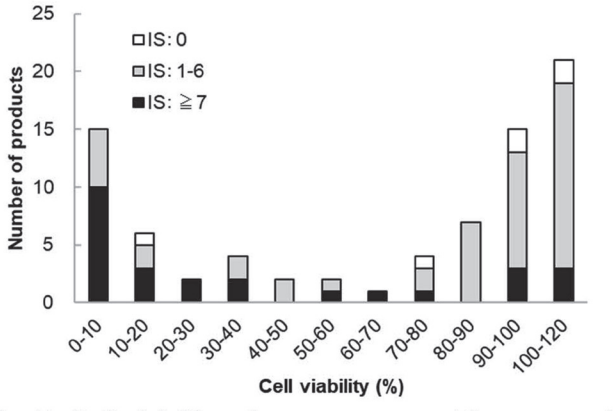

Fig. 5. Cell viability after exposure to 79 commercial skin lotions by the modified STE test method. SIRC cells were exposed for 10 min to 40% dilutions of the test products, and cell viability was measured. The bars represent the numbers of products according to cell viability range. (Products with an IS of 0: white bars; products with an IS between 1 and 6: gray bars; products with an IS of 7 or more: black bars.)

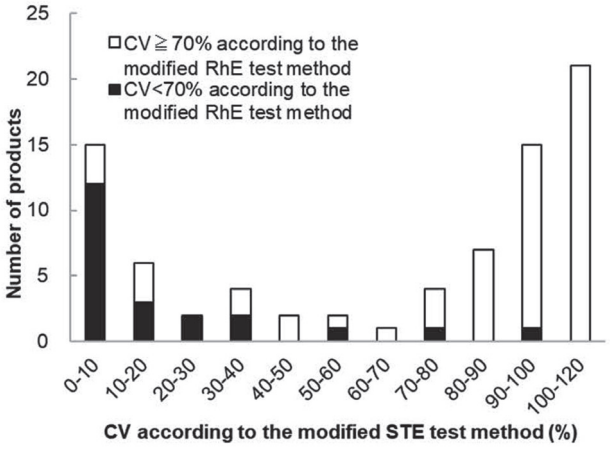

Fig. 6. Results of the evaluations by the modified STE test method according to the results obtained by the modified RhE test method. SIRC cells were exposed for 10 min to 40% dilutions of the test products, and cell viability was measured. The bars represent the numbers of products according to cell viability ranges. (Cell viability ≥70% according to the modified RhE test method; white bars; cell viability >70% according to the modified RhE test method; black bars.)

ence in cell viability between Group A and Group C was significant, confirming the high sensitivity and specificity of the modified STE test method. At the 40% exposure concentration, the cell viability of Group A was over 80% after excluding L54. Products that did not exhibit toxicity in the RhE model had a higher probability of being distributed over 80% in this condition. it is important to reduce false negatives in the safety evaluation, so we therefore tentatively defined 80% or more cell viability when exposed to a 40% concentration of a skin lotion for 10 min as the criterion for non-irritation by the modified STE test method.

Evaluation of the skin irritation potential of products by means of the modified STE test method

We evaluate cell viability after exposure to the 79 skin lotions according to the modified STE test method. The results for cell viability and the ISs of the test products are shown in Fig. 5 and Table 3.

Exposure to most of the products with an IS of 7 or more resulted in less than 40% cell viability. We calculated the predictive performance of IS and cell viability according to the results of the modified STE test method when products having an IS of 7 or more were assumed to be potentially irritating products. The modified STE test method had a sensitivity of 76.9% (20/26), specificity of 69.8% (37/53), and accuracy of 72.2% (57/79) for predicting skin irritation when products with an IS of 7 or more were assumed to be irritant and a cell viability of 80% was used as the criterion for irritation (Fig. 5). The sensitivity of the modified STE test method in relation to ISs was almost the same as the sensitivity of the modified RhE test method.

Next, the results of modified STE test method versus the modified RhE test method are shown in Fig. 6. Many of the products that resulted in a cell viability 70% or more by the modified RhE test method resulted in a cell viability 80% or more by the modified STE test method. On the other hand, many of the products that resulted in a cell viability less than 70% by the modified RhE test method resulted in a cell viability less than 40% by the modified STE test method. When the products that resulted in a cell viability of under 70% by modified RhE test method were regarded as potentially irritating products and the irritation criterion for the modified STE test method was set at 80% cell viability, the sensitivity, specificity, and accuracy of the modified STE test method were 95.5% (21/22), 73.7% (42/57), and 79.7% (63/79), respectively. Thus, the results obtained by the modified STE test method were found to closely correspond to the results obtained by the modified RhE test method.

