

Fragrance Skin Sensitization Evaluation and Human Testing: 30-Year Experience

Mihwa Na, PhD,* Gretchen Ritacco, PhD,* Devin O'Brien, PhD,* Maura Lavelle, PhD,* Anne Marie Api, PhD,* and David Basketter, DSc, FRCPATH†

Background: The human repeated insult patch test (HRIPT) has a history of use in the fragrance industry as a component of safety evaluation, exclusively to confirm the absence of skin sensitization at a defined dose.

Objective: The aim of the study was to document the accumulated experience from more than 30 years of conducting HRIPTs.

Methods: A retrospective collation of HRIPT studies carried out to a consistent protocol was undertaken, with each study comprising a minimum of 100 volunteers.

Conclusions: The HRIPT outcomes from 154 studies on 134 substances using 16,512 volunteers were obtained. Most studies confirmed that at the selected induction/challenge dose, sensitization was not induced. In 0.12% of subjects ($n = 20$), there was induction of allergy. However, in the last 11 years, only 3 (0.03%) of 9854 subjects became sensitized, perhaps because of improved definition of a safe HRIPT dose from the local lymph node assay and other skin sensitization methodologies, as well as more rigorous application of the standard protocol after publication in 2008. This experience with HRIPTs demonstrates that de novo sensitization induction is rare and becoming rarer, but it plays an important role as an indicator that toxicological predictions from nonhuman test methods (in vivo and in vitro methods) can be imperfect.

In the middle of the last century, Shelanski^{1,2} conceived “a new technique of human patch tests” as a means to expose the skin sensitizing activity of substances. More or less in parallel, Schwartz^{3,4} reported similar types of investigations. These works laid the foundation of the human repeated insult patch test (HRIPT). Neither Shelanski nor Schwartz published much more on the HRIPT (the authors could only identify a single subsequent publication made more than 45 years after the original work).⁵ Consequently, it fell to others to develop a more consistent protocol for the HRIPT and to establish its scientific foundation.^{6–11} On this basis were derived a limited number of publications detailing the

application of the HRIPT to specific use categories of substances, notably preservatives, and of particular relevance in this present article, fragrances.^{12,13} The HRIPT used by the Research Institute for Fragrance Materials (RIFM) is a repeated patch test that is used to confirm the no-observed-effect level for the induction of skin sensitization in a normal human population, under exaggerated exposure conditions. Statistically, when no reactions occur in 100 test subjects, then the rate of positive reactions in a larger population is unlikely to exceed 2.9%, with a confidence level of 95%, under identical conditions. That upper level of 2.9% positive reactions should not be confused with an expected rate of 2.9% in the general population, not least because the test conditions in the HRIPT are not identical to real-life scenarios.^{14,15}

Over several decades, the approach taken by the RIFM for the evaluation of skin sensitization potential used the HRIPT as a final step to confirm the absence of this activity at the dose level determined from a preceding risk assessment to be nonsensitizing.¹⁶ For this article, it is not necessary to detail the history of the development and evolution of that risk assessment process, because that methodology has been fully detailed elsewhere.^{15–18} Published at the same time was a critical review of how to perform and interpret the HRIPT.^{19,20} Nevertheless, it is also fair to note that there remains a significant concern surrounding the ethics, effectiveness, and human safety of the HRIPT.^{21–23} In response to these concerns, the RIFM has undertaken an extensive retrospective review of its HRIPT portfolio, which is reported herein. Obviously, ethical questions associated with the HRIPT must be the remit of a properly constituted, independent, and transparent ethical review committee

From the *Research Institute for Fragrance Materials, Inc, Woodcliff Lake, NJ; and †DABMEB Consultancy, Ltd, Gloucestershire, United Kingdom.

Address reprint requests to Anne Marie Api, PhD, Research Institute for Fragrance Materials, 50 Tice Blvd, Woodcliff Lake, NJ 07667. E-mail: amapi@rifm.org.

At the time of this work, M.N., G.R., D.O., M.L., and A.M.A. were all full-time employees of the Research Institute for Fragrance Materials. D.B. was compensated for his work in the preparation of this manuscript. D.B. received financial compensation for the time spent in the preparation of this publication. All other authors are fully paid employees of Research Institute for Fragrance Materials.

DOI: 10.1097/DER.0000000000000684

© 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Contact Dermatitis Society. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

(institutional review board). With this in mind, the primary focus here is the risk of induction of contact allergy in those who participate in any HRIPT.

MATERIALS AND METHODS

Substances

All the fragrance ingredients tested were commercial quality samples identified by their Chemical Abstracts Service number.

