Organisation for Economic Co-operation and Development

ENV/CBC/MONO(2025)2/ANN4

Unclassified

English - Or. English

11 September 2025

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

Annex 4: Report of the Human Data Sub-Group on the Curation and Evaluation of the Human Reference Data and the Derivation of Associated Substance Classifications

Series on Testing and Assessment
No. 336

(Second edition)

E-mail: ehscont@oecd.org.

JT03570710

Please cite this publication as:

OECD (2025), *Supporting document to Test Guideline 497 on Defined Approaches for Skin Sensitisation, Annex 4: Report of the Human Data Sub-Group on the Curation and Evaluation of the Human Reference Data and the Derivation of Associated Substance Classifications*, OECD Series on Testing and Assessment, No. 336, OECD Environment, Health and Safety, Paris, https://one.oecd.org/document/ENV/CBC/MONO(2025)2/ANN4/en/pdf

Contact us

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

E-mail: ehscont@oecd.org

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several Partner countries and from interested international organisations attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (https://www.oecd.org/en/topics/chemical-safety-and-biosafety.html).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank, Basel, Rotterdam and Stockholm Conventions and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

Table of contents

1 Summary

As described in more detail in section 2 of this report, the human data sub-group (HDSG) of the OECD Expert Group on Defined Approaches for Skin Sensitisation (OECD EG DASS) has performed a review of the available human predictive patch test (HPPT) database, in order to

- describe the variability and uncertainties associated with these data,
- explore their usefulness for classifying chemicals for their skin sensitisation potential, and
- use these data – as far as available – to classify a set of 200 reference substances with respect to their skin sensitisation potential according to the rules of the United Nations' Globally Harmonized System of Classification and Labelling of Chemicals (GHS) in its current version[1].

Figure 1 provides a brief overview of the chronology of this activity and the subsequent review and Quality Control (QC) process.

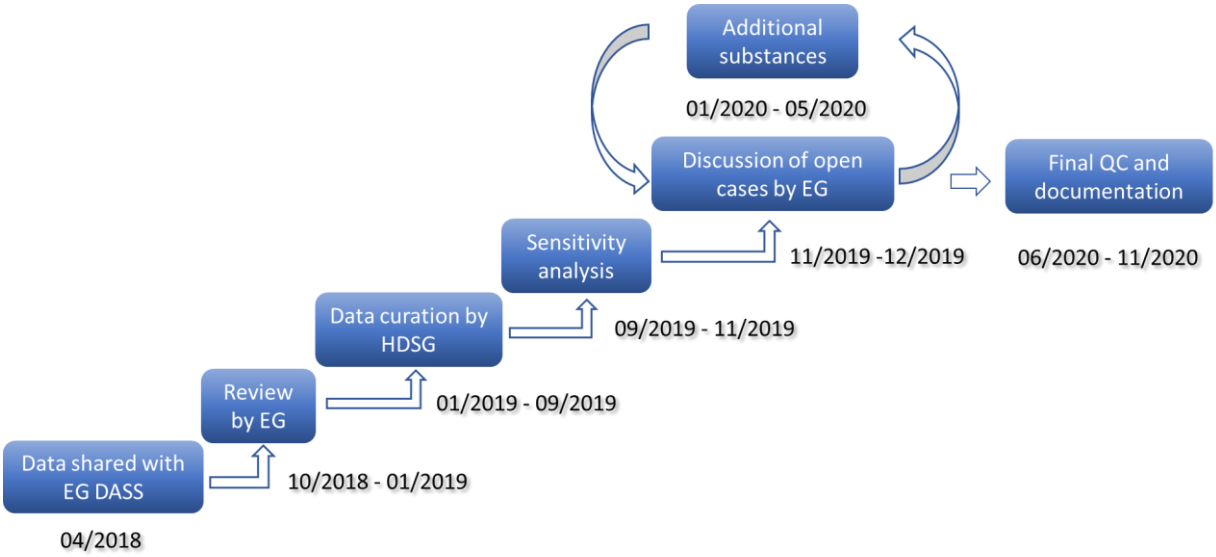

Figure 1: Chronology of the work of the EG DASS on the human reference data

In total, NICEATM and BfR have compiled a database of currently[2] 2277 human predictive patch test (HPPT) results referenced in more than 1700 publications from the late 1950s until and including December 2019. Human diagnostic patch test data, which – in principle – can also be used for classification under the GHS, were not included due to considerations reported in section 2 of this report. The database contains information on substance identity, test design, and test results and provides bibliographic information on the (mostly unpublished) original test reports as well as on later publications citing these primary sources. A system for categorising the HPPT results with respect to their relative reliability (relative reliability score, RRS, ranging from 1/high relative reliability to 5/unreliable) has been developed (cf. 3.3.1); 2255 test results with RRS < 5 were taken forward for further analyses. These activities are reported in detail in section 3 of this report, while section 4 provides an overview of the database including basic descriptive statistics.

In section 5, the multiple factors potentially affecting HPPT data variability and uncertainty are qualitatively discussed in some detail. For a more quantitative assessment, either the necessary data

[1] Rev. 8, https://www.unece.org/trans/danger/publi/ghs/ghs_rev08/08files_e.html, last accessed 2021-07-10

[2] The reference substances used by the EG DASS for DA performance assessment make up only a small part of the whole database. The complete database will be published after a final round of editorial quality control. As a consequence, numbers for the final published version of the database may slightly deviate from those provided here.

do not exist or more work is needed, which is beyond the scope of the current phase of the OECD project on the DA guideline (cf. section 0 for potential follow-up activities).

When compared with the respective GHS classification criteria, the vast majority (2188, or 97%) of the available HPPT results with sufficient reliability did not translate into an unambiguous classification, including sub-categorisation, i.e. GHS Skin Sens. 1A, 1B, or Not Classified (NC)[3]. About 73% (1642) of the test results did not even allow for an unambiguous binary classification decision (Skin Sens. 1 vs. NC)[4]. This is mainly due to the fact that HPPT are normally only performed at one dose level which is often either too high or too low[5] to reliably exclude the need for a stricter classification at a different concentration/dose level. As a consequence, the HDSG developed new concepts for handling ambiguity for at least part of the available data, which are reported in detail in section 6 of this report.

Section 7 of the report describes how the individual test results were translated into so-called "extrapolated" classifications by applying the concepts developed in the previous sections to the 453 test results available for the EG DASS reference chemical list. Furthermore, it explains how, in the case of discordant individual test results, the overall classification was obtained by means of a weight-of-evidence (WoE) assessment.

In section 8.2, the reproducibility of HPPT-based reference classifications is analysed and characterised quantitatively. Although the results are based on only 14 - 34 substances with multiple test results (and therefore conclusions should be drawn with care), HPPT-based GHS reference classifications were found to be highly reproducible on average.

Finally, the resulting reference classifications are provided in an overview table, with more detailed documentation provided in Appendix 2 to this report. Data suitable for classification (albeit not always unambiguous) could be obtained for 104 of the 196 EG DASS reference substances, while for 92 substances either no or only unreliable primary HPPT data were available.

2 About this project

2.1 Introduction

"Defined approaches (DAs) to testing and assessment" (OECD, 2016b) consist of a fixed data interpretation procedure (DIP) used to interpret data generated from a defined set of information sources. In recent years, based on the Adverse Outcome Pathway (AOP) published by the OECD (OECD, 2012), DAs for skin sensitisation have been published, which rely on new methodologies for skin sensitisation testing and assessment, such as *in vitro* and *in silico* methods (OECD, 2016a). Previous work under a project led by Cosmetics Europe provided a preliminary demonstration that many of the DAs under OECD consideration had superior performance to the LLNA when compared to expert-derived human classifications (Hoffmann et al., 2018; Kleinstreuer et al., 2018).

To judge the performance of these methods and, therefore, their usefulness for regulatory purposes such as classification & labelling, a suitable reference standard is required. Up to now, animal testing,

[3] 59/2255 test results (2.6%) were positive at a dose per skin area (DSA) ≤ 500 µg/cm² (clear GHS 1A) and 8/2255 test results (0.4%) were negative with a test concentration of 100% (clear NC). For details on the current GHS classification criteria, cf. section 6.2.

[4] 1642/2255 test results (72.8%) were negative, but these substances were tested at concentrations < 100% (i.e. it was uncertain whether a higher test concentration might have given a positive result) or the test concentration applied was not reported.

[5] This is not a criticism of the original data (or their authors), since these experiments were usually designed to test the safety of in-use concentrations << 100% and were not performed with the goal of classification in mind. However, the fact needs to be mentioned here because it strongly affects the usability of the historical HPPT database for classification purposes.

in particular the Local Lymph Node Assay (LLNA), still represents the "gold standard" in many regulatory domains.

In recent years, however, the variability of the LLNA has been critically discussed in the published literature (Dimitrov et al., 2016; Dumont et al., 2016; Hoffmann, 2015)[6]. It is unclear whether replacing the LLNA with any of the currently available alternative methodologies would significantly reduce variability. Nevertheless, both the desire to keep animal testing to an indispensable minimum based on ethical reasons and the so-called "species barrier" are strong arguments for trying to replace the LLNA with animal-free testing or prediction methods.

In this context, some developers of new skin sensitisation test methodology, e.g. Urbisch et al. (2015), have claimed that the results of their test methods/strategies were "more relevant to humans" than those of the LLNA, because of their use of human cell lines and because of their higher concordance with human test data. However, many such comparisons rely on a human reference dataset consisting of studies which mainly have been performed in the 1960s - 80s, i.e. up to almost 60 years ago, under non-guideline protocols, with unclear quality control and scarce, if any, reporting beyond a general reference to a test design (vehicle, number of total subjects tested, and number of subjects considered sensitised as a result of the test). On the other hand, such data were accepted as part of the human reference set when the LLNA was validated for regulatory use, unfortunately with incomplete documentation of how their adequacy and reliability for this purpose was assessed.

As a consequence of these considerations, the OECD "Expert Group on Defined Approaches for Skin Sensitisation" (OECD EG DASS), which took up work in 2018, has recognised the need for an in-depth analysis of the reliability and robustness of the potential benchmark reference data (i.e. both LLNA and human data) for measuring DA performance. Sub-groups were formed to further refine the reference data and create robust datasets that can be used to a) benchmark DA performance and b) analyse and characterise the sources of variability and uncertainty in the reference test methods. This report describes the work of the human data sub-group (HDSG).

Patch testing is the technique of choice in human skin sensitisation testing and assessment. Two basic types of patch test data can be distinguished: human diagnostic patch tests (HDPT) and human predictive patch tests (HPPT). The HPPT can be further divided into two study designs: the human maximisation test (HMT) and the human repeated insult patch test (HRIPT), cf. section 4, Table 1 for further details on the protocols.

The HDSG performed a detailed analysis of the available reference data only for HPPT, due to the following considerations regarding the limitations of HDPT data for classification and labelling:

- HDPT cover elicitation (response to challenge with allergen) and not induction of sensitisation.
- In the case of HDPT data, previous exposure to the chemical often cannot be established with enough certainty to evaluate its potency for induction of sensitisation; even if it is established qualitatively, exposure can almost never be quantified in a satisfying manner (there might be exceptions at the workplace, but for consumers it is practically not possible). The situation is further complicated because patients might have had contact with other chemicals cross-reacting with the chemical in question.
- For adverse toxicological effects in general, both the incidence in the population (the number of affected individuals) and the magnitude of effect in affected individuals increase with increasing dose/exposure (cf. WHO IPCS (2017) for a more detailed elaboration on this general concept). For skin sensitisation, magnitude of effect in affected individuals could be expressed by the dose required to elicit an allergic reaction as well as by the severity of that reaction. While little seems to be known about the correlation of induction dose/pattern and severity of the allergic response,

[6] Note that it is beyond the scope of the present report to discuss potential shortcomings of these papers or the validity of these authors' conclusions.

the dose needed for elicitation has been shown to be inversely related to the induction dose, cf. section 5.5.3. In addition, the number of induction exposures, the intervals between exposures, and the overall duration of exposure will influence the magnitude of sensitisation in a given individual. As a consequence, to compare two chemicals with respect to their potency would require good knowledge about exposure levels and patterns. For a realistic comparison of the examined populations, exposure levels and patterns should be comparable for both chemicals, which is not the case for HDPTs, as this information is not readily available.

- Published HDPT data comprise reports from dermatological clinics, from institutions dedicated to workplace safety, or, less frequently, private practitioners with an academic inclination. Data recorded in companies at the workplace often remain in-house. HDPT data are published in the form of single case reports, epidemiological case studies, or in aggregated form by dermatology clinic networks reporting the results from standard series testing in large cohorts, such as the German "Informationsverbund Dermatologischer Kliniken" (IVDK), the "European Surveillance System on Contact Allergies" (ESSCA), or the "North American Contact Dermatitis Group" (NACDG). Both latter formats will generally only reflect well-known sensitisers. In contrast, single case reports are often published to document the first case(s) of sensitisation by a chemical that previously had not been on the record, with potentially much smaller interest on the publishers' side to publish the third, fourth, or fifth case. From all of this it follows that the number of publications available for a skin sensitiser does not constitute a reliable indicator of its sensitisation potential or incidence in the population. Or, in other words: absence of reports does not equate to absence of effects.
- HDPTs may be carried out in unselected, consecutive dermatitis patients, in selected dermatitis patients (e.g. consumers, where targeted testing might be directed to the ingredients of a specific cosmetic product), or in workers known to be exposed to a given chemical at their workplace. Test results in these different settings have a different meaning and cannot be compared directly.
- Nevertheless, HDPT data - beyond their immediate use in elucidating the origin of a patient's allergy - can provide important information about "allergy trends" in society, thereby helping to set regulatory priorities (EAACI, 2018). Depending on how clearly a positive result can be traced back to exposure to an individual chemical, they can also provide general supportive evidence that a chemical is capable of sensitisation. In this way, they may also factor into weight-of-evidence classifications in cases where the "gold standard reference data" are not capable of producing unambiguous results. Pooled HDPT data can also be used for classification under the GHS in case of a "substantial incidence of reactions in a defined population" (GHS, rev. 8).

With respect to HPPT data, it is somewhat striking that the relevant questions which the HDSG tries to address in the present report have already been asked more than half a century ago by Albert Kligman who developed the HMT:

> *"How do authors of these tests know that they are predictive? How well do they perform in warning against the risk of sensitization? Are known allergens readily detected by these technics? Hundreds, perhaps thousands, of preparations have been screened by these procedures, but the results have not been published. The literature contains no study which demonstrates the comparative effectiveness of these tests for even one substance. The issue can be framed in a single question. Have the tests been tested?"* (Kligman, 1966a)

Even more strikingly, while many test results have been published since 1966, it appears that a fundamental review of the potential as well as the limitations of these tests in the context of classification and labelling has not been performed. This report can be considered a first step in that direction.

2.2 Remit of the HDSG

The following specific tasks have been identified by the HDSG and confirmed by the OECD EG DASS:

- curation of the HPPT skin sensitisation database,
- analysis of the variability and uncertainty in the HPPT database,
- development of a framework for using HPPT data to classify chemicals with respect to their skin sensitisation potential based on the GHS, and
- classification – as far as possible - of the reference substances used by the OECD EG DASS for DA performance assessment with respect to their skin sensitisation potential.

2.3 Members of the HDSG (in alphabetical order)

Anne-Marie Api	Research Institute for Fragrance Materials (RIFM), on behalf of BIAC
John Gordon	United States Consumer Product Safety Commission (CPSC)
Matthias Herzler	German Federal Institute for Risk Assessment (BfR)
Nicole Kleinstreuer	United States National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), representing the leads of the OECD project on developing a guideline for Defined Approaches for skin sensitisation
Hon-Sum Ko	United States Food and Drug Administration (FDA)
Joanna Matheson	United States Consumer Product Safety Commission (CPSC)
Judy Strickland	Integrated Laboratory Systems (ILS), on behalf of NICEATM
Hermann-Josef Thierse	German Federal Institute for Risk Assessment (BfR)

Disclaimer: The views expressed in this report do not necessarily represent an official view of the institutions to which its authors are affiliated, the OECD, or of the EG DASS as a whole."

2.4 Outlook

The work reported here has provided many insights into the nature of the available HPPT database, along with its merits and limitations. A pragmatic approach was taken to provide a robust and consistent way of classifying the reference substances for the purposes of the OECD EG DASS. However, from a scientific perspective, many questions remain open which could not be addressed in greater detail due to the constraints in time and scope defined by the OECD project on DAs for skin sensitisation. Therefore, members of the HDSG plan to address a number of open questions in follow-up activities outside of/beyond the OECD project:

- NICEATM and BfR intend to make the HPPT database available to the public to serve as an information source for those interested in specific chemicals, for assisting with activities such as *in silico* modelling and DA development and also – by means of documenting citation chains - as a means of tracing HPPT results reported in the secondary literature back to the original reports.
- Given the time and resources available for this project, it was only possible to provide a qualitative overview of the factors affecting the variability and uncertainty in the HPPT data, citing exemplary references where they were readily available. More work is needed and members of the HDSG have expressed their intention to perform a more thorough review of these issues to be published in the peer-reviewed literature. This review will as much as possible utilise the HPPT database

compiled within this project, e.g. to help in better understanding dose-response relationships in human predictive patch testing.

- In addition, potential deficiencies of the current GHS classification scheme (as far as it refers to the use of HPPT data) have been identified and ways for improving this scheme should be explored. This work could directly factor into the process of adapting the GHS system for use with non-animal test results currently ongoing at the UN. As a matter of fact, the rather simplistic GHS approach seems deficient in a number of ways, e.g. by using a dose metric (µg/cm^2) which may fall short of the complexity of HPPT procedures by not accounting for the number of repetitions of the induction step, challenge concentrations, or the number of repetitive challenge exposures.
- Moreover, under the GHS, HMT and HRIPT results are taken as equivalent, which would need further critical review in this context, as would previously published approaches for categorisation such as the ones published in Basketter et al. (2014) or Api et al. (2017).
- Finally, given the variability in the HPPT data, the uncertainties associated with the new "extrapolated" classification approach taken in this report, the fact that for many substances only one test was available, and the low number of unambiguous classifications, the HDSG recommends further work to compare the classifications based on LLNA and HPPT data to identify possible systematic deviations that could be due to the chosen classification methodology rather than to differences between the test methods themselves. Such an effort, which could also include additional data sources such as the Guinea Pig Maximisation Test (pending curation efforts to ensure equivalent data quality), could point out possible ways to refine and improve the methodology used in this report.

3 Curation of the HPPT database

3.1 History

The HPPT database reported here was built starting from a database collated previously at the United States National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). The original NICEATM database comprised a total of 884 entries (i.e. individual HPPT results) as well as a list of the respective source literature references. Most of these references were so-called “monographs” on fragrance ingredients, published by the Research Institute for Fragrance Materials (RIFM) since the early 1970s in the journal “Food and Chemical Toxicology” (previously: “Food and Cosmetics Toxicology”). Additional references from the published literature were also included.

A first review of the NICEATM database at the German Federal Institute for Risk Assessment (BfR) revealed that not all the references provided in the NICEATM database referred to the original test reports, resulting in several previously unnoticed duplicates. Moreover, it was noted that a considerable number of additional HPPT data, particularly in the form of further RIFM monographs, were available in the published literature.

Considering that the database should be as broad as possible to serve the intended review of variability and uncertainty in the HPPT data, BfR decided to perform the following tasks:

- verify the data compiled by NICEATM,
- add additional test data from all RIFM fragrance material monographs as well as further publications[7],
- add additional data fields for absolute number of test subjects with a positive test result, chemical identity (ID), primary reference, and test protocol,
- exclude substances with undefined/poorly characterised composition or test results with missing key information,
- capture in detail all different test designs as far as reported, and
- annotate each entry with all later 2nd (3rd, 4th, 5th...) order citations to make the test results used in these papers traceable to their original source.

In turn, NICEATM volunteered to perform quality control on all entries newly added by BfR. This procedure ensured that the essential information for each entry in the final database was double-checked, i.e. reviewed for correctness by both institutions.

3.2 Sources

In a first step, bibliographic information and full texts of all RIFM monographs on fragrance ingredients as published in the peer-reviewed journal Food and Cosmetics Toxicology (January 1973 - December 1981) and its successor Food and Chemical Toxicology (January 1982 – December 2019) were gathered and stored in an EndNote™ (v. X9.2) reference library. Next, this database was complemented with further journal articles, book chapters, and reports from the published literature which had been provided to BfR alongside the original NICEATM database. As evaluation work progressed, further relevant publications were noted and added to the reference library. As of 3 March 2021, the library comprised 1942 references with available full texts.

3.3 Reliability

An overview of the database is given in section 4. A detailed listing of all database fields is provided in Appendix 1. As of 3 March 2021, the database comprised 2277 individual HPPT test results[8], most of which were generated a long time ago under non-harmonised conditions with no/unreported quality control, using non-guideline, non-validated test protocols. Often the original test reports are not available to the public. To gain information about these tests one has to resort to published summaries which are mostly scant in reporting and do not reveal information with respect to important sources of experimental variability.

3.3.1 Relative Reliability Score (RRS)

While the aforementioned problems may apply to all published HPPT data alike, there are nevertheless differences in quality, rendering some results more reliable than others. In this section, the term "reliability" is used as defined by Klimisch and co-workers with respect to the quality assessment of (eco)toxicological data:

> *"…the inherent quality of a test report or publication relating to preferably standardised methodology and the way the experimental procedure and results are described to give evidence of the clarity and plausibility of the findings."* (Klimisch et al., 1997)

[7] The HDSG is grateful to Drs. Annekatrin Aue, Ana-Maria Florea, Lisa Heiserich, and Frauke Hoffmann of BfR for their help in capturing the HPPT data from the RIFM monographs.

[8] This number is not to be mistaken for the number of substances in the database which is lower, since for many substances more than one test result was found (cf. section 4).

Note that this definition focuses on quality of reporting rather than on the scientific quality of the experiment itself. Along these lines, the HDSG defined a minimum set of information ("essential information") required for including a test result into their assessments and a second set of information ("relevant information"), the absence of which would not preclude inclusion of the respective test result in the assessment, but would negatively affect its reliability.

Essential information:

- the primary report is clearly identified and it is clear whether an HMT or HRIPT was performed; mixed designs, e.g. HRIPT with SLS pre-treatment, are excluded;
- substance identity (ID) is sufficiently characterised[9];
- it is known whether positive reactions occurred in the test or not; AND
- the number of test subjects is available[10].

Relevant information:

- the vehicle used for induction,
- the test concentration (CONC),
- the dose per skin area (DSA) applied for induction (or information allowing to calculate it, i.e. test volume and patch size),
- the positive incidence (i.e. the fraction of test subjects with a positive reaction, or the absence of any positive reaction) is available or can be calculated.

Following this concept, the HDSG then defined a "Relative[11] Reliability Score" (RRS) which was applied to the individual test results and associated references as follows:

- High relative reliability (RRS = 1): All essential and all relevant information is available from the original test report (primary reference, referred to as "REF1"). Currently, 318 (14.0%) of the 2277 test results in the database have this score.

  Example: In Kligman (1966c), the concept of the HMT is introduced, providing all necessary experimental details as well as a number of test results obtained under these conditions.

- Sufficiently reliable (RRS = 2): REF1 does not include all essential and all relevant information, but all missing essential or relevant information is available from references cited by REF1. Currently 91 (4.0%) of the 2277 test results in the database have this score.

  Example: In Greif (1967), the author directly refers to the design as provided in Kligman (1966c).

- Relatively reliable, with some additional uncertainty (RRS = 3): REF1 and the references cited by or citing REF1 together include all essential, but not all relevant information. However, all missing relevant information can be inferred from further available sources, e.g. from publications by

[9] Normally, this means that a CASRN or EC no. is available (or, in the case of mixtures, CASRN/EC nos. for all constituents). For some natural materials/extracts, e.g. ethereal oils commonly used in cosmetics, no - or several different - CASRN/EC nos. may be available without one single identifier being clearly assignable to the test material. In such circumstances, the substance ID is seen as sufficiently described by the name of the source material (e.g. "alantroot oil") or another commonly used trivial name ("Peru balsam").

[10] In some cases, the number of test subjects was given as a range, i.e. with a defined minimum/maximum number of test subjects. All other conditions fulfilled, the results from such tests could be used if the resulting classification outcome did not change, regardless of whether the calculation was based on the minimum or maximum number given.

[11] The term "relative" acknowledges that uncertainties pertaining to all available test results or inferences made for most/all test results alike are not included in the score. Rather, this score expresses the relative reliability of the respective test result on top of the baseline uncertainties associated with the HPPT data as such.

authors with access to REF1, but with less confidence compared to RRS = 2. Currently 1630 (71.6%) of the 2277 test results in the database have this score.

Example: In Kligman and Epstein (1975), the authors revisited their 1966 approach and reported that in the meantime they had changed some of the test design parameters, e.g. the ratio of test volume to patch size changed from ca. 69 µL/cm^2 to 75 µL/cm^2. However, with respect to the test design used, many RIFM monographs from the 1970s and 1980s cite both references and it can only be inferred that the test volume/patch size ratio was either 69 or 75 µL/cm^2. For practical reasons, since the difference was not big, the average value (72 µL/cm^2) was used for further calculations based on these monographs, but the results obtained in this way are associated with higher uncertainty.

- Relatively reliable in part (RRS = 4): REF1 and the references cited by REF1 include all essential, but not all relevant information. In addition, not all missing relevant information can be inferred from further available sources (cf. RRS = 3), but as a minimum, either CONC or DSA are available or can be inferred. Currently 216 (9.5%) of the 2277 test results in the database have this score.

  Example: For methyl heptine carbonate (CASRN 111-12-6/EC 203-836-6), results of an HRIPT have been published. However, only secondary publications are available which do not report the details of the test design. Most prominently they do not give the test volume and patch size used. For this reason, the dose applied per skin area (DSA) cannot be calculated. However, other data are provided and therefore some of the calculations in this report, such as the positive incidence, could be performed.

- Not reliable (RRS = 5): All essential information is there, but neither CONC, nor DSA are available. Currently 22 (1.0%) of the 2277 test results in the database have this score. Test results with RRS = 5 or were excluded from further analysis in this report.

  Example: Gad et al. (1986) report a negative HRIPT result for benzoic acid (CASRN 65-85-0/EC 200-618-2), but do not provide the test concentration or the dose per skin area (DSA). It is therefore not possible to tell whether a sufficiently high test concentration/DSA was used to at least exclude that this substance is a strong skin sensitiser (GHS 1A).

- 576 further test results were not included in the database because of missing essential information or if there was reason to assume that the vehicle used might have contained surface-active, irritant and/or sensitising constituents.

  Example: Basketter et al. (1999) report sulfanilic acid (CASRN 121-57-3/EC 204-482-5) as not being a skin sensitiser in humans, but of the primary sources containing HPPT data cited in this publication, none contained a result for this substance. It is unclear, whether this (non-)classification really goes back to an HMT or HPPT result (possibly it is based on HDPT data), and therefore essential information is missing.

As of 3 March 2021, this left 2255 test results with RRS 1 – 4 for further analysis. The decision tree for assigning an RRS to an individual HPPT result is summarised in Figure 2.

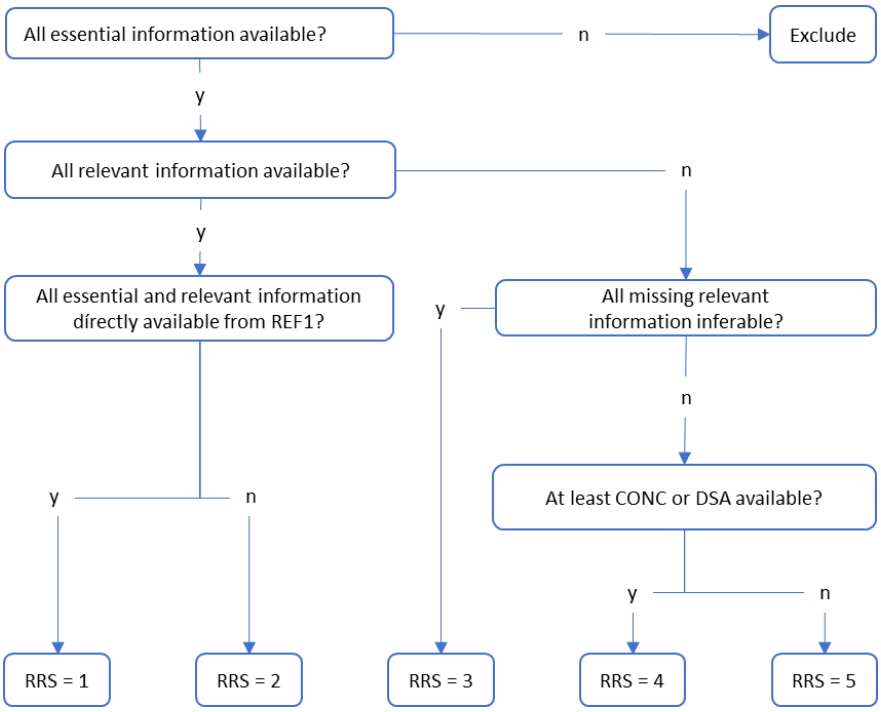

Figure 2: Decision tree for assigning Relative Reliability Scores (RRS) to individual HPPT results

4 Database overview

The 2255 test results with RRS < 5 relate to a total of 1366 substances with a unique ID[12]. The number of test results per substance ranges from 1 to 29 as shown in Figure 3.

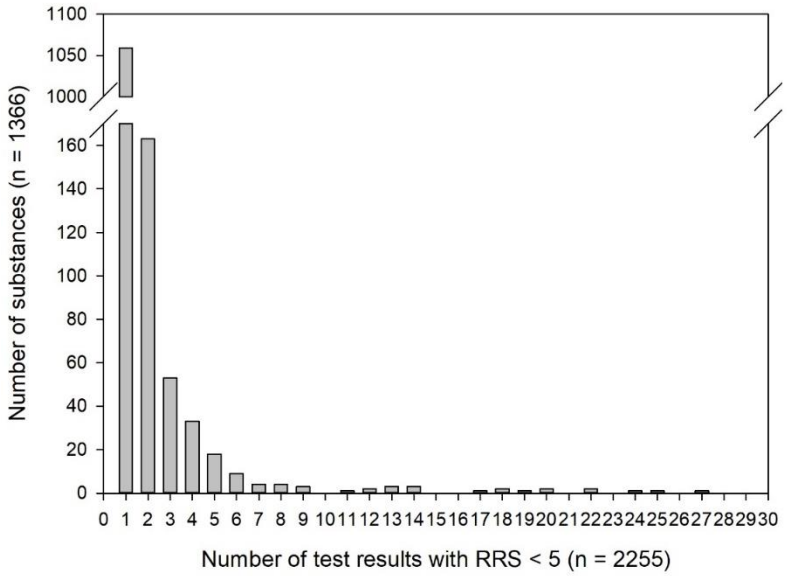

Figure 3: Distribution of the number of tests available per substance

For 1059 (77.5%) of the 1366 substances, only one test report was available; 163, 53, 33, 18, and 9 substances had 2, 3, 4, 5, and 6 test results, respectively. Between 7 and 29 test results were available for the remaining 32 substances.

[12] cf. 3.3.1 on essential information, but also 5.2 regarding variability between nominally identical test substances

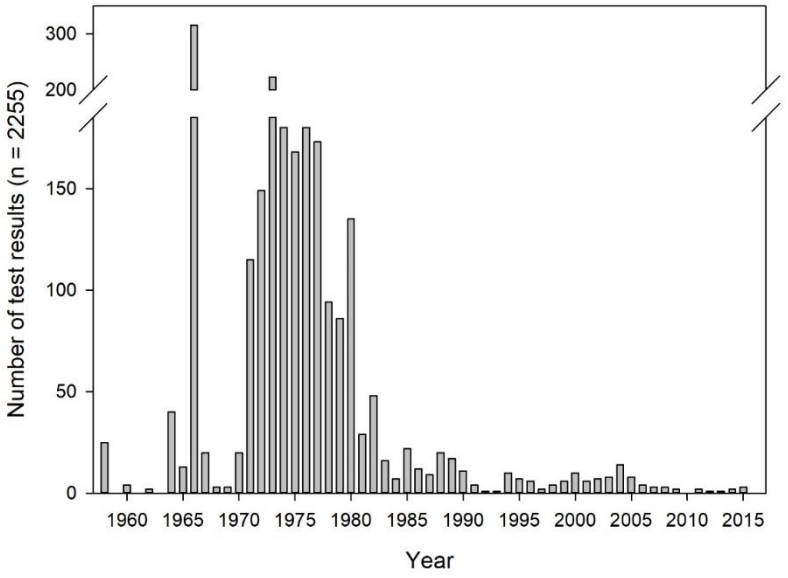

Figure 4: Distribution of the original test reports over the years 1958-2015 (n = 2255)

As mentioned in section 3.2, most of the original test reports date from the 1970s/1980s. Figure 4 shows the distribution of the 2255 test results with RRS < 5 over the years from 1958 to 2015[13]. Of these 2255 test results, 426 (18.9%) were generated before 1970, 1814 (80.4%) before 1980, 2129 (94.4%) before 1990, and 2181 (96.7%) before 2000.