DISCUSSION

In order for customers to use cosmetics safely, the human skin irritation potential of cosmetic products needs to be evaluated. While, in principle, the raw ingredients in products are used within concentration ranges that ensure safety, unexpected adverse effects may occur as a result of interactions between them. For example, there have been reports that the combination of chemicals increase cytotoxicity, skin irritation and skin sensitization (Zanoni *et al*., 2018; Rentiis *et al*., 2021; Wu *et al*., 2001). An alternative method of predicting the safety of combinations of ingredients or products is needed during the development stage prior to proceeding to human testing.

Establishing an *in vitro* test for end-products that predicts the response in humans requires a large amount of cosmetic formulation data in order to be able to analyze correlations with the results obtained by alternative methods and in humans. Since, in actual practice, it is difficult to obtain such large amounts of human data, we analyzed review statements on a cosmetic website and utilized the ISs of commercial products instead. In order to analyze the review statements in relation to skin irritation by commercial products objectively, we selected keywords from indexes used in human patch tests and actual use tests, and used the information in the statements to calculate the Irritation Scores (ISs) shown in Table 1. Because some chemicals that cause sensory irritation have been reported to exhibit cytotoxicity (Inoue *et al*., 1996), we also selected keywords that seemed to refer to sensory irritation to use to calculate ISs.

Our analysis of review statements revealed that words related to skin irritation were used in regard to 84 of the 148 commercial products we adopted as the subject of our study. This suggests that some commercial products may be irritating to humans when applied to the skin on a daily basis. Studies of alternative test methods to assess the safety of cosmetic products have been reported, including skin irritation evaluations by means of RhE models (Kose *et al*., 2018) and SENS-IS assays (Cottrez *et al*., 2020) and eye irritation evaluations by the STE test method (Abo *et al*., 2018). However, further studies are needed to compare the safety of different formulations and to define safety criteria.

Although OECD TG439 has been validated as a method of evaluating skin irritation in humans following a 4-hr exposure to cosmetic ingredients, since it may not be suitable for predicting skin irritation when cosmetics are actually used, it is necessary to increase the sensitivity of the evaluation system to accurately predict the irritancy of the product. Skin irritation has been reported to increase with exposure time and the concentration of the chemical (Cruzan *et al*., 1986; Gilman *et al*., 1978), and increasing exposure concentrations in OECD TG439 has been reported to be an effective means of predicting skin irritation following 24-hr exposure in humans (Sugiyama *et al*., 2018). We attempted to increase the sensitivity of the RhE test method by extending exposure time, and the results showed that exposure to many of the products with a high IS resulted in low cell viability rates, and that exposure to many of the products with a low IS resulted in high cell viability rates. These findings indicated that it might be possible to predict skin irritation by cosmetic products in humans based on evaluations by the modified RhE test method. Exposure to some products with a low IS resulted in low cell viability. These products may have been detected to be a potential skin irritant.

On the other hand, exposure to several products resulted in high cell viability despite a high IS. Since these products included wiping cosmetics, the physical stimulation during cosmetic application may have resulted in a high IS. Moreover, certain conditions such as pH, skin permeability, polarity, osmotic pressure, and other physical properties of products and ingredients that promote percutaneous absorption may also affect cell viability. Since our objective in the present study was to establish a product evaluation system, we did not evaluate ingredients by the modified RhE test method. Because the specific formulations of the products tested in our study were unknown, the individual characteristics and interactions of the ingredients have not been discussed in this report. Correlations between the IS irritancy criteria set at 5, 7, 13, or 18 and the irritancy criteria of 50%~80% cell viability in the RhE model are shown in Table 4. The false-negative rate improved as the IS criterion increased, but there was no significant change in accuracy. Products with ISs of 7-13 that resulted in 50%-70% cell viability in the modified RhE model were found to have lower correlation values between ISs and cell viability results. Because the results for products with an IS around the irritant criteria were highly variable and their ISs were review-based data, they may not have been properly evaluated by the modified RhE test method. It will be necessary to improve the accuracy of the modified RhE test method by increasing the number of products with an IS around the irritant criterion and evaluating them.

The STE test method can be applied quickly if the SIRC cells are retained, and the relative simplicity of the test procedures enables evaluation of multiple samples. Since the eyes do not have a stratum corneum barrier, if the eye irritation test is negative, skin irritation is assumed

Table 4. Correlations between the irritancy criteria of the IS and irritancy criteria of cell viability