Test Protocol

In brief, 0.3 mL (liquid) or 0.3 g (solid) of the selected concentration of the test fragrance material is applied in a vehicle of 3:1 diethyl phthalate/ethanol (or on 7 occasions 1:3 diethyl phthalate/ethanol) using occlusive 25-mm Hill Top Chamber patches; saline (128 studies) and/or vehicle (154 studies) control patches are applied in parallel. The induction patches are applied to the skin between the scapula and spinal midline for 24 hours, followed by a 24-hour rest period and retreatment of the same site for a total of 9 induction applications over 3 weeks. This induction phase is followed by a 2-week rest period and then the challenge phase. Challenge is made by a single 24-hour patch to a naive test site; the site is scored 24, 48, and 72/96 hours after application by a trained evaluator. Rechallenge may be made to confirm the nature of any skin reaction. Normally, at least 100 subjects must finish the test. More than a dozen inclusion/exclusion criteria were used to identify appropriate volunteers, and they are described by Politano and Api.¹⁹ The test fragrance material concentration depends on detailed preceding toxicological evaluation and is always built on a weight of evidence, but for most substances reported herein, it has depended on relative potency information from the local lymph node assay (LLNA).¹⁷⁻¹⁹ The amount of fragrance material per unit area of skin is used to quantify the dosage in these studies, as it has been previously shown to be the most relevant metric to skin sensitization.²⁴ The dose per unit area can be easily calculated by dividing the amount of test material by the size of the patch used. For instance, in a study with α -amylcinnamaldehyde (Table 2), 0.3 mL (approximately equal to $3.0 \times 10^5 \mu\text{g}$) of 20% fragrance material was applied using a Hill Top Chamber. An area of 2.54 cm^2 is covered by the fragrance material using this patch system. The dose per unit area in this study was calculated as follows:

$$\frac{0.2 \times (3.0 \times 10^5 \mu\text{g})}{2.54 \text{ cm}^2} = 23,622 \mu\text{g}/\text{cm}^2.$$

To determine the likelihood that skin sensitization (contact allergy) has been induced in an individual, the data set was inspected for evidence that reactions to test material were greater than those to vehicles, persisted/increased during the observation period, and/or were reproducible upon rechallenge. The scoring scale used for skin reactions is shown in Table 1. A skin reaction of at least 1E, erythema combined with edema, was, in the absence of confounding irritation, taken to indicate induction of contact allergy.

TABLE 1. The HRIPT Scoring Scheme

Reaction Grade	Description
0	No visible skin reaction
±	Faint, minimal erythema
1	Erythema
2	Intense erythema, induration
3	Intense erythema, induration, vesicles
4	Severe reaction with erythema, induration, vesicles, pustules
E	Edema
DR	Dryness
P	Papule—red, solid, pinpoint elevation

HRIPT, human repeated insult patch test.

Human Safety Considerations

Every study was conducted with the approval of an independent institutional review board. The test concentration selected is based on a careful toxicological examination, such that the exposure level should not be associated with any other adverse health effects, including local toxicity (irritation, depigmentation, hyperpigmentation, etc), genotoxicity, or other systemic adverse reactions. Analysis of the predicted nonsensitizing dose is based on a weight of evidence from LLNA data and other sensitization assays, including in silico and in vitro assays. When available, historical human tests, such as human maximization studies, are also considered in predicting the nonsensitizing dose. This dose may not be the highest level that can be achieved in humans. In addition, toxicology predictions inevitably contain a degree of uncertainty, so despite the low risk, volunteers who develop skin sensitization reactions are notified of what they are allergic to, are examined by a dermatologist, receive follow-up care until the allergic skin reaction subsides, and are provided information on how to avoid future cases of dermatitis.^{19,20,24}

Limitations

The HRIPT is only used by the RIFM to confirm a no-effect level, which is established through rigorous preclinical investigation. For this reason, the HRIPT described here is limited in discovering a threshold of a material that induces skin sensitization. The statistical limitation may arise because of the number of subjects. According to previous analysis, the sensitization rates of less than 1.0% are not likely to be detected when the test is conducted with a group of approximately 100 subjects.^{14,20} It should also be noted that the test is conducted under the exaggerated exposure scenario that is unlike the real-life situation. Although this exaggerated test condition may increase the sensitivity of the test in detecting the possible skin sensitization reaction, it also means that HRIPT cannot be used to precisely predict skin sensitization potency of a material.^{20,21} Other limitations include the challenge in gathering a volunteer population with diversity in terms of age, sex, and ethnicity.²⁰ It can also be difficult to ensure compliance of all participating subjects to the

TABLE 2. (Continued)