Primary references were available to the HDSG for only 535 (23.7%) of the 2255 test results with RRS < 5.

Two generic HPPT types with a comparatively large number of studies available and with a comparatively high level of standardisation were included in the analysis: HMT and HRIPT. In addition, these are the two test designs specifically recognised by the GHS for potency sub-categorisation. Cases where it was unclear whether the test was a HMT or HRIPT or where only one induction exposure was applied were excluded for lack of comparability. Some experimental details for the main sub-types of these designs are given in Table 1 (next page), along with the number of test results (with RRS < 5) obtained with the respective generic design/subtype.

The general scheme is similar for all designs: a defined volume of the test substance dissolved or dispersed in a suitable vehicle at a defined concentration is applied to a patch, usually some type of woven fabric, which is then topically applied to the extremities or back of the test subjects. Patches are generally held in place by occlusive adhesive tape (this is not reported in all cases, however). This procedure is repeated for a defined number of times (induction phase). After a rest phase of usually 10 - 14 days, the test subjects are again patched with the test substance, followed (after a defined time) by an assessment of any skin reactions indicative of an allergic response (challenge or elicitation phase). In the HRIPT design, usually 50 - 200 test subjects receive nine to ten induction exposures over a period of three weeks, whereas the HMT usually only uses 25 test subjects and five induction exposures over ten days. However, the sensitivity of the HMT is increased for non-irritant substances by pre-treatment of the exposure sites with sodium lauryl sulfate (SLS) in order to enhance skin permeability and sensitivity of the test subjects by creating a state of mild irritation (Kligman, 1966c).

As of 3 March 2021, the database contains 1619 (71.8%) HMT and 636 (28.2%) HRIPT results with RRS < 5. Of those 2255 test results, 605 (26.8%) are positive and 1650 (73.2%) negative.

[13] Date of the original test reports. The most recent original test report was from 2015, while the most recent literature references citing original reports were from 2019.

Table 1: Overview of HPPT sub-types found in the HPPT database; data are from the original publications, unless otherwise noted

Test type[14]	Sub-type[14]	Typical number of test subjects	Induction						Typical rest phase (d)	Challenge		Reference(s)
			SLS[15] pre-treatment	Test volume (µL)	Patch size (cm²)	Contact per exposure (h)	No. of exposures	Site of induction		No. of exposures	Contact time per exposure (h)	
HMT (1619)[16]	Kli66 (178)	25	yes	1000	14.5	48	5	forearm or lower leg	10[17]	1	48	(Kligman, 1966c; Kligman and Epstein, 1975)
	Kli6675 (528)			1000 or 300[18]	14.5 or 4.0[18]			forearm, back, or lower leg				
	Kli75 (785)			300	4			forearm or back				
HRIPT (636)[17]	Shel53 (11)	200[19]	no	not reported	not reported	24	10-15	not reported	14-21		48	(Shelanski and Shelanski, 1953)
	Voss58 (24)	50-60		500	4.2		9	upper arm	10		24	(Voss, 1958)
	Drai59 (73)	200		500	6.5[20]		10	arms or back	10-14		24	(Draize, 1959)
	Gri69 (3)	60-70			3.9		9	upper arm	17			(Griffith, 1969)
	MM73 (91)	200			6.5[21]	48-72	10	arm	14		72	(Marzulli and Maibach, 1973)
	JK77 (10)	150		500	4		9	upper back		2	48	(Jordan Jr. and King, 1977)
	MM80 (45)	200		200	2.54[21]		10	arm or upper back		1-2	72	(Marzulli and Maibach, 1980a)
	RIFM08 (67)	≥ 100		300		24	9	back	10-14		24	(Politano and Api, 2008)

[14] The number of test results (RRS ≤ 4) following the respective design is given in parentheses.

[15] SLS = sodium lauryl sulfate, an irritant used for pre-treatment in the HMT in order to maximize sensitization outcome

[16] The discrepancy between this figure and the sum of the figures given for the subtypes is explained by the presence of test results for which the test design could not be assigned to one of the subtypes in this table (the test type of such test results is given as "other" in the database).

[17] According to Greif (1967)

[18] The original HMT (Kli66) was published in Kligman (1966c). However, later (Kligman and Epstein, 1975), the authors reported a number of adaptations to that design (Kli75). Most prominently, they reduced patch size from 14.5 to 4 cm^2 and test volume from 1 to 0.3 mL. For the tests performed in their lab between 1967 and 1974 (Kli6675), there is now uncertainty about the degree to which the original design had already been modified. For the calculations in this report, a ratio of test volume/patch size of 72 $\mu L/cm^2$ was used, in-between that of Kli66 (69 $\mu g/cm^2$) and that of Kli75 (75 $\mu g/cm^2$). It is noted that only a small error in the order of 4% is introduced by using the average value vs. a potential greater error of ca. 9% when erroneously assigning Kli66 or Kli75.

[19] Not reported in original publication, but in Kligman (1966a) and Marzulli and Maibach (1976a).

[20] The exact value used in the calculations reported here is 6.4516 cm^2, i.e. one square inch.

[21] Not reported in the original papers, but inferred by the HDSG after communications with RIFM and Dr. Maibach (cf. also 5.4.3).

5 Variability and uncertainty in the HPPT database

This section lists a number of factors affecting the overall variability of HPPT results. Exemplary references are provided for illustration, but a comprehensive literature review could not be carried out given the available time and resources. Likewise, it was not possible to quantitatively evaluate the relative impact of the individual factors on the overall variability of the test results. A qualitative impression can be obtained for the substances from the reference chemical list, for which more than one test result was available (cf. Table 19 in section 8 and Table 21 in Appendix 2).

5.1 About variability and uncertainty

For the purpose of this document the term "variability" is used as defined in WHO IPCS (2017), i.e.:

> ***"Variability:*** *Heterogeneity of values over time, space or different members of a population, including stochastic variability and controllable variability. Variability implies real differences among members of that population. For example, different individual persons have different intake and susceptibility. In relation to human exposure assessment, differences over time for a given individual are referred to as intraindividual variability; differences over members of a population at a given time are referred to as interindividual variability."* (WHO IPCS, 2017)

In relation to the HPPT database, variability affects all parameters with a possible impact on the test result that were not kept constant for all tests performed. It can be subdivided as follows:

- variability of nominally identical test items[22],
- variability within and between[23] test populations,
- variability of the exposure regime, and
- variability in scoring the test results.

The term "uncertainty" is used in the following sense:

> ***"Uncertainty:*** *Uncertainty in risk assessment in the general sense is defined by IPCS (2004) as 'imperfect knowledge concerning the present or future state of an organism, system, or (sub)population under consideration'. In relation to the specific topic of this monograph, it can be further defined as lack of knowledge regarding the "true" value of a quantity, lack of knowledge regarding which of several alternative model representations best describes a system of interest, or lack of knowledge regarding which probability distribution function and its specification should represent a quantity of interest."* (WHO IPCS, 2017)

With respect to the HPPT-associated variabilities noted above, a (non-exhaustive) list of associated uncertainties includes:

- imperfect knowledge about whether all relevant parameters with a possible impact on the test result have been identified and documented (i.e. whether there were relevant differences in the design of two tests performed under completely identical conditions with respect to all known relevant parameters),
- imperfect knowledge about whether test designs referenced in a test report have been followed to the necessary detail (i.e. whether two tests really were performed under identical conditions as claimed),

[22] The term „test item" denotes a given combination of test substance and vehicle, i.e. the material which is actually put on the skin of the test subject.

[23] Variability between test populations can be considered under the above WHO definition, if test populations are seen as – more or less representative - samples of the general population, for which HPPT data claim to provide a prediction.

- imperfect knowledge about the exact test conditions (e.g. due to insufficient reporting of one or more relevant parameters), and
- imperfect knowledge about the representativeness of the test populations used in the HPPT database as compared to the target population being considered (i.e. in the case of this report, the general population in OECD member countries, despite variability between them).

In the following subsections, the different types of variability in the HPPT database are discussed along with their associated uncertainties. For the moment, given available time and resources, this can only be a general discussion of the main factors of variability and uncertainty in the HPPT database. However, members of the HDSG are currently planning on publishing a more fundamental review of these issues outside the time and scope constraints of the OECD EG DASS (cf. section 0).

5.2 Variability and uncertainty associated with the test item

Nominally identical test materials as identified by CASRN or EC number do not necessarily have to be fully identical in terms of their composition. Due to different production processes, a different quality of the source chemicals used in their production, or sensitivity to abiotic decomposition by autoxidation or photolysis (which may work both ways, i.e. the breakdown products may be more (in the case of pre-haptens) or less sensitising than the parent), this composition may be variable, potentially resulting in variable HPPT test results. Specifically, this holds for substances produced from biological materials, e.g. plant extracts.

Example - 1974 RIFM monograph on vetiver acetate:

"*A maximization test* […] *was carried out on 25 volunteers. The material* […] *was tested at a concentration of 20%, in petrolatum and produced sensitization reactions in three of the 25* […]. *A maximization test* […] *was carried out on 25 volunteers using a sample of better quality (i.e. one with a lower acid value). The material was tested at a concentration of 8% in petrolatum and produced no sensitization reactions* […]. *Vetiver acetate was then specially prepared from the alcohol and tested in a maximization test* […] *carried out on 25 volunteers. The material* […] *was tested at a concentration of 20% in petrolatum and produced no sensitization reactions* […]". (Opdyke, 1974a)

Where – as in the example above – multiple HPPT results are available for a given substance, the possible variability is to some degree covered in the database. However, as shown in section 4 above, for 77.5% of the 1366 substances with test results of RRS < 5, only one test result is available, while only 10.5% have three or more test results with RRS < 5.

5.3 Variability and uncertainty associated with the test population

For the HMT, Kligman (1966c) investigated the reproducibility of sensitisation rate and grade of the allergic reaction when performing tests in triplicate for ten substances (Table 2).

With the exception of the test panels (naturally, each new replicate had to be performed in a new, previously unexposed group), all other parameters can be regarded as constant to the degree technically possible. As can be seen, there was some variability in both sensitisation rates and grades. With respect to sensitisation rate, differences between replicates of up to ca. 20% were found; in one case (hexachlorophene) even the overall conclusion (positive/negative) varied between replicates. Notably, this variability includes both interindividual variability and general reproducibility. For the HRIPT, comparable data do not seem to be available.

Table 2: Reproducibility of the HMT when performed in triplicate*

Substance	Induction concentration (%)	Sensitisation rates test no.			Grades test no.		
		1	2	3	1	2	3
p-Aminobenzoic acid	25	0/23	0/24	0/24	1	1	1
Hexachlorophene	25	0/24	1/25	0/23	1	1	1
Tetramethylthiuramdisulfide	25	4/22	1/23	1/25	2	1	1
Benzocaine	25	5/24	3/24	5/22	2	2	2
Chloroquine diphosphate	25	9/24	6/25	11/23	3	2	3
Penicillin G	25	16/23	12/23	13/25	4	3	3
Ammoniated mercury	25	13/25	11/24	15/25	3	3	4
Monobenzyl ether of hydroquinone	25	22/22	25/25	24/25	5	5	5
Apresoline®	5	25/25	24/24	25/25	5	5	5
Tetrachlorosalicylanilide	5	24/24	22/23	24/25	5	5	5

* Table reproduced from Kligman (1966c)

5.3.1 Size of the test population

As all tests in biological systems, HPPT are carried out in a test panel representing a small sample of a much larger population. While the interindividual variability in the test panels used for HMT and HRIPT or their representativeness for the general population are largely unknown, even between different samples from an otherwise homogeneous population statistical variability will occur. This can be expected to be inversely correlated with test panel size.

The issue of statistical resolution in skin sensitisation testing has already been noted more than 50 years ago:

> *"Suppose further that we can tolerate no more than one case of sensitization per thousand users. How many subjects must be tested with a negative result before we can be certain that this tolerance is not exceeded? The answer is 29,978 test subjects among whom there will not even be one reactor!"* (Kligman, 1966a)

While it is not the task of this document to discuss or judge the correctness of all aspects of this statement, it highlights the lack of statistical power of HPPT designs as an important issue. Notably, in the HPPT database used for this report (test results with RRS < 5), the number of test subjects ranged from 7 to 225 (with a median of 25 for the HMT and a median of 51 for the HRIPT results).

5.3.2 Sex

Prospective HPPT studies that have directly examined sex-dependent effects show that females could be more likely to be induced than males and may have a more severe reaction under the same test conditions (Jordan Jr. and King, 1977; Rees et al., 1989). A retrospective study of HMT results for 73 allergens found no sex-related difference (Leyden and Kligman, 1977a). One prospective study showed that the severity of the allergic response was similar for males and females when the sensitisation exposure in females of reproductive age was performed at least five days before or after the start of the menstrual cycle (Morrissey et al., 2008). HPPT practitioners have typically regarded the sensitivity of the sexes to be comparable (Griffith, 1969; Kligman, 1966b; McNamee et al., 2008). Although primary references in the published scientific literature often reported the sex(es) tested, RIFM monographs published before the year 2000 typically did not. Thus, the HPPT database does not report the sex of the subjects tested and this parameter may potentially be responsible for some of the variability in HPPT responses.

5.3.3 Age

HPPT studies are performed in adult subjects. Elderly subjects, typically 65 years or older, are usually avoided due to diminished immunological responses (Robinson, 1999; Waldorf et al., 1968). In HPPT

studies, fewer subjects 70 years of age or older were sensitised to DNCB (Waldorf et al., 1968) compared with younger subjects. Age did not affect the severity of allergic response to DNCB in adult subjects 63 years of age or younger (Morrissey et al., 2008). Although primary references in the published scientific literature typically reported the age range or mean age of the subjects tested, RIFM monographs in general did not. The available HPPT database therefore does not report the age of the subjects tested, and it cannot be ruled out that differences in age distribution might have contributed to the variability in HPPT response between test results.

5.3.4 Skin characteristics

Some of the earlier publications on human predictive patch testing, e.g. Kligman (1966b) or Leyden and Kligman (1977a) have discussed variability in skin pigmentation and other skin characteristics between human subpopulations, and its influence on skin sensitisation testing results, in terms of "race". From today's perspective this concept appears scientifically unjustified, since the suitability of "race" as a concept for separating human subpopulations according to geno- and/or phenotypic sub-categorisations has been the subject of critical debate[24]. However, even as late as 2003, Wesley and Maibach have reviewed available papers on differences in skin characteristics between different "racial" and "ethnic"[25] groups (Wesley and Maibach, 2003). For some properties, such as trans-epidermal water loss (TEWL) or skin surface pH, they claim to have found significant differences between e.g. "black" and "white" skin, while for other properties no significant differences were found; overall, however, these conclusions appear ill-founded since neither classification criteria used by the authors of the original publications to assign people to "races" or "ethnic groups", nor their impact on the comparability of different studies were analysed in-depth (nor questioned, for that matter).

Apart from such specific investigations of "racial" or "ethnic" differences, HPPT investigators recruit subjects without restrictions on skin characteristics (McNamee et al., 2008). RIFM monographs typically did not report the skin phenotype of the subjects tested, and therefore the HPPT database does not contain any such information.

It must be concluded that any potential impact of certain skin characteristics on skin sensitisation remains largely unknown. On the technical side, potential difficulties in detecting erythema in test subjects/patients with pigmented skin have been recognised. While nowadays scoring methodology has been refined to circumvent these problems, cf. e.g. Zhao et al. (2017), a discussion of this issue is generally absent, possibly due to a lack of awareness, from the test reports of the 1960s-1990s, which constitute almost all of the available HPPT database.

5.3.5 Previous exposure to the test substance

Some of the HPPT designs employ a pre-testing step to exclude already sensitised individuals from the test panel, but for most HPPT results in the database this is not the case, or it is not known whether such a pre-test has been performed. This can impact the reliability of the test result:

- If undetected, the number of sensitised individuals at the test concentration will be exaggerated.
- If detected, different protocols seem to apply different strategies. Excluding pre-sensitised subjects from the panel may underestimate skin sensitisation rate (by potentially removing sensitive subjects shown to be sensitisable by the test substance). Some designs, e.g. the HMT according to Kligman and Epstein (1975), explicitly use pre-tests to single out pre-sensitised individuals before adding them to the test panel. However, where suitable replacements are not available, this may lead to a smaller test panel and, thus, lower statistical power (see section 5.3.1).

[24] cf. e.g. https://en.wikipedia.org/wiki/Race_(human_categorization), last accessed 2021-07-10

[25] They use "race" as denoting genetic variability, whereas "ethnicity" additionally includes cultural and societal aspects.

In addition, there are examples where test subjects showing skin reactions in the late phase of induction have been considered pre-sensitised, but due to insufficient reporting it is not always clear on what grounds a possible new sensitisation was excluded and, moreover, whether this assessment was made by the original authors themselves or rather by the authors of the secondary publication reporting the finding. If applied, the pre-testing methodology or its validity generally are not presented, which may cast additional uncertainty on the data from the included sample.

5.3.6 Other aspects of interindividual variability

Both genetic pre-disposition (e.g. in the form of atopy) and epigenetic modulation via environmental exposure (the "exposome") play an important role in determining an individual's susceptibility to be sensitised by a chemical allergen, see e.g. Cecchi et al. (2018), Han et al. (2017), or Moggs et al. (2012). Thus, even between healthy, previously unexposed test subjects of the same sex, age, and skin phenotype, significant interindividual variability must be expected.

In general HPPT study authors claim to have recruited "healthy" test subjects, but it is unclear whether individuals with pathological changes of the immune status would have been detected in all cases and the respective subjects removed from the test panel. For example, an inverse correlation between contact allergy and autoimmune disease has been reported (Bangsgaard et al., 2011).

5.4 Variability and uncertainty associated with the induction exposure regime

5.4.1 Test volume

As noted above, in many cases the actual test volume used in a given test is not reported in the original report (REF1), or a secondary report citing REF1. Instead it can be inferred from the generic test design referred to, but only with some uncertainty.

Example: Politano and Api (2008) give a volume of 300 µL for their standard RIFM HRIPT design according to which 200 out of a historical database of 1000 HRIPT tests were performed. In RIFM HRIPTs from the 2000s, on the other hand, frequently other volumes have been used, e.g. 200 or 500 µL.

Another example was provided in section 3.3 to illustrate the criteria used for assigning an RRS of 3 based on additional uncertainty due to a change of test volume for the HMT from Kligman (1966c) to Kligman and Epstein (1975).

Variability in test volume has a direct impact on the calculated DSA. All other parameters being constant, a two-fold change in test volume will result in a two-fold change in DSA. On the other hand, in most cases changes in test volume historically were accompanied by corresponding changes in patch size. For the test designs described in Table 1 (section 4), test volumes varied from 200 to 1000 µL (i.e. by five-fold) and patch sizes from 2.54 to 14.5 cm^2 (i.e. > six-fold), but test volume/patch size ratios only varied from ca. 69 to 128 $\mu L/cm^2$, i.e. by less than two-fold. McNamee et al. (2008) calculated DSA values for using different patch types at typical volumes and obtained a greater, i.e. almost six-fold DSA range (of 300-1770 $\mu g/cm^2$), however, their calculation included a very small patch (8 mm Finn chamber) of 0.5 cm^2 in combination with a very small test volume (15 µL) not commonly used in the designs present in the NICEATM/BfR HPPT database.

5.4.2 Patch type

Due to evaporation and leakage, for test substances with a high vapour pressure, occlusive vs. non- or semi-occlusive induction exposure could make a significant difference with respect to the actual dose received by the test subject, potentially resulting in false negative results (Marzulli and Maibach, 1974). On the other hand, the use of occlusive patches may exaggerate real-life exposure. The HDSG is not aware of any publication systematically investigating the extent of this difference. In any case, when comparing different test results, it should be known which patch type was used. While most published

test designs seem to use occlusive patches, information on this aspect is, however, often not provided in the available HPPT reports, introducing uncertainty with respect to the pooling and comparison of different HPPT results.

5.4.3 Patch size

Even in the original papers introducing HPPT test designs (cf. Table 1), the size of the patch used is not always given or different values are reported, even though this parameter has a direct impact on DSA calculation (reducing the patch size by 50% will result in a two-fold increase in DSA, but see also the note on the variability of the test volume/patch size ratio in section 5.4.1). Furthermore, some authors, e.g. Kligman (1966b) used different patch sizes for the induction and challenge phases.

Example: In the early publications about their modification of the Draize HRIPT, Marzulli and Maibach report the test volume (500 µL or mg), and the patch type (Johnson & Johnson square BandAid™), but not the patch size (Marzulli and Maibach, 1973; Marzulli and Maibach, 1974; Marzulli and Maibach, 1976a). In a later paper (Marzulli and Maibach, 1980a), the same authors reported that now they used 200 µL as test volume. They also stated that they still used the same patch type, but again did not report patch size. The HDSG contacted one of the authors, Dr. Maibach, about this. His answers in combination with the published details now lead to the (still somewhat uncertain) assumption that their group started off using 500 µL or mg in combination with 1 inch x 1 inch BandAid patches and later switched to 200 µL or mg and so-called 25 mm HillTop chambers.[26]

Kligman (1966b) investigated the influence of patch size as well as location and number of patch test sites on sensitisation rates (with the DSA kept constant) for four test substances. In many cases sensitisation rates increased with increasing patch size, but the author concluded that patch size was not a *"very influential factor"* unless very small patches were used (Table 3).

Table 3: Influence of patch size on sensitisation rates (number sensitised/number of test subjects, percent rates are given in parentheses) from Kligman (1966b)

Substance	Induction concentration (%)	Challenge concentration (%)	Patch test area (cm²/one extremity)*			
			1 x 0.4	1 x 14.5	4 x 14.5 (58.1)	1 x 58.1
Ammoniated mercury	25	10	6/22 (27)	12/24 (50)	12/23 (52)	10/25 (40)
Monobenzyl ether of hydroquinone	25	5	10/22 (45)	20/24 (83)	21/23 (91)	23/25 (92)
Nickel sulfate	1	5	2/22 (9)	4/24 (17)	5/23 (22)	7/25 (28)
Neomycin sulfate	25	10§	0/22 (0)	5/24 (21)	8/23 (35)	5/25 (20)

* Values differ slightly from those in the original publication which seems to have used a slightly different conversion factor from inches to centimetres; § SLS was also applied upon challenge

In addition, it is not always clear from the available HPPT reports whether the dimensions reported really refer to the area available for skin contact or also include the outer, adhesive part of the patch.

5.4.4 Vehicle

Different vehicles have been used in the different test designs: the HMT normally uses petrolatum, whereas HRIPT designs often rely on ethanol or mixtures of ethanol and diethyl phthalate (DEP). Other, more rarely used vehicles include e.g. acetone, corn oil, or glycerin.

[26] Notably, the HDSG has received different information regarding the inner diameter of 25 mm HillTop chambers. Dr. Anne-Marie Api of RIFM informed the HDSG that they had measured this inner diameter and calculated an area of 2.54 cm2 available for skin contact of the test substance. However, on request by the HDSG the HillTop company gave the inner diameter as 20 mm, which would result in an inner area of 3.14 cm2. The HDSG decided to use 2.54 cm2 for the calculations in this report but notes that this adds uncertainty in the form of a possible systematic error to some of the dose per skin area values in the HPPT database.

However, the nature of the vehicle used can have a profound impact on the sensitisation rate observed in an HPPT experiment, e.g. due to a penetration-enhancing or –interfering effect, or by interacting with the test substance partitioning the substance into phases. On the other hand, using the same vehicle for all test substances can result in solubility problems. Kligman investigated the impact of different vehicles on sensitisation rates combining six hydro- and five lipophilic test substances with six different vehicles in three different dilutions each (using between five and nine test subjects per test group). For some (but not all) of the vehicles investigated, clear differences in sensitisation rates (e.g. sensitisation capacity was blocked with water-soluble sensitisers) between test substances were observed (Kligman, 1966b).

Moreover, different vehicles have different densities which, again, directly affects DSA calculation. An increase in density by 10% will lead to an increase of the DSA by the same relative amount. Normally, DSA values in the published literature are calculated assuming a density of 1 g/mL. To achieve slightly more exact results, the HDSG used data from the REACH registration dossiers found in the ECHA database (Table 4) in order to approximate the densities used for DSA calculation, with the assumption that the REACH density data apply to the test formulation at skin temperature.

Table 4: Densities for different vehicles as found in the ECHA database[27] and values used in the HPPT database

Vehicle	Value(s) provided in ECHA database (mg/µL)	Value used in HPPT database (mg/µL)
Acetone	0.79 at 20 °C	0.8
Diethyl phthalate (DEP)	1.118 at 20 °C	1.1
Dimethyl phthalate (DMP)	1.194 at 20 °C	1.2
Dimethyl sulfoxide (DMSO)	1.1 at 20 °C	1.1
Ethanol (EtOH)	0.78933 at 20 °C	0.8
Polyethylene glycol (PEG)	1.116 at 20 °C	1.1
Petrolatum	0.77 – 0.96 at 15 °C	0.9
Triacetin	1.161 at 20 °C	1.2
Water	-	1.0
Others (density unknown)	-	1.0

For mixtures of the above, individual densities were weighted according to their relative proportion, e.g. DEP/EtOH 3:1 = (3 x 1.1 + 0.8)/4 = 1.0. Nevertheless, it must be acknowledged that these values are still estimates and the true density of the test items used in HPPT testing is not known (also considering the possible impact of dissolving the test item in the vehicle), adding further uncertainty.

5.4.5 Induction concentration/DSA

One of the major problems associated with the use of HMT and HRIPT data lies in the fact that only one concentration per test is used. This prevents a dose-response assessment whereby e.g. the highest ineffective dose and the lowest effective dose could be reliably determined or a benchmark dose could be calculated. In the case of negative results, often the test concentration used is not high enough to reliably exclude that a positive result might have occurred at a higher test concentration. This is discussed in more detail below, in section 6.3.1.

Conversely, if a positive test result was obtained at a comparatively high concentration pointing at a moderate sensitisation potential, a test using a lower concentration might have been positive as well, perhaps even resulting in the conclusion that the substance should be considered a strong sensitiser. This is discussed in section 6.3.1.

Kligman (1966b) investigated the influence of induction concentration on sensitisation rate, changing the concentration in the vehicle (from 0.1% to 50%), while keeping all other variables constant (Table 5). A clear dose-related increase of sensitisation rates with increasing doses was seen for all substances

[27] https://echa.europa.eu, as of 2019-06-22

up to the maximum concentration tested, with the exception of two substances, for which sensitisation rates were clearly lower at 50% than at 25%. This was attributed by the author to "viscosity changes" of the test items at 50%. Similar results were demonstrated by Marzulli and Maibach (1974), at least for stronger sensitisers such as benzocaine, mafenide, or p-phenylenediamine.

Furthermore, for many of the substances listed in Appendix 2 of this report, test results obtained at different test concentrations/DSA values are available illustrating the influence induction concentration has on calculated DSA, notwithstanding that direct comparison is not always possible due to differences in test design. For example, in tests with formaldehyde using induction concentrations ranging from 0.037% to 3.7%, 100-fold differences in the calculated DSA (29-2868 µg/cm^2) were observed.

Table 5: Influence of induction concentration on sensitisation rates (number sensitised/number of test subjects, percent rates are given in parentheses), from Kligman (1966b)

Substance	Concentration (%)						
	0.1	0.2	1.0	5.0	10	25	50
Furacin®	-	0/22 (0)	3/25 (12)	-	7/25 (28)	14/21 (67)	2/25 (8)
Penicillin G*		2/22 (9)	4/22 (18)	7/25 (28)	11/25 (44)	16/25 (64)	-
Streptomycin*	1/24 (4)	-	8/22 (36)	-	18/23 (78)	20/23 (87)	
Tetrachlorosalicylanilide	6/25 (24)		17/24 (71)	21/24 (88)	21/21 (100)	-	
Technical malathion®	2/25 (8)		8/25 (32)	-	25/25 (100)	25/25 (100)	
Ammoniated mercury	-		0/25 (0)		8/23 (35)	13/24 (54)	1/25 (4)
Neomycin*			0/22 (0)		4/25 (16)	7/25 (28)	-
Thephorin®*	2/23 (9)		6/24 (25)		11/25 (44)	25/25 (100)	
Monobenzyl ether of hydroquinone	3/24 (13)		7/23 (30)		16/25 (64)	22/22 (100)	
p-Phenylenediamine	5/24 (21)		17/25 (68)		24/24 (100)	20/20 (100)	

*Challenge was performed including SLS treatment ("provocative SLS challenge test")

5.4.6 Site of induction

Induction exposures are typically applied to the forearm or the lower leg for the HMT (Kligman, 1966c; Kligman and Epstein, 1975) and to the upper lateral portion of the arm (Griffith, 1969; Marzulli and Maibach, 1973; Marzulli and Maibach, 1974; McNamee et al., 2008) or the upper back (Marzulli and Maibach, 1980a; Politano and Api, 2008) for the HRIPT. That distinct allergen patterns per body site may play a role in patients with allergic contact dermatitis has been shown and underlined recently (Oosterhaven et al., 2019).

Kligman (1966b) compared sensitisation rates when four successive induction exposures were confined to one extremity (same regional lymph node) to those when four induction exposures were applied four extremities (different regional lymph nodes), also in succession.

Table 6: Influence on sensitisation rates (number sensitised/number of test subjects, percent rates are given in parentheses) of applying four successive induction patches to the same vs. once to each of four different extremities, from Kligman (1966b)

Substance	Induction concentration (%)	Challenge concentration (%)	Same extremity	Four extremities	Significant (?²)
Extract of Krameria	50	50	17/20 (85)	11/23 (48)	y
Thephorin®*	5	10	7/20 (35)	2/23 (9)	n
Nickel sulfate		2.0	6/20 (30)	1/23 (4)	y
Epoxy resin		10	10/20 (50)	4/23 (17)	

*Challenge was performed including SLS treatment ("provocative SLS challenge test")

Applying all of the induction exposures to one extremity increased the rate of sensitisation for each of four substances which underwent this test procedure comparison. The differences were statistically significant for three of the four substances.

For HMT and HRIPT, the entire series of induction patches is typically applied to exactly the same site (Kligman, 1966c; Kligman and Epstein, 1975; Marzulli and Maibach, 1973; Marzulli and Maibach, 1974). Kligman (1966b) compared sensitisation rates when a single site on one extremity was used for induction to that when five different sites were used for induction exposures (Table 7). Applying induction exposures to one site on a single extremity increased the rate of sensitisation for all five substances tested. The differences were statistically significant for three of the five substances.

Table 7: Influence on sensitisation rates (number sensitised/number of test subjects, percent rates are given in parentheses) of applying five successive induction patches to the same vs. five different sites on the same extremity, from Kligman (1966b)[$]

Substance	Induction concentration (%)	Challenge concentration (%)	One site	Multiple sites	Significant (χ^2)
Butyn sulfate	25	10	8/24 (33)	5/22 (23)	n
Streptomycin*	10		24/24 (100)	6/22 (27)	y
Mercaptobenzothiazole			5/24 (21)	0/22 (0)	
Cobaltous sulfate		50	10/24 (42)	4/22 (18)	

* Challenge was performed including SLS treatment ("provocative SLS challenge test"); [$] results for diethyl fumarate are not shown due to unreliable reporting (sensitisation rates > 100%)

5.4.7 Duration of one induction exposure

The duration of a single induction exposure is typically 48 hours for the HMT (Kligman, 1966c; Kligman and Epstein, 1975) and 24 hours for the HRIPT (Draize, 1959; McNamee et al., 2008; Politano and Api, 2008). Some HRIPT protocols had induction patches applied for 48 or 72 hours (Marzulli and Maibach, 1973; Marzulli and Maibach, 1974; Marzulli and Maibach, 1980a).

