Food and Chemical Toxicology 159 (2022) 112715

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

The RIFM approach to evaluating Natural Complex Substances (NCS)

A.M. Api [a], D. Belsito [b], D. Botelho [a], M. Bruze [c], G.A. Burton Jr. [d], J. Buschmann [e], M. A. Cancellieri [a], M.L. Dagli [f], M. Date [a], W. Dekant [g], C. Deodhar [a], A.D. Fryer [h], L. Jones [a], K. Joshi [a], M. Kumar [a], A. Lapczynski [a], M. Lavelle [a], I. Lee [a], D.C. Liebler [i], H. Moustakas [a], M. Na [a], T.M. Penning [j], G. Ritacco [a], J. Romine [a], N. Sadekar [a], T.W. Schultz [k], D. Selechnik [a], F. Siddiqi [a], I.G. Sipes [l], G. Sullivan [a,*], Y. Thakkar [a], Y. Tokura [m]

[a] *Research Institute for Fragrance Materials, Inc., 50 Tice Boulevard, Woodcliff Lake, NJ, 07677, USA*
[b] *Member Expert Panel, Columbia University Medical Center, Department of Dermatology, 161 Fort Washington Ave., New York, NY, 10032, USA*
[c] *Member Expert Panel, Malmo University Hospital, Department of Occupational & Environmental Dermatology, Sodra Forstadsgatan 101, Entrance 47, Malmo, SE, 20502, Sweden*
[d] *Member Expert Panel, School of Natural Resources & Environment, University of Michigan, Dana Building G110, 440 Church St., Ann Arbor, MI, 58109, USA*
[e] *Member Expert Panel, Fraunhofer Institute for Toxicology and Experimental Medicine, Nikolai-Fuchs-Strasse 1, 30625, Hannover, Germany*
[f] *Member Expert Panel, University of Sao Paulo, School of Veterinary Medicine and Animal Science, Department of Pathology, Av. Prof. dr. Orlando Marques de Paiva, 87, Sao Paulo, CEP 05508-900, Brazil*
[g] *Member Expert Panel, University of Wuerzburg, Department of Toxicology, Versbacher Str. 9, 97078, Würzburg, Germany*
[h] *Member Expert Panel, Oregon Health Science University, 3181 SW Sam Jackson Park Rd., Portland, OR, 97239, USA*
[i] *Member Expert Panel, Vanderbilt University School of Medicine, Department of Biochemistry, Center in Molecular Toxicology, 638 Robinson Research Building, 2200 Pierce Avenue, Nashville, TN, 37232-0146, USA*
[j] *Member of Expert Panel, University of Pennsylvania, Perelman School of Medicine, Center of Excellence in Environmental Toxicology, 1316 Biomedical Research Building (BRB) II/III, 421 Curie Boulevard, Philadelphia, PA, 19104-3083, USA*
[k] *Member Expert Panel, The University of Tennessee, College of Veterinary Medicine, Department of Comparative Medicine, 2407 River Dr., Knoxville, TN, 37996- 4500, USA*
[l] *Member Expert Panel, Department of Pharmacology, University of Arizona, College of Medicine, 1501 North Campbell Avenue, P.O. Box 245050, Tucson, AZ, 85724-5050, USA*
[m] *Member Expert Panel, The Journal of Dermatological Science (JDS), Department of Dermatology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu, 431-3192, Japan*

ARTICLE INFO

Handling editor: Dr. Jose Luis Domingo

ABSTRACT

The Research Institute for Fragrance Materials, Inc. (RIFM) has evaluated safety data for fragrance materials for 55 years. The safety assessment of Natural Complex Substances (NCS) is similar to that of discrete fragrance materials; all of the same endpoints are evaluated. A series of decision trees, reflecting advances in risk assessment approaches of mixtures and toxicological methodologies, follows a tiered approach for each endpoint using a 4-step process with testing only as a last resort: 1) evaluate available data on NCS; 2) verify whether the Threshold of Toxicological Concern (TTC) can be applied; 3) verify whether the NCS risk assessment can be achieved on a component basis; and 4) determine whether data must be generated. Using *in silico* tools, RIFM examined NCS similarities based on the plant part, processing, and composition of materials across 81 plant families to address data gaps. Data generated from the Creme RIFM Aggregate Exposure Model for over 900 fragrance NCS demonstrate that dermal exposure is the primary route of human exposure for NCS fragrance uses. Over a third of materials are below the most conservative TTC limits. This process aims to provide a comprehensive Safety Assessment of NCS used as a fragrance ingredient.

* Corresponding author.
E-mail address: gsullivan@rifm.org (G. Sullivan).

https://doi.org/10.1016/j.fct.2021.112715
Received 7 July 2021; Received in revised form 20 October 2021; Accepted 24 November 2021
Available online 27 November 2021

1. Introduction

The Research Institute for Fragrance Materials, Inc. (RIFM) has evaluated the safety data for fragrance materials for 55 years. The process of evaluating fragrance materials and available data has evolved as new scientific approaches emerge. Since 2013, RIFM has focused its efforts to evaluate discrete, chemically defined materials, but in 2019 and beyond, the primary focus has shifted to evaluating NCS. Plants and derivatives have been used for various purposes, including perfuming, for centuries. It is well known that some plants and derivatives can be highly toxic, yet it is a common misconception that "natural" necessarily means "safe" while "synthetic chemical" necessarily means "dangerous." Determining the safety profile of NCS, applying today's scientific standards, is an issue faced by a number of industry sectors. To do so for fragrance materials, RIFM is leveraging advances in the risk assessment of complex substances and extensive knowledge compiled over the years on components of NCS as discrete chemicals.