Chemical Name	CAS No.	Subjects	Dose µg/cm ²	Induction and Challenge			Rechallenge	Conclusion
				Reactions,* Test Material	Reactions,* Vehicle	Reactions,* Saline		
2,6-Octadienal, 3,7-dimethyl-, reaction products with ethyl alcohol	147060-73-9	103	1535	0	0	0	0	Negative
2,6,10-Trimethylundeca-5,9-dienal	54082-68-7	108	10,039	0	0	0	0	Negative
2,6,6-Trimethylcyclohexa-1,3-dienyl methanal	116-26-7	105	29	0	0	0	0	Negative
2,6,6-Trimethylcyclohexa-1,3-dienyl methanal	116-26-7	99	59	1E: 1 subject induction #1; 1E: 1 subject induction #6; 1E: 1 subject induction #8; 1E: 1 subject induction #9; 1E: 1 subject challenge #2; 1E: 1 subject challenge #3	1E: 1 subject, induction #6	0	0	Sensitization in 1 subject
3 and 4-(4-hydroxy-4-methylpentyl)-3-cyclohexene-1-carboxaldehyde (HICC)	31906-04-4	201	4000	1E: 1 subject induction #4	1E: 1 subject induction #3	0	0	Negative
3-(2-Oxopropyl)-2-pentylcyclopentanone	40942-73-2	112	2362	0	0	0	0	Negative
3-Decen-2-one	10519-33-2	107	118	0	0	0	0	Negative
3-Methyl-2-(pentyl-oxo)cyclopent-2-en-1-one	68922-13-4	107	1181	0	0	1 subject	0	Negative
3-Phenylbutanal	16251-77-7	102	5905	0	0	1E: 1 subject induction #4	0	Negative
3-Phenylpropyl cinnamate	122-68-9	105	2716	0	0	0	0	Negative
3-Propylidenephthalide	17369-59-4	109	945	0	0	0	0	Negative
3,3-Dimethyl-5-(2,2,3-trimethyl-3-cyclopenten-1-yl)-4-penten-2-ol	107898-54-4	104	2598	0	0	0	0	Negative
3,7-Dimethyl-2-methylenocta-6-enal	22418-66-2	107	590	0	0	0	0	Negative
4-(2,6,6-Trimethyl-2-cyclohexen)-2-methylbutanal	65405-84-7	107	1181	0	0	0	1E: 1 subject, induction #7	Negative
4-(3,4-Methylenedioxyphenyl)-2-butanone	55418-52-5	106	2362	0	0	0	N/A	Negative
4-Hydroxy-2,5-dimethyl-3(2H)-furanone	3658-77-3	108	591	0	0	0	0	Negative
4-Hydroxy-2,5-dimethyl-3(2H)-furanone	3658-77-3	110	1181	1E: 1 subject induction #6; 1E: 1 subject induction reading #8; 1E: 1 subject, challenge reading #2	0	0	0	Sensitization in 1 subject
4-Hydroxy-3-pentenoic acid lactone	591-12-8	110	236	0	0	0	0	subject
4-Methoxy-α-methylbenzene propanal	5462-06-6	104	5906	0	0	0	0	Negative
4-Tricyclodecylidene butanal	30168-23-1	105	1181	0	0	0	0	Negative
5-Methyl-5-phenyl-3-hexanone	4927-36-0	113	1890	0	0	0	0	Negative
5,8-Methano-2H-1-benzopyran, 6(or 7)-ethylideneoctahydro-, [4aR,5S,8S,8aS(or 4aR,5R,8S,8aR)]-rel-	943723-15-7	105	8267	0	0	0	0	Negative
6-Acetyl-1,1,2,4,4,7-hexamethyltetraline	21145-77-7	111	11,811	0	0	0	N/A	Negative
6-Methoxy-2,6-dimethylheptan-1-ol	62439-41-2	106	5905	0	0	0	N/A	Negative

TABLE 2. (Continued)