Kligman (1966b) compared 24, 48, and 72 hour induction patch exposures in the typical HMT series of five exposures and found that the proportion of subjects sensitised increased with increasing patch exposure duration (Table 8).

Table 8: Influence on sensitisation rates (number sensitised/number of test subjects, percent rates are given in parentheses) of applying induction patches for 24, 48, or 72 h, from Kligman (1966b)

Substance	Induction concentration (%)	Challenge concentration (%)	Duration		
			24 h	48 h	72 h
p-Phenylenediamine	0.2	2.5	7/24 (29)	8/24 (33)	10/23 (43)
Thioglycerol	20	5	11/24 (46)	10/24 (42)	9/23 (39)
Chloroquine diphosphate	25	10	2/24 (8)	6/24 (25)	48/23[$]
Streptomycin*	5.0	10	6/24 (25)	12/24 (50)	14/23 (61)

* Challenge was performed including SLS treatment ("provocative SLS challenge test"); [$] likely this is a typing error in the original publication and should be 8/23 or 18/23

He concluded that:

> *"…for stronger allergens, p-phenylenediamine and thioglycerol, one day exposures are as effective as the longer periods. With less potent allergens, two day exposure periods yield higher sensitization rates than one. There is some indication that lengthening the exposure to three days increases the sensitization rate, though perhaps not very materially."* (Kligman, 1966b)

Table 9: Influence of the number of induction exposures on sensitisation rates (number sensitised/number of test subjects, percent rates are given in parentheses), from Kligman (1966b)

Substance	Induction concentration (%)	Number of exposures			
		3	5	10	15
Hexachlorophene	10	0/25 (0)	0/20 (0)	1/22 (5)	3/24 (13)
Benzocaine*	5	0/23 (0)	1/22 (5)	3/25 (12)	6/25 (24)
Ammoniated mercury		0/24 (0)	4/22 (18)	7/22 (32)	14/20 (70)
Tetramethylthiuramdisulfide	10	0/25 (0)	0/25 (0)	2/22 (9)	6/18 (33)
Vioform®*		0/24 (0)	0/23 (0)	0/22 (0)	0/18 (0)
Aluminium chloride	20	0/25 (0)	0/22 (0)	0/23 (0)	0/22 (0)
Furacin®*	10	1/24 (4)	4/24 (17)	5/22 (23)	12/20 (60)
Neomycin sulfate*		0/24 (0)	1/25 (4)	4/23 (17)	10/21 (48)
Penicillin G*		1/25 (4)	5/25 (20)	10/21 (48)	16/21 (76)
Sodium sulfathiazole	25	0/24 (0)	1/25 (4)	1/22 (5)	3/23 (13)

* Challenge was performed including SLS treatment ("provocative SLS challenge test")

5.4.8 Number of induction exposures

The number of induction exposures was typically five for the HMT (Kligman, 1966c; Kligman and Epstein, 1975) and nine (McNamee et al., 2008; Politano and Api, 2008) or ten (Marzulli and Maibach, 1973; Marzulli and Maibach, 1974; Marzulli and Maibach, 1980a) for the HRIPT.

Kligman (1966b) showed that increasing the number of exposures increases the proportion of subjects sensitised. Ten substances were tested using three, five, ten, and 15 induction exposures (Table 9).

With the exception of two substances known to very rarely produce skin sensitisation (which did not sensitise the test subjects of any group), the number of induction exposures had a profound effect on the observed sensitisation rates.

5.4.9 Interval between induction exposures

For many patch test methods, the most common interval between induction exposures is 24 hours. However, the Stotts HRIPT exposure paradigm (Stotts, 1980) consists of exposures three times a week (Monday, Wednesday, Friday) for three weeks, resulting in a 72-hour break between weekly exposures.

To address whether constant stimulus of the draining lymph nodes affected the sensitisation rate, Kligman (1966b) looked at consecutive day exposures (no break between five doses), as well as two day and six day rest periods between five 24-hour exposures. Sensitisation rates increased markedly from zero- to six-day intervals, with the exception of diethyl fumarate, for which the maximum sensitisation rate of 100% was reached with a two-day interval (Table 10).

Table 10: Impact of different intervals between induction exposures, from Kligman (1966b)

Substance	Induction concentration (%)	Challenge concentration (%)	Intervals		
			0 d	2 d	6 d
Mercuric chloride	2.0	0.05	6/23 (26)	11/25 (44)	15/20 (75)
Diethyl fumarate		0.2	18/23 (78)	25/25 (100)	21/21 (100)
Monobenzyl ether of hydroquinone	25.0	5.0	4/23 (17)	9/25 (36)	12/21 (57)
Potassium dichromate	3.0	0.25	3/22 (14)	18/25 (72)	18/21 (86)

5.4.10 Rest phase between induction and challenge

Typically, a rest phase of between ten and 21 days is allowed between the last induction and the (first) challenge exposure (cf. Table 1, section 4). It is conceivable that the duration of this rest phase might affect the sensitisation rate and/or the magnitude of sensitisation, but the HDSG is not aware of any

publication systematically investigating this issue. Moreover, the duration of the rest phase is practically never mentioned in the RIFM monographs.

5.4.11 HMT vs. HRIPT

The major categories of human predictive testing involve the HMT and the HRIPT. Their main difference lies with the enhancement of skin reactions with irritants, notably SLS, in the HMT. The irritant impairs skin barrier function and thus enhances penetration, which potentially facilitates delivery of the allergen to the cells involved in the sensitisation process or in the reaction to challenge. This allows a reduction in the number of patches to be used, although there is an increase in the time of each patch exposure with HMT (5 patches, each of 48 hours) as compared to most HRIPT procedures (usually 9 to 10 patches, for 24 hours, and sometimes 72 hours on weekends). This also allows for a smaller sample size for HMT (typically 25).

The use of irritant for enhancement not only introduces variability between test procedures, it may also cause uncertainty when there is disparity in concentration of the irritant for induction and for challenge. When a negative reaction is observed, it could be due to a number of factors besides non-sensitisation, and adequacy of the irritant concentration at challenge may come into consideration in an HMT setting.

The original HMT also used patches with greater area and higher volumes as compared to later modifications, but the ratio of test article volume to exposure area remained approximately constant. For both HMT and HRIPT, the volume for application is constrained by the patch area, which determines the size of the receptacle for the test article. In addition, the applicable volume is affected by the dosage form and whether a non-woven cotton pad is present in the patch area. These are all sources of variability (and uncertainty, if not reported).

5.5 Variability and uncertainty associated with the challenge exposure regime

Exposure to the test substance in the challenge phase allows elicitation of the sensitisation immune response. Variability and uncertainty of the challenge regime will directly impact the interpretation of the skin reaction, if any, and the conclusion on whether the test subject has indeed been sensitised.

5.5.1 Number of challenge exposures

In terms of patch application for the challenge phase, there is no major difference among study protocols, as the study participants generally receive one application of the test substance in both the HMT and the HRIPT procedure. The HRIPT procedure allows re-challenge when the initial testing produces an ambiguous result. In the HRIPT design according to Jordan Jr. and King (1977), two consecutive challenge exposures were employed.

5.5.2 Volume, patch type and size used for challenge

Essentially the same considerations apply as for the induction phase. See discussions in sections 4.3.1 to 4.3.3. The investigators often use the same kind of patch for induction and challenge, and this poses the same constraints for applicable volume because of the patch size involved. Sometimes patch size and test substance volume are reduced for challenge, e.g. in Kligman's HMT (Kligman, 1966c).

5.5.3 Challenge concentration

Regulatory frameworks such as the GHS and REACH assume that in general elicitation thresholds are generally lower than induction thresholds:

> *"3.4.1.4 Usually for skin and respiratory sensitization, lower levels are necessary for elicitation than are required for induction"* (GHS, rev. 8[28]/Regulation (EC) 1272/2008)

This may go back to the work of Friedmann and co-workers on 2,4-dinitrochlorobenzene which produced elicitation thresholds that were inversely related to induction doses in newly sensitised subjects (i.e. when using HPPT protocols) (Friedmann et al., 1983). It may also refer to the work of Griem et al. who reported that elicitation thresholds in subjects with previously established allergies to 12 substances were lower than induction thresholds, with the most potent allergens having the smallest induction:elicitation threshold ratio (Griem et al., 2003). The HDSG notes that further work should be undertaken to confirm these findings.

This is important because based on the assumption that elicitation thresholds are generally lower than induction thresholds, the concentration of test substance used for challenge exposure in HPPTs is often reduced compared to that used for induction, *inter alia* to protect sensitised individuals from suffering from severe skin reactions. For six test substances, Kligman (1966b) has shown the (profound) impact of challenge concentrations (under constant induction conditions) on sensitisation rates (Table 11).

Table 11: Impact of different challenge concentrations on sensitisation rates (in %, n = 10 in all cases), from Kligman (1966b)

Substance	**Concentration (%)**			
	25	**10**	**1.0**	**0.1**
Penicillin G	100		60	10
Streptomycin			70	30
Neomycin			30	0
	Concentration (%)			
	5	**1.0**	**0.1**	**0.01**
Nickel sulfate	100	90	60	20
Monobenzyl ether of hydroquinone		100	50	30
Epoxy resin		90	40	0

As the choice of challenge concentrations in the various HPPT test designs does not seem to follow a systematic approach and challenge concentrations are rarely, if ever, reported in the published literature, variability in this parameter is likely high and may have a significant impact on the identification of sensitisers and their potencies in terms of the sensitisation rate at a given induction concentration/DSA.

This issue is also discussed in a number of further publications, e.g. (Friedmann et al., 1983; Griffith and Buehler, 1977; Marzulli and Maibach, 1974; Marzulli and Maibach, 1980b; Marzulli and Maibach, 1976b; McNamee et al., 2008).

5.5.4 Site of challenge

5.5.4.1 Location

Challenge tests are often on the lower back, but the upper extremities can be used for challenge as well. It is not known whether the location for challenge testing has a significant impact on data variability. However, it has been also suggested by European dermatologists that *"the possibility of linking positive patch test reactions to relevance, along with affected body sites, should be a useful addition to patch test documentation systems"* (Oosterhaven et al., 2019).

[28] https://www.unece.org/trans/danger/publi/ghs/ghs_rev08/08files_e.html, last accessed 2021-07-10

5.5.4.2 Site used upon repetition

Most study protocols require that challenge tests be conducted at a new patch-test site which has not been tested previously for irritation.

5.5.5 Duration of challenge exposure

Challenge exposure is usually 48 hours of patch application with test substance for HMT. However, it may vary between 24 to 72 hours for patch removal and investigator evaluation for HRIPT (cf. Table 1 in section 4). Thus, test results may present uncertainty due to this variability in exposure time. In addition, time points for readout of the skin reactions after challenge are variable between test designs.

5.6 Variability and uncertainty associated with scoring the test results

Scoring can significantly contribute to both variability and uncertainty due to the subjective nature of the approach. The scoring systems for skin reactions to contact allergens are not standardised. Investigators use their own systems and often base the scoring on erythema with an ordinal scale varying from four levels (0 for "none" to 3 for "severe", as described in Politano and Api (2008)) to more, e.g., including a $\pm$ score and/or a "very severe" score. Other manifestations beside erythema, such as oedema, papules, spreading, etc. may be assigned a numerical value to be added to the erythema score or included as a letter separately. Some studies include other manifestations in the scoring such as 1 = erythema, 2 = erythema and induration, 3 = vesiculation, and 4 = bulla formation, e.g. Marzulli and Maibach (1973). These observations are semi-objective and require training of the reader, in particular regarding the delineation of skin reactions of different origin, such as irritation. Most studies do not present information on intra-rater reproducibility or inter-rater reliability. For the HMT studies, almost all were rated by members of one group (Kligman, 1966c; Kligman and Epstein, 1975); despite this, as the studies spanned a number of years, there remains uncertainty of consistency in the scoring of skin reactions due to staff changes and gains in experience over time.

It should be noted that critical scoring of the test is in the challenge phase. Despite the many ways a skin reaction can be scored and recorded, for the purpose of classification, the most important part of scoring is the criterion for a positive reaction. There is no general agreement on the definition of a positive sensitisation reaction, and this may vary from showing any manifestation, however slight, to requiring visible vesiculation. The timing of the readout could affect the conclusion substantially, because the skin reaction is not a static process, and at any point in time, it may be in a progressive or a regressive state. As the challenge patch may remain at the test site between 24 to 72 hours, this can be a determining factor affecting the conclusion on positivity. Most studies involve a challenge patch exposure duration of 48 hours, and the test site is examined serially to ensure adequate observation time for detection of a reaction.

5.7 Generic uncertainties

It has been noted above that the HPPT results evaluated in this report were mostly designed for safety testing of chemicals at their typical in-use concentrations. They were not meant to be used for classification and labelling or for establishing potency on an absolute scale. It should be a general caveat in risk assessment that using data for purposes other than those they were generated for always bears the risk of overinterpretation. On the other hand, most of the variabilities and uncertainties regarding general shortcomings of these studies, e.g. the use of only one test concentration, have been discussed earlier in this section (and also in sections 1.1 and 2.3).

It should not be forgotten that HPPT designs represent a model of human skin sensitisation in a similar way as GPMT or LLNA test designs do. Beyond possible individual shortcomings of these models, all of them share an inherent generic uncertainty with regard to how much they adequately reflect reality.

For instance, there is a generic uncertainty about how well an HPPT result obtained after five induction exposures at a given test concentration is able to indicate the risk of becoming sensitised when a consumer is chronically exposed to a lower concentration, e.g. once a week but for a considerably longer time. This model uncertainty is currently unquantifiable, since not even a defined regulatory reference point is available specifying thresholds in terms of magnitude of effect, incidence in the general population, and exposure timeframe below/above which a substance is considered a sensitiser of regulatory relevance. Even if such a reference point could be developed, which would require a much deeper understanding of dose-response relationships for skin sensitisation in humans than currently available, measuring model performance and deriving a measure for model performance would need standard reference data generated under a design developed with that reference point in mind. As such data are currently not available, the HPPT data discussed in this report, with all their deficiencies and uncertainties, still represent the largest database of human sensitisation data providing induction metrics under controlled experimental conditions.

6 Using HPPT data to classify chemicals with respect to their skin sensitisation potential

6.1 Introduction

Within the scope of this project, the goal of the HDSG is to provide – as far as possible - a reliable classification for the EG DASS reference chemicals list regarding their human skin sensitisation potential, as a basis for evaluating the performance of various DAs for skin sensitisation submitted to OECD.

In this context, the GHS defines the classification scheme to be used, because a) all DAs refer to this scheme, either for hazard characterisation or potency sub-categorisation, and b) all OECD member countries taking part in this project have endorsed the GHS.

The GHS scheme foresees that substances considered non-sensitisers do not require classification (not classified, NC). Substances considered "strong" or "extreme" sensitisers receive a classification as Skin Sens. 1A (1A subsequently), while substances with "moderate" or “weak” sensitisation potency are classified as Skin Sens. 1B (1B)[29]. Where sub-categorisation is not required by regional legislation or data do not allow for sub-categorisation, sensitisers are classified as Skin Sens. 1.

The GHS scheme for classification based on HPPT data is described in more detail in the following section. The rules followed in its application to the EG DASS reference chemical list as well as the results of the classification are reported in section 7.

6.2 Current system according to the GHS

The GHS, Rev. 8 of 2019, part 3 defines the way in which HPPT results can be used to classify chemicals into sub-categories for their sensitisation potential:

> *"3.4.2.2.2.1 Human evidence for sub-category 1A can include: (a) positive responses at ≤ 500 µg/cm^2 (HRIPT, HMT – induction threshold) […]. 3.4.2.2.2.2 Human evidence for sub-category 1B can include: (a) positive responses at > 500 µg/cm^2 (HRIPT, HMT – induction threshold) […]"*[30]

[29] Note that the terms "extreme", "strong", "moderate", or "weak", although common in regional legislative frameworks e.g. under Regulation (EC) 1272/2008 (CLP Regulation), are not mentioned in the GHS text.

[30] https://www.unece.org/trans/danger/publi/ghs/ghs_rev08/08files_e.html, last accessed 2021-07-10

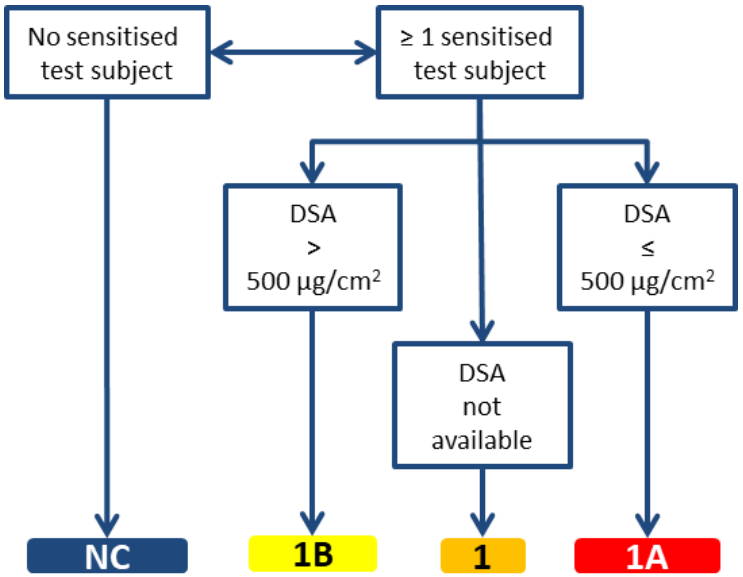

Figure 5: Schematic representation of the decision logic for the classification of substances for their sensitisation potential according to the GHS, based on HPPT data

In other words, any positive test result from an HMT or HRIPT will be classified as at least Skin Sens. 1B. If the positive result was obtained at a DSA ≤ 500 µg/cm^2, the substance is classified as Skin Sens. 1A. If the DSA is unknown, sub-categorisation is not possible and the substance is classified as Skin Sens. 1. Figure 5 shows a schematic representation of this decision logic.

As already mentioned in section 5.4.5, usually each individual HPPT is performed using one test concentration for induction only. Therefore, under the GHS scheme, in many cases one cannot be certain whether a test at a different concentration would not have resulted in a stricter classification. The resulting uncertainty is summarised in Table 12 for all the test results in the database (a total of 2289 test results with RRS < 5 available for 1385 substances).

Table 12: Possible classifications compatible with 605 positive and 1650 negative HMT and HRIPT test results based solely on the GHS classification criteria for HPPT test data from Figure 5 (numbers provided in the coloured boxes mark the number of affected test results)

Test outcome	Test concentration*	1A	1B	NC
Positive (n = 605)	$c_{Test} \leq c_{500}$	59		
	$c_{500} < c_{Test} \leq 100\%$ or DSA unknown	546		
Negative (n = 1650)	$c_{Test} < c_{500}$ or unknown whether $c_{Test} < c_{500}$	142		
	$c_{500} \leq c_{Test} < 100\%$		1500	
	$c_{Test} = 100\%$			8

*c_{Test} = tested concentration (%); c_{500} = test concentration (%) equivalent to a dose per skin area (DSA) of 500 µg/cm^2; the coloured bars cover the range of possible classifications compatible with the scenario represented by the respective combination of test outcome and test concentration; the numbers in the coloured bars represent the number of HPPT results (out of a total of 2289 test results available for 1385 substances) affected by that scenario.

The following scenarios can be distinguished in Table 12 (from top to bottom):

- Positive test results at DSA ≤ 500 µg/cm^2 (59 or 2.6% of the test results in the HPPT database with RRS < 5) unambiguously result in classification as Skin Sens. 1A.
- Positives at DSA > 500 µg/cm^2 (517 or 22.9%) point at a classification as Skin Sens. 1B, but a need for classification as Skin Sens. 1A cannot be ruled out with certainty because a lower test

concentration may have produced a positive result. Positives with unknown DSA (29 or 1.3%) identify the test substance as a skin sensitiser, but do not allow for sub-categorisation. Both result types (546 or 24.2%), however, would be sufficient to identify the chemical as a sensitiser (Skin Sens. 1, without sub-category).

- Negatives at DSA < 500 µg/cm^2 (55 or 2.4%) signal that a classification for skin sensitisation might not be needed at all, however, a need for classification as Skin Sens. 1A or 1B cannot be ruled out with sufficient certainty, because the concentration tested was not high enough to exclude these possibilities. The same holds for test results for which it is unknown whether the test concentration corresponded to a DSA < 500 µg/cm^2 (87 or 3.9%).
- Negatives tested at DSA ≥ 500 µg/cm^2 (1500 or 66.5%) suggest no need for classification, however, while classification as Skin Sens. 1A can be ruled out, classification as Skin Sens. 1B cannot, because a higher test concentration might have resulted in a positive test result.
- Negative test results at a concentration of 100% (8 or 0.4%) unambiguously result in no classification (NC).

Figure 6 shows the distribution of test concentrations for the HPPT tests. For the HMT test results (n = 1619), the median concentration was 8%, with the 75th percentile at 10%, the 95th percentile at 25%, and the 99th percentile at 30%. For the HRIPT results (n = 592[31]), these values were mostly lower (with the median at % and the 75th, 95th, and 99th percentiles at 10, 20, and 63.6%, respectively).

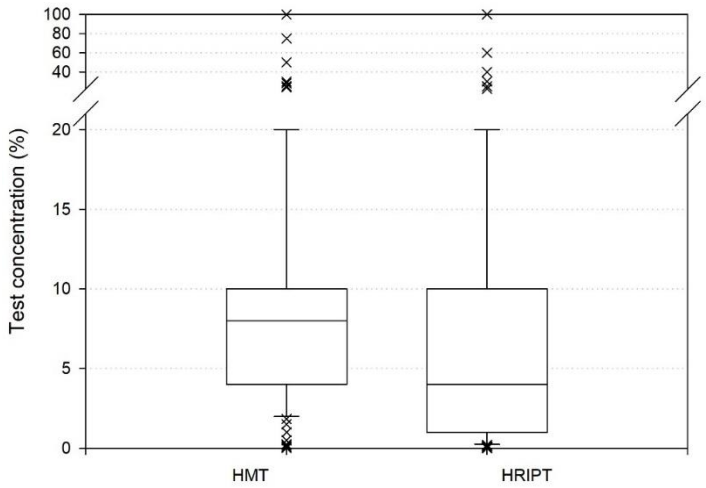

Figure 6: Distribution of test concentrations used in the HMT (n = 1619) and HRIPT (n = 636) test results with relative reliability score (RRS) < 5. Lines within the boxes mark the medians, box lower and upper boundaries the 25th/75th, and whiskers the 10th/90th percentiles, respectively. Fourty-four HRIPT results from the HPPT database are left out because the test concentration was available as a range only or not available at all.

In addition, Figure 7 shows the distribution of the DSA values for both test types. Only a small fraction (39/1619 or 2.4%) of the HMT results and less than one eighth (75/636 or 11.8%) of the HRIPT results with RRS < 5 were generated with a DSA ≤ 500 µg/cm^2.

These findings show that most of the available test results would not lead to an unambiguous classification, if the GHS criteria were applied as they stand. In fact, as Table 12 above shows, when applying the rules of the GHS directly to the 2255 test results in the HPPT database with RRS < 5, an unambiguous classification including sub-categorisation (1A, 1B, or NC) could be obtained in only 67 or ca. 3% of the cases. If only a binary decision (sensitiser yes/no) is required, 613 or ca. 27% of the test results can give an unambiguous classification.

[31] For 44 HRIPT results, no or no exact test concentration was available.

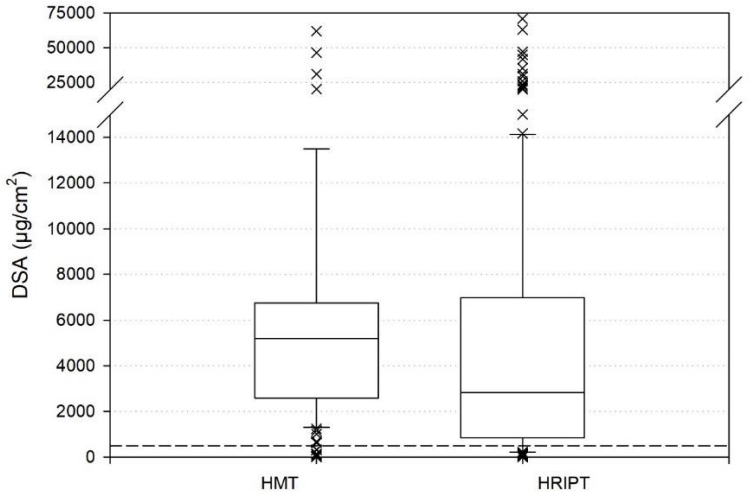

Figure 7: Distribution of the dose per skin area (DSA, in µg/cm²) used in the HMT (n = 1616) and HRIPT (n = 521) test results with relative reliability score (RRS) < 5. Lines within the boxes mark the medians, box lower and upper boundaries the 25th/75th, and whiskers the 10th/90th percentiles, respectively. The dashed horizontal reference line marks the GHS threshold value of 500 µg/cm². Three HMT and 115 HRIPT results from the HPPT database are left out because the DSA was available as a range only or not available at all.

Consequently, a methodology had to be developed to answer the following two questions:

- If a positive test result is obtained at DSA > 500 µg/cm², (how) can the likelihood of the same outcome at ≤ 500 µg/cm² be determined?
- Is there a test concentration at or above which a negative can be accepted as such and how can this concentration be determined?

Both questions are related to the definition of a positive vs. a negative GHS classification of the substance, which is straightforward for HPPT results: any test resulting in at least one sensitised person is a positive; test results with no sensitised individuals are negatives.

6.3 Resolution of ambiguous classifications and definition of borderline cases

6.3.1 Positive test result at DSA > 500 µg/cm²

Question 1 in section 6.2 translates into: "If a positive test result was obtained at > 500 µg/cm², is the number of sensitised test subjects expected at 500 µg/cm² ≥ 1?". This question could either be answered by estimating the number of sensitised test subjects at DSA = 500 µg/cm², or by determining whether the DSA estimated to cause exactly one sensitised subject was smaller or greater than (or equal to) 500 µg/cm². The latter approach was chosen here.

But how can the (hypothetical) DSA causing exactly one sensitised test subject (called "DSA1+" for the purposes of this report) be estimated? If test data at different DSAs were available, the number of sensitised individuals (POSNR) could be plotted vs. the DSA and a benchmark dose could be derived. However, this is not possible for HPPT data which are generated using only one test concentration, i.e. where only one (POSNR/DSA) data point is available.

In this situation one possibility to estimate the DSA1+ is by linear extrapolation from the POSNR observed in the test to POSNR = 1[32]. In other words:

$$DSA1+ = DSA/POSNR$$

[32] Note this implicitly assumes zero sensitisation at zero concentration which seems justified in general, but might be wrong e.g. if subjects in the test panel show an allergic reaction to the vehicle.

Of course, using linear extrapolation constitutes an oversimplification. In reality, a plot of number of positives vs. DSA based on an experiment using more than one test concentration most certainly would not result in a linear dose-response relationship. However, in spite of more than 60 years of applying HPPT concepts, apparently the underlying dose-response relationships have not been systematically investigated and are thus poorly understood.

Nevertheless, the database described in the present report may form a good basis for increasing this understanding. While it was not possible for the HDSG to perform this work within the time and capacity constraints of the OECD project on characterising DA performance, members of the HDSG are nevertheless currently planning to prepare a comprehensive review paper on diverse aspects of using HPPT data for regulatory purposes, which will also include a deeper analysis of dose-response correlations. For the time being, however, the HDSG decided not to speculate about the "true" shape of the dose-response curve, but rather to use a linear model for extrapolation, acknowledging the additional uncertainty introduced in this way (for an individual test result, this uncertainty increases as more sensitised individuals are observed, since the distance to the concentration causing exactly one sensitised subject increases).

Example: If a test with DSA = 3000 µg/cm^2 yielded three sensitised individuals (POSNR = 3), then linear extrapolation would lead to the “best guess” that one third of that DSA, i.e. 1000 µg/cm^2, would have sufficed to generate one third of the number of sensitised individuals, i.e. one sensitised individual.

Notably, in the published literature dealing with skin sensitisation classification based on HPPT data, frequently a parameter called "DSA05" has been calculated, i.e. the dose per skin area resulting in 5% of the test panel being sensitised. This parameter is obtained by linear extrapolation in much the same way as the DSA1+ (only in this case, extrapolation can be to both higher and lower DSA values).

Arguably, if potency categorisation follows the concept of "equipotent doses", i.e. two substances are of equal potency if under identical test conditions the same dose will lead to the same magnitude and incidence of effect (WHO IPCS, 2017), the DSA05 would appear to be a better choice for the direct comparison of the relative potencies of two substances tested under the same HPPT design.

However, in the present report the criteria of the GHS are applied, which does not use a percent incidence threshold (such as 5% sensitised). Instead, classification of the test substance as a skin sensitiser results from the occurrence of (at least) one sensitised member of the test panel. It is obvious that this may relate to very different sensitisation rates depending on panel size, e.g. 1/25 or 4% in the case of the standard HMT designs (Kligman, 1966b; Kligman and Epstein, 1975) and from 1/50 (2%) to 1/200 (0.5%) in the standard HRIPT designs (Draize, 1959; Griffith, 1969; Jordan Jr. and King, 1977; Marzulli and Maibach, 1973; Marzulli and Maibach, 1980a; Politano and Api, 2008; Shelanski and Shelanski, 1953; Voss, 1958).

The DSA1+ can then be used for classification under the GHS in the same way as the DSA:

- If DSA1+ ≤ 500 µg/cm^2, then this test result should be interpreted as Skin Sens. 1A.
- If DSA1+ > 500 µg/cm^2, then classification as Skin Sens. 1B is acceptable.

However, since these classifications are derived by extrapolation and to distinguish them from the "standard" GHS classifications, they will be called "extrapolated" classifications.

In addition, given the variability and uncertainty associated with the HPPT data, a borderline range should be defined around the cut-off of 500 µg/cm^2 to indicate that here the decision between Skin Sens. 1A or 1B is associated with a higher likelihood of misclassification than for DSA1+ values in a greater distance from the 500 µg/cm^2 cut-off. However, since variability and uncertainty around the HPPT data cannot be reliably quantified (cf. section 5), the width of this borderline range naturally can only be chosen in a somewhat arbitrary manner. Still, this allows for a uniform, transparent, and reproducible classification mechanism and is therefore preferable over a subjective, "expert judgement" case-by case approach. In addition, the borderline range could be adjusted easily, e.g. to

better match extrapolated classifications with reference classifications obtained in other ways (e.g. via LLNA or GPMT data), if necessary.

For the DSA1+, the HDSG chose a borderline range of ± 25% around the 500 µg/cm^2 cut-off, i.e. from 375 to 625 µg/cm^2. For the practical implementation of this decision, cf. section 7.1.1 below.

6.3.2 Negative test result at CONC < 100%

Question 2 in section 6.2 translates into: "Are there ways to estimate a minimum concentration/DSA, above which it is unlikely that a negative test result is a false negative?" This question cannot be answered on a per substance basis, but an estimate can be attempted by evaluating the full database of positive test results.

In a first step, for each positive result in the database, the normalised concentration hypothetically leading to exactly one positively tested test subject was determined by linear extrapolation from the actual observed number of positives at the test concentration, in analogy to the approach chosen to calculate the DSA1+ in the previous section:

$$CONC1+ = CONC/POSNR$$

Example: If a test yields three sensitised individuals using a test concentration of, say, 12%, then linear extrapolation would lead to the “best guess” that one third of that concentration, i.e. 4%, would have sufficed to generate one sensitised individual.

Regarding the application of linear extrapolation, the same considerations apply as discussed in the previous subsection.

Next, the distribution of the CONC1+ values was analysed (Figure 8).