		Modified RhE test method							
		<80%	≥80%	<70%	≥70%	<60%	≥60%	<50%	≥50%
	≥5	21	16	21	16	19	18	17	20
IS	1-4	3	33	1	35	1	35	1	35
	0	1	5	0	6	0	6	0	6
sensitivity (%)		56.8		56.8		51.4		45.9	
specificity (%)		90.5		97.6		97.6		97.6	
false-positive rate (%)		9.5		2.4		2.4		2.4	
false-negative rate (%)		43.2		43.2		48.6		54.1	
accuracy (%)		74.7		78.5		75.9		73.4	
		Modified RhE test method							
		<80%	≥80%	<70%	≥70%	<60%	≥60%	<50%	≥50%
	≥7	20	6	20	6	18	8	16	10
IS	1-6	4	43	2	45	2	45	2	45
	0	1	5	0	6	0	6	0	6
sensitivity (%)		76.9		76.9		69.2		61.5	
specificity (%)		90.6		96.2		96.2		96.2	
false-positive rate (%)		9.4		3.8		3.8		3.8	
false-negative rate (%)		23.1		23.1		30.8		38.5	
accuracy (%)		86.1		89.9		87.3		84.8	
		Modified RhE test method							
		<80%	≥80%	<70%	≥70%	<60%	≥60%	<50%	≥50%
	≥13	11	1	11	1	10	2	10	2
IS	1-12	13	48	11	50	10	51	8	53
	0	1	5	0	6	0	6	0	6
sensitivity (%)		91.7		91.7		83.3		83.3	
specificity (%)		79.1		83.6		85.1		88.1	
false-positive rate (%)		20.9		16.4		14.9		11.9	
false-negative rate (%)		8.3		8.3		16.7		16.7	
accuracy (%)		81		84.8		84.8		87.3	
		Modified RhE test method							
		<80%	≥80%	<70%	≥70%	<60%	≥60%	<50%	≥50%
	≥18	8	0	8	0	8	0	8	0
IS	1-17	16	49	14	51	12	53	10	55
	0	1	5	0	6	0	6	0	6
sensitivity (%)		100		100		100		100	
specificity (%)		76.1		80.3		83.1		85.9	
false-positive rate (%)		23.9		19.7		16.9		14.1	
false-negative rate (%)		0		0		0		0	
accuracy (%)		78.5		82.3		84.8		87.3	

to be negative. We therefore thought that it might be easier to predict skin irritancy by using the eye irritation test. Although there are differences in physiological functions, including the barrier function, between RhE models and SIRC cells, both test methods use cell viability as an indicator to predict irritation by test substances. The results of the product evaluations obtained by the modified RhE test method and the modified STE test method have generally been correlated. However, the modified STE test method is more likely to yield false-positive results. Exposure to some skin lotions, including L9, L36, L37, L49, and L61, resulted in cell viability of 70% or more by the modified RhE test method, but much lower cell viability by the modified STE test method, and these results of the modified STE test method might be interpreted as false-positive. Conversely, exposure to one skin lotion (L23) that resulted in cell viability less than 70% by the modified RhE test method resulted in higher cell viability by the modified STE test method, and this result by the modified STE test method might be interpreted as a false-

negative. These conflicting results may be attributable to differences in the reactivity of the cell types and exposure conditions. Furthermore, since the RhE model has a three-dimensional structure, and the product is evaluated under conditions in which it can retain its original form on the cell surface, differences such as skin metabolism and percutaneous absorption may affect the results. Moreover, some formulations may contain ingredients intended to relieve irritation, and the dilution may have led to loss of the relieving effect, resulting in false positives in the modified STE test method.

The results obtained with the modified RhE test method, and the modified STE test method were comparable in terms of sensitivity to IS. When the modified STE test method is performed, it is desirable to judge the outcome comprehensively by making a comparison with a similar formulation as a benchmark and evaluating it in combination with the modified RhE test method. We performed a human patch test to evaluate the four skin lotions that were classified as non-irritant based on the results of the modified RhE test method and the modified STE test method, and the results for all four lotions were "non-irritating" (data not shown). In other words, these test methods may be useful for evaluating products before proceeding to the development of formulations and human testing. However, some issues remain unresolved, and it is necessary to analyze the correlations between the results of evaluations by these test methods and the results of human patch tests, and to analyze the extent to which individual ingredients and combinations of ingredients contribute to the results of the evaluation. It will be necessary to deal with these issues and improve the precision of the test methods. We would also like to investigate application of these modified test methods to other dosage forms, including to creams and formulations containing active ingredients.

We concluded that further study will contribute to the development of safer cosmetics.

ACKNOWLEDGEMENT

We wish to thank Istyle Co., Ltd. (Tokyo, Japan) for allowing us to use the posted review information.

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

REFERENCES

Abo, T., Hilberer, A., Wagner, C.B., Watanabe, M., Cameron, D., Kirst, A., Nukada, Y., Yuki, T., Araki, D., Sakaguchi, H. and Itagaki, H. (2018): Predictive performance and inter-laboratory reproducibility in assessing eye irritation potential of water- and oil-soluble mixtures using the Short Time Exposure test method. Toxicol. In Vitro, **48**, 78-85.