Chemical Name	CAS No.	Subjects	Dose µg/cm ²	Induction and Challenge			Rechallenge	Conclusion
				Reactions,* Test Material	Reactions,* Vehicle	Reactions,* Saline		
Benzyl alcohol	100-51-6	110	8858	2: 1 subject (subject #34, reaction reduced to + at later readings); 2: 1 subject (subject #52, also showed level 3 reactions during induction); 2: 1 subject (subject #101, reaction reduced to + at a later reading)	0	0	Rechallenge for subjects #34, 52, 101 were done: negative for subjects #34 and 101, positive for subject #52	Sensitization reaction in 1 subject
Benzyl benzoate	120-51-4	108	59,050	2E: 1 subject challenge #2 only; 3: 2E: 1 subject challenge #2 only; 3: 1 subject challenge #3 and #4 (subject #28)	0	0	Subject #28 rechallenged; positive reactions sustained at rechallenge for both test material and vehicle	Negative
Benzyl cinnamate	103-41-3	102	4724	0	0	0	N/A	Negative
Benzyl salicylate	118-58-1	101	17,715	0	0	0	N/A	Negative
Cedrene	11028-42-5	111	3543	0	0	0	N/A	Negative
Cedrol	77-53-2	106	2008	0	0	0	N/A	Negative
Cinnamaldehyde	104-55-2	94	591	0	0	0	0	Negative
Cinnamic aldehyde dimethyl acetal	4364-06-1	92	827	0	0	0	0	Negative
Cinnamyl acetate	103-54-8	101	3424	0	0	0	0	Negative
Cinnamyl alcohol	104-54-1	106	2953	1E: 5 subjects induction #2; 1E: 3 subjects induction #4; 1E: 1 subject induction #5; 1E: 2 subjects induction #6; 1E: 1 subject induction #7; 1E: 1 subject induction #8; 1E: 2 subjects challenge #3 and during induction	0	0	N/A	Negative
Cinnamyl nitrite	1885-38-7	118	1063	0	0	0	0	Negative
Citral	5392-40-5	101	1417	0	0	0	0	Negative
Citronellal	106-23-0	110	7086	0	0	0	N/A	Negative

TABLE 2. (Continued)

Chemical Name	CAS No.	Subjects	Dose µg/cm ²	Induction and Challenge			Reactions,* Saline	Rechallenge	Conclusion
				Reactions,* Test Material	Reactions,* Vehicle	Reactions,* Saline			
Geraniol	106-24-1	109	5905	1E: 1 subject challenge #2 (subject #88)	0	0	0	Subject #88 was rechallenged; the positive reaction was not sustained in the rechallenge	Negative
Geraniol	106-24-1	110	2362	0	1E: 1 subject, induction #1	1E: 1 subject, induction #1	0		Negative
Geranyl acetate	105-87-3	111	5019	0	0	0	0		Negative
Hexen-2-al	6728-26-3	106	23.6	1E: 1 subject challenge #3; 1E: 1 subject challenge #4	0	0	0	Subject #47 rechallenged; positive reaction not sustained in the rechallenge	Negative
Hexen-2-al	6728-26-3	109	18.0	0	0	0	0		Negative
Hexyl 2-methylbutyrate	10032-15-2	109	7086	0	0	0	0		Negative
Hexyl salicylate	6259-76-3	103	35,430	0	0	0	0		Negative
Hydroxycitronellal	107-75-5	100	4960	1E: 1 subject induction #3	0	1E: 1 subject challenge #3	0	Subject #92 - negative at rechallenge	Negative
Hydroxycitronellal	107-75-5	110	1181	4E-1E: 1 subject (subject #18, showed 1E-2 level reactions during induction as well, refused rechallenge); 2E: 1 subject (subject #39); 4: 1 subject (subject #127 also showed positive reactions during induction)	4E-1E: 1 subject (subject #18, showed 1E-3 level reactions during induction as well)	N/A	0	For 2 subjects: #127 negative at rechallenge @ 1% and 5%, #39 negative at rechallenge @ 1% and 5%	Negative
Isobornyl acetate	125-12-2	99	6496	0	0	0	0		Negative
Isobornyl acetate	37677-14-8	108	5905	0	0	0	0		Negative
Isobornyl acetate	23787-90-8	110	9093	0	0	0	0		Negative
Jasmine	8022-96-6	114	1476	0	1E: 1 subject challenge #3	1E: 1 subject challenge #3	0	Subject #105: negative at rechallenge	Negative
Jasmine sambac	1034798-23-6	109	8858	1E: 1 subject induction #4; 1E: 1 subject induction #5	0	0	0		Negative

Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ced.14100 by University of California, San Diego on 09/05/2020