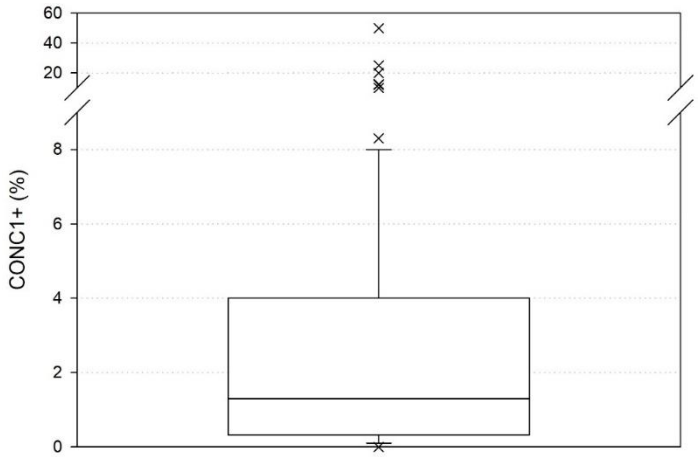

Figure 8: Distribution of the interpolated CONC1+ values dose obtained from 592 positive HMT and HRIPT test results. The line within the box marks the median, and the box lower and upper boundaries give the 25th/75th, and whiskers the 10th/90th percentiles, respectively. Thirteen positive results from the HPPT database were left out because CONC1+ was available as a range only or not available at all.

For 592 of the 605 positive test results with RRS < 5, a CONC1+ value was available. The following quantiles were obtained for this distribution: the median CONC1+ was 1.3%, the 95th percentile was found at CONC1+ = 10%, the 97th percentile at 20%, and the 99th percentile at 25%. The highest CONC1+ in the dataset characterising the test result for the least potent sensitiser (in HPPT assays) was 50% (two test results).

Since CONC1+ is a potency indicator (the less potent a sensitiser, the higher CONC1+), these values allow for the following conclusions:

- If a negative test result was obtained at > 50%, the test substance could still be a sensitiser, but its potency would be lower than that of all positive test substances in the database.
- If a negative test result was obtained at > 25%, the substance could still be a sensitiser, but its potency would be lower than that of 99% of the substances in the database with a positive result.
- If a negative test result was obtained at > 20%, the substance could still be a sensitiser, but its potency would be lower than that of 97% of the substances in the database with a positive result.

Based on these findings, the HDSG decided to use 25% (the 99th percentile) as the cut-off for conditional acceptance of negative test results. For the practical implementation of this decision, cf. section 7.1.1.

As for the DSA1+, setting a cut-off and defining a borderline range has the advantage that clear and transparent rules are put in place as a basis for all further discussions. Using the 99th percentile as the boundary for the borderline range might be regarded as fairly conservative here. On the other hand, it is obvious that both the cut-off and borderline range chosen, based to some extent on the CONC1+ distribution for the available data, are associated with considerable uncertainties. For instance, Figure 6 shows that test concentrations used for HPPT were generally quite low – testing higher concentrations more often might have shifted the CONC1+ distribution to higher values.

7 Methods

7.1 Classification of the EG DASS reference chemicals based on the available HPPT data

The reference database submitted along with the first draft of the guideline document on defined approaches for skin sensitisation initially consisted of the so-called “Cosmetics Europe” reference dataset for 128 substances (CosEU128) which has been described in detail in (Hoffmann et al., 2018). In the course of its work, the EG DASS removed some of these substances, e.g. because of variable or ill-defined composition, and a number of further substances were added, mainly to broaden the set of LLNA-negative reference chemicals. The final EG DASS reference data set comprises 200 reference substances with *in vitro* data (DPRA, KeratinoSens, h-CLAT) and varying degrees of coverage for LLNA and HPPT data.

The results of classifying these substances by applying the concepts explained in the previous and subsequent sections to the available HPPT data are reported in Table 19 (section 8). The underlying 453 individual test results evaluated for that purpose are documented in more detail in Appendix 2. In order to illustrate how the overall classifications were generated, the rules followed in evaluating individual test results, accounting for uncertainty/variability, and combining multiple test results are explained in the following subsection.

7.1.1 Rules followed in applying the GHS scheme to the EG DASS reference list

7.1.1.1 Evaluation of (relatively) reliable individual test results

The methods described in section 6.3 were applied to the data available for the EG DASS reference substances (Appendix 2). Since these assessments were directly (DSA1+) or indirectly (the test concentration (CONC) compared to the CONC1+ distribution) based on extrapolations to the threshold of one sensitised individual, the resulting classifications will be called “extrapolated classifications” hereafter.

In detail, the following classification rules were applied to individual test results:

- For negative test results, the classification outcome was *NC*[33] (not classified) if CONC ≥ 25%, regardless of the DSA value. This means that according to the rules of the GHS, this test result does not call for a classification of the test substance as a skin sensitiser; it does, however, not mean that this test result proves that the substance is not a sensitiser.
  - Negative test results with CONC < 25% were considered NC/1B, if they were obtained at a DSA ≥ 625 µg/cm2, i.e. above the upper boundary of the borderline range around the cut-off between sub-categories 1A and 1B.
  - It is important to understand that NC/1B is not a newly invented classification category; on the contrary it denotes an ambiguous classification outcome, in which it is not possible to assign one of two GHS sub-categories (NC and 1B) with sufficient certainty.
  - However, for test results with the ambiguous classification outcome NC/1B, the likelihood that a required classification as 1A had been missed was considered very low (whereas the need for classification as 1B could not be excluded).
  - In the case of negative test results obtained with CONC < 25% and a DSA ≤ 375 µg/cm2, i.e. below the lower boundary of the borderline range around the 500 µg/cm2 cut-off, the ambiguous classification outcome NC/1 was assigned. This effectively means that the test result could not provide any decisive information on the skin sensitisation potential of the test chemical and therefore had to be excluded from the overall classification. Nevertheless, such results were reported as additional information in parentheses in Table 19 (section 8) and Table 21 (Appendix 2).
- For positive test results, the extrapolated classification was *1B* if DSA1+ > 625 µg/cm^2, and *1A* if DSA1+ ≤ 375 µg/cm^2.
  - Positive test results with 500 µg/cm2 < DSA1+ ≤ 625 µg/cm2 received an extrapolated classification as "1B+". These test results were interpreted as showing a moderate sensitisation potential (1B), but with some (non-quantifiable) likelihood of under-classification.
  - For some of the positive test results, a DSA1+ value was not available. To such cases the ambiguous classification outcome “POS” was assigned, in order to reflect that a reliable GHS sub-categorisation (1A or 1B) was not possible. This is the third ambiguous classification outcome used in this report (in addition to NC/1B and NC/1, cf. above).
  - Positive test results with 375 µg/cm2 < DSA1+ ≤ 500 µg/cm2 were classified as "1A-". These test results were interpreted as showing a strong sensitisation potential (1A), but with some (non-quantifiable) likelihood of over-classification.
  - Finally, positive test results with DSA1+ ≤ 375 µg/cm2 were classified as 1A.

7.1.1.2 Evaluation of (relatively) unreliable individual test results

Test results with RRS = 5 were not considered for classification. Nevertheless, the outcome of these studies is reported in Table 21 (Appendix 2), but marked as additional information by putting the classification result(s) in parentheses.

Most test results with RRS = 4 do not allow for calculating a DSA1+ value and therefore cannot distinguish between Skin Sens. *1A* and *1B* classifications. However, they provide the test concentration

[33] Classification results are set in italics to indicate they represent individual classification outcomes.

used for induction (CONC). As a result, the possible extrapolated classification outcomes were: *NC*, *NC/1*, and *NC/1B* for negative, and *POS* for positive test results.

7.1.1.3 General decision logic for individual test results – "extrapolated classifications"

The general decision tree for establishing "extrapolated" classifications for individual study results as developed in the previous subsections is summarised in Figure 9.

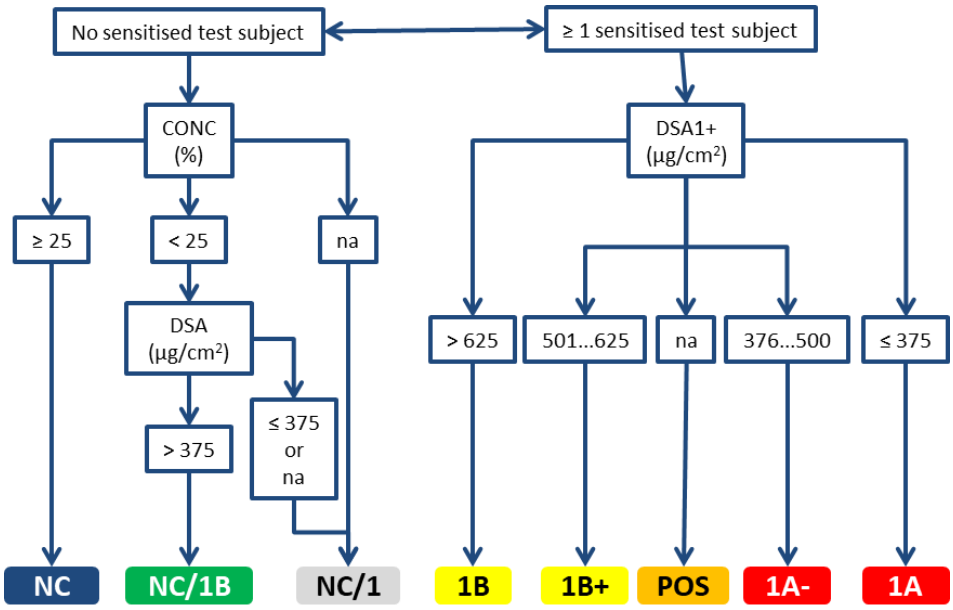

Figure 9: Schematic representation of the decision logic for the "extrapolated classification" of the EG DASS reference substances for their sensitisation potential according to the GHS, based on HPPT data

7.1.1.4 Weight-of-evidence (WoE) classification in the presence of diverging multiple test results

The original approach applied by the HDSG to combine multiple test results into an overall classification result for a given substance was an arithmetic mean-based approach called the "WoE score method". When the results from the parallel activity on curating the LLNA data for the EG DASS reference became available, it turned out that the LLNA sub-group (LSG) had used a median-based method, the so-called "Median-Like Location Parameter" (MLLP) approach. Furthermore, during the discussions in the LSG, a slightly modified median-based approach, the so-called "Median Sensitisation Potency Estimate" (MSPE) approach was applied in addition.

Since it was felt that the two datasets should be analysed with the same methodology, in this report, as in the final LSG report, all three methods are applied (cf. subsequent sub-sections).

7.1.1.4.1 Overall WoE score method

If for a given substance, diverging test results (each with RRS < 5) were available, first the individual extrapolated classification outcomes for these test results were determined. Next, each outcome received a numerical score based on the scheme given in Table 13 below. Note that the ambiguous outcome *NC/1* did not receive a numerical score, nor did it contribute to the WoE classification.

Table 13: Numerical scores assigned to different (extrapolated) classification outcomes from individual test results

Extrapolated classification	NC	NC/1B*	NC/1	1B	1B+	POS*	1A-	1A
Score	0	0.5	NA	1	1.25	1.5	1.75	2

* For ambiguous outcomes, the average score is used, i.e. *NC/1B* receives a score of (0 + 1)/2 = 0.5 and *POS* a score of (1+2)/2 = 1.5.

Next, all individual scores were added up and divided by the number of test results to give an overall WoE score.[34] This overall WoE score was then rounded to the second decimal and the corresponding "overall reference classification based on HPPT data" determined from Table 14. In Table 14, the overall WoE scores are translated into three different classification modes:

- GHS_{BIN}: Binary classification, i.e. *1*[35] (sensitiser) or *NC* (not classified);
- GHS_{SUB}: Ternary GHS classification including sub-categorisation, i.e. *1A* (strong/extreme sensitiser), *1B* (moderate/weak sensitiser), or *NC* (not classified);
- GHS_{BORDER}: Same as GHS_{SUB}, with two additional ambiguous classification outcomes, i.e. *1* (sensitiser, but sub-categorisation not possible) and *NC/1B* (cannot decide whether the substance is a sensitiser or not, but a strong/extreme sensitisation potential can be ruled out). Again, both *1* and *NC/1B* are not classification sub-categories; they characterise a limited data situation, where the assignment of the test substance to a GHS sub-category is associated with high uncertainty.

The main purpose of introducing the GHS_{BORDER} mode is to support DA performance assessment later in the process. This performance assessment will primarily use GHS_{BIN} and GHS_{SUB}, and the respective DA predictions will be judged to be "true" or "false" on that basis. However, the GHS_{BORDER} information will then point out substances, for which this decision is associated with high uncertainty or may even be wrong.

Table 14: Overall reference classification based on HPPT data as derived from the overall WoE score (na = not applicable)

Overall WoE score	Classification mode		
	GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}
1.76-2	1	1A[$]	1A
1.51-1.75			1[$]
1.50		na	
1.26-1.49		1B[$]	
0.76-1.25			1B
0.26-0.75	na	na	NC/1B
0-0.25[§]	NC	NC	NC

[$] For GHS_{SUB}, individual test results with *1A* and *1A-* are both considered as *1A* (and both *1B* and *1B+* are considered as *1B*), while for GHS_{BORDER}, *1A-* and *1B+* are considered as *1*. [§] In the special case that one or more test results yield the outcome *NC* and in addition only test results with the outcome *NC/1B* are present, the overall WoE score is set to zero.

Examples (for details regarding the database behind these examples, cf. Table 19 and Table 21):

1.) Penicillin G (CASRN 61-33-6/EC 200-506-3)

A total of 21 HPPT test results with RRS < 5 are available.

- There are four test results resulting in an extrapolated classification as *1A* (score 2)
- Two test results yield an extrapolated classification as *1A-* (score 1.75)
- Three test results yield an extrapolated classification as *1B+* (score 1.25).

[34] Note that this approach assumes equal weight of all test results with RRS < 5, where arguably a case-by-case analysis would show that certain test results should factor into the overall WoE score with higher weight than others. However, for the present report such case-by-case analyses could not be performed given the absence of clear-cut criteria as well as the fact that ca. 500 test results had to be evaluated.

[35] Note that while the outcome *POS* is assigned to individual positive study outcomes without sub-categorisation, the overall classification for a substance in this situation is designated *1*.

- Ten test results give an extrapolated classification as *1B* (score 1).
- Finally, two test results result in the ambiguous outcome *NC/1B* (score 0.5).

The overall WoE score is then calculated as:

[(4 x 2) + (2 x 1.75) + (3 x 1.25) + (10 x 1) + (2 x 0.5)]/21 = 26.25/21 = 1.25

This value is rounded to the second decimal, i.e. to an overall WoE score of 1.25. According to Table 14 above, this results in GHS_{BIN} = *1*, GHS_{SUB} = *1B*, and GHS_{BORDER} = *1B* as well.

2.) 2-Hexylidenecyclopentanone (CASRN 17373-89-6/EC 241-411-7)

Three HPPT test results with RRS < 5 are available.

- There is one test result resulting in a classification as *1A*, already without extrapolation (score 2).
- One test result yields an extrapolated classification as *1A-* (score 1.75).
- One test result gives an extrapolated classification as *1B* (score 1).

The overall WoE score is then calculated as:

[(1 x 2) + (1 x 1.75) + (1 x 1))]/3 = 4.75/3 = 1.58333

This value is rounded to the second decimal, i.e. to an overall WoE score of 1.58. According to Table 14 above, this results in GHS_{BIN} = *1*, GHS_{SUB} = *1A*, but GHS_{BORDER} = *1*.

This result can be interpreted as follows: under normal GHS conditions, this substance would be considered a strong skin sensitiser (based on the HPPT data); however, this sub-categorisation has to be considered uncertain, as demonstrated by an overall WoE score of 1.58, which is very close to the border between *1A* and *1B* (score 1.50). As a result, one should be conscious that setting *1A* as the overall reference classification for DA performance assessment bears a considerable risk of error.

3.) α-Amylcinnamaldehyde (CASRN 122-40-7/EC 204-541-5)

Three HPPT test results with RRS < 5 are available. No test subject was sensitised in any of the studies, but in two cases the test concentration was 6%, while in the third study it was 20%, i.e., still below 25%. These studies are sufficient to rule out that (based on the HPPT data) the substance should be classified under the GHS as a strong/extreme skin sensitiser; however, it is not possible to rule out that it should not be classified for skin sensitisation at all. As a result, neither GHS_{BIN}, nor GHS_{SUB} are applicable, and GHS_{BORDER} is *NC/1B*.

7.1.1.4.2 MLLP method

The LSG used a different WoE approach for combining multiple test results into an overall reference classification, based on the so-called "Median-Like Location Parameter" (MLLP) described by Hoffmann and co-workers:

> *"This parameter was defined as the median for substances with repeat studies with an EC3 in more than 50% of the repeats. For substances with at least 50% negative repeat studies, i.e. no EC3 value was available, the parameter was defined as the modified median. The first step in deriving the modified median was to review the negative studies in detail: when the maximum concentration tested in a given study was lower than the median EC3 of the positive studies for the same chemical, the respective negative study was excluded, because it was considered a limited validity as tested concentrations were too low. From the remaining negative and all positive studies, the median was used as a location parameter (modified median). In the case of*

50% of repeat studies being negative and 50% being positive, the highest EC3 value was defined as the modified median." (Hoffmann et al., 2018)[36]

Since both datasets preferably should be evaluated according to the same methodology, it was decided that also this MLLP approach should be applied to the HPPT data, in order to find out whether this would change any of the overall reference classifications.

Comparison of the approaches would then also allow for a sensitivity analysis: If different WoE approaches would result in different overall reference classifications, this would then either indicate ambiguous or borderline data situations susceptible to small changes in the evaluation methodology, or weaknesses of that methodology itself.

The MLLP approach published in Hoffmann et al. (2018) was further interpreted by the LSG[37] and this interpretation was adapted to the HPPT data (based on the individual test results as determined by the decision logic given in section 7.1.1.3) as follows:

- Test results with the ambiguous classification outcomes *NC/1* or *NC/1B* are included as negatives in the WoE, but only if the database for the respective chemical also includes studies with a positive outcome (*1A*, *1B*, or *POS*) and the DSA applied in the *NC/1B* study was higher than or equal to the median DSA1+ of the positive studies (for which a DSA1+ was available)[38].
- If the majority of the remaining studies is positive, the substance is considered a sensitiser under GHS_{BIN}, if it is negative, the overall GHS_{BIN} reference classification is *NC*.
- For GHS sub-categorisation (GHS_{SUB}), test results with a positive, but ambiguous classification outcome *POS* are excluded (in addition to the *NC/1* and *NC/1B* studies with too low test concentrations, see first bullet). Then the MLLP of the remaining study results is calculated. In case the median falls between two test results with DSA1+ values, the MLLP is the average of those two values. If it falls between the highest negative study and the lowest test result with a DSA1+ value, that DSA1+ value is the MLLP.
- As under the WoE score method, if the available individual test result outcomes are only *NC* or *NC/1B*, the overall MLLP is *NC*.

Table 15 shows how the MLLP is then translated into the overall reference classifications.

Table 15: Translation of MLLP values into overall GHS reference classifications (na = not applicable)

MLLP[§]	**Classification mode**		
	GHS_{BIN}	**GHS_{SUB}**	**GHS_{BORDER}**
≤ 375	1	1A	1A
375 < MLLP ≤ 500			1
500 < MLLP ≤ 625		1B	
> 625			1B
NC/1B	na	na	NC/1B
NC	NC	NC	NC

[§] Numerical values are (i.e. if the MLLP was a numerical DSA1+ value) in µg/cm^2

7.1.1.4.3 MSPE method

During the analysis of the LLNA data provided by the LSG it was noted that the MLLP approach had certain weaknesses and in a few cases produced WoE results which either appeared insufficiently conservative or counter-intuitive as compared to how different test results would be brought together

[36] For a more detailed description of this process, cf. the LSG report.

[37] Cf. LLNA sub-group report

[38] This rule also implies that if there are only *NC/1B* outcomes, no MLLP is available.

in a WoE assessment by a regulator tasked with classifying the respective substance for skin sensitisation.

As a result, the MLLP approach presented in the previous section was slightly modified, as follows:

- *NC/1* test results were completely excluded from the assessment, since they do not add any relevant information (but add noise to the median determination).
- Positive test results with a *POS* outcome (i.e. without an available DSA1+ value) are included when determining the position of the median.
- All test outcomes, whether numerical (DSA1+ values) or non-numerical (*POS*, *NC/1B*) are called "Sensitisation Potency Estimates" (SPEs)[39] and the median test result is therefore called the "Median Sensitisation Potency Estimate" (MSPE).
- The MSPE is then calculated by sorting all SPE values from low to high potency in the following order:

  NC → *NC/1B*[40] → Numerical SPE (DSA1+) values > 500 µg/cm^2 in descending order → *POS* → Numerical SPE values ≤ 500 µg/cm^2 in descending order.

The value of the MSPE is then determined as follows:

- If there are one or more positive results in addition to one or more NC/1B results, but there is no clear NC result, the median DSA1+ of the positive results with numerical values is taken as the MSPE. However, in all of these cases in which the number of *1A* (incl. *1A*-) study results equals that of the *1B* (incl. *1B+*) results, the MSPE is *POS*.
- If there are one or more *NC* results and all other test outcomes are *NC/1B*, the MSPE is *NC*.
- In all other cases, i.e. all cases in which both, clear positives and negatives are present, where the median falls between a numerical and a non-numerical result, the numerical result is taken as the MSPE.

Table 16 shows how the MSPE is then translated into the overall reference classifications.

Table 16: Translation of MSPE values into overall GHS reference classifications (na = not applicable)

<table>
<tr><th rowspan="2">MSPE[§]</th><th colspan="3">Classification mode</th></tr>
<tr><th>GHS$_{BIN}$[$]</th><th>GHS$_{SUB}$[$]</th><th>GHS$_{BORDER}$[$]</th></tr>
<tr><td>≤ 375</td><td rowspan="5">1</td><td rowspan="2">1A</td><td>1A</td></tr>
<tr><td>375 < MSPE ≤ 500</td><td rowspan="3">1</td></tr>
<tr><td>POS</td><td>na</td></tr>
<tr><td>500 < MSPE ≤ 625</td><td rowspan="2">1B</td></tr>
<tr><td>> 625</td><td>1B</td></tr>
<tr><td>NC/1B</td><td>na</td><td>na</td><td>NC/1B</td></tr>
<tr><td>NC</td><td>NC</td><td>NC</td><td>NC</td></tr>
</table>

§ Numerical values (i.e. if the MSPE was a numerical DSA1+ value) are in µg/cm^2

7.1.2 Determination of the overall classification result

The above calculations resulted in three classification sets (GHS$_{BIN}$, GHS$_{SUB}$, and GHS$_{BORDER}$) for each of the WoE score, MLLP, and MSPE approaches. The latter parameters as well as the resulting overall classifications are reported alongside an overview of the available studies in Table 19 in section 8.1.

[39] In analogy to the "Acute Toxicity Estimate" (ATE) in the GHS.

[40] Only those with a sufficiently high test concentration were included, cf. previous subsection.

The overall reference classifications resulting from the MLLP, the MSPE, and the WoE approaches were identical in the great majority of cases. Nevertheless, in some cases diverging classification results were noted between these three approaches, as shown in Table 17 below.

Table 17: Overview of those reference substances for which the classifications resulting from the MLLP, MSPE, and WoE score approaches were different from each other and/or from the overall reference classification (na = not available)

Name	CAS no.	EC no.	MLLP			MSPE			WoE score			Overall ref. classification		
			GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}
Benzaldehyde	100-52-7	202-860-4	na		NC/1B	1	1A	1	1	1B		1	1A	1
Benzyl salicylate	118-58-1	204-262-9	NC			na		NC/1B	na		NC/1B	na		NC/1B
L-Carvone	6485-40-1	229-352-5	1	1B		1	1B		na		NC/1B	1	1B	
Cinnamaldehyde	104-55-2	203-213-9	1	1A		1	1A		1	1B	1	1	1A	1
Citral	5392-40-5	226-394-6	1	1B	1	1	1A		1	1B	1	1	na	1
Dibenzoyl peroxide	94-36-0	202-327-6	1	1B	1	1	1B	1	1	1B	1B	1	1B	1
Diethyl maleate	141-05-9	205-451-9	1	1A		1	1A		1	1A	1	1	1A	
3,4-Dihydrocoumarin	119-84-6	204-354-9	1	1A	1	1	1A	1	1	na	1	1	na	1
Ethyl acrylate	140-88-5	205-438-8	1	1B		1	1A	1	1	1B		1	1B	
Eugenol	97-53-0	202-589-1	na		NC/1B	1	1B		na		NC/1B	1	1B	
Farnesol	4602-84-0	225-004-1	na		NC/1B	1	1B		na		NC/1B	1	1B	
Geraniol	106-24-1	203-377-1	na		NC/1B	1	1B		1	1B		1	1B	
trans-Hex-2-enal	6728-26-3	229-778-1	1	1A		1	1A		1	1B		1	1A	1
Hydratropaldehyde	93-53-8	202-255-5	1	1B		1	1B		na		NC/1B	1	1B	
Imidazolidinyl urea	39236-46-9	254-372-6	1	1B		1	1B		na		NC/1B	1	1B	
Isobergamate	68683-20-5	272-066-0	1	1B		1	1B		na		NC/1B	1	1B	
Isoeugenol	97-54-1	227-678-2	1	1B	1	1	1B	1	1	1B		1	1B	1
α-Isomethylionone	127-51-5	204-846-3	NC			NC			na		NC/1B	NC		
Lilial	80-54-6	201-289-8	1	1B		1	1B		na		NC/1B	1	1B	
Methyl non-2-ynoate	111-80-8	203-909-2	na		NC/1B	1	1A		1	1B		1	na	1
Methyl oct-2-ynoate	111-12-6	203-836-6	1	1A		1	1A		1	1A	1	1	1A	
OTNE	54464-57-2	259-174-3	NC			NC			na		NC/1B	NC		
Perillaldehyde	2111-75-3	218-302-9	1	1B		1	1B		na		NC/1B	1	1B	
Phenylacetaldehyde	122-78-1	204-574-5	1	1A		1	1B		1	1B		1	1A	1

In all of these cases, these questions were resolved via an expert decision by the EG DASS. The overall decision process for obtaining the HPPT reference classifications is shown in Figure 10.

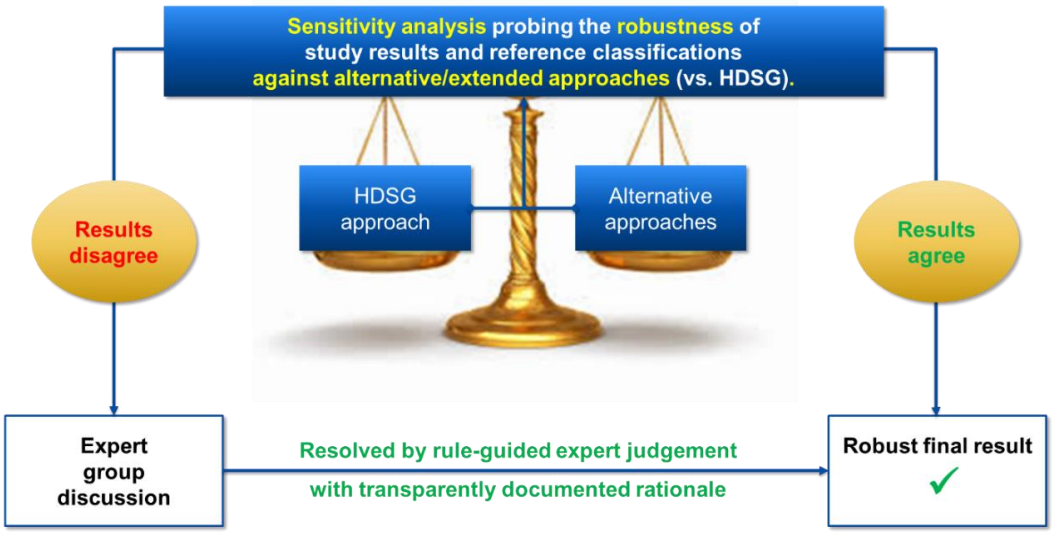

Figure 10: Schematic representation of the overall decision process for determining the HPPT reference classifications

The original HDSG approach (WoE score method) was challenged with slightly modified/extended, alternative evaluation strategies for sensitivity analysis (MLLP, MSPE approaches). If the results agreed, they were considered robust. If not, the respective reference substances were brought forward to the EG DASS for further discussion. In the majority of cases, such substances had either borderline potency or the test results by themselves did not provide for an unambiguous classification result. They were then resolved by rule-guided expert judgement and the rationale was documented. If no consensus could be achieved by the group, such substances were excluded from the reference data set.

7.2 Reproducibility of the HPPT-based reference classifications

If it were possible to perform an infinite number of HPPTs with a given substance, this would allow for a determination of the "true" reproducibility of the HPPT-based overall reference classification for that substance. Unfortunately, this is not possible and reproducibility must be estimated from a limited number of test results. As a result, sampling error must be taken into account.

Obviously, sampling error can be reduced by increasing the number of test results per substance, but on the other hand excluding substances with a "low" number of test results would leave fewer and fewer substances to include in the assessment. However, for obvious reasons, substances with only one test result cannot be used and had to be excluded from these calculations. For 58 substances with multiple HPPT results, reproducibility was determined for each classification mode in the following way:

- GHS_{BIN}: Reproducibility is calculated as the fraction of all HPPT results (with RRS < 5) for a given chemical that yielded an unambiguous classification result (1 or NC) correctly predicting the overall call. Studies resulting in an SPE of NC/1B were excluded from this evaluation, since for them GHS_{BIN} was not applicable.

- GHS_{SUB}: Reproducibility was calculated as the fraction of all HPPT results for a given chemical that yielded an unambiguous classification result (1A, *1B*, or *NC*) correctly predicting the overall call. Studies resulting in an SPE of *1* were excluded from this evaluation, since for them GHS_{SUB} was not applicable. For the same reason, studies resulting in *NC/1B* were omitted if the overall classification was *1B* or *NC.* They were, however, counted as contradictory, if it was *1A*.

It is important to understand that these reproducibilities refer to the HPPT-based GHS classifications. They do not represent the reproducibility of the DSA/DSA1+ values. Those can be quite variable, however, this is comparatively less important for clear 1A or 1B results, where even e.g. a two-fold increase or decrease in the DSA1+ might still result in the same sub-categorisation. However, as shown by the results given in Table 20 in section 8.2, a low reproducibility often indicates that the substance may have a skin sensitisation potential in the borderline area between 1A and 1B or that it is borderline between a weak sensitiser and a substance not requiring classification.

The results obtained in this way can be seen as a surrogate probability that a new (valid) HPPT result will confirm a classification previously established based on other HPPT results.

When this approach was presented to the EG DASS, some experts criticised that by not including all test results, but only those producing a classification result (whether binary or sub-categorised), reproducibility would appear higher than it actually was. However, the logic behind the approach is essentially the same as for any other assessment in the regulatory arena: only valid test results (in the sense of fit for the purpose of giving a classification result) are included in the assessment.

Example: Geraniol (CASRN 106-24-1/EC 203-377-1)

For geraniol, nine study results with RRS < 5 are available in total. The classification outcomes based on the individual studies were *NC/1B* (6 x), *1B* (2 x), and *1A* (1 x).

The MLLP was NC/1B, the MSPE was 4050 µg/cm^2, and the overall WoE score was 0.78. Overall, this resulted in GHS_{BIN} = *1* and GHS_{SUB} = *1B* (and GHS_{BORDER} = *1B*, too).

The reproducibility of GHS_{BIN} was calculated as follows:

- The six *NC/1B* results were not considered, since they did not give an unambiguous GHS_{BIN} result. The other three test results predict this substance to be a sensitiser and hence the reproducibility of the overall GHS_{BIN} = 1 is 100%.

However, as is obvious on first view, GHS_{SUB} = *1B* is not reproduced equally well:

- Again, the six *NC/1B* results are not considered, since they do not provide an unambiguous GHS_{SUB} outcome.
- Of the remaining three study results, only two reproduce GHS_{SUB} = *1B*, while one result predicts *1A*. Therefore reproducibility of GHS_{SUB} = *1B* is only (2/3) x 100% = 67%

Reproducibility results for the two classification modes GHS_{BIN} and GHS_{SUB} are reported in section 8.2.