Basketter, D.A., Griffiths, H.A., Wang, X.M., Wilhelm, K.P. and Mcfadden, J. (1996): Individual, ethnic and seasonal variability in irritant susceptibility of skin: the implications for a predictive human patch test. Contact Dermat., **35**, 208-213.

Cottrez, F., Boitel, E., Gomez, M.P., Dalhuchyts, H., Bidan, C., Rattier, S., Ferret, P.J. and Groux, H. (2020): *In vitro* measurement of skin sensitization hazard of mixtures and finished products: results obtained with the SENS-IS assays. Toxicol. In Vitro, **62**, 104644.

Cruzan, G., Dalbey, W.E., D'aleo, C.J. and Singer, E.J. (1986): A composite model for multiple assays of skin irritation. Toxicol. Ind. Health, **2**, 309-320.

Gilman, M.R., Evans, R.A. and De Salva, S.J. (1978): The influence of concentration, exposure duration, and patch occlusivity upon rabbit primary dermal irritation indices. Drug Chem. Toxicol., **1**, 391-400.

Inoue, K., Nakazawa, K., Inoue, K., Fujimori, K., Ohno, Y., Takanaka, A., Itagaki, H., Kato, S., Kobayashi, T. and Kuroiwa, Y. (1996): Evaluation of Stinging-inducing Chemicals using Cultured Neuronal Cells: an Electrophysiological Approach. Toxicol. In Vitro, **10**, 455-462.

Kanto, H., Washizaki, K., Ito, M., Matsunaga, K., Akamatsu, H., Kawai, K., Katoh, N., Natsuaki, M., Yoshimura, I., Kojima, H., Okamoto, Y., Okuda, M., Kuwahara, H., Sugiyama, M., Kinoshita, S. and Mori, F. (2013): Optimal patch application time in the evaluation of skin irritation. J. Dermatol., **40**, 363-369.

Kose, O., Erkekoglu, P., Sabuncuoglu, S. and Gumusel, B.K. (2018): Evaluation of skin irritation potentials of different cosmetic products in Turkish market by reconstructed human epidermis model. Regul. Toxicol. Pharmacol., **98**, 268-273.

Ma, X., Wang, F. and Wang, B. (2021): Application of an *in vitro* reconstructed human skin on cosmetics in skin irritation tests. J. Cosmet. Dermatol., **20**, 1933-1941.

OECD. (2020): OECD guideline for the testing of chemicals, No. 491, Short Time Exposure *In Vitro* Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage. Organisation for Economic Cooperation and Development.

OECD. (2021): OECD guideline for the testing of chemicals, No. 439, *In Vitro* Skin Irritation: Reconstructed Human Epidermis Test Methods. Organisation for Economic Cooperation and Development.

Rentiis, A.M., Pink, M., Verma, N. and Spanke, S.S. (2021): Assessment of the different skin sensitization potentials of irritants and allergens as single substances and in combination using the KeratinoSens assay. Contact Dermat., **84**, 317-325.

Sasaki, M., Kondo, M., Sato, K., Umeda, M., Kawabata, K., Takahashi, Y., Suzuki, T., Matsunaga, K. and Inoue, S. (2014): Rhododendrol, a depigmentation-inducing phenolic compound, exerts melanocyte cytotoxicity via a tyrosinase-dependent mechanism. Pigment Cell Melanoma Res., **27**, 754-763.

Sugiyama, M., Akita, M., Alépée, N., Fujishiro, M., Hagino, S., Handa, Y., Ikeda, H., Imai, N., Jitsukawa, S., Katoh, M., Kurihara, K., Kyotani, D., Nomura, S., Okamoto, Y., Okumura, H., Omori, T., Sugibayashi, K., Todo, H., Toyoda, A. and Ohno, Y. (2018): Comparative assessment of 24-hr primary skin irritation test and human patch test data with *in vitro* skin irritation tests according to OECD Test Guideline 439 (for quasi-drugs in

Japan). J. Toxicol. Sci., **43**, 751-768.

Wolf, R., Wolf, D., Tüzün, B. and Tüzün, Y. (2001): Contact Dermatitis to Cosmetics. Clin. Dermatol., **19**, 502-515.

Wu, P.C., Chang, J.S., Huang, Y.B., Chai, C.Y. and Tsai, Y.H. (2001): Evaluation of percutaneous absorption and skin irritation of ketoprofen through rat skin: *in vitro* and *in vivo* study. Int. J. Pharm., **222**, 225-235.

Zanoni, T.B., Pedrosa, T.N., Catarino, C.M., Spiekstra, S.W., Oliveira, D.P., Hartog, G.D., Bast, A., Hagemann, G., Gibbs, S., Barros, S.B. and Engler, S.S. (2018): Allergens of permanent hair dyes induces epidermal damage, skin barrier loss and IL-1α increase in epidermal *in vitro* model. Food Chem. Toxicol., **112**, 265-272.

@cosme; https://www.cosme.net/