L-Carvone	6485-40-1	93	18,896	1E: 2subjects induction #6; 1E: 5 subjects induction #7; 1E: 4 subjects induction #8; 1E: 1 subject induction #9; 3; 1 subject induction #7; 1E: 3 subjects challenge #2; 1E: 3 subjects challenge #3; 1E: 3 subjects challenge #4	0	0	N/A	Sensitization reaction in 4 subjects
L-Carvone	6485-40-1	99	2675	1E: 1 subject induction #8	0	0	N/A	Negative
Linalool	78-70-6	119	14,999	0	0	0		Negative
Linalyl acetate	115-95-7	99	2362	0	0	0		Negative
Longifolene	475-20-7	105	3543	0	0	N/A		Negative
Menthadiene-7-methyl formate	68683-20-5	101	1063	0	0	0		Negative
Methyl 2,6,10-trimethylcyclodeca-2,5,9-trien-1-yl ketone	28371-99-5	106	4724	0	0	N/A		Negative
Methyl atrarate	4707-47-5	100	11,810	0	1E: 1 subject, challenge #3; 1E: 1 subject, challenge #4	0		Negative
Methyl cinnamate	103-26-4	105	2953	0	0	N/A		Negative
Methyl hexadecanoate	112-39-0	103	2480	0	0	0		Negative
Methyl octanoate	111-11-5	103	4724	0	0	0		Negative
Methyl <i>p</i> -methylbenzoate	99-75-2	112	4133	0	0	0		Negative
Musk ketone	81-14-1	107	6023	0	0	0		Negative
Octahydro-4,7-methano-1 <i>H</i> -indenecarbaldehyde	30772-79-3	102	1181	0	0	N/A		Negative
Octahydro-5,5-dimethylnaphthalene-2-carbaldehyde	68738-96-5	110	5078	0	0	0		Negative
Octahydro-7-methyl-1,4-methanonaphthalen-6(2 <i>H</i>)-one	41724-19-0	103	5315	0	0	0		Negative
Oxacyclohexadecen-2-one	34902-57-3	111	7559	0	0	0		Negative
<i>p</i> -Isobutyl- <i>m</i> -ethyl hydrocinnamaldehyde	6658-48-6	104	2362	0	0	0		Negative
<i>p</i> -Mentha-1,3-diene	99-86-5	110	2244	0	0	0		Negative
<i>p</i> -Mentha-1,8-dien-7- <i>al</i>	2111-75-3	116	709	0	0	0		Negative
<i>p</i> -Methoxybenzaldehyde	123-11-5	102	3543	0	0	0		Negative
<i>p</i> -Methoxybenzaldehyde	123-11-5	109	2363	0	0	0		Negative

(Continued on next page)

Piperonyl acetate	326-61-4	104	4724	0	0	0	Negative
Styrax (G)	8024-01-9	114	1500	0	0	N/A	Negative
Styrax (H)	8024-01-9	105	2000	0	0	N/A	Negative
trans-2-Decenal	3913-81-3	105	236	0	0	0	Negative
Tridecene-2-nitrile	22629-49-8	108	6967	0	0	0	Negative
Triethyl citrate	77-93-0	106	25,982	0	0	N/A	Negative
Vanillin	121-33-5	114	5314	0	0	0	Negative
Vanillin	121-33-5	105	1181	0	0	0	Negative
Vanillin isobutyrate	20665-85-4	109	590	0	0	0	Negative
Vanillyl butyl ether	82654-98-6	104	3543	0	0	0	Negative
Vetiveryl acetate	117-98-6	112	2362	0	0	0	Negative
Ylang-ylang	8006-81-3	109	1772	0	0	0	Negative

*Scoring scheme is provided.

CAS, Chemical Abstracts Service; HRIPT, human repeated insult patch test; N/A, not applied.

test protocol throughout the whole study. It is important to bear in mind that a number of subjects may drop out during the course of the test.²⁰

Study Curation

The RIFM has a long history of conducting human studies as part of the assessment of skin-sensitizing activity. For this publication, only those HRIPTs conforming to the fully defined published protocol, which has now been used as a matter of routine for the last 30 years, were selected. The HRIPT studies conducted by the RIFM were collected from the RIFM Database in December 2019 (<https://rifmdatabase.rifm.org/>). The study reports were manually examined to assess whether they conform to the currently published protocol. The studies compiled include both shared and exclusive panels of subjects. Of 345 HRIPTs, a little fewer than half met the criteria of fully conforming to the published protocol. Reasons that an HRIPT did not conform included failure to have at least 90 participants to complete the study, use of a nonstandard vehicle, or noninclusion of control patches. It was also required that a given fragrance material is not tested along with other fragrance materials on the same group of subjects. It is important to note that among the 345 HRIPTs, some were done on 50 subjects, and a second study was undertaken later to produce a combined study total of at least more than 90 participants.

RESULTS

The RIFM records contained 154 HRIPTs fulfilling the quality criteria (see hereinabove), conducted on the 134 fragrance substances. These are reported in Table 2. Note that column 3 details the number of subjects who completed the study; a greater number would have been enrolled, but approximately 10% on average fail to complete any study. Dermal responses to fragrance material occurred in 27 (18%) of the 154 studies during the induction and/or challenge phases; vehicle reactions occurred in 14 (9.1%) of the 154; reactions to saline occurred in 12 (9.4%) of 128. Only in 4 subjects were there reactions to both vehicle and saline, such that 22 (14%) of 154 showed a degree of skin response. Finally, in 10 of the 154 studies, subjects who reacted to test materials also reacted to vehicle and/or saline. Thus, there is considerable overlap and nonspecificity, associated with minor skin reactions to a 24-hour occlusive patch application (all of which resolves rapidly), which has to be filtered out of detailed analysis of the data set in terms of potential allergic responses.