8 Results

8.1 Classification of the EG DASS reference dataset

For 92/196 reference substances (47%) no or no reliable HPPT data were found. The results for the remaining 104 substances can be summarised as follows:

- GHS_{BIN}: An unambiguous classification could be obtained for only 66 (63.5%) of the 104 chemicals, of which 55 were classified as sensitisers (GHS_{BIN} = *1*) and eleven were not classified (GHS_{BIN} = *NC*).
- GHS_{SUB}: For 63/66 substances with an unambiguous GHS_{BIN} classification, also GHS_{SUB} classifications could be obtained. Of the 52 sensitisers among these 63 substances, 21 were classified as *1A* and 31 as *1B*.
- GHS_{BORDER}: Of the 21 substances with GHS_{SUB} = *1A*, seven had GHS_{BORDER} = *1*, i.e. the sub-categorisation is somewhat uncertain. The same was true for two of the 31 substances with GHS_{SUB} = 1B. As for the two other classification modes, 11 substances had GHS_{BORDER} = *NC*. For the

remaining 38 of the 104 substances with available reliable data, GHS_{BORDER} was NC/1B, i.e. it was not clear from the available HPPT data whether the substance was a sensitiser or not, but the need for classification as Skin Sens. 1A could be ruled out.

The classification results for the EG DASS reference substances are summarised in Table 18 and reported in detail in Table 19.

Table 18: Distribution of HPPT reference classifications over the GHS hazard classes/sub-categories (N = 104 substances with at least one test result with an RRS < 5)

Mode	GHS class/sub-category				
	1			NC/1B	NC
	1A	1	1B		
GHS_{BIN}	55			na	11
GHS_{SUB}	21	na	31		
GHS_{BORDER}	14	12	29	38	

Table 19: Classification of the EG DASS list of reference substances based on HPPT data[41].

Substance	CASRN	EC	"Extrapolated" classifications							MLLP	MSPE	WoE score	Overall classification			Reproducibility	
			NC	NC/1B	1B	1B+	1	1A-	1A				GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}
Abietic acid	514-10-3	208-178-3															
Acetanisole	100-06-1	202-815-9		1						na	NC/1B	0.50	na		NC/1B	na	
2-Acetylcyclohexanone	874-23-7	212-858-5															
4-Allylanisole	140-67-0	205-427-8		1						na	NC/1B	0.50	na		NC/1B	na	
Allyl phenoxyacetate	7493-74-5	231-335-2		1						na	NC/1B	0.50	na		NC/1B	na	
4-Aminobenzoic acid	150-13-0	205-753-0	4							NC		0.00	NC			100	
4-Amino-m-cresol	2835-99-6	220-621-2															
5-Amino-o-cresol	2835-95-2	220-618-6															
2-Aminophenol	95-55-6	202-431-1															
3-Aminophenol	591-27-5	209-711-2															
α-Amylcinnamaldehyde	122-40-7	204-541-5		3						na	NC/1B	0.50	na		NC/1B	na	
α-Amylcinnamic alcohol	101-85-9	202-982-8		2						na	NC/1B	0.50	na		NC/1B	na	
Anethole	104-46-1	203-205-5															
Aniline	62-53-3	200-539-3			1					1773		1.00	1	1B		na	
Anisyl alcohol	105-13-5	203-273-6		1						na	NC/1B	0.50	na		NC/1B	na	
2-(p-Anisyl)propanal	5462-06-6	226-749-5		1						na	NC/1B	0.50	na		NC/1B	na	
Applelide	478695-70-4	639-080-2															
BADGE	1675-54-3	216-823-5															
Bandrowski's base	20048-27-5	na															
Benzaldehyde	100-52-7	202-860-4		3				1		NC	492	0.81	1	1A[42]	1	na	
1,2-Benzisothiazol-3(2H)-one	2634-33-5	220-120-9															
Benzocaine	94-09-7	202-303-5		4	8	1				3103		0.87	1	1B		100	
Benzoic acid	65-85-0	200-618-2		2						na	NC/1B	0.50	na		NC/1B	na	
p-Benzoquinone	106-51-4	203-405-2															
Benzyl alcohol	100-51-6	202-859-9		2	3					3543		0.80	1	1B		100	
Benzyl benzoate	120-51-4	204-402-9	1							NC		0.00	NC			na	
Benzyl bromide	100-39-0	202-847-3															
Benzyl butyl phthalate	85-68-7	201-622-7															
Benzyl cinnamate	103-41-3	203-109-3		3						na	NC/1B	0.50	na		NC/1B	na	
Benzyl salicylate	118-58-1	204-262-9	3	2	2					NC	NC/1B	0.43	na		NC/1B	na	
Benzylidene acetone	122-57-6	204-555-1							2	231		2.00	1	1A		100	
BGE	2426-08-6	219-376-4							1	327		2.00	1	1A		na	
Bis-GMA	1565-94-2	216-367-7															
Bourgeonal	18127-01-0	242-016-2		1						na	NC/1B	0.50	na		NC/1B	na	
1-Bromobutane	109-65-9	203-691-9															

[41] More details on all test results (including those with RRS = 5 not included here) are provided in Appendix 2.

[42] EG DASS consensus based on the decision that the clear positive result should outrule the NC/1B results.

Substance	CASRN	EC	"Extrapolated" classifications							MLLP	MSPE	WoE score	Overall classification			Reproducibility	
			NC	NC/1B	1B	1B+	1	1A-	1A				GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}
1-Bromohexane	111-25-1	203-850-2															
Bromothalonil	35691-65-7	252-681-0															
Butan-1-ol	71-36-3	200-751-6															
2-Butoxyethyl acetate	112-07-2	203-933-3															
Butyl acrylate	141-32-2	205-480-7															
L-Carvone	6485-40-1	218-827-2		1	1					4724		0.75	1	1B		100	
CD-3	25646-71-3	247-161-5															
Chloramine-T	127-65-1	204-854-7															
3-Chloro-p-anisaldehyde	1097074	225-532-2															
Chlorobenzene	108-90-7	203-628-5															
Chlorothalonil	1897-45-6	217-588-1															
Chlorpromazine	50-53-3	200-045-8			1					862		1.00	1	1B		na	
Cinnamaldehyde	104-55-2	203-213-9		4	2	1			9	343	338	1.45	1	1A	1	100	75
Cinnamic alcohol	104-54-1	203-212-3		6	20		1			2575	2250	0.91	1	1B		100	
Cinnamonitrile	1885-38-7	217-552-5		1	2					1000	887.5	0.83	1	1B		100	
Citral	5392-40-5	226-394-6		2	2	1	4		4	620	163	1.40	1	na[43]	1	100	na
Citronellol	106-22-9	203-375-0	1	1						NC		0.25	NC			na	
Clofibrate	637-07-0	211-277-4															
Coumarin	91-64-5	202-086-7		1	2					5039	3937	0.83	1	1B		100	
Cyclamen aldehyde	103-95-7	203-161-7		2						na	NC/1B	0.50	na		NC/1B	na	
trans-Dec-2-enal	3913-81-3	223-474-2		1						na	NC/1B	0.50	na		NC/1B	na	
DEET	134-62-3	205-149-7															
Diacetyl	431-03-8	207-069-8		1						na	NC/1B	0.50	na		NC/1B	na	
2,5-Diaminotoluene sulfate	615-50-9	210-431-8															
Dibenzoyl peroxide	94-36-0	202-327-6				1				567		1.25	1	1B	1	na	
Dibenzyl ether	103-50-4	203-118-2		1						na	NC/1B	0.50	na		NC/1B	na	
Dibutyl phthalate	84-74-2	201-557-4															
N,N-Dibutylaniline	613-29-6	210-335-6															
1,1-Dichloroethene	75-35-4	200-864-0															
Diethyl maleate	141-05-9	205-451-9			1				3	179.5		1.75	1	1A		100	75
Diethyl phthalate	84-66-2	201-550-6		1						na	NC/1B	0.50	na		NC/1B	na	
Diethyl sulfate	64-67-5	200-589-6															
Diethylenetriamine	111-40-0	203-865-4							1	296		2.00	1	1A		na	
3,4-Dihydrocoumarin	119-84-6	204-354-9				1		1		480.5		1.50	1	na	1	100	na
Dihydroeugenol	2785-87-7	220-499-0		1						na	NC/1B	0.50	na		NC/1B	na	
Dihydromyrcenol	18479-58-8	242-362-4		1						na	NC/1B	0.50	na		NC/1B	na	
Dimethyl fumarate	624-49-7	210-849-0															

[43] Not suitable for sub-categorisation – there are eight studies in favour of or not contradicting *1A*, and nine studies in favour of or not contradicting *1B*

Substance	CASRN	EC	"Extrapolated" classifications							MLLP	MSPE	WoE score	Overall classification			Reproducibility	
			NC	NC/1B	1B	1B+	1	1A-	1A				GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}
3-Dimethylaminoopropylamine	109-55-7	203-680-9															
Diphenylcyclopropenone	886-38-4	212-948-4															
Disodium 2-[4-[[2-cyano-3-[4-[methyl(2-sulfonatoethyl)amino]-phenyl]-1-oxoallyl]amino]phenyl]-6-methylbenzothiazole-7-sulfonate	2498-95-5	219-694-3															
DMSO	67-68-5	200-664-3	1							NC		0.00	NC			na	
DNBS, sodium salt	885-62-1	212-943-7															
DNCB	97-00-7	202-551-4							6	1.5		2.00	1	1A		100	
EGDMA	97-90-5	202-617-2															
Ethyl acrylate	140-88-5	205-438-8		2	1			1	1	709	420	1,15	1	1B		100	33
Ethyl benzoylacetate	94-02-0	202-295-3															
Ethyl vanillin	121-32-4	204-464-8		1						na	NC/1B	0.50	na		NC/1B	na	
2-Ethylbutanal	97-96-1	202-623-5															
Ethylene brassylate	105-95-3	203-347-8	1	22	2					na	NC/1B	0.52	na		NC/1B	na	
Ethylene diamine (free base)	107-15-3	203-468-6															
2-Ethylhexyl acrylate	103-11-7	203-080-7															
Eugenol	97-53-0	202-589-1		3	1					NC/1B	2520	0.63	1	1B		na	
Farnesal	502-67-0	242-957-9															
Farnesol	4602-84-0	225-004-1		9	5					NC/1B	1688	0.68	1	1B		na	
2-Fluoro-5-nitroaniline	369-36-8	206-720-3															
Formaldehyde	50-00-0	200-001-8							5	172		2.00	1	1A		100	
Furil	492-94-4	207-766-7															
Geraniol	106-24-1	203-377-1		6	2				1	NC/1B	3150	0.78	1	1B		100	
Glutaraldehyde	111-30-8	203-856-5						1		498		1.75	1	1A	1	na	
Glycerol	56-81-5	200-289-5															
Glyoxal	107-22-2	203-474-9							1	259		2.00	1	1A		na	
Helional	1205-17-0	214-881-6															
Hepta-2,4-dienal	5910-85-0	227-627-4		1						na	NC/1B	0.50	na		NC/1B	na	
n-Hexane	110-54-3	203-777-6	1							NC		0.00	NC			na	
trans-Hex-2-enal	6728-26-3	229-778-1		1					1	39		1.25	1	1A	1	na	
Hexyl salicylate	6259-76-3	228-408-6	1	1						NC		0.00	NC			na	
α-Hexylcinnamaldehyde	101-86-0	202-983-3		2						na	NC/1B	0.50	na		NC/1B	na	
2-Hexylidenecyclopentanone	17373-89-6	241-411-7			1			1	1	422		1.58	1	1A[44]	1	100	67
HHPA	85-42-7	201-604-9															
Hydratropaldehyde	93-53-8	202-255-5		1	1					1550		0.75	1	1B		na	

[44] EG DASS consensus: 1A

Substance	CASRN	EC	"Extrapolated" classifications							MLLP	MSPE	WoE score	Overall classification			Reproducibility	
			NC	NC/1B	1B	1B+	1	1A-	1A				GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}
Hydrocortisone	50-23-7	200-020-1	1							NC		0.00	NC			na	
Hydroquinone	123-31-9	204-617-8															
4-Hydroxybenzoic acid	99-96-7	202-804-9															
Hydroxycitronellal	107-75-5	203-518-7		7	12					4650	3937.5	0.82	1	1B		na	
2-Hydroxyethyl acrylate	818-61-1	212-454-9															
2-Hydroxypropyl methacrylate	923-26-2	213-090-3															
Imidazolidinyl urea	39236-46-9	254-372-6		1	1					1250		0.75	1	1B		na	
Iodocarb	55406-53-6	259-627-5															
1-Iodohexane	638-45-9	211-339-0															
Isobergamate	68683-20-5	272-066-0		3	3					2250		0.7511	1	1B		100	
Isobornyl acetate	125-12-2	204-727-6		1						na	NC/1B	0.50	na		NC/1B	na	
p-Isobutyl-α-methyl-hydrocinnamaldehdye	6658-48-6	229-695-0															
2-Isobutyl-4-methyl-tetrahydro-2H-pyran-4-ol	63500-71-0	613-238-0		2						na	NC/1B	0.50	na		NC/1B	na	
Isoeugenol	97-54-1	227-678-2		1		1				560		0.88	1	1B	1	na	
α-Isomethylionone	127-51-5	204-846-3	2	3						NC		0.30	NC			na	
Isopropanol	67-63-0	200-661-7															
Isopropyl myristate	110-27-0	203-751-4		1						na	NC/1B	0.50	na		NC/1B	na	
Kanamycin	59-01-8	200-411-7			1					1411		1.00	1	1B		na	
Kathon CG	55965-84-9	611-341-5; 911-418-6							6	0.002		2.00	1	1A		100	
Lactic acid	50-21-5	200-018-0															
Lauryl gallate	1166-52-5	214-620-6															
Lilial	80-54-6	201-289-8		2	1					29528		0.67	1	1B		na	
D-Limonene	5989-27-5	227-813-5		1						na	NC/1B	0.50	na		NC/1B	na	
Linalool	78-70-6	201-134-4		4						na	NC/1B	0.50	na		NC/1B	na	
Maleic anhydride	108-31-6	203-571-6															
2-Mercaptobenzothiazole	149-30-4	205-736-8		1	2					1724	1482.5	0.83	1	1B		100	
2-Methoxy-p-cresol	93-51-6	202-252-9															
1-(4-Methoxyphenyl)-pent-1-en-3-one	104-27-8	203-190-5							1	338		2.00	1	1A		na	
Methyl acrylate	96-33-3	202-500-6															
Methyl 3-bromoproprionate	3395-91-3	222-247-5															
Methyl dihydrojasmonate	24851-98-7	246-495-9		5						na	NC/1B	0.50	na		NC/1B	na	
Methyl methacrylate	80-62-6	201-297-1		1						na	NC/1B	0.50	na		NC/1B	na	
Methyl methanesulfonate	66-27-3	200-625-0															

Substance	CASRN	EC	"Extrapolated" classifications							MLLP	MSPE	WoE score	Overall classification			Reproducibility	
			NC	NC/1B	1B	1B+	1	1A-	1A				GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}
Methyl non-2-ynoate	111-80-8	203-909-2		2					1	NC/1B	20	1.00	1	na[45]	1	na	
Methyl oct-2-ynoate	111-12-6	203-836-6		2	1		1	1	9	270		1.66	1	1A		100	
Methyl pyruvate	600-22-6	209-987-4															
Methyl o-toluate	89-71-4	201-932-2															
1-(3-Methyl-2-benzofuranyl)ethanone	23911-56-0	429-100-6															
α-Methylcinnamaldehyde	101-39-3	202-938-8		1						na	NC/1B	0.50	na		NC/1B	na	
6-Methylcoumarin	92-48-8	202-158-8		1						na	NC/1B	0.50	na		NC/1B	na	
2-Methyldecanenitrile	69300-15-8	273-960-3															
6-Methylhepta-3,5-dien-2-one	1604-28-0	216-507-7						1		459		2.00	1	1A	1	na	
5-Methylhexane-2,3-dione	13706-86-0	237-241-8			2					2531.5		1.00	1	1B		100	
Methylisothiazolinone	2682-20-4	220-239-6							2	22.5		2.00	1	1A		100	
4-Methyl-2-nitroanisole	119-10-8	204-296-4															
Methylparaben	99-76-3	202-785-7															
2-Methylundecanal	110-41-8	203-765-0		1						na	NC/1B	0.50	na		NC/1B	na	
Metol	55-55-0	200-237-1															
1-Naphthol	90-15-3	201-969-4															
Neomycin sulfate	1405-10-3	215-773-1	1	2	14	1			1	1936		0.96	1	1B		94	89
4-Nitrobenzylbromide	100-11-8	202-820-6															
2-Nitro-p-phenylenediamine	5307-14-2	226-164-5															
cis-6-Nonenal	2277-19-2	218-900-9		1						na	NC/1B	0.50	na		NC/1B	na	
Octanoic acid	124-07-2	204-677-5		1						na	NC/1B	0.50	na		NC/1B	na	
OTNE	54464-57-2	259-174-3	1	5						NC		0.42	NC			na	
Oxalic acid	144-62-7	205-634-3															
Oxazolone	15646-46-5	239-713-9															
Penicillin G	61-33-6	200-506-3		2	10	3		2	4	970	690	1.25	1	1B		100	68
Pentachlorophenol	87-86-5	201-778-6															
Perillaldehyde	2111-75-3	218-302-8		2	2					1350	1125	0.75	1	1B		100	
3-Phenoxypropanenitrile	3055-86-5	3055-86-5															
Phenyl benzoate	93-99-2	202-293-2			1					10393		1.00	1	1B		na	
Phenylacetaldehyde	122-78-1	204-574-5		3	1				3	648	221	1.21	1	1A	1	100	75
p-Phenylenediamine	106-50-3	203-404-7			1				12	18		1.92	1	1A		100	92
1-Phenylpropane-1,2-dione	579-07-7	209-435-2															
Phthalic anhydride	85-44-9	201-607-5															
Propyl gallate	121-79-9	204-498-2															
Propylene glycol	57-55-6	200-338-0	2	1						NC		0.17	NC			100	
3-Propylidenephthalide	17369-59-4	241-402-8			1					900		1.00	1	1B		na	
Propylparaben	94-13-3	202-307-7															

[45] HDSG expert consensus: sub-categorisation is not possible, because MLLP results in NC/1B, MSPE results in 1A, and WoE score results in 1B, but there is also a clear (i.e. not extrapolated) 1A result. Overall, only the classification outcome 1 (no sub-categorisation) is compatible with all test results.

Substance	CASRN	EC	"Extrapolated" classifications							MLLP	MSPE	WoE score	Overall classification			Reproducibility	
			NC	NC/1B	1B	1B+	1	1A-	1A				GHS_{BIN}	GHS_{SUB}	GHS_{BORDER}	GHS_{BIN}	GHS_{SUB}
Pyridine	110-86-1	203-809-9			1					31034		1.00	1	1B		na	
Resorcinol	108-46-3	203-585-2		1						na	NC/1B	0.50	na		NC/1B	na	
Saccharin	81-07-2	201-321-0															
Safranal	116-26-7	204-133-7							1	60		2.00	1	1A		na	
Salicylic acid	69-72-7	200-712-3		1						na	NC/1B	0.50	na		NC/1B	na	
Sodium lauryl sulfate	151-21-3	205-788-1	1							NC		0.00	NC^{46}			na	
Squaric acid	2892-51-5	220-761-4															
Sulfanilamide	63-74-1	200-563-4			1					3103		1.00	1	1B		na	
Sulfanilic acid	121-57-3	204-482-5															
L-Tartaric acid	87-69-4	201-766-0															
Tetrachlorosalicylanilide	1154-59-2	214-576-8							13	83		2.00	1	1A		100	
2,2,6,6-Tetramethyl-heptane-3,5-dione	1118-71-4	214-268-3															
Thioglycerol	96-27-5	202-495-0			4				1	1241		1.20	1	1B		100	80
Thiram	137-26-8	205-286-2		2	6					3879	9698	0.88	1	1B		100	
α-Tocopherol	59-02-9	200-412-2															
Triethanolamine	102-71-6	203-049-8															
Trimellitic anhydride	552-30-7	209-008-0															
Tropolone	533-75-5	208-577-2															
Undec-10-enal	112-45-8	203-973-1		1						na	NC/1B	0.50	na		NC/1B	na	
Vanillin	121-33-5	204-465-2		2						na	NC/1B	0.50	na		NC/1B	na	
4-Vinylpyridine	100-43-6	202-852-0															

[46] Expert group consensus: NC due to excessive irritancy at higher concentrations

8.2 Reproducibility

Table 20: Reproducibility of HPPT-based GHS_{BIN} and GHS_{SUB} classification modes in relation to the number of available test results per substance

Classification mode	Number of test results available	No. of substances	Reproducibility (%)		
			Mean	Median	SD
GHS_{BIN}	≥ 2	34	99.8	100	1.0
	≥ 3	23	99.7	100	1.2
	≥ 4	20	99.7	100	1.3
	≥ 5	16	99.6	100	1.5
GHS_{SUB}	≥ 2	32	91.0	100	16
	≥ 3	22	86.9	96.2	18
	≥ 4	17	90.9	100	12
	≥ 5	14	92.5	100	11

These results show that classifications based on the HPPT are highly reproducible. For GHS_{SUB}, there was a trend towards lower (but still high) reproducibilities, when the minimum number of test results was increased (the significance of the apparent minimal trend observed for GHS_{BIN} cannot be assessed).

Taking only the results with at least three tests per substance, on average, binary classifications ("1"/"NC"; GHS_{BIN}) were almost 100% reproducible, while GHS classifications including sub-categorisation ("1A"/"1B"/"NC"; GHS_{SUB}) had a reproducibility of ca. 87-93%.

Another way to calculate reproducibility based on the available data is by calculating a weighted average. This means that for each substance, first the individual reproducibility of the result was multiplied by the number of tests available for that substance. These products were added up for all substances and that sum was then divided by the total number of test results (for all substances). In that way, a reproducibility value based on e.g. six test results received twice the weight of a reproducibility value calculated based on only three experiments.

The weighted mean reproducibilities for all three classifications (considering all substances with at least two test results) were found to be 99.5% for GHS_{BIN} and 90.3% for GHS_{SUB}, i.e. about the same as the respective non-weighted reproducibilities for substances with at least 4 or 5 test results.

It is concluded that where multiple HPPT results were available for substances from the EG DASS reference list, HPPT-based GHS_{BIN} classifications were almost 100% and GHS_{SUB} classifications ca. 90% reproducible.

These figures are still remarkably high, but they have to be taken with a grain of salt.

- First, for many substances these reproducibility calculations are still based on only a handful of results. Therefore, the estimates of "true" reproducibility given here might still be somewhat off-target.
- Second, the median reproducibilities and standard deviations reported in Table 20 above indicate that a closer look into those substances with lower reproducibility might be useful. This might allow for each individual case to identify whether the observed reproducibility < 100% can be attributed to substance-specific issues (the substance could be a "true borderliner" with a potency located directly at the border between two sub-categories or it could be a substance with known variability issues for the test material) or rather to a weakness of the test design itself.

Starting with the GHS_{BIN} classifications, only 1/33 substances with more than one test result had a reproducibility < 100% (neomycin sulfate, CASRN 1405-10-3/EC 215-773-1; 94%). For this substance, one test results in *NC*, two test results give *NC/1B*, 14 tests give *1B*, one test results in *1B+*, and one test suggests *1A*. After excluding the two *NC/1B* results (unsuitable for binary classification, 16 of the

17 remaining studies reproduce the positive result. The negative result was obtained in one of a series of experiments with different patch sizes which were reported in Table 3 (page 18) above. As evident from that table, this was obtained with the smallest patch used, while all experiments with larger patches were clearly positive. It can be concluded that in this case, reproducibility < 100% can most likely be attributed to variability in test design.

For the GHS_{SUB} classification, 6/12 *1A* chemicals with more than one unambiguous test result had a reproducibility of 100%, the other six were:

- Cinnamaldehyde (CASRN 104-55-2/EC 203-213-9): Out of the twelve test results suitable for GHS_{SUB}, nine confirmed the overall classification as 1A (75% reproducibility). The first of the three remaining test results was from an HMT predicting 1B, with 3/25 test subjects sensitised at a test concentration of 3% and a DSA1+ barely above the borderline range (648 µg/cm^2). One more sensitised test subject would have made this result a 1A. The second test result was from an HRIPT predicting 1B+, with 1/55 test subjects sensitised at a test concentration of 1% (DSA = 620 µg/cm^2). Another test under the same conditions (including the same vehicle) sensitised 3/107 test subjects, resulting in classification as 1A, which would also be the result, if the two tests were pooled (4/162 sensitised). The third of the remaining tests was an HRIPT predicting 1B, also at 3% test concentration and a DSA1+ of 886 µg/cm^2, sensitising 4/28 test subjects. This is an unusually low number of test subjects for an HRIPT, and therefore the reliability of this test result might seem debatable. All in all, the potency of this chemical seems to be in the 1A range, but not too far remote from the 1A/1B border. This is also represented by GHS_{BORDER} = *1*, and therefore a lower reproducibility of the GHS_{SUB} reference classification does not come as a surprise.
- Diethyl maleate (CASRN 141-05-9/EC 205-451-9): One out of four available tests gives *1B*, i.e. it does not confirm the overall reference classification of *1A* (75% reproducibility). In this HRIPT, 2% of the test substance in petrolatum sensitised 2/52 test subjects at a DSA1+ of 709 µg/cm^2. With one more sensitised subject this would have resulted in *1A*, which was also the outcome of a second study, in which 8/52 test subjects were sensitised by the same test concentration, but using ethanol as vehicle. The deviation here seems to be due to experimental variability rather than due to a weakness of the test itself.
- 2-Hexylidenecyclopentanone (CASRN 17373-89-6/EC 241-411-7): There are three tests resulting in *1B*, *1A*-, and *1A*, respectively (the latter even without extrapolation). This is a borderline 1A/1B chemical, with GHS_{BORDER} = 1 (note that the one available LLNA result for this chemical is also in the borderline 1A/1B range).
- Methyl oct-2-ynoate (CASRN 111-12-6): All ten HMT tests suitable for GHS_{SUB} have been performed at a test concentration of 2%, using 21-25 test subjects and resulting in sensitisation rates between 4% and 40%. Eight of these tests result in *1A*, one in *1A*- (DSA1+ = 422 µg/cm^2), and one in *1B*, and therefore the overall reference classification is confirmed with 90% reproducibility, while the one *1B* result can be attributed to normal experimental variability.
- Phenylacetaldehyde (CASRN 122-78-1/EC 204-574-5): All four test results suitable for sub-categorisation were obtained at the same DSA of 1296 µg/cm^2, with 2, 4, 11 and 12 positive reactions, respectively. When determining the extrapolated classification[47] this resulted in DSA1+ values of 648, 324, 118, and 108 µg/cm^2, i.e. three extrapolated classification outcomes were *1A* and one was *1B*. Without extrapolation, all tests would have resulted in *1B*. Notably, two negative (*NC/1B*) test results were available for this substance, which achieved a DSA of 1550 µg/cm^2, i.e. higher than that used in the positive tests, and were therefore included in MLLP and MSPE calculation. The overall classification was 1A, but it is clear that this is a borderline result, and so the reproducibility of only 75% is not astonishing. Furthermore, all positive test results came from

[47] cf. 6.3.1

HMTs, while the negatives originated from HRIPTs, hence test design might have played an additional role.

- p-Phenylenediamine (CASRN 106-50-3/EC 106-50-3): Only one of the 13 available tests did not reproduce the overall reference classification as 1A (92% reproducibility). This 1B result was obtained at a test concentration of 25%, where all 20 test subjects were sensitised. The DSA1+ of 776 µg/cm^2 (i.e. not very far from the 1A/1B borderline zone beginning at 625 µg/cm^2) was therefore obtained by linear extrapolation from a 20-fold higher DSA, and the deviating result is most likely explained by the inherent error of the linear extrapolation.

For 12 of the 17 substances with a GHS_{SUB} of "1B" and more than one unambiguous test result, classifications were 100% reproducible.

- Ethyl acrylate (CASRN 140-88-5/EC 205-438-8): There are two *NC/1B* and one *1B* test result excluding and two 1A results supporting *1A*. While it is clear that this chemical is a borderline 1A/1B substance, the overall reference classification is 1B. However, the two *NC/1B* studies are not suitable for sub-categorisation and are therefore discounted in the reproducibility calculations. Two of the remaining studies are in favour of *1A* and therefore the overall classification as 1B is only reproduced by 33% of the available studies.
- Geraniol (CASRN 106-24-1/EC 203-377-1): Six *NC/1B* test results are discounted. Only one of the three positive tests results in *1A* and does not confirm the overall reference classification as 1B (which therefore has 67% reproducibility). This particular test is not well-documented and therefore this result is attributed to possible variability in the experimental conditions.
- Neomycin sulfate (CASRN 1405-10-3/EC 215-773-1): Cf. above. The reproducibility of the GHS_{SUB} classification is 89%, i.e. slightly lower than that of the GHS_{BIN} classification, because in addition to the one available *NC* result, 1/16 positive test results is *1A*, i.e. it does not confirm the overall GHS_{SUB} = *1B*.
- Penicillin G (CASRN 61-33-6/EC 200-506-3): Two NC/1B test results are discounted. Six out of the remaining 19 tests result in *1A* and therefore do not confirm the overall reference classification as 1B (68% reproducibility). Since all these results come from the original HMT publications in which Albert Kligman investigated the influence of a number of experimental parameters on the sensitisation outcome, this is likely attributable to experimental variability.
- Thioglycerol (CASRN 96-27-5/EC 202-495-0): Four test results give *1B*, while one result gives *1A*, resulting in an 80% reproducibility of the overall *1B* reference classification, which might be attributable to normal variability. It is however noted that with one sensitised test subject less, the one *1A* result would already have turned into a *1A*- result, i.e. it would have fallen into the borderline 1A/1B zone.

Also for GHS_{SUB}, all five NC classifications with more than one test result were 100% reproducible.

All in all, it seems that under the conditions of this project, i.e. using rigorously curated HPPT data and following standardised assessment routines, the reproducibility of HPPT-based GHS classifications on average is almost 100% for GHS_{BIN} and around 90% for GHS_{SUB}, with the exception of a few substances the potency of which must be considered borderline or for which the specific data situation leads to different outcomes from the three overall WoE methods (MLLP, MSPE, WoE). More specifically, the latter applies to substances for which the overall reference classification was changed by including ambiguous test results in the overall WoE reference classification assessment. On the other hand, one should not draw too far-reaching conclusions from these findings, based on the evaluation of 14 - 34 substances only. A broader evaluation based on the larger database to be published by NICEATM and BfR will show whether these reproducibility numbers stay robust when a larger number of substances are evaluated. Nevertheless, to a certain degree, these results are found to build trust in the overall HPPT reference classifications, including those where only one test result is available.

Appendix 1 – Documentation of the database structure

Actual database field names (as of 3 March 2021) are printed in *ITALIC CAPITALS*.

Fields related to substance ID

NAMES Names/synonyms of the test compound

CASRN CAS registry number(s); if more than one: separated by ";", if substance contains more than one constituent: separated by "+"; if unclear, different possible CASRN are separated by "or"; if no CASRN was found: "na" (for "not available").

ECNR EC number(s); if more than one: separated by ";", if substance contains more than one constituent: separated by "+"; if unclear, different possible EC numbers are separated by "or"; if no CASRN was found: "na" (for "not available")

Fields related to generic test design parameters

TYPE Basic HPPT type: "HMT" or "HRIPT"

SUBTYPE HMT: "Kli66"; "Kli6675"; "Kli75"; "other" (i.e. none of the previous or not clearly assignable); HRIPT: "Shel53"; "Drai59"; "Grif69"; "MM73"; "MM80"; "JK77"; "RIFM08"; "other"

VOLUME Volume in µL of the test item (test substance + vehicle) as applied to the induction patch.

AREA Size of the induction patch in cm^2, i.e. the area of the patch used to bring the test item into contact with the skin.