Overall, 20 of 16,512 volunteers in 9 studies exhibited de novo skin sensitization induction. What follows is a brief commentary on substances where there was evidence of sensitization reaction. Presensitization, nonspecific reactions, and intermittent skin reactions during the induction that were not confirmed at the challenge phase were not considered as evidence of de novo skin sensitization induction. They are referred to in their order of first appearance in Table 2, except where substances have been grouped.

2,6,6-Trimethylcyclohexa-1,3-dienyl methanal was negative at a lower dose of 29 $\mu\text{g}/\text{cm}^2$. At the higher concentration of 59 $\mu\text{g}/\text{cm}^2$, it gave 4 grade 1E reactions during induction, including 1 at the first induction, indicating a preexisting contact allergy to this material. Two grade 1E reactions were observed at the challenge, and a single subject had the same reaction to vehicle during induction. In 1 subject, the pattern of reactions, appearing during the later induction steps and repeated at the challenge, persisting to the 72-/96-hour time point, indicated contact allergy had been induced.

4-Hydroxy-2,5-dimethyl-3(2H)-furanone gave grade 1E reactions in 2 subjects, 1 only during the later stages of induction and 1 upon induction and challenge. The absence of a response at the challenge phase in 1 subject who reacted during induction is inconsistent with allergy. The other subject with low-level reactions during the later induction stages exhibited a grade 1E reaction during the 48-hour challenge reading, which subsided to grade 1 at the later readings. However, the positive response was confirmed during the rechallenge, confirming the induction of contact allergy.

6-Methyl-3,5-heptadien-2-one produced minimal responses in several subjects during the induction phase; at challenge, 3 individuals presented grade 1E responses at all later readings, and two-thirds showed a response at the ninth induction and at rechallenge, thereby confirming the induction of contact allergy. In an addition of 2 subjects, low-grade erythema reactions in the absence of edema arose at the challenge, in 1 case being preceded by a slight reaction at the ninth induction.

Benzaldehyde at 590 $\mu\text{g}/\text{cm}^2$ was negative, but at 10 times the concentration, reactions occurred, some clear evidence of the induction of contact allergy. Minor nonspecific irritation reactions occurred during the induction. However, in 12 subjects, grade 1E and occasional 2E reactions occurred at the challenge, often after dermal responses during the latest stages of induction. Skin reactions in 6 of 12 subjects subsided on the last challenge reading. Rechallenge was deemed unnecessary as the sensitization response was clear; 6 of 104 subjects were sensitized to 5900 $\mu\text{g}/\text{cm}^2$ of benzaldehyde.

Benzyl alcohol was tested at 3 concentrations. At 3543 $\mu\text{g}/\text{cm}^2$, it was negative, whereas at 5905 and 8858 $\mu\text{g}/\text{cm}^2$, the picture was less clear. At the mid concentration, almost no irritation was apparent, but 2 subjects showed grade 1E or 2E responses at the first induction patch, indicating that they may be allergic to benzyl alcohol. Both subjects had no further induction patches; they reacted more strongly at challenge, confirming that they already had contact allergy. No other subject showed evidence of the induction of contact allergy. At the highest dose, 1 subject exhibited a grade 3 reaction during the fourth induction. The induction was continued at a different site, but a grade 2E reaction was observed. Because of the strong reactions, the rest of the induction patches were omitted. This subject showed reactions at challenge, which were confirmed at rechallenge, leading to the conclusion that 1 of 110 developed allergy to 8858 $\mu\text{g}/\text{cm}^2$ of benzyl alcohol.

L-Carvone was evaluated at 18,896 and 2675 $\mu\text{g}/\text{cm}^2$. It gave rise to significant evidence of minor skin irritation during the induction, with 1 in 3 subjects at 18,896 $\mu\text{g}/\text{cm}^2$ experiencing multiple responses. In 12 subjects, the reactions were sufficient to dictate that challenge patching was not appropriate. This does not indicate sensitization induction because many had very early induction phase skin reactions, and other subjects who had experienced similar induction reactions were negative at challenge. Ultimately, 4 subjects experienced reactions consistent with the induction of contact allergy. When L-carvone was tested at a lower dose, 2675 $\mu\text{g}/\text{cm}^2$, 1 subject exhibited a 1E reaction in response to 1 of the 9 induction patchings. At challenge, no subject exhibited skin reactions, confirming that contact allergy was not induced at this lower dose.