Fields related to experiment-specific test design parameters

VEHICLE Vehicle used for application of the test substance to the skin

DENSITY Density of the vehicle (in mg/µL), estimated by using data from the ECHA registration database

N Number of test subjects

CONC Concentration of the test substance in the test vehicle in %

DSA Dose per skin area (in $\mu g/cm^2$) received by the test subjects for induction, calculated as *DSA* = *CONC*/100 x *VOLUME*/AREA x *DENSITY* x 1000 $\mu g/cm^2$

Fields related to test results

POSNR Number of test subjects with a positive test result; accepted as evaluated by the original authors, except when sensitisation was attributed to impurities or autoxidation products

POSINC Incidence of test subjects with positive test results in %, calculated as *POSINC = POSNR/N*

POSNEG Overall binary test outcome (positive = p, negative = n)

CONC1+ Hypothetical induction concentration (in %) resulting in exactly one test subject with a positive result (*POSNR* = 1) under the conditions of the test, calculated as:

CONC1+ = *CONC*/*POSNR*

i.e. by linear extrapolation from the actual number of positives observed in the test. As this can only be calculated for positive test results, it is “na” (not available”) for negative test results.

DSA1+ Hypothetical dose per skin area (in µg/cm^2) resulting in exactly one test subject with a positive test result (*POSNR* = 1) under the conditions of the test, calculated as:

DSA1+ = *DSA*/*POSNR*

i.e. by linear extrapolation from the actual number of positives observed in the test. As this can only be calculated for positive test results, it is “na” (not available”) for negative test results.

Administrative fields

RRS Relative reliability score (ranging from 1-5, for details, cf. section 3.3).

REMARKS Any remarks regarding the test result

Reference section

REF1 Original test report (primary reference)

AV1 Availability of the original test report (yes/no) to the HDSG

YR1 Year of the primary reference ("na", if unknown or unclear)

REF2 Reference(s) citing the original test report directly, in chronological order, by first author within a year. Individual references receive the prefix "#*a*#", where *a* marks the position in the list (i.e. the first 2nd order reference receives the prefix "#1#", the second "#2#", and so on.

REF3 Reference(s) citing the 2nd order reference(s) directly, ordered by 2nd order reference cited, in chronological order within those 3rd order references citing the same 2nd order reference, and then by first author within a year. Individual references receive the prefix "#*a*#b#", where *a* refers to the prefix of the 2nd order reference cited and *b* to the position in the list (i.e. the first 3rd order reference citing the first 2nd order reference receives the prefix "#1#1#", the second "#1#2#", and so on.

REF4 (REF5,…) Reference(s) citing the 3rd (4th, 5th…) order reference(s) directly, references are ordered and prefixes are assigned in analogy to the 2nd and 3rd order references. For example, the prefix "#2#2#1#" marks the first 4th order reference citing the second 3rd order reference of those which cite the second 2nd order reference.

Appendix 2 - Detailed documentation of the individual test results

Table 21: Overview of the available HPPT test results for the CosEU128 reference chemicals, sorted by CASRN, in ascending order[48]

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Abietic acid	514-10-3	208-178-3												
Acetanisole	100-06-1	202-815-9	HMT	Kli6675	pet	25	6	3888	0	0	na	NC/1B	3	*(Kligman, 1973f)*
2-Acetylcyclohexanone	874-23-7	212-858-5												
4-Allylanisole	140-67-0	205-427-8	HMT	Kli6675	pet	25	3	1944	0	0	na	NC/1B	3	*(Kligman, 1972e)*
Allyl phenoxyacetate	7493-74-5	231-335-2	HMT	Kli6675	pet	26	1	648	0	0	na	NC/1B	3	*(Kligman, 1974d)*
4-Aminobenzoic acid	150-13-0	205-753-0	HMT	Kli66	pet	23 24[51] 25	25	15517	0	0	na	NC	1	(Kligman, 1966c)
4-Amino-m-cresol	2835-99-6	220-621-2												
5-Amino-o-cresol	2835-95-2	220-618-6												
2-Aminophenol	95-55-6	202-431-1												
3-Aminophenol	591-27-5	209-711-2												
α-Amylcinnamaldehyde	122-40-7	204-541-5	HMT	Kli66	pet	25	6	3724	0	0	na	NC/1B	2	(Greif, 1967)
				Kli6675		71		3888					3	*(Letizia and Api, 2002)*

[48] Results in parentheses are only given as additional information. They have not been included in the overall classification calculation due to insufficient relative reliability (RRS = 5) and/or because of the overall outcome NC/1 (cf. section 7.1.1 for details).

[49] Borderline DSA1+ values (375 µg/cm2 ≤ DSA1+ < 625 µg/cm2) are printed red.

[50] References printed in italics were not directly available to the HDSG, these results were taken from secondary sources.

[51] Two test results

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
α-Amylcinnamaldehyde, ctd.	122-40-7	204-541-5	HRIPT	RIFM08	na	95	20	23622	0	0	na	NC/1B	4	*(RIFM, 2001c)*
α-Amylcinnamic alcohol	101-85-9	202-982-8	HMT	Kli6675	pet	25	8	5184	0	0	na	NC/1B	3	*(Kligman, 1973c)*
			HRIPT	RIFM08	DEP:EtOH 3:1	105	3	3543						*(RIFM, 2004e)*
				MM80	EtOH	78	8	5669	1[52]	1.3	5669	(1B)	5	(Marzulli and Maibach, 1980a)
Anethole	104-46-1	203-205-5												
Aniline	62-53-3	200-539-3	HMT	Kli66	pet	25	20	12414	7	28	1773	1B	1	(Kligman, 1966c)
Anisyl alcohol	105-13-5	203-273-6	HMT	Kli6675	pet	25	5	3240	0	0	na	NC/1B	3	*(Kligman, 1971f)*
2-(p-Anisyl)propanal	5462-06-6	226-749-5	HMT	Kli75	pet	26	2	1350	0	0	na	NC/1B	3	*(Epstein, 1980b)*
Applelide	478695-70-4	639-080-2												
BADGE	1675-54-3	216-823-5												
Bandrowski's base	20048-27-5	na												
Benzaldehyde	100-52-7	202-860-4	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1973e)*
														(Kligman, 1973a)
			HRIPT	RIFM08	DEP:EtOH 3:1	107	0.5	590						*(RIFM, 2009)*
						104	5	5900	12	11.5	492	1A-		*(RIFM, 2008)*
1,2-Benzisothiazol-3(2H)-one	2634-33-5	220-120-9												

[52] This positive result was obtained with a test material designated as "semi-pure" by the authors, but differences vs. the HMT result could also be explained by the different vehicle used or the lower number of test subjects, while the other HRIPT was performed at a lower concentration. It cannot be judged whether the purity issue is relevant (the authors apparently did not think so) and therefore the Relative Reliability Score (RRS) was set to 5.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Benzocaine	94-09-7	202-303-5	HMT	Kli66	pet	23	5	3103	0	0[53]	na	NC/1B	1	(Kligman, 1966b)
						22			1	4.5	3103	1B		
						25			3	12[54]	1035			
									6	24[55]	517	1B+		
						24	25	15517	3	12.5	5172	1B		(Kligman, 1966c)
					pet				5	20.8	3103			
						23				21.7				
						22				22.7				
			HRIPT	Shel53		200	2	1395	0	0	na	NC/1B		(Kligman, 1966a)
				Drai59										
				MM73		92							2	(Marzulli and Maibach, 1974)
						173	10	6975	2	1.2	3488	1B		
						99	20	13950	6	6.1	2325			
Benzoic acid	65-85-0	200-618-2	HMT	Kli75	pet	25	2	1350	0	0	na	NC/1B	3	*(Kligman, 1977e)*
					hydrophilic pet	10	5	3750					2	(Leyden and Kligman, 1977b)
			HRIPT	other	na	50	na	na				(NC/1B)	5	(Gad et al., 1986)
p-Benzoquinone	106-51-4	203-405-2												

[53] Only three exposures (compared to five in the standard HMT design)

[54] Ten exposures (compared to five in the standard HMT design)

[55] Fifteen exposures (compared to five in the standard HMT design)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Benzyl alcohol	100-51-6	202-859-9	HMT	Kli6675	pet	25	10	6480	0	0	na	NC/1B	3	*(Kligman, 1970c)*
			HRIPT	RIFM08	DEP:EtOH 3:1	107	3	3543						*(RIFM, 2004h)*
						110	7.5	8858	3	0.9 - 2.7	2953 - 8858	1B	4	*(RIFM, 2004i)*[56]
						46	15	17717	5	8.7 - 10.9	3543 - 4430			*(RIFM, 2003b)*[57]
						56	20	23622		1.8 - 8.9	4724 - 23622			*(RIFM, 2002b)*[58]
Benzyl benzoate	120-51-4	204-402-9	HMT	Kli6675	pet	25	30	19440	0	0	na	NC	3	*(Kligman, 1970a)*
Benzyl bromide	100-39-0	202-847-3												
Benzyl butyl phthalate	85-68-7	201-622-7												
Benzyl cinnamate	103-41-3	203-109-3	HMT	Kli6675	pet	25	8	5184	0	0	na	NC/1B	3	*(Kligman, 1972a)*
				Kli75				5400						*(RIFM, 1975g)*
			HRIPT	RIFM08	DEP:EtOH 3:1	101	4	4724						*(RIFM, 2005c)*

[56] Three positive reactions were counted in RIFM assessments published by December 2019 (Scognamiglio et al., 2012b), which was the deadline for including study results in this report. According to Na et al. (2020), only one positive reaction should be counted. The original report was not available to the HDSG, therefore it was not possible to resolve this issue here. The lower bound of the DSA1+ range was used for median calculation.

[57] Five positive reactions were counted in RIFM assessments published by December 2019 (Scognamiglio et al., 2012b), which was the deadline for including study results in this report. According to Na et al. (2020), only four positive reactions should be counted here. The original report was not available to the HDSG, therefore it was not possible to resolve this issue. The lower bound of the DSA1+ range was used for median calculation.

[58] Five positive reactions were counted in RIFM assessments published by December 2019 (Scognamiglio et al., 2012b), which was the deadline for including study results in this report. According to Na et al. (2020), only one positive reaction should be counted. The original report was not available to the HDSG, therefore it was not possible to resolve this issue here. The lower bound of the DSA1+ range was used for median calculation.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Benzyl salicylate	118-58-1	204-262-9	HMT	Kli75	pet	25	20	13500	1	4	13500	1B	3	*(RIFM, 1979b)*
									2	8	6750			*(RIFM, 1980h)*
				Kli6675		25	30	19440	0	0	na	NC		*(Kligman, 1970c)*
				Kli75		22		20250						*(RIFM, 1975h)*
						25								*(RIFM, 1975g)*
			HRIPT	other	DMP	52	5	na				(NC/1)	4	*(RIFM, 1968b)*
					SDA 39C	35	10	6200				NC/1B	3	*(RIFM, 1975a)*
				RIFM08	DEP:EtOH 3:1	101	15	17717						*(RIFM, 2004j)*
Benzylidene acetone	122-57-6	204-555-1	HMT	Kli6675	pet	25	2	1296	12	48	108	1A	3	*(Kligman, 1972b)*
			HRIPT	MM80		62	3	2126	6	9.7	354		2	(Marzulli and Maibach, 1980a)
BGE	2426-08-6	219-376-4	HMT	Kli66	pet	24	10	6207	19	79.2	327	1A	1	(Kligman, 1966c)
Bis-GMA	1565-94-2	216-367-7												
Bourgeonal	18127-01-0	242-016-2	HMT	Kli75	pet	29	6	4050	0	0	na	NC/1B	3	*(Epstein, 1980a)*
1-Bromobutane	109-65-9	203-691-9												
1-Bromohexane	111-25-1	203-850-2												
Bromothalonil	35691-65-7	252-681-0	HRIPT	na	corn oil	52	0.3	na	0	0	na	(NC/1)	4	(Mathias, 1983)
Butan-1-ol	71-36-3	200-751-6												
2-Butoxyethyl acetate	112-07-2	203-933-3												
Butyl acrylate	141-32-2	205-480-7												

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
L-Carvone	6485-40-1	229-352-5	HMT	Kli6675	pet	25	1	648	0	0	na	NC/1B	3	*(Kligman, 1971f)*
			HRIPT	RIFM08	DEP:EtOH 3:1	93	16	18896	4	4.3	4724	1B	3	*RIFM (2008), unpublished*[59]
CD-3	25646-71-3	247-161-5												
Chloramine-T	127-65-1	204-854-7												
3-Chloro-p-anisaldehyde	4903-09-7	225-532-2												
Chlorobenzene	108-90-7	203-628-5												
Chlorothalonil	1897-45-6	217-588-1												
Chlorpromazine	50-53-3	200-045-8	HMT	Kli66	pet	24	25	15517	18	75	862	1B	1	(Kligman, 1966c)
Cinnamaldehyde	104-55-2	203-213-9	HMT	Kli75	pet	25	0.5[60]	338	1	4.0	338	1A	3	*(Kligman, 1977d)*
									2	8.0	169			
				Kli6675			2	1296	11	44.0	118			*(Kligman, 1973d)*
							3	1944	3	12.0	648	1B		*(Kligman, 1974e)*
			HRIPT	other	EtOH	41	0.125	77	0	0	na	(NC/1)		*(RIFM, 1964d)*
				RIFM08	EtOH:DEP 3:1[61]	19	0.5	591				NC/1B		*(RIFM, 2002c)*
					DEP:EtOH 3:1[61]	22								*(RIFM, 2002c)*
				na	EtOH	38		349				(NC/1)		*(RIFM, 1965b)*

[59] Published in Na et al. (2020) after the literature search for this report was closed (December 2019)

[60] Four tests were performed at this dose level with 25 test subjects each, varying the number of induction exposures: 5, 10, and 15 (one subject sensitized in each), and 20 exposures (2 sensitized).

[61] α-Tocopherol was added for stabilisation.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Cinnamaldehyde, ctd.	104-55-2	203-213-9	HRIPT	RIFM08	EtOH:DEP 3:1[62]	94	0.5	591	0	0	na	NC/1B	3	*(RIFM, 2004b)*
				MM73	pet	53	1	698	0[63]				2	(Marzulli and Maibach, 1976b)
					95% EtOH	55		620	1	1.8	620	1B+		
					pet	107		698	2[64]	1.9	348	1A		
					95% EtOH			620	3	2.8	207			
				other	SDA 39C	41		620	5	12.2	124		3	*(RIFM, 1973b)*
					EtOH	10	1.25	775		50.0	155			*(RIFM, 1964d)*
				RIFM08	DEP:EtOH 3:1	28	3	3543	4	14.3	886	1B		*(RIFM, 2003c)*
Cinnamic alcohol	104-54-1	203-212-3	HMT	Kli75	pet	25 - 30	4	2400	1	3.3 - 4.0	2400	1B	3	(Jordan Jr. and King, 1977)
				Kli66		25		2483	0	0	na	NC/1B	2	(Greif, 1967)
				Kli75				2700					3	*(RIFM, 1979b)*
						24	10	6750	1	4.2	6750	1B		*(RIFM, 1977b)*
						11				9.1				*(RIFM, 1976a)*
						25			3	12.0	2250			*(RIFM, 1977b)*
														(RIFM, 1975e)
									5	20.0	1350			*(RIFM, 1976c)*

[62] α-Tocopherol was added for stabilisation.

[63] Result after first challenge

[64] Result after rechallenge

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Cinnamic alcohol, ctd.	104-54-1	203-212-3	HMT	Kli75	pet	25	10	6750	5	20.0	1350	1B	3	*(RIFM, 1977b)*
									7	28.0	964			*(RIFM, 1976b)*
									9	36.0	750			*(RIFM, 1976c)*
						33			10	30.3	675			*(RIFM, 1977c)*
					DEP	24		8250	0	0	na	NC/1B		*(RIFM, 1980d)*
						35								*(RIFM, 1980e)*
						21								*(RIFM, 1980g)*
						28			1	3.6	8250	1B		*(RIFM, 1980f)*
						23			2	8.7	4125			*(RIFM, 1981c)*
						22				9.1				*(RIFM, 1981a)*
														(RIFM, 1980e)
						27			3	11.1	2750			*(RIFM, 1982b; RIFM, 1982c)*
						28			4	14.3	2063			*(RIFM, 1982c)*
									6	21.4	1375			*(RIFM, 1979b)*
						26				23.1				*(RIFM, 1979c)*
			HRIPT	Drai59	EtOH:DEP 3:1	54	4	4724	2	3.7	2362			*(RIFM, 2001a; RIFM, 2002a)*
				JK77	pet	150		4500	0	0	na	NC/1B		(Jordan Jr. and King, 1977)
					EtOH			4000	4	2.7	1000	1B		
				na	DMP	54	6	na	2	3.7	na	POS	4	(Steltenkamp et al., 1980a)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Cinnamonitrile	1885-38-7	217-552-5	HMT	Kli75	pet	25	4	2700	0	0	na	NC/1B	3	*(Kligman, 1975b)*
			HRIPT	na	EtOH	41	2.5	1550	2	4.9	775	1B		*RIFM (1965), unpublished*[65]
					SDA 39C	54		1000	1	1.9	1000			*RIFM (1980), unpublished*[65]
Citral	5392-40-5	226-394-6	HMT	Kli6675	pet	25	0.1	65	1	4.0	65	1A	3	*(Kligman, 1972c; Kligman, 1972d; Kligman, 1972f; Kligman, 1972g)*[66]
									3	12.0	22			
							0.5	324	2	8.0	162			
						24	2	1296		8.3	648	1B		
			HRIPT	other	EtOH	82	0.5	na	0	0	na	(NC/1)	4	*(Steltenkamp et al., 1980b)*
					SDA 39C	40	1	775				NC/1B	3	*(RIFM, 1965a)*
					EtOH	84	1	na	2	2.4		POS	4	*(Steltenkamp et al., 1980b)*
						41	2		4	9.8				
				Drai59	na	50	4	3100	0	0		NC/1B		*(Shelanski, 1971)*
						10			5	50.0	620	1B+		*(Majors, 1971b)*
				other		49	5	na	16	32.7	na	POS		*(Steltenkamp et al., 1980b)*
						105	4 - 8[67]		na	> 0				*(Blau and Kanof, 1971)*

[65] The EG DASS decided to include these results, although they were not published by December 2019. They are from the commercial RIFM database.

[66] From Opdyke (1979), the RIFM monograph citing these reports, it was not clear which result came from which report.

[67] "8%, later reduced to 4%" (Opdyke, 1979)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Citral, ctd.	5392-40-5	226-394-6	HRIPT	Drai59	na	40	4 - 8[68]	3100 - 6200	19	47.5	163-326	1A		*(Majors, 1971a; Majors, 1971c)*
						56	8	6200	6	10.7	1033	1B		*(Maibach, 1971a; Maibach, 1971b)*
Citronellol	106-22-9	203-375-0	HMT	Kli66	pet	25	6	3724	0	0	na	NC/1B	2	(Greif, 1967)
			HRIPT	RIFM08	DEP:EtOH 3:1	101	25	29528				NC	3	*(RIFM, 2005d)*
Clofibrate	637-07-0	211-277-4												
Coumarin	91-64-5	202-086-7	HMT	Kli66	pet	25	8	4966	0	0	na	NC/1B	2	(Greif, 1967)
			HRIPT	MM80	EtOH	73		5039	1	1.4	5039	1B		(Marzulli and Maibach, 1980a)
						104		5669	2	1.9	2835			
Cyclamen aldehyde	103-95-7	203-161-7	HMT	Kli6675	pet	25	3	1944	0	0	na	NC/1B	3	*(Kligman, 1971g)*
			HRIPT	other	SDA 39C	64	4	4724						*(RIFM, 1980c)*
trans-Dec-2-enal	3913-81-3	223-474-2	HMT	Kli75	pet	25	4	2700	0	0	na	NC/1B	3	*(Kligman, 1977f)*
DEET	134-62-3	205-149-7												
Diacetyl	431-03-8	207-069-8	HMT	Kli75	pet	29	2	1350	0	0	na	NC/1B	3	*(Epstein, 1976a)*
2,5-Diaminotoluene sulfate	615-50-9	210-431-8												
Dibenzoyl peroxide	94-36-0	202-327-6	HRIPT	Gri69	1% sulfur in PEG	69	10	14177	25	36.2	567	1B+	3	(Poole et al., 1970)
Dibenzyl ether	103-50-4	203-118-2	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1974a)*

[68] "8%, later reduced to 4%" (Opdyke, 1979)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Dibutyl phthalate	84-74-2	201-557-4												
N,N-Dibutylaniline	613-29-6	210-335-6												
1,1-Dichloroethene	75-35-4	200-864-0												
Diethyl maleate	141-05-9	205-451-9	HMT	Kli75	pet	25	4	2700	25	100	108	1A	3	*(Kligman, 1975e)*
			HRIPT	MM80	pet	52	2	1417	2	3.8	709	1B		(Marzulli and Maibach, 1980b)
					EtOH			1260	8	15.4	157	1A		
					pet	187	4	2835	14	7.5	202		2	*(Marzulli and Maibach, 1980a)*
Diethyl phthalate	84-66-2	201-550-6	HMT	Kli66	pet	25	10	6207	0	0	na	NC/1B	2	(Greif, 1967)
Diethyl sulfate	64-67-5	200-589-6												
Diethylenetriamine	111-40-0	203-865-4	HMT	Kli66	pet	25	10	6207	21	84	296	1A	1	*(Kligman, 1966c)*
3,4-Dihydrocoumarin	119-84-6	204-354-9	HMT	Kli6675	pet	25	20	12960	25	100	518	1B+	3	*(Kligman, 1972i)*
			HRIPT	MM80		62		14173	32	51.6	443	1A-	2	(Marzulli and Maibach, 1980a)
Dihydroeugenol	2785-87-7	220-499-0	HMT	Kli75	pet	25	8	5400	0	0	na	NC/1B	3	*(Kligman, 1977a)*
Dihydromyrcenol	18479-58-8	242-362-4	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1973f)*
Dimethyl fumarate	624-49-7	210-849-0												
3-(Dimethylamino)propylamine	109-55-7	203-680-9												
Diphenylcyclopropenone	886-38-4	212-948-4												
Disodium 2-[4-[[2-cyano-3-[4-[methyl(2-sulfonatoethyl)amino]-phenyl]-1-oxoallyl]amino]phenyl]-6-methylbenzothiazole-7-sulfonate	2498-95-5	219-694-3												

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
DMSO	67-68-5	200-664-3	HMT	Kli66	pet	23	75	46552	0	0	na	NC	1	(Kligman, 1966c)
DNBS, sodium salt	885-62-1	212-943-7												
DNCB	97-00-7	202-551-4	HRIPT	other	acetone	24	0.0625	7	2	8.3	3.5	1A	1	(Friedmann et al., 1983)
						40	0.125	14	25	62.5	0.6			
						22			20	90.9	0.7			(Rees et al., 1989)
						30	0.25	28	25	83.3	1.1			(Friedmann et al., 1983)
							0.5	57	30	100.0	1.9			
						8	1	113	8		14.1			
EGDMA	97-90-5	200-617-2												
Ethyl acrylate	140-88-5	205-438-8	HMT	Kli6675	pet	24	4	2592	10	41.7	259	1A	3	*(Epstein, 1974a)*
			HRIPT	MM80		27		2835	0	0	na	NC/1B	2	(Marzulli and Maibach, 1980a)
						28								
						70			4	5.7	709	1B		
					EtOH	78			6	7.7	420	1A-		
Ethyl benzoylacetate	94-02-0	202-295-3												
Ethyl vanillin	121-32-4	204-464-7	HMT	Kli6675	pet	25	2	1296	0	0	na	NC/1B	3	*(Kligman, 1970b)*
2-Ethylbutanal	97-96-1	202-623-5												
Ethylene brassylate	105-95-3	203-347-8	HMT	Kli6675	pet	25	30	19440	0	0	na	NC	3	*(Kligman, 1973g)*
			HRIPT	RIFM08	DEP	99	10	12292				NC/1B		*(RIFM, 1993)*
						22								*(RIFM, 1994a)*
						95								*(RIFM, 1994c)*

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Ethylene brassylate, ctd.	105-95-3	203-347-8	HRIPT	RIFM08	DEP	91	10	12292	0	0	na	NC/1B	3	*(RIFM, 1994b)*
						109								*(RIFM, 1995c)*
						125								*(RIFM, 1995b)*
						65								*(RIFM, 1995a)*
						107			2[69]	1.9	6496	1B		*(RIFM, 1991a)*
						197	20	21260	0	0	na	NC/1B		*(RIFM, 1987)*
						58								*(RIFM, 1988c)*
						36								*(RIFM, 1989a)*
						71								*(RIFM, 1990e)*
						64								*(RIFM, 1990f)*
						50								*(RIFM, 1990c)*
						48								*(RIFM, 1990b)*
						34								*(RIFM, 1988e)*
						28								*(RIFM, 1988d)*
						38								*(RIFM, 1988f)*
						108								*(RIFM, 1989f)*
						93								*(RIFM, 1990a)*
						103								*(RIFM, 1989b)*

[69] Published RIFM assessments (Belsito et al., 2011; McGinty et al., 2011) of ethylene brassylatereport this test result as positive. The HDSG was informed that RIFM now suspects that these two reactions could have been caused by a so-called "spill-over" effect, i.e. a reaction to another test substance applied on another patch nearby. However, this could not be verified by the HDSG and therefore the test result is maintained as previously reported in the published literature.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Ethylene brassylate, ctd.	105-95-3	203-347-8	HRIPT	RIFM08	DEP	109	20	21260	0	0	na	NC/1B	3	*(RIFM, 1988b)*
						106								*(RIFM, 1988a)*
					EtOH:DEP 3:1	67			2[70]	3	10630	1B		*(RIFM, 1990d)*
Ethylenediamine (free base)[71]	107-15-3	203-468-6												
2-Ethylhexyl acrylate	103-11-7	203-080-7												
Eugenol	97-53-0	202-589-1	HMT	Kli66	pet	25	8	4966	0	0	na	NC/1B	2	(Greif, 1967)
			HRIPT	MM80		104		5669						(Marzulli and Maibach, 1980a)
					EtOH	73		5039	2	2.7	2520	1B		
				RIFM08	DEP:EtOH 3:1	108	5	5905	0	0	na	NC/1B	3	*(RIFM, 2001b)*
Farnesal	502-67-0	242-957-9												
Farnesol	4602-84-0	225-004-1	HMT	Kli75	pet	25	10	6750	0	0	na	NC/1B	3	*(RIFM, 1976b)*
									4	16	1688	1B		*(RIFM, 1977b)*
									6	24	1125			*(RIFM, 1976b)*
							12	8100	0	0	na	NC/1B		*(RIFM, 1977d)*
														(RIFM, 1975b)
														(RIFM, 1975c)

[70] Published RIFM assessments (Belsito et al., 2011; McGinty et al., 2011) for ethylene brassylate report this test result as positive. The HDSG was informed that RIFM now suspects that one of the two reactions could have been caused by a so-called "spill-over" effect, i.e. a reaction to another test substance applied on another patch nearby. According to RIFM, the other reaction was considered "transient" only. However, this could not be verified by the HDSG and therefore the test result is maintained as previously reported in the published literature.

[71] The database contains a „POS“ HRIPT result for ethylenediamine dihydrochloride (CASRN 333-18-6, EC no. 206-369-6).

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Farnesol, ctd.	4602-84-0	225-004-1	HMT	Kli75	pet	26	12	8100	0	0	na	NC/1B	3	*(RIFM, 1977d)*
						35								*(RIFM, 1978c)*
						25			2	8.0	4050	1B		*(RIFM, 1975d)*
				Kli6675				7776	4		1944			*(RIFM, 1974)*
				Kli75				8100	7	28.0	1157			*(RIFM, 1975f)*
			HRIPT	other		101	5	1529	0	0	na	NC/1B		*(RIFM, 2000b)*
						103								*(RIFM, 2000a)*
					DEP:EtOH 3:1	108		2865						*(RIFM, 2004c)*
2-Fluoro-5-nitroaniline	369-36-8	206-720-3												
Formaldehyde[72]	50-00-0	200-001-8	HMT	Kli66	pet	25	1.85	1148	18	72	64	1A	1	(Kligman, 1966c)
			HRIPT	MM73	water	45	0.037	29	0	0	na	(NC/1)	2	(Marzulli and Maibach, 1974)
						89	0.37	287	4	4.5	72	1A		
						88	1.11	860	5	5.7	172			
						52	1.85	1434	4	7.7	358	1A	1	(Marzulli and Maibach, 1973)
						102	3.7	2868	8	7.8				
Furil	492-94-4	207-766-7												

[72] Formaline (37.5% aqueous formaldehyde) was applied in these experiments. The concentrations given in the table, however refer to the actual formaldehyde concentrations (i.e. 0.375 x the formaline concentration).