p-Methoxybenzaldehyde was negative at 2363 and 3543 $\mu\text{g}/\text{cm}^2$, but at 4724 $\mu\text{g}/\text{cm}^2$, several subjects exhibited grade 1E reactions during induction. However, scattered \pm reactions indicated a slight irritant response to test material. One subject had a grade 2E reaction to the final induction treatment. At challenge, several minor responses were seen; a sole subject displayed a pattern of late-developing response during induction, also seen at challenge, and thus consistent with induction of contact allergy. However, the reaction could not be reproduced at rechallenge, questioning whether it was a true-positive response. The subject with a grade 2 response at the final induction stage had only \pm reactions at challenge, which, given the many scattered irritant responses seen throughout the study, is unlikely to be sensitization. Taking a conservative view, a single subject may have been sensitized in this study, but it remains doubtful. A study at a higher concentration, 6496 $\mu\text{g}/\text{cm}^2$, was carried out. There were very few reactions during induction (or challenge) indicative of irritation. One subject presented a grade 1 response at all challenge readings, and at the same time, the induction site displayed the same degree of reaction. Another subject exhibited grade 1E reactions during the 72- and 96-hour readings. This induction of contact allergy was confirmed by a positive rechallenge.

p-*t*-Butyl- α -methylhydrocinnamic aldehyde applications led to evidence of minor irritant reactions, matched in intensity, exceeded in number by those to the vehicle control in 1 subject, and thus attributed to nonspecific rather than allergenic effects. Sensitization effects were seen in 2 other subjects, who showed reactions to the fragrance material without exhibiting reactions to vehicle or saline.

DISCUSSION

In the material reported herein, the outcomes of the 154 studies involving 16,512 human volunteers are detailed. As expected from the conservative approach to the prediction of the no expected sensitization induction level (NESIL), most HRIPTs conducted (76%) did not lead to any type of dermal reaction, whether irritant or allergic. Allergy induction was evidenced in 20 subjects (0.12% of those tested out of 16,512 total subjects) across 9 of the 154 studies. In

other words, in fewer than 6% of HRIPTs was there any evidence of the induction of allergy. The reactions were seen with 8 (6.0%) of the 134 substances. Furthermore, with the introduction of the LLNA and a more standardized approach, both to risk assessment and to the procedure of the HRIPT over the last decade, the proportion of those in whom sensitization was induced fell significantly, from 17/6658 to 3/9854 ($P < 0.001$; Fisher exact test, 2-sided), by a factor of more than 4, to 0.03%. Subjects who exhibited skin reactions were advised (and fully funded) to follow up with a dermatologist. There is a mechanism for a long-term follow-up with the volunteers who reacted, and none of the subjects subsequently reported any adverse effects from their daily use of consumer products.

At the heart of this article is the essential debate concerning the balance of the procedure of a human skin sensitization test, the wider interests of human safety, and matters of ethics in toxicological risk assessment. Implicit in a commentary by one of the authors of this article, some years ago, was the principle that an HRIPT could not be considered ethical if the study did not have scientific credibility.²¹ The HRIPTs carried out by RIFM are often given the epithet “confirmatory,” but it is essential to be clear that the expected negative outcome is not certain. The analysis carried out before the initiation of an HRIPT focuses on the definition of a NESIL.^{17,25} However, to ensure that prediction is accurate, the human test is also performed, because there is evidence that albeit infrequently, the potency of a skin sensitizer differs between mice and humans.²⁶ Thus, this usage passes the essential criterion that it has scientific merit. Given that there is no appropriate method that can accurately replace human testing at this time, HRIPTs remain necessary for minimizing the risk of skin sensitization for a larger normal human population. However, in doing so, it puts the panelists involved at (low) risk of the induction of sensitization. It is for this reason that some commentators regard the implementation of the test to be unethical.²² Although it might be argued that the question of ethics should be left to a properly constituted and wholly independent ethical review committee, it is entirely reasonable to ask, “What is the level of risk to which HRIPT panelists are exposed?” That main purpose to quantify the level of risk is complete—it is very low.

This comprehensive retrospective analysis confirms the underlying rationale for the management of a human study in which an in vitro/in vivo prediction of a NESIL remains imperfect. It substantiates earlier conclusions from a smaller data set.²⁷ Although this continues to be the status quo, the only alternative to the application of an additional uncertainty factor to all NESIL predictions, the great majority of which are adequately accurate, is to carry out a human confirmatory study. Furthermore, to distinguish this type of study from the general product testing that is undertaken in the HRIPT, we propose in the future to refer to the study via the acronym CNIH (confirmation of no induction in humans).

ACKNOWLEDGMENTS

The authors thank the many thousands of people who volunteered to participate in the HRIPTs reported in this publication.

REFERENCES

- Shelanski HA. Experience with and considerations of the human patch test method. *J Soc Cosmet Chem* 1951;2:324–331.
- Shelanski HA, Shelanski MV. A new technique of human patch tests. *Proc Sect Toilet Goods Assoc* 1953;19:46–49.
- Schwartz L. The skin testing of new cosmetics. *J Soc Cosmet Chem* 1951;2:321–324.
- Schwartz L. Twenty-two years' experience in the performance of 200,000 prophetic patch tests. *South Med J* 1960;53:478–483.
- Shelanski MV, Phillips SB, Potts CE. Evaluation of cutaneous reactivity to recently marketed dermatologic products. *Int J Dermatol* 1996;35:137–140.
- Marzulli FN, Maibach HI. The use of graded concentrations in studying skin sensitizers: experimental contact sensitization in man. *Food Cosmet Toxicol* 1974;12:219–227.
- Marzulli FN, Maibach HI. Effects of vehicles and elicitation concentration in contact dermatitis testing. I. Experimental contact sensitization in humans. *Contact Dermatitis* 1976;2:325–329.
- Atkinson JC, Rodi SB. Effect of vehicles and elicitation concentration in contact dermatitis testing. II. Statistical analysis of data. *Contact Dermatitis* 1976;2:330–334.
- Marzulli FN, Maibach HI. Contact allergy: predictive testing in man. *Contact Dermatitis* 1976;2:1–17.
- Marzulli FN, Maibach HI. Further studies of effects of vehicles and elicitation concentration in experimental contact sensitization testing in humans. *Contact Dermatitis* 1980;6:131–133.
- Stotts J. Planning, conduct and interpretation of human predictive sensitization patch tests. In: Drill VA, Lazar P, eds. *Current Concepts in Cutaneous Toxicity*. St. Louis, MO: Academic Press (Elsevier); 1980:41–53.
- Marzulli FN, Maibach HI. Antimicrobials: experimental contact sensitization in man. *J Soc Cosmet Chem* 1973;24:399–421.
- Marzulli FN, Maibach HI. Contact allergy: predictive testing of fragrance ingredients in humans by Draize and maximization methods. *J Environ Pathol Toxicol* 1980;3:235–245.
- Henderson CR, Riley EC. Certain statistical considerations in patch testing. *J Invest Derm* 1945;6:227–232.
- Api AM. The use of human data when conducting dermal sensitization quantitative risk assessments for fragrance ingredients. *Regul Toxicol Pharmacol* 2008;50:163–165.
- Api AM, Belsito D, Bruze M, et al. Criteria for the Research Institute for Fragrance Materials, Inc. (RIFM) safety evaluation process for fragrance ingredients. *Food Chem Toxicol* 2015;82(suppl):S1–S19.
- Api AM, Basketter DA, Cadby PA, et al. Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. *Regul Toxicol Pharmacol* 2008;52:3–23.
- Api AM, Vey M. Implementation of the dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. *Regul Toxicol Pharmacol* 2008;52:53–61.
- Politano VT, Api AM. The Research Institute for Fragrance Materials' human repeated insult patch test protocol. *Regul Toxicol Pharmacol* 2008;52:35–38.
- McNamee PM, Api AM, Basketter DA, et al. A review of critical factors in the conduct and interpretation of the human repeat insult patch test. *Regul Toxicol Pharmacol* 2008;52:24–34.
- Basketter DA. The human repeated insult patch test in the 21st century: a commentary. *Cutan Ocul Toxicol* 2009;28:49–53.
- Scientific Committee on Consumer Safety. Memorandum on use of human data in risk assessment of skin sensitisation. Adopted December 158, 2015.

Available at: https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_010.pdf. Accessed October 8, 2019.

23. Menné T, Thyssen JP, Boman A, et al. The critical review of methodologies and approaches to assess the inherent skin sensitization potential (skin allergies) of chemicals. 2009. Available at: https://ec.europa.eu/health/sites/health/files/scientific_committees/docs/service_contract_20096104_en.pdf. Accessed October 8, 2019.
24. Kimber I, Dearman RJ, Basketter DA, et al. Dose metrics in the acquisition of skin sensitization: thresholds and importance of dose per unit area. *Regul Toxicol Pharmacol* 2008;52(1):39–45.
25. Api AM, Basketter DA, Bridges J, et al. Updating exposure assessment for skin sensitization quantitative risk assessment (QRA2) for fragrance materials. *Regul Toxicol Pharmacol*. Available online: 17 October 2020.
26. Api AM, Parakhia R, O'Brien D, et al. Fragrances categorized according to relative human skin sensitization potency. *Dermatitis* 2017;28:299–307.
27. Ritacco G, O'Brien D, Na M, et al. Retrospective analysis of a subset of human repeat insult patch tests conducted by the Research Institute for Fragrance Materials over the past 30 years. *J Invest Derm* 2019;139, number 5S, supplement 1, abstract 171.