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Geraniol	106-24-1	203-377-1	HMT	Kli6675	pet	25	5	3240	20	80.0	162	1A	3	*(Opdyke, 1974b)*
				Kli66			6	3724	0	0	na	NC/1B	2	(Greif, 1967)
				Kli75		24		4050					3	*(RIFM, 1979b)*
						26			1	3.8	4050	1B		*(RIFM, 1979e)*
			HRIPT	RIFM08	DEP:EtOH 3:1	110	2	2362	0	0	na	NC/1B		*(RIFM, 2000c)*
				other	SDA 39C	40	5	3100						*(RIFM, 1964c)*
				MM80	pet	104	10	7087					2	(Marzulli and Maibach, 1980b)
					EtOH	73		6299	2	2.7	3150	1B		
				Drai59		41	12.5	7750	0	0	na	NC/1B	3	*(RIFM, 1964b)*
Glutaraldehyde	111-30-8	203-856-5	HRIPT	MM73	pet	102	0.1	70	0	0	na	(NC/1)	2	(Marzulli and Maibach, 1974)
						30	5	3488	7	23.3	498	1A-		
Glycerol	56-81-5	200-289-5	HRIPT	other	na	50	na	na	0	0	na	(NC/1)	5	(Gad et al., 1986)
Glyoxal	107-22-2	203-474-9	HMT	Kli66	pet	24	10	6207	24	100	259	1A	1	(Kligman, 1966c)
Helional	1205-17-0	214-881-6												
Hepta-2,4-dienal	5910-85-0	227-627-4	HMT	Kli75	pet	22	1	675	0	0	na	NC/1B	3	*(Epstein, 1980b)*
n-Hexane	110-54-3	203-777-6	HMT	Kli66	none	25	100	62069	0	0	na	NC	1	(Kligman, 1966c)
trans-Hex-2-enal	6728-26-3	229-778-1	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1973a)*
			HRIPT	RIFM08	EtOH:DEP 3:1		0.2	236	6	24	39	1A	2	*RIFM (1990), unpublished*[73]

[73] This result was referenced in the literature published by December 2019, but essential details were lacking. The EG DASS received these details via private communication from RIFM.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Hexyl salicylate	6259-76-3	228-408-6	HMT	Kli75	pet	22	3	2025	0	0	na	NC/1B	3	*(Epstein, 1975f)*
			HRIPT	RIFM08	DEP:EtOH 3:1	103	30	35433				NC		*(RIFM, 2004k)*
α-Hexylcinnamaldehyde	101-86-0	202-983-3	HMT	Kli6675	pet	25	12	7776	0	0	na	NC/1B	3	*(Kligman, 1973a)*
				na		81						(NC/1B)	5	(Letizia and Api, 2002)
			HRIPT	RIFM08	na	324	5-20	5906-23622						
					DEP	91	20	23622				NC/1B	3	*(RIFM, 1994d)*
2-Hexylidenecyclopentanone	17373-89-6	241-411-7	HMT	Kli75	pet	25	5	3375	3	12.0	1125	1B	3	*(RIFM, 1979d)*
						24			8	33.3	422	1A-		*(RIFM, 1981b)*
			HRIPT	other	DEP:EtOH 3:1	102	0.6	300[74]	0	0	na	(NC/1)		*(RIFM, 2005b)*
					SDA 39C	51	1	111	4	7.8	28	1A		*(RIFM, 1982a)*
HHPA	85-42-7	201-604-9												
Hydratropaldehyde	93-53-8	202-255-5	HMT	Kli6675	pet	25	2	1296	0	0	na	NC/1B	3	*(Kligman, 1971a)*
			HRIPT	other	EtOH	7	2.5	1550	1	14.3	1550	1B	3	*(Hill Top Research Institute, 1991)*
Hydrocortisone	50-23-7	200-020-1	HMT	Kli66	pet	25	25	15517	0	0	na	NC	1	(Kligman, 1966c)
Hydroquinone	123-31-9	204-617-8												

[74] The DSA values are reproduced as provided in Scognamiglio et al. (2012a), the secondary source citing these studies.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
4-Hydroxybenzoic acid	99-96-7	202-804-9												
Hydroxycitronellal	107-75-5	203-518-7	HMT	Kli6675	pet	25	5	3240	0	0	na	NC/1B	3	*(Kligman, 1973h)*
				Kli75		26		3375						*(Epstein, 1976c)*
						25	10	6750						*(Kligman, 1976b)*
									2	8	3375	1B		
				Kli66			12	7448	0	0	na	NC/1B	2	(Greif, 1967)
				Kli75				8100					3	*(Kligman, 1978c)*
					DEP	26		9900	1	3.8	9900	1B		*(Epstein, 1980c)*
					pet			8100	4	14.8	2025			*(Epstein, 1978e)*
									6	23.1	1350			*(Epstein, 1979c)*
						26			7	26.9	1157			*(Epstein, 1979c)*
			HRIPT	RIFM08	EtOH:DEP 3:1	65	2.5	2657	0	0	na	NC/1B	2	(Ford et al., 1988)
				JK77	pet	150	4	4500	1	0.7	4500	1B	3	(Jordan Jr. and King, 1977)
				other	EtOH	39	5	3100	0	0	na	NC/1B		(Steltenkamp et al., 1980c)
				RIFM08	EtOH:DEP 3:1	66	5	5315	1	1.5	5315	1B	2	(Ford et al., 1988)
				MM73	EtOH	38	7.5	4650		2.6	4650		3	(Steltenkamp et al., 1980c)
				RIFM08	EtOH:DEP 3:1	66		7972		1.5	7972		2	(Ford et al., 1988)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Hydroxycitronellal, ctd.	107-75-5	203-518-7	HRIPT	other	EtOH	40	10	6200	6	15	1033	1B	3	(Steltenkamp et al., 1980c)
				MM80	pet	99	20	14173	1	1	14713			(Marzulli and Maibach, 1980b)
					EtOH	73		12598	14	19.2	900			
2-Hydroxyethyl acrylate	818-61-1	212-454-9												
2-Hydroxypropyl methacrylate	923-26-2	213-090-3												
Imidazolidinyl urea	39236-46-9	254-375-6	HMT	other	water	25 - 30	2	2500	0	0	na	NC/1B	3	(Jordan Jr. and King, 1977)
			HRIPT	JK77		150			2	1.3	1250	1B		
Iodocarb	55406-53-6	259-627-5												
1-Iodohexane	638-45-9	211-339-0												
Isobergamate	68683-20-5	272-066-0	HMT	Kli75	pet	25	1	675	0	0	na	NC/1B	3	*(Kligman, 1979a; Kligman, 1979b)*[75]
							1.5	1013						*(Dragoco, 1979a; Dragoco, 1979b; Dragoco, 1979c; Dragoco, 1979d; Dragoco, 1979e)*[76]
							10	6750	1	4.0	6750	1B		*(Epstein, 1978a)*
									3	12.0	2250			*(Kligman, 1978b)*
						29			4	13.8	1688			*(Epstein, 1978d)*

[75] Results from two tests with 25 test subjects each

[76] From the secondary source citing this result, it was not clear in which of the cited primary reports it was contained.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Isobornyl acetate	125-12-2	204-727-6	HMT	Kli6675	pet	25	10	6480	0	0	na	NC/1B	3	*(Kligman, 1970a)*
p-Isobutyl-α-methyl-hydrocinnamaldehdye	6658-48-6	229-695-0												
2-Isobutyl-4-methyl-tetrahydro-2H-pyran-4-ol	63500-71-0	613-238-0	HRIPT	RIFM08	DEP: EtOH 3:1	110	3.7	4408	0	0	na	NC/1B	3	*(RIFM, 2004d)*
				na	na	57	8	na					4	*(RIFM, 1985)*[77]
Isoeugenol	97-54-1	202-590-7	HMT	Kli6675	pet	25	8	5184	0	0	na	NC/1B	3	*(Kligman, 1971g)*
			HRIPT	na	SDA 39C	53	0.5	250				(NC/1)	3	*(RIFM, 1980a)*
				MM80	EtOH	73	8	5039	9	12.3	560	1B⁺	2	(Marzulli and Maibach, 1980a)
α-Isomethylionone	127-51-5	204-846-3	HRIPT	other	DMP	52	2	1860	0	0	na	NC/1B	3	*(RIFM, 1968a)*
					EtOH	28	10	6200						*(RIFM, 1964e)*
					na	37	12.5	9688					4	*(RIFM, 1964a)*
				RIFM08	DER:EtOH 3:1	23	60	70866					3	*(RIFM, 2004g)*
						106								*(RIFM, 2004f)*
Isopropanol	67-63-0	200-661-7												
Isopropyl myristate	110-27-0	203-751-4	HMT	Kli6675	pet	25	20	12960	0	0	na	NC/1B	3	*(Kligman, 1974a)*
Kanamycin	59-01-8	200-411-7	HMT	Kli66	pet	24	25	15517	11	45.8	1411	1B	3	(Kligman, 1966c)

[77] The DSA is not known, but the test concentration was 8% and for a RIFM-conducted test in the mid-1980s this would have corresponded to a DSA > 500 µg/cm².

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Kathon CG	55965-84-9	611-341-5; 911-418-6	HRIPT	other	na	184	0.00075	na	1	0.5	na	1A	4	*(Rohm and Haas, 1990b)*
					water	175	0.001	0.8-1.4	0	0		(NC/1)	1	*(Cardin et al., 1986)*
						109	0.0015	na					4	*(Curtis, 1989)*
					na	189	0.0015		> 0	> 0	na	1A		*(Rohm and Haas, 1990a)*
					water	45	0.002	2.9	2	4.4	1.5		1	*(Cardin et al., 1986)*
						216	0.005 - 0.01	na	40	18.5	0.00013 - 0.00025		4	*(CTFA, 1989)*
					1.: pet 2.: water	196	1.: 0.015 2.: 0.03		7	3.6	0.002-0.004			*(Maibach, 1980)*
				MM80	pet or 5% aqueous Tween 85	96	0.05	39	0	0	na	(NC/1)	4	*(Maibach, 1985)*
						104	0.1	79	2	1.9	39	1A		
Lactic acid	50-21-5	200-018-0												
Lauryl gallate	1166-52-5	214-620-6												
Lilial	80-54-6	201-289-8	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1972e)*
							5	3240						*(Kligman, 1971e)*
			HRIPT	na	EtOH/DEP	225	25	29528	1	0.4	29528	1B		(Cocchiara and Api, 2003)
D-Limonene	5989-27-5	227-813-5	HMT	Kli66	pet	25	8	4966	0	0	na	NC/1B	2	(Greif, 1967)
Linalool	78-70-6	201-134-4	HMT	Kli66	pet	25	8	4966	0	0	na	NC/1B	2	(Greif, 1967)
				Kli6675				5184					3	*(Kligman, 1972h)*

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Linalool, ctd.	78-70-6	201-134-4	HMT	Kli6675	pet	25	20	12960	0	0	na	NC/1B	3	*(Kligman, 1970c)*
			HRIPT	RIFM08	DEP:EtOH 3:1	119	12.7	15000						*(RIFM, 2005e)*
Maleic anhydride	108-31-6	203-571-6												
2-Mercaptobenzothiazole	149-30-4	205-736-8	HMT	other	pet	22	10	6207	0	0	na	NC/1B	1	(Kligman, 1966b)
						24			5	20.8	1241	1B		
				Kli66			25	15517	9	37.5	1724			(Kligman, 1966c)
2-Methoxy-p-cresol	93-51-6	202-252-9												
1-(p-Methoxyphenyl)pent-1-en-3-one	104-27-8	203-190-5	HMT	Kli75	pet	24	8	5400	16	66.7	338	1A	3	*(Kligman, 1973f)*
Methyl acrylate	96-33-3	202-500-6												
Methyl 3-bromoproprionate	3395-91-3	222-247-5												
Methyl dihydrojasmonate	24851-98-7	246-495-9	HMT	Kli75	pet	25	20	13500	0	0	na	NC/1B	3	*(Kligman, 1976a)*
			HRIPT	na	SDA 39C	23	2.42	1500						*(RIFM, 1971a)*
							10	6200						*(RIFM, 1971b)*
					DEP:EtOH 3:1	111	20	10000						*(RIFM, 2005a)*
					DEP	100		11000						*(RIFM, 2003a)*
Methyl methacrylate	80-62-6	201-297-1	HRIPT	MM80	na	184	10	7874	0	0	na	NC/1B	4	(Marzulli and Maguire, 1983)
Methyl methanesulfonate	66-27-3	200-625-0												
Methyl non-2-ynoate	111-80-8	203-909-2	HMT	Kli6675	pet	25	2	1296	0	0	na	NC/1B	3	*(Kligman, 1973i)*
				Kli75				1350						*(Kligman, 1975f)*

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Methyl non-2-ynoate, ctd.	111-80-8	203-909-2	HRIPT	na	EtOH:DEP 3:1	34	na	24	0	0	na	(NC/1)	4	*(RIFM, 1989g)*
						66							3	*(RIFM, 1990i)*
						138		118	6	5.1	20	1A	4	*(RIFM, 1989c; RIFM, 1989e)*
Methyl oct-2-ynoate	111-12-6	203-836-6	HMT	Kli6675	pet	25	2	1296	0	0	na	NC/1B	3	*(Epstein, 1974b)*
				Kli75		21		1350						*(Epstein, 1975a; Epstein, 1975b; Epstein, 1975c; Epstein, 1975d; Epstein, 1975e)[78]*
						24			1	4.2	1350	1B		
				Kli6675		25		1296	3	12	422	1A-		*(Kligman, 1974b; Kligman, 1974c)*
				Kli75				1350	4	16	338	1A		*(Epstein, 1975a; Epstein, 1975b; Epstein, 1975c; Epstein, 1975d; Epstein, 1975e)[78]*
						24				16.7				
						26	2		5	19.2	270			*(Epstein, 1976b)*
						22				22.7				*(Epstein, 1975a; Epstein, 1975b; Epstein, 1975c; Epstein, 1975d; Epstein, 1975e)[57]*
				Kli6675		25		1296	7	28	185			*(Kligman, 1972d)*
				Kli75				1350			193			*(Kligman, 1975d)*
														(Epstein, 1976b)

[78] From the secondary source citing these reports, it was not clear which result came from which report.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Methyl oct-2-ynoate, ctd.	111-12-6	203-836-6	HMT	Kli75	pet	25	2	1350	8	32	169	1A	3	*(Kligman, 1977c)*
									10	40	135			*(Kligman, 1975a)*
			HRIPT	other	EtOH:DEP 3:1	33	0.1	118	0	0	na	(NC/1)	3	*(RIFM, 1989d; RIFM, 1989e)*
						71								*(RIFM, 1990g; RIFM, 1990h)*
					EtOH	53	0.0002	na					4	*(IFRA, 1978)*
						10	0.001							
						40	0.05							
						42	0.1							
						41	0.25		2	4.9	na	POS		
Methyl pyruvate	600-22-6	209-987-4												
Methyl o-toluate	89-71-4	201-932-2												
1-(3-Methyl-2-benzo-furanyl)ethanone	23911-56-0	429-100-6												
α-Methylcinnamaldehyde	101-39-3	202-938-8	HMT	Kli6675	pet	25	8	5184	0	0	na	NC/1B	3	*(Kligman, 1973b)*
6-Methylcoumarin	92-48-8	202-158-8	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1973f)*
2-Methyldecanenitrile	69300-15-8	273-960-3												
6-Methylhepta-3,5-dien-2-one	1604-28-0	216-507-7	HRIPT	RIFM08	na	48	2	2754	6	12.5	459	1A-	4	*RIFM (1986), unpublished*[79]

[79] The EG DASS decided to include this result, although it was not published by December 2019. It is from the commercial RIFM database.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
5-Methylhexane-2,3-dione	13706-86-0	237-241-8	HMT	Kli75	pet	25	5	3375	2	8	1688	1B	3	*(Kligman, 1978a)*
						29			1	3.4	3375			*(Epstein, 1979a)*
Methylisothiazolinone[80]	2682-20-4	220-239-6	HRIPT	other	water	98	0.01	na	0	0	na	(NC/1)	4	*(Rohm and Haas, 2000a)*
						100	0.02	10					3	*(Rohm and Haas, 2000b)*
						98	0.03	15						*(Rohm and Haas, 2000c)*
						116	0.04	20	1	0.9	20	1A	3	*(Rohm and Haas, 2000d)*
						210	0.05	25		0.5	25			*(Rohm and Haas, 2000e)*
						214	0.06	30	0	0	na	(NC/1)		*(Rohm and Haas, 2000f)*
4-Methyl-2-nitroanisole	119-10-8	204-296-4												
Methylparaben	99-76-3	202-785-7												
2-Methylundecanal	110-41-8	203-765-0	HMT	Kli6675	pet	25	4	2592	0	0	na	NC/1B	3	*(Kligman, 1971d)*
Metol	55-55-0	200-237-1												
1-Naphthol	90-15-3	201-969-4												

[80] In Cosmetic Ingredient Review Expert Panel (1992), numerous HRIPT results are reported, however, all of them were generated with a mixture of methylisothiazolinone and methylchloroisothiazolinone.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Neomycin sulfate	1405-10-3	215-773-1	HMT	Kli66	pet	24	10	6207	0	0	na	NC/1B	1	(Kligman, 1966b)
						25			1	4.0	6207	1B		
						23			4	17.4	1552[81]			
						24			5	20.8	1241			
						21			10	47.6	621	1B+		
				other		22	25	15500	0	0	na	NC		
						25		15491	5	20.0	3098	1B		
				Kli66		24		15517		20.8	3103			
				other		23			7	30.4	2217			
								15491	8	34.8	1936			
						25		15517	11	44.0	1411			
			HRIPT	MM73		54	0.5	349	0	0	na	(NC/1)		(Marzulli and Maibach, 1973)
									3	5.6	116	1A		
						186	5	3488		1.6	1163	1B		
									5	2.7	698			
						42	20	13950	0	0	na	NC/1B		
									7	16.7	1993	1B		
						25			12	48.0	1163			

[81] Two test results with 23 test subjects each, with slightly different test designs

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
4-Nitrobenzylbromide	100-11-8	202-820-6												
2-Nitro-p-phenylenediamine	5307-14-2	226-164-5												
cis-6-Nonenal	2277-19-2	218-900-9	HMT	Kli75	pet	29	1	675	0	0	na	NC/1B	3	*(Epstein, 1978d)*
Octanoic acid	124-07-2	204-677-5	HMT	Kli75	pet	25	1	675	0	0	na	NC/1B	3	*(Kligman, 1977b)*
OTNE	54464-57-2	259-174-3	HRIPT	other	SDA 39C	44	2.5	1240	0	0	na	NC/1B	3	*(RIFM, 1977a)*
						36		1550						*(RIFM, 1973a)*
				na		42								*(RIFM, 1978a; RIFM, 1978b)*
				other		51	12.5	1378						*(RIFM, 1979a)*
						28	12.5	10000						*(RIFM, 1980b)*[82]
					SDA 39C: DEP 3:1	53	22.5	11160						*(RIFM, 1999)*
				RIFM08	DEP:EtOH 3:1	101	40	47250				NC		*(RIFM, 2004l)*
Oxalic acid	144-62-7	205-634-3												
Oxazolone	15646-46-5	239-713-9												
Penicillin G	61-33-6	200-506-3	HMT	Kli66	pet	23	0.1	62	0	0	na	(NC/1)	1	(Kligman, 1966b)
						22			2	9.1	31	1A		
						24			6	25	10			
						22	0.2	124	2	9.1	62			

[82] Only six induction exposures (2/wk, over 6 wk) and using a reduced number of test subjects (1 M, 27 F)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Penicillin G, ctd.	61-33-6	200-506-3	HMT	Kli66	pet	22	1	621	4	18.2	155	1A	1	(Kligman, 1966b)
						25	5	3103	7	28	443	1A-		
							10	6207	1	4	6207	1B		
									5	20	1241			
									9	36	690			
						21			10	47.6	621	1B+		
						24			11	44	564			
				other		25				45.8				
				Kli66		21			16	76.2	388	1A-		
				other		24	25	15517	10	41.7	1552	1B		(Kligman, 1966c)
				Kli66		25			13	52	1194			
						23			12	52.2	1293			(Kligman, 1966b)
				other		25			16	64	970			
				Kli66		24				66.7				(Kligman, 1966c)
						23				69.6				
			HRIPT	Shel53		200	1	692	0	0	na	NC/1B		(Kligman, 1966a)
				Drai59				698						
Pentachlorophenol	87-86-5	201-778-6												

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Perillaldehyde	2111-75-3	218-302-8	HMT	Kli75	pet	25	1	675	0	0	na	NC/1B	3	*(Epstein, 1979b)*
						29	4	2700	2	6.9	1350	1B		*(Epstein, 1978b; Epstein, 1978c)*[83]
						25			3	12.0	900			
			HRIPT	RIFM08	DEP:EtOH 3:1	116	0.6	709	0	0	na	NC/1B		*(RIFM, 2007)*
3-Phenoxypropanenitrile	3055-86-5	221-278-1												
Phenyl benzoate	93-99-2	202-293-2	HRIPT	other	DEP	107	8	10393	1	0.9	10393	1B	3	*(RIFM, 1991b)*
Phenylacetaldehyde	122-78-1	204-574-5	HMT	Kli6675	pet	25	2	1296	2	8.0	648	1B	3	*(Maibach, 1971b)*
									4	16.0	324	1A		*(Kligman, 1971c; Kligman, 1971e)*
									11	44.0	118			
						23			12	52.2	108			*(Epstein, 1973a; Epstein, 1973b)*
			HRIPT	Drai59	na	56		1550	0	0	na	NC/1B	4	*(Maibach, 1971b)*
				other		50								*(Shelanski, 1971)*[84]
				na	EtOH:DEP 3:1	110	0.5	531					3	*(RIFM, 2004a)*

[83] From the secondary source citing these reports, it was not clear which result came from which report.

[84] This test used 15 inductions.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
p-Phenylenediamine	106-50-3	203-404-7	HMT	Kli66	pet	24	0.1	62	5	20.8	12	1A	1	(Kligman, 1966b)
							0.2	124	7	29.2	18			
									8	33.3	16			
						23			10	43.5	12			
						25	1	621	17	68.0	37			
						24	5	3103	22	91.7	141			
						25			24	96.0	129			
						24	10	6207		100	259			(Kligman, 1966c)
						20	25	15517	20		776	1B		
			HRIPT	MM73		97	0.01	7	7	7.2	1	1A	2	(Marzulli and Maibach, 1974)
						98	0.1	70	11	11.2	6			
						88	1	698	47	53.4	15			
1-Phenylpropane-1,2-dione	579-07-7	209-435-2												
Phthalic anhydride	85-44-9	201-607-5												
Propyl gallate	121-79-9	204-498-2												
Propylene glycol	57-55-6	200-338-0	HMT	Kli66	pet	24	25	15517	0	0	na	NC	1	(Kligman, 1966c)
			HRIPT	MM73	cream	204	12	9300				NC/1B		(Marzulli and Maibach, 1973)
					pet	88	60	41850				NC	2	(Marzulli and Maibach, 1974)

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
3-Propylidenephthalide	17369-59-4	241-402-8	HMT	Kli75	pet	25	4	2700	3	12.0	900	1B	3	*(Kligman, 1975c)*
Propylparaben	94-13-3	202-307-7												
Pyridine	110-86-1	203-809-9	HMT	Kli66	pet	24	50	31034	1	4.2	31034	1B	1	(Kligman, 1966c)
Resorcinol	108-46-3	203-585-2	HMT	Kli66	pet	22	15	9310	0	0	na	NC/1B	1	(Kligman, 1966c)
Saccharin	81-07-2	201-321-0												
Safranal[85]	116-26-7	204-133-7	HRIPT	other	na	54	0.05	25	0	0	na	(NC/1)	4	*(RIFM, 1998)*
					DMP	53	0.5	300	5	9.4	60	1A	4	*(RIFM, 1996)*
Salicylic acid	69-72-7	200-712-3	HMT	Kli66	pet	25	20	12414	0	0	na	NC/1B	1	(Kligman, 1966c)
Sodium lauryl sulfate	151-21-3	205-788-1	HMT	Kli66	pet	22	10	6207	0	0	na	NC[86]	1	(Kligman, 1966c)
Squaric acid	2892-51-5	220-761-4												
Sulfanilamide	63-74-1	200-563-4	HMT	Kli66	pet	25	25	15517	5	20	3103	1B	1	(Kligman, 1966c)
Sulfanilic acid	121-57-3	204-482-5												
L-Tartaric acid	87-69-4	201-766-0												
Tetrachlorosalicylanilide	1154-59-2	214-576-8	HMT	other	pet	23	0.05	31	3	13	10	1A	1	(Kligman, 1966b)
						22			6	27.3	5			
						23			13	56.5	2			
						25	0.1	62	6	24	10			

[85] These test results, but not their DSA values, were published by December 2019. DSA values were provided to the HDSG by RIFM via private communication.

[86] Based on expert knowledge the EG DASS decided that 10% sodium lauryl sulfate was the highest concentration testable without severe irritation hampering the outcome of the experiment. Hence, this result was considered a valid negative test result.

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Tetrachlorosalicylanilide, ctd.	1154-59-2	214-576-8	HMT	other	pet	24	1	621	17	70.8	37	1A	1	(Kligman, 1966b)
							5	3103	21	87.5	148			
						25			22	88.0	141			(Kligman, 1966c)
						23				95.7				
						25			24	96.0	129			
						24				100				
						21	10	6207	21		296			(Kligman, 1966b)
			HRIPT	Drai59		200	0.2	140	0	0	na	(NC/1)		(Kligman, 1966a)
				Shel53					9	4.5	16	1A		
2,2,6,6-Tetramethyl-heptane-3,5-dione	1118-71-4	214-268-3												
Thioglycerol	96-27-5	202-495-0	HMT	other	pet	23	20	12414	9	39.1	1379	1B	1	(Kligman, 1966b)
						24			10	41.7	1241			
									11	45.8	1129			
							50	31034	24	100	1293			(Kligman, 1966c)
			HRIPT	Voss58	na	52	1.23	1453	5-6	10 - 11.5	242-291	1A	4	(Voss, 1958)[87]
Thiram	137-26-8	205-286-2	HMT	Kli66	pet	25	10	6207	0[88]	0	na	NC/1B	1	(Kligman, 1966b)
						22			2	9.1	3103	1B		

[87] The lower DSA value was used for median calculation.

[88] Two individual test results with 25 test subjects each, with slightly different test design details

Substance	CASRN	EC no.	Test type	Subtype	Vehicle	No. of subjects tested	Induction conc. (%)	DSA (µg/cm²)	Number of positive reactions	Incidence of positive reactions (%)	DSA1+ (µg/cm²)[49]	"Extrapolated" classification	RRS	Primary ref.[50]
Thiram, ctd.	137-26-8	205-286-2	HMT	Kli66	pet	18	10	6207	6	33.3	1034	1B	1	(Kligman, 1966b)
						23	25	15517	1	4.0	15517			(Kligman, 1966c)
						25				4.3				
						25			4	16.0	3879			
						22				18.2				
α-Tocopherol	59-02-9	200-412-2												
Triethanolamine	102-71-6	203-049-8												
Trimellitic anhydride	552-30-7	209-008-0												
Tropolone	533-75-5	208-577-2												
Undec-10-enal	112-45-8	203-973-1	HMT	Kli6675	pet	25	1	648	0	0	na	NC/1B	3	*(Kligman, 1971b)*
Vanillin	121-33-5	204-465-2	HMT	Kli66	pet	25	2	1241	0	0	na	NC/1B	2	(Greif, 1967)
				Kli6675			5	3240					3	*(Kligman, 1970c)*
4-Vinylpyridine	100-43-6	202-852-0												

Abbreviations

BfR	Bundesinstitut für Risikobewertung (German Federal Institute for Risk Assessment)
CASRN	Chemical Abstracts Service Registry Number
CosEU128	List of 128 reference chemicals originally compiled by Cosmetics Europe which formed the original basis of the reference substance dataset for DA performance standards in this OECD project
CPSC	United States Consumer Product Safety Commission
DA	Defined approach (to testing and assessment)
DEP	Diethyl phthalate
DMP	Dimethyl phthalate
DMSO	Dimethyl sulfoxide
DSA	Dose applied per skin area (in µg/cm^2)
EtOH	Ethanol
FDA	United States Food and Drug Administration
GHS	(United Nations) Globally Harmonized System (of classification and labelling of chemicals)
GHS_{BIN}	Binary GHS classification mode: 1 (sensitiser) or NC (not classified)
GHS_{SUB}	Ternary GHS classification mode: 1A (strong/extreme sensitiser), 1B (moderate/weak sensitiser), or NC (not classified)
GHS_{BORDER}	Ternary GHS classification mode with additional ambiguous classification outcomes: 1A (strong/extreme sensitiser), 1 (sensitiser, sub-categorisation not possible), 1B (moderate/weak sensitiser), NC/1B (not possible to decide whether substance is a sensitiser or not, but 1A can be ruled out), or NC (not classified)
GPMT	Guinea pig maximisation test
HDPT	Human diagnostic patch test
HDSG	Human data sub-group
HMT	Human maximization test
HPPT	Human predictive patch test
HRIPT	Human repeated insult patch test
ID	Identity
LLNA	Local lymph node assay
LSG	LLNA sub-group (of the EG DASS)

MLLP	Median-like location parameter
MSPE	Median sensitisation potency estimate
na	not available
NC	not classified
NICEATM	United States National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
OECD	Organization for Economic Co-operation and Development
OECD EG DASS	OECD Expert Group for Defined Approaches for Skin Sensitisation
PEG	Polyethylene glycol
pet	Petrolatum
REF1	Primary reference, original test report
RIFM	Research Institute for Fragrance Materials
RRS	Relative reliability score
SDA	Specially denatured alcohol (SDA 39C: 99% ethanol + 1% DEP)
SPE	Sensitisation potency estimate
TEWL	Trans-epidermal water loss
UN	United Nations
UVCB	(Substances of) unknown or variable composition or biological origin

References

Api A.M. et al. (2017): Fragrances categorized according to relative human skin sensitization potency. Dermatitis 28 (5), 299-307. DOI: 10.1097/der.0000000000000304

Bangsgaard N. et al. (2011): Impaired hapten sensitization in patients with autoimmune disease. Clinical and experimental immunology 165 (3), 310-317. DOI: 10.1111/j.1365-2249.2011.04428.x

Basketter D.A. et al. (2014): Categorization of chemicals according to their relative human skin sensitizing potency. Dermatitis 25 (1), 11-21. DOI: 10.1097/der.0000000000000003

Basketter D.A. et al. (1999): Threshold for classification as a skin sensitizer in the local lymph node assay: A statistical evaluation. Food and Chemical Toxicology 37 (12), 1167-1174. DOI: 10.1016/S0278-6915(99)00112-X

Belsito D. et al. (2011): A toxicological and dermatological assessment of macrocyclic lactone and lactide derivatives when used as fragrance ingredients. Food and Chemical Toxicology 49, S219-S241. DOI: 10.1016/j.fct.2011.07.052

Blau S. and Kanof N. (1971): Report to RIFM, 7 May, unpublished

Cardin C.W. et al. (1986): Dose-response assessments of Kathon® biocide (II) Threshold prophetic patch testing. Contact Dermatitis 15 (1), 10-16. DOI: doi:10.1111/j.1600-0536.1986.tb01254.x

Cecchi L. et al. (2018): External exposome and allergic respiratory and skin diseases. Journal of Allergy and Clinical Immunology 141 (3), 846-857. DOI: 10.1016/j.jaci.2018.01.016 (last accessed 2019/08/20)

Cocchiara J. and Api A.M. (2003): A dermal safety evaluation of p-(t-butyl)-α-methylhydrocinnamic aldehyde (BMHCA). Toxicologist 72 (S1), 301

Cosmetic Ingredient Review Expert Panel (1992): Final report on the safety assessment of methylisothiazolinone and methylchloroisothiazolinone. Journal of the American College of Toxicology 11 (1), 75-128. DOI: 10.3109/10915819209141993

CTFA (1989): Summary of repeat insult patch test data on 2335 subjects with 28 formulations containing 7.5 ppm Kathon-CG, unpublished

Curtis H. (1989): Repeat insult patch tests of hand and body lotions, date: 1989-11-16, unpublished

Dimitrov S. et al. (2016): Accounting for data variability, a key factor in in vivo/in vitro relationships: Application to the skin sensitization potency (in vivo LLNA versus in vitro DPRA) example. Journal of Applied Toxicology 36 (12), 1568-1578. DOI: 10.1002/jat.3318

Dragoco (1979a): Private communication to RIFM, 1 June, unpublished

Dragoco (1979b): Private communication to RIFM, 1 March, unpublished

Dragoco (1979c): Private communication to RIFM, 1 May, unpublished

Dragoco (1979d): Private communication to RIFM, 8 May, unpublished

Dragoco (1979e): Private communication to RIFM, 23 April, unpublished

Draize J.H. (1959): Dermal toxicity. In: Appraisal of the safety of chemicals in foods, drugs and cosmetics, chapter 6, pp. 46-59. The Association of Food & Drug Officials of the United States, Austin, Texas, USA. https://babel.hathitrust.org/cgi/pt?id=uc1.b3596550;view=1up;seq=5 (last accessed 2021-07-10)

Dumont C. et al. (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. Toxicology in Vitro 34, 220-228. DOI: 10.1016/j.tiv.2016.04.008

EAACI (2018): EAACI white paper on research, innovation and quality care. European Academy of Allergy and Clinical Immunology. http://webcast.eaaci.cyim.com/mediatheque/media.aspx?mediaId=60234&channel=8518 (last accessed 2021-07-10)

Epstein W.L. (1973a): Report to RIFM, 9 October, unpublished

Epstein W.L. (1973b): Report to RIFM, 29 June, unpublished

Epstein W.L. (1974a): Report to RIFM, 7 October, unpublished

Epstein W.L. (1974b): Report to RIFM, 23 July, unpublished

Epstein W.L. (1975a): Report to RIFM, 5 March, unpublished

Epstein W.L. (1975b): Report to RIFM, 11 July, unpublished

Epstein W.L. (1975c): Report to RIFM, 16 April, unpublished

Epstein W.L. (1975d): Report to RIFM, 22 December, unpublished

Epstein W.L. (1975e): Report to RIFM, 28 March, unpublished

Epstein W.L. (1975f): Report to RIFM, 31 January, unpublished

Epstein W.L. (1976a): Report to RIFM, 8 October, unpublished

Epstein W.L. (1976b): Report to RIFM, 23 July, unpublished

Epstein W.L. (1976c): Unpublished data

Epstein W.L. (1978a): Report to RIFM, 1 July, unpublished

Epstein W.L. (1978b): Report to RIFM, 21 November, unpublished

Epstein W.L. (1978c): Report to RIFM, 25 August, unpublished

Epstein W.L. (1978d): Report to RIFM, 28 April, unpublished

Epstein W.L. (1978e): Unpublished data

Epstein W.L. (1979a): Report to RIFM, 24 October, unpublished

Epstein W.L. (1979b): Report to RIFM, 31 August, unpublished

Epstein W.L. (1979c): Unpublished data

Epstein W.L. (1980a): Report to RIFM, 25 June, unpublished

Epstein W.L. (1980b): Report to RIFM, 26 August, unpublished

Epstein W.L. (1980c): Unpublished data

Ford R.A. et al. (1988): Allergic contact sensitization potential of hydroxycitronellal in humans. Food and Chemical Toxicology 26 (11), 921-926. DOI: 10.1016/0278-6915(88)90090-7

Friedmann P.S. et al. (1983): Quantitative relationships between sensitizing dose of DNCB and reactivity in normal subjects. Clinical and experimental immunology 53 (3), 709-715

Gad S.C. et al. (1986): Development and validation of an alternative dermal sensitization test: The mouse ear swelling test (MEST). Toxicology and Applied Pharmacology 84 (1), 93-114. DOI: 10.1016/0041-008X(86)90419-9 (last accessed 2016-09-20)

Greif N. (1967): Cutaneous safety of fragrance material as measured by the maximization test. American Perfumer and Cosmetics 82, 54-57

Griem P. et al. (2003): Proposal for a risk assessment methodology for skin sensitization based on sensitization potency data. Regulatory Toxicology and Pharmacology 38 (3), 269-290. DOI: 10.1016/j.yrtph.2003.07.001

Griffith J.F. (1969): Predictive and diagnostic testing for contact sensitization. Toxicology and Applied Pharmacology 14, 90-102. DOI: 10.1016/S0041-008X(69)80014-1

Griffith J.F. and Buehler E.V. (1977): Prediction of skin irritancy and sensitizing potential by testing with animals and man. In: Cutaneous Toxicity (Drill L.A. and Lazar P., eds.), pp. 155-174. Academic Press, New York

Han H. et al. (2017): The atopic march: current insights into skin barrier dysfunction and epithelial cell-derived cytokines. Immunological reviews 278 (1), 116-130. DOI: 10.1111/imr.12546

Hill Top Research Institute (1991): Letter submitting five enclosed studies on alpha-methyl benzeneacetaldehyde with attachments (sanitized). NTIS/OTS0535053, unpublished

Hoffmann S. (2015): LLNA variability: An essential ingredient for a comprehensive assessment of non-animal skin sensitization test methods and strategies. Altex 32 (4), 379-383. DOI: 10.14573/altex.1505051

Hoffmann S. et al. (2018): Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database. Critical Reviews in Toxicology 48 (5), 344-358. DOI: 10.1080/10408444.2018.1429385

IFRA (1978): Unpublished communication. International Fragrance Association,

Jordan Jr. W.P. and King S.E. (1977): The development of allergic contact dermatitis in females during the comparison of two predictive patch tests. Contact Dermatitis 3 (1), 19-26. DOI: 10.1111/j.1600-0536.1977.tb03582.x

Kleinstreuer N.C. et al. (2018): Non-animal methods to predict skin sensitization (II): an assessment of defined approaches. Critical Reviews in Toxicology 48 (5), 359-374. DOI: 10.1080/10408444.2018.1429386

Kligman A.M. (1966a): The identification of contact allergens by human assay: I. A critique of standard methods. Journal of Investigative Dermatology 47 (5), 369-374. DOI: 10.1038/jid.1966.158

Kligman A.M. (1966b): The identification of contact allergens by human assay: II. Factors influencing the induction and measurement of allergic contact dermatitis. Journal of Investigative Dermatology 47 (5), 375-392. DOI: 10.1038/jid.1966.159

Kligman A.M. (1966c): The identification of contact allergens by human assay: III. The maximization test: A procedure for screening and rating contact sensitizers. Journal of Investigative Dermatology 47 (5), 393-409. DOI: 10.1038/jid.1966.160

Kligman A.M. (1970a): Report to RIFM, 1 June, unpublished

Kligman A.M. (1970b): Report to RIFM, 2 December, unpublished

Kligman A.M. (1970c): Report to RIFM, 7 October, unpublished

Kligman A.M. (1971a): Report to RIFM, 2 April, unpublished

Kligman A.M. (1971b): Report to RIFM, 3 November, unpublished

Kligman A.M. (1971c): Report to RIFM, 9 June, unpublished

Kligman A.M. (1971d): Report to RIFM, 17 June, unpublished

Kligman A.M. (1971e): Report to RIFM, 20 April, unpublished

Kligman A.M. (1971f): Report to RIFM, 24 May, unpublished

Kligman A.M. (1971g): Report to RIFM, 25 March, unpublished

Kligman A.M. (1972a): Report to RIFM, 1 June, unpublished

Kligman A.M. (1972b): Report to RIFM, 2 May, unpublished

Kligman A.M. (1972c): Report to RIFM, 13 October, unpublished

Kligman A.M. (1972d): Report to RIFM, 14 March, unpublished

Kligman A.M. (1972e): Report to RIFM, 18 February, unpublished

Kligman A.M. (1972f): Report to RIFM, 18 May, unpublished

Kligman A.M. (1972g): Report to RIFM, 19 October, unpublished

Kligman A.M. (1972h): Report to RIFM, 22 November, unpublished

Kligman A.M. (1972i): Report to RIFM, 25 August, unpublished

Kligman A.M. (1973a): Report to RIFM, 9 October, unpublished

Kligman A.M. (1973b): Report to RIFM, 10 December, unpublished

Kligman A.M. (1973c): Report to RIFM, 10 July, unpublished

Kligman A.M. (1973d): Report to RIFM, 10 October, unpublished

Kligman A.M. (1973e): Report to RIFM, 11 October, unpublished

Kligman A.M. (1973f): Report to RIFM, 13 August, unpublished

Kligman A.M. (1973g): Report to RIFM, 13 July, unpublished

Kligman A.M. (1973h): Report to RIFM, 13 June, unpublished

Kligman A.M. (1973i): Report to RIFM, 23 August, unpublished

Kligman A.M. (1974a): Report to RIFM, 4 June, unpublished

Kligman A.M. (1974b): Report to RIFM, 8 January, unpublished

Kligman A.M. (1974c): Report to RIFM, 18 October, unpublished

Kligman A.M. (1974d): Report to RIFM, 19 November, unpublished

Kligman A.M. (1974e): Report to RIFM, 22 August, unpublished

Kligman A.M. (1975a): Report to RIFM, 8 January, unpublished

Kligman A.M. (1975b): Report to RIFM, 14 February, unpublished

Kligman A.M. (1975c): Report to RIFM, 16 June, unpublished

Kligman A.M. (1975d): Report to RIFM, 18 January, unpublished

Kligman A.M. (1975e): Report to RIFM, 19 May, unpublished

Kligman A.M. (1975f): Report to RIFM, 27 March, unpublished

Kligman A.M. (1976a): Report to RIFM, 11 November, unpublished

Kligman A.M. (1976b): Unpublished data

Kligman A.M. (1977a): Report to RIFM, 3 November, unpublished

Kligman A.M. (1977b): Report to RIFM, 9 May, unpublished

Kligman A.M. (1977c): Report to RIFM, 14 February, unpublished

Kligman A.M. (1977d): Report to RIFM, 17 June, unpublished

Kligman A.M. (1977e): Report to RIFM, 24 May, unpublished

Kligman A.M. (1977f): Report to RIFM, 25 July, unpublished

Kligman A.M. (1978a): Report to RIFM, 11 December, unpublished

Kligman A.M. (1978b): Report to RIFM, 26 October, unpublished

Kligman A.M. (1978c): Unpublished data

Kligman A.M. (1979a): Report to RIFM, 7 December, unpublished

Kligman A.M. (1979b): Report to RIFM, 11 September, unpublished

Kligman A.M. and Epstein W. (1975): Updating the maximization test for identifying contact allergens. Contact Dermatitis 1 (4), 231-239. DOI: 10.1111/j.1600-0536.1975.tb05389.x

Klimisch H.J. et al. (1997): A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regulatory Toxicology and Pharmacology 25 (1), 1-5. DOI: 10.1006/rtph.1996.1076

Letizia C.S. and Api A.M. (2002): α-Amylcinnnamaldehyde (α-ACA) and α-hexylcinnamic aldehyde (α-HCA) do not produce dermal sensitization or cross-sensitization in humans. Toxicologist 66 (1-5), 163-164

Leyden J.J. and Kligman A.M. (1977a): Allergic contact dermatitis: Sex differences. Contact Dermatitis 3 (6), 333-336. DOI: 10.1111/j.1600-0536.1977.tb03698.x

Leyden J.J. and Kligman A.M. (1977b): Contact sensitization to benzoyl peroxide. Contact Dermatitis 3 (5), 273-275. DOI: doi:10.1111/j.1600-0536.1977.tb03674.x

Maibach H.I. (1971a): Report to RIFM, 7 July, unpublished

Maibach H.I. (1971b): Report to RIFM, September, unpublished

Maibach H.I. (1980): Modified Draize sensitization study - Kathon. Unpublished data. Received September 13, 1989

Maibach H.I. (1985): Diagnostic patch test concentration for Kathon CG. Contact Dermatitis 13 (4), 242-245. DOI: 10.1111/j.1600-0536.1985.tb02557.x

Majors P.A. (1971a): Report to RIFM, 13 April, unpublished

Majors P.A. (1971b): Report to RIFM, 21 March, unpublished

Majors P.A. (1971c): Report to RIFM, 23 August, unpublished

Marzulli F.N. and Maguire H.C. (1983): Validation of guinea pig tests for skin hypersensitivity. In: Dermatotoxicology (Marzulli F.N. and Maibach H.I., eds.), ed. 2, chapter 10, pp. 237-250. Hemisphere, New York

Marzulli F.N. and Maibach H.I. (1973): Antimicrobials: Experimental contact sensitization in man. Journal of the Society of Cosmetic Chemists 24 (7), 399-421

Marzulli F.N. and Maibach H.I. (1974): The use of graded concentrations in studying skin sensitizers: Experimental contact sensitization in man. Food and Cosmetics Toxicology 12 (2), 219-227. DOI: 10.1016/0015-6264(74)90367-8

Marzulli F.N. and Maibach H.I. (1976a): Contact allergy: Predictive testing in man. Contact Dermatitis 2 (1), 1-17. DOI: 10.1111/j.1600-0536.1976.tb02972.x

Marzulli F.N. and Maibach H.I. (1980a): Contact allergy: Predictive testing of fragrance ingredients in humans by Draize and Maximization methods. Journal of Environmental Pathology and Toxicology 3 (5-6), 235-245

Marzulli F.N. and Maibach H.I. (1980b): Further studies of effects of vehicles and elicitation concentration in experimental contact sensitization testing in humans. Contact Dermatitis 6 (2), 131-133. DOI: 10.1111/j.1600-0536.1980.tb03921.x

Marzulli F.N. and Maibach H.J. (1976b): Effects of vehicles and elicitation concentration in contact dermatitis testing I. Experimental contact sensitization in humans. Contact Dermatitis 2 (6), 325-329. DOI: 10.1111/j.1600-0536.1976.tb03069.x (last accessed 2019/04/26)

Mathias C.G.T. (1983): Contact dermatitis to a new biocide (Tektamer 38®) used in a paste glue formulation. Contact Dermatitis 9 (5), 418-418. DOI: 10.1111/j.1600-0536.1983.tb04440.x

McGinty D. et al. (2011): Fragrance material review on ethylene brassylate. Food and Chemical Toxicology 49, S174-S182. DOI: https://doi.org/10.1016/j.fct.2011.07.023

McNamee P.M. et al. (2008): A review of critical factors in the conduct and interpretation of the human repeat insult patch test. Regulatory Toxicology and Pharmacology 52 (1), 24-34. DOI: 10.1016/j.yrtph.2007.10.019

Moggs J.G. et al. (2012): Regulation of allergic responses to chemicals and drugs: Possible roles of epigenetic mechanisms. Toxicological Sciences 130 (1), 60-69. DOI: 10.1093/toxsci/kfs207 (last accessed 8/20/2019)

Morrissey K. et al. (2008): Age and gender effects on contact sensitization and photoimmune suppression in young and middle-aged adults. Photodermatology, Photoimmunology & Photomedicine 24 (1), 46-48. DOI: 10.1111/j.1600-0781.2008.00325.x

Na M. et al. (2020): Fragrance Skin Sensitization Evaluation and Human Testing: 30-Year Experience Publish Ahead of Print. DOI: 10.1097/der.0000000000000684

OECD (2012): The adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins. Document no. ENV/JM/MONO(2016)29. Organization for Economic Co-Operation and Development, Paris, France. https://read.oecd.org/10.1787/9789264221444-en?format=pdf (last accessed 2021-07-10)

OECD (2016a): Guidance document on the reporting of defined approaches and individual information to be used within integrated approaches to testing and assessment (IATA) for skin sensitisation. Document no. ENV/JM/MONO(2016)29. Organization for Economic Co-Operation and Development, Paris, France. https://www.oecd-ilibrary.org/guidance-document-on-the-reporting-of-defined-approaches-and-individual-information-sources-to-be-used-within-integrated-approaches-to-testing-and-assessment-iata-for-skin-sensitisation_5jfrvgk27bmx.pdf?itemId=%2Fcontent%2Fpublication%2F9789264279285-en&mimeType=pdf (last accessed 2021-07-10)

OECD (2016b): Guidance document on the reporting of defined approaches to be used within integrated approaches to testing and assessment (IATA). Document no. ENV/JM/MONO(2016)28. Organization for Economic Co-Operation and Development, Paris, France. https://www.oecd-ilibrary.org/guidance-document-on-the-reporting-of-defined-approaches-to-be-used-within-integrated-approaches-to-testing-and-assessment_5jfxdqj2n99x.pdf?itemId=%2Fcontent%2Fpublication%2F9789264274822-en&mimeType=pdf (last accessed 2021-07-10)

Oosterhaven J.A.F. et al. (2019): European Surveillance System on Contact Allergies (ESSCA): Contact allergies in relation to body sites in patients with allergic contact dermatitis. Contact Dermatitis 80 (5), 263-272. DOI: 10.1111/cod.13192

Opdyke D.L.J. (1974a): Fragrance raw materials monographs: Vetiver acetate. Food and Cosmetics Toxicology 12 (7), 1011-1012. DOI: 10.1016/0015-6264(74)90229-6

Opdyke D.L.J. (1974b): Geraniol, personal communication

Opdyke D.L.J. (1979): Monographs on fragrance raw materials. Food and Cosmetics Toxicology 17 (3), 241-275. DOI: 10.1016/0015-6264(79)90288-8

Politano V.T. and Api A.M. (2008): The Research Institute for Fragrance Materials' human repeated insult patch test protocol. Regulatory Toxicology and Pharmacology 52 (1), 35-38. DOI: 10.1016/j.yrtph.2007.11.004

Poole R.L. et al. (1970): Experimental contact sensitization with benzoyl peroxide. Archives of Dermatology 102 (4), 400-404. https://jamanetwork.com/journals/jamadermatology/article-abstract/531831 (last accessed 2021-07-10)

Rees J.L. et al. (1989): Sex differences in susceptibility to development of contact hypersensitivity to dinitrochlorobenzene (DNCB). British Journal of Dermatology 120 (3), 371-374. DOI: doi:10.1111/j.1365-2133.1989.tb04162.x

RIFM (1964a): Repeated insult patch test in human subjects. Unpublished report from IFF Incorporated, 28 April. Report no. 12544. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1964b): Repeated insult patch test with geraniol. Unpublished report from IFF. RIFM report no. 14094. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1964c): Repeated insult patch test with geraniol. Unpublished report from IFF. 30 April. RIFM report no. 51135. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1964d): Repeated insult patch test. Unpublished report from IFF Incorporated. Report no. 12510, date: 1964-07-29/1964-11-25. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1964e): Sensitization study of α-iso-methylionone. Unpublished report from IFF Incorporated, 3 July. Report no. 50426. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1965a): Repeated insult patch test of citral in human subjects. Unpublished report from IFF Incorporated. Report no. 14577, date: 1965-10-01. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1965b): Repeated insult patch test. Unpublished report from IFF Incorporated. Report no. 12508, date: 1965-10-01. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1968a): Sensitization and irritation studies of α-iso-methylionone in human subjects. Unpublished report from Givaudan, 18 November. Report no. 15453, date: 1968-11-18. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1968b): Sensitization and irritation studies of fragrance materials in human subjects. Unpublished report from Givaudan. Report no. 15453, date: 1968-11-18. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1971a): Repeated insult patch test with methyl dihydrojasmonate. Report no. 51179, date: 1971-12-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1971b): Repeated insult patch test with methyl dihydrojasmonate Report no. 51178, date: 1971-12-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1973a): Repeated insult patch test in humans with OTNE. Unpublished report from IFF Incorporated, 26 September. Report no. 19978. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1973b): Repeated insult patch test. Unpublished report from IFF Incorporated. Report no. 12509, date: 1973-01-23. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1974): Report on human maximization studies, 12 September. RIFM report no. 1779. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975a): Repeated insult patch test of benzyl salicylate on human subjects. Unpublished report from IFF Incorporated. Report no. 24190, date: 1975-06-18. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975b): Report on human maximization studies, 9 April. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975c): Report on human maximization studies, 10 December. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975d): Report on human maximization studies, 12 December. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975e): Report on human maximization studies, 15 December. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975f): Report on human maximization studies, 15 January. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975g): Report on human maximization studies, 27 March. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1975h): Report on human maximization studies, 28 March. RIFM report no. 1799. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1976a): Report on human maximization studies, 1 November. RIFM report no. 1797. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1976b): Report on human maximization studies, 11 November. RIFM report no. 1797. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1976c): Report on human maximization studies, 12 July. RIFM report no. 1796. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1977a): Repeated insult patch test in humans with OTNE. Unpublished report from IFF Incorporated, 15 November. Report no. 19981. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1977b): Report on human maximization studies, 7 February. RIFM report no. 1702. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1977c): Report on human maximization studies, 7 October. RIFM report no. 1691. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1977d): Report on human maximization studies, 15 December. RIFM report no. 1691. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1978a): Repeated insult patch test in humans with OTNE. Unpublished report from IFF Incorporated, 15 August. Report no. 19984. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1978b): Repeated insult patch test in humans with OTNE. Unpublished report from IFF Incorporated, 29 December. Report no. 19984. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1978c): Report on human maximization studies, 2 July. RIFM report no. 1698. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1979a): Evaluation of potential hazards of fragrance materials by dermal contact in humans. Unpublished report from IFF Incorporated, 17 September. Report no. 19986. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1979b): Report on human maximization studies, 6 July. RIFM report no. 1697. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1979c): Report on human maximization studies, 13 July. RIFM report no. 1697. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1979d): Report on human maximization studies, 14 September. RIFM report no. 1697. RIFM report no. 1697. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1979e): Report on human maximization studies, 18 June. RIFM report no. 1697. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980a): Evaluation of potential hazards by dermal contact of methyl 2-nonynoate, isoeugenol and cinnamyl nitrile in human subjects. Unpublished report from IFF Incorporated. Report no. 1982. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980b): Photoallergic/phototoxic study in humans with fragrance materials. Unpublished report from IFF Incorporated, 16 December. Report no. 19989. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980c): Repeated insult patch test of 2-methyl-3-(-p-isopropylphenyl)propionaldehyde in human subjects. Unpublished report from IFF, Incorporated. Report no. 15036. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980d): Report on human maximization studies, 14 April. RIFM report no. 1790. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980e): Report on human maximization studies, 18 January. RIFM report no. 1790. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980f): Report on human maximization studies, 26 August. RIFM report no. 1790. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980g): Report on human maximization studies, 26 February. RIFM report no. 1790. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1980h): Sensitization studies with benzyl salicylate in human subjects. RIFM report no. 3305, date: 1980-01-21. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1981a): Report on human maximization studies, 10 July. RIFM report no. 1792. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1981b): Report on human maximization studies, 18 March. RIFM report no. 1792. RIFM report no. 1697. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1981c): Report on human maximization studies, 22 September. RIFM report no. 1792. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1982a): Repeated insult patch test of 2-hexylidene cyclopentanone in human subjects. RIFM report no. 15002, date: 2005-07-21. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1982b): Report on human maximization studies, 17 May. RIFM report no. 1643. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1982c): Report on human maximization studies, 28 June. RIFM report no. 1643. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1985): Repeated insult patch test with 2-isobutyl-4-methyltetrahydro-2H-pyran-4-ol (Florol) in human subjects. Unpublished report from Firmenich Incorporated. RIFM report no. 39890. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1987): Report on human repeated insult patch test. Report no. 7973, date: 1987-07-20. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1988a): Repeat insult patch test of ethylene brassylate in human subjects, December 16. Report no. 8881, date: 1988-12-16. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1988b): Repeat insult patch test of ethylene brassylate in human subjects. December 7. Report no. 8517, date: 1988-12-07. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1988c): Repeat insult patch test of ethylene brassylate, methyl 2-nonynoate, 6,7-dihydrogeraniol and cyclohexyl methyl pentanone in humans. Report no. 6063, date: 1988-12-13. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1988d): Repeated insult patch test in human subjects, January 29 Report no. 27674, date: 1988-01-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1988e): Repeated insult patch test in human subjects, January 29. Report no. 27673, date: 1988-01-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1988f): Repeated insult patch test in human subjects, March 16. Report no. 27675, date: 1988-03-16. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1989a): Human repeated insult patch test of ethylene brassylate, November 16. Report no. 12366, date: 1989-11-16. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1989b): Human repeated insult patch test of ethylene brassylate, October 31. Report no. 12359, date: 1989-10-31. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1989c): Human repeated insult patch test of methyl 2-octynoate and methyl 2-nonynoate RIFM Report no. 12369. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1989d): Human repeated insult patch test of methyl 2-octynoate and methyl 2-nonynoate RIFM report no. 12368, unpublished

RIFM (1989e): Human repeated insult patch test of methyl 2-octynoate and methyl 2-nonynoate. RIFM Report no. 12367. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1989f): Repeated insult patch test in human subjects. Report no. 9433, date: 1989-06-07. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1989g): Repeated insult patch test of methyl octine carbonate in human subjects. Report to RIFM. RIFM Report no. 27820. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990a): Human repeat insult patch test of ethylene brassylate. Report no. 12381, date: 1990-03-01. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990b): Repeat insult patch test of ethylene brassylate in human subjects, 26 November Report no. 14119, date: 1990-11-26. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990c): Repeat insult patch test of ethylene brassylate in human subjects, 26 November. Report no. 14121, date: 1990-11-26. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990d): Repeat insult patch test of ethylene brassylate in human subjects, 27 April Report no. 12455, date: 1990-04-27. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990e): Repeat insult patch test of ethylene brassylate in human subjects, 27 April. Report no. 12453, date: 1990-04-27. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990f): Repeat insult patch test of ethylene brassylate in human subjects, 27 April Report no. 12457, date: 1990-04-27. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990g): Repeated insult patch test of methyl 2-octynoate in human subjects. RIFM Report no. 12456. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990h): Repeated insult patch test of methyl 2-octynoate in human subjects. RIFM Report no. 12452. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1990i): Repeated insult patch test of methyl octine carbonate and t-2-hexenal in human subjects. Report to RIFM. RIFM Report no. 27822. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1991a): Repeated insult patch test of ethylene brassylate on human subjects. Report no. 33718, date: 1991-11-25. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1991b): Repeated insult patch test of fragrance materials in human subjects, 25 November. RIFM report no. 33718. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1993): Repeated insult patch test of ethylene brassylate on human subjects, 12 November. Report no. 25756, date: 1993-11-12. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1994a): Repeated insult patch test of ethylene brassylate on human subjects, 25 July. Report no. 25752, date: 1994-07-25. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1994b): Repeated insult patch test of ethylene brassylate on human subjects, 29 July Report no. 26501, date: 1994-07-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1994c): Repeated insult patch test of ethylene brassylate on human subjects, 29 July. Report no. 26499, date: 1994-07-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1994d): Repeated insult patch test on hexylcinnamaldehyde in humans. RIFM report no. 26500, date: 1994-07-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1995a): Repeated insult patch test of ethylene brassylate. Report no. 41232, date: 1995-12-22. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1995b): Repeated insult patch test of ethylene brassylate on human subjects, 20 June. Report no. 25754, date: 1995-06-20. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1995c): Repeated insult patch test of ethylene brassylate on human subjects, 21 June. Report no. 25750, date: 1995-06-21. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1996): Repeated insult patch test. Unpublished report from the Givaudan-Roure Corporation. Report no. 29893, date: 1996-07-03. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1998): Repeated insult patch test of 2,6,6-trimethylcyclohexa-I,3-dienyl methan al in human subjects): Unpublished report from the Givaudan-Roure Corporation. Report no. 32089, date: 1998-03-03. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (1999): Repeated insult patch test with OTNE. Unpublished report from IFF Incorporated, 26 April. Report no. 47085. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2000a): Repeated insult patch test of farnesol. Unpublished report from Dragoco Inc., 28 April. RIFM report no. 35518. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2000b): Repeated insult patch test of farnesol. Unpublished report from Dragoco Inc., 28 April. RIFM report no. 35519. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2000c): Repeated insult patch test of geraniol in humans. RIFM report no. 36679. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2001a): Clinical safety evaluation - repeated insult patch test of cinnamic alcohol. RIFM report no. 40696. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2001b): Repeated insult patch test of eugenol in human subjects. RIFM report no. 39081. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2001c): Repeated insult patch test on alpha-amylcinnamaldehyde in humans. RIFM report no. 26498, date: 2001-07-29. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2002a): Clinical safety evaluation - repeated insult patch test of cinnamic alcohol. RIFM report no. 40698. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2002b): Repeated insult patch test (RIPT) with benzyl alcohol. RIFM report no. 44247. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2002c): Repeated insult patch test of cinnamaldehyde. RIFM report no. 41692, date: 2003-08-27. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2003a): Repeated insult patch study of methyl dihydrojasmonate at 20.0% in diethyl phthalate (DEP). RIFM report no. 43009, date: 2003-06-16. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2003b): Repeated insult patch test (RIPT) with benzyl alcohol. RIFM report no. 44246. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2003c): Repeated insult patch test with cinnamaldehyde. RIFM report no. 43502, date: 2003-01-01. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004a): Human repeated insult patch test with phenylacetaldehyde (modified Draize procedure). RIFM report no. 45132, date: 2004-03-31. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004b): Repeated insult patch test of cinnamaldehyde. RIFM report no. 47158, date: 2004-11-22. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004c): Repeated insult patch test of farnesol. Unpublished report from Symrise Inc., 126 December. RIFM report no. 35519. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004d): Repeated insult patch test with 2-isobutyl-4-methyltetrahydro-2H-pyran-4-ol. Unpublished report from International Flavors and Fragrances. RIFM report no. 54089. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004e): Repeated insult patch test with α-amylcinnamic alcohol. RIFM report no. 46097, date: 2004-07-07. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004f): Repeated insult patch test with α-iso-methylionone, 10 March. RIFM report no. 47278. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004g): Repeated insult patch test with α-iso-methylionone, 10 March. RIFM report no. 47279. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004h): Repeated insult patch test with benzyl alcohol. RIFM report no. 47046. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004i): Repeated insult patch test with benzyl alcohol (modified Draize procedure). RIFM report no. 44247. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004j): Repeated insult patch test with fragrance materials. RIFM report no. 45129. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004k): Repeated insult patch test with fragrance materials, 3 May. RIFM report no. 45130. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2004l): Repeated insult patch test, 28 June. RIFM report no. 45124, date: 2004-11-22. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2005a): Repeated insult patch study of methyl dihydrojasmonate. RIFM report no. 49735, date: 2005-06-22. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2005b): Repeated insult patch test study of 2-hexylidene cyclopentanone at 0.6% in 75% diethyl phthalate (DEP)/25% ethanol. RIFM report no. 48712, date: 2005-03-17. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2005c): Repeated insult patch test with benzyl cinnamate. RIFM report no. 49109, date: 2005-06-23. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2005d): Repeated insult patch test with DL-citronellol. RIFM Report no. 47277, date: 2005-01-06. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2005e): Repeated insult patch test with linalool. RIFM Report no. 49469. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2007): Repeated insult patch test with p-mentha-1,8-dien-7-al. RIFM report no. 53802. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2008): Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. RIFM report no. 55663. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

RIFM (2009): Repeated insult patch test with benzaldehyde. RIFM report no. 57360. Research Institute for Fragrance Materials, Woodcliff Lake, NJ, USA, unpublished

Robinson M.K. (1999): Population differences in skin structure and physiology and the susceptibility to irritant and allergic contact dermatitis: Implications for skin safety testing and risk assessment. Contact Dermatitis 41 (2), 65-79. DOI: 10.1111/j.1600-0536.1999.tb06229.x

Rohm and Haas (1990a): A double-blind study to determine the topical contact sensitization potential of three test products. Study no. 743 RN 1289, date: 1990-05-17. International Research Services, Inc., unpublished

Rohm and Haas (1990b): Modified Draize skin sensitization study. Report no. HIM 89-R-R&H-D-1&2, unpublished

Rohm and Haas (2000a): A patch test to determine the skin irritation and sensitization propensities of Kordek TM 50C (study conducted at 100 ppm active ingredient). Report no. 99RC-0138, unpublished

Rohm and Haas (2000b): Repeat insult patch study with 2-methylisothiazolin-3-one at an aqueous concentration of 200 ppm active ingredient. Report no. 00RC-099A, unpublished

Rohm and Haas (2000c): Repeat insult patch study with 2-methylisothiazolin-3-one at an aqueous concentration of 300 ppm active ingredient. Report no. 00RC-099B, unpublished

Rohm and Haas (2000d): Repeat insult patch study with 2-methylisothiazolin-3-one at an aqueous concentration of 400 ppm active ingredient. Report no. 00RC-099C, unpublished

Rohm and Haas (2000e): Repeat insult patch study with 2-methylisothiazolin-3-one at an aqueous concentration of 500 ppm active ingredient. Report no. 00RC-099D, unpublished

Rohm and Haas (2000f): Repeat insult patch study with 2-methylisothiazolin-3-one at an aqueous concentration of 600 ppm active ingredient. Report no. 00RC-099E, unpublished

Scognamiglio J. et al. (2012a): Fragrance material review on 2-hexylidene cyclopentanone. Food and Chemical Toxicology 50, S631-S640. DOI: 10.1016/j.fct.2012.03.023

Scognamiglio J. et al. (2012b): Fragrance material review on benzyl alcohol. Food and Chemical Toxicology 50, S140-S160. DOI: https://doi.org/10.1016/j.fct.2011.10.013

Shelanski H.A. and Shelanski M.V. (1953): A new technique of human patch tests. Proceedings of the Scientific Section of the Toilet Goods Association 19, 46-49

Shelanski M.V. (1971): Report to RIFM, 30 August, unpublished

Steltenkamp R.J. et al. (1980a): Cinnamic alcohol: A survey of consumer patch-test sensitization. Food and Cosmetics Toxicology 18 (4), 419-424

Steltenkamp R.J. et al. (1980b): Citral: A survey of consumer patch-test sensitization. Food and Cosmetics Toxicology 18 (4), 413-417

Steltenkamp R.J. et al. (1980c): Hydroxycitronellal: A survey of consumer patch-test sensitization. Food and Cosmetics Toxicology 18 (4), 407-412

Stotts J. (1980): Planning, conduct and interpretation of human predictive sensitization patch tests. In: Current Concepts in Cutaneous Toxicology (Drill V.A. and Lazar P., eds.). Academic Press (Elsevier), St. Louis, MO, USA

Urbisch D. et al. (2015): Assessing skin sensitization hazard in mice and men using non-animal test methods. Regulatory Toxicology and Pharmacology 71 (2), 337-351. DOI: 10.1016/j.yrtph.2014.12.008

Voss J.G. (1958): Skin sensitization by mercaptans of low molecular weight. Journal of Investigative Dermatology 31 (5), 273-279. DOI: 10.1038/jid.1958.120

Waldorf D.S. et al. (1968): Impaired delayed hypersensitivity in an aging population: Association with antinuclear reactivity and rheumatoid factor. JAMA 203 (10), 831-834. DOI: 10.1001/jama.1968.03140100013003 (last accessed 8/13/2019)

Wesley N.O. and Maibach H.I. (2003): Racial (Ethnic) Differences in Skin Properties. American Journal of Clinical Dermatology 4 (12), 843-860. DOI: 10.2165/00128071-200304120-00004

WHO IPCS (2017): Guidance document for evaluating and expressing uncertainty in hazard characterization. Harmonization document 11. World Health Organization, International Programme for Chemical Safety, Geneva, Switzerland. http://apps.who.int/iris/bitstream/handle/10665/259858/9789241513548-eng.pdf (last accessed 2021-07-10)

Zhao C.Y. et al. (2017): A comparison study of clinician-rated atopic dermatitis outcome measures for intermediate- to dark-skinned patients. British Journal of Dermatology 176 (4), 985-992. DOI: 10.1111/bjd.15271