The basic premise for the safety assessment of NCS is the same as for discrete fragrance materials in that all of the same endpoints (Genotoxicity, Repeat Dose Toxicity, Reproductive Toxicity, Skin Sensitization, Photoirritation/Photoallergenicity, Local Respiratory Toxicity, and Environmental Risk and Hazard) will be addressed (Api et al., 2015). A series of decision trees, reflecting advances in approaches to the risk assessment of mixtures as well as classical toxicological methodologies will follow a similar 4-step process with testing only as a last resort: 1) evaluate the available data on the whole NCS; 2) verify if the TTC can be applied; 3) verify if the NCS risk assessment can be achieved on a component basis; and 4) determine if data must be generated. For each endpoint, a tiered approach was developed based on this 4-step premise. Using *in silico* tools, RIFM examined NCS similarities based on the plant part, processing (i.e., distillation, mechanical extraction, and solvent extraction), and composition of materials across 81 plant families in an effort to address data gaps.

A requirement of the RIFM Safety Assessment program is to complete a Safety Assessment up to current scientific standards on all substances in the RIFM fragrance material inventory. This inventory includes all materials used as fragrance ingredients that are reported in the IFRA Volume of Use Surveys at the exclusion of materials for which no exposure data are available or a sample could not be provided. In these cases, the material is on the RIFM "Not Supported List" and is removed from the RIFM Fragrance Material Safety Assessment program.

There are over 900 NCS within the RIFM inventory. The NCS evaluation process begins with its identification, understanding the botanical origin, plant taxonomy, and plant part from which the complex substance is extracted. The International Organization for Standardization has described aromatic natural raw materials as they are processed and used (ISO 9235:2013). The most common of these raw materials include essential oils, concretes and absolutes, oleoresins and resinoids, CO_2 extracts and infusions, and solvent extracts. All of these are referred to as NCS. While the major constituents of a given NCS are determined by the botany and taxonomy, the region of growth and variations in climate can contribute to compositional differences. Furthermore, the type of extraction and processing of the material also impacts the composition as chemical modifications may occur, such as hydrolysis, hydration, dehydration, or decarboxylation (Góra et al., 2002; Johnson et al., 2004; Novak et al., 2006; Salgueiro et al., 2010). Within the fragrance industry, many companies will often store raw materials as part of a "communal batch." Many NCS derived from the same plant, plant part, and region using the same processing methodology will be combined into a single, "communal" batch to minimize compositional variation over time and growing seasons (Canter et al., 2005; Salgueiro et al., 2010; Vallat et al., 2005; Waimer et al., 2007; RIFM, 2021).

It is important to underscore that most NCS used in perfumery do not exist in the plant as such; the plant must first be processed in order to extract the aromatic raw material. Aromatic refers to the defining odor of the material. The aromatic feature of a plant is due to the volatile fraction of the plant or its parts (e.g., flowers, leaves, roots, stems, trunks, barks, fruits, peel, and seeds). Plants synthesize an enormous variety of volatile components that do not need to be present in large quantities in order to produce a detectable odor (Dudareva et al., 2013). In addition, for some NCS, chemical modifications post-processing may occur if exposed to prolonged light, heat, and/or oxygen. These changes must be under the control of the manufacturer to maintain the stability and composition of the NCS. The fragrance industry typically mitigates post-processing modifications through various methods (e.g., light and temperature controls, keeping batches of materials under nitrogen, or using antioxidants or stabilizers). This aids in maintaining the stability of NCS over time. Chemical modifications that indicate material integrity has been compromised are easily detected through color change and odor; however, headspace testing and gas chromatography-mass spectrometry (GC-MS) can be used in confirmatory testing of sample material stability (Bernal et al., 2020; Geng et al., 2019).

By leveraging scientific knowledge of plant taxonomy, NCS processing, and industry expertise of sample composition and sample handling, RIFM has crafted a stepwise process to begin evaluating these complex mixtures: 1) evaluate the available data on the whole NCS; 2) verify whether the TTC can be applied; 3) verify whether the NCS risk assessment can be achieved on a component basis; and 4) determine what data may need to be generated on the NCS or NCS component(s). This paper aims to provide the process to develop a comprehensive and robust Safety Assessment of NCS used as a fragrance ingredient.

2. Exposure of Natural Complex Substances (NCS)

Every Safety Assessment requires knowledge about the exposure. The exposure and risk assessment of any fragrance material is an iterative process that incorporates the available hazard data for the key toxicological endpoints coupled with the exposure assessment. Since exposure is critical to the Safety Assessment process, RIFM and scientific modeling, data analytics, and computing company Creme Global (Cremeglobal.com) partnered to develop an aggregate exposure model for fragrance materials (i.e., the total exposure coming from all different sources). This model looks at the exposure resulting from different fragrance materials used across a range of cosmetic, personal, household, and air care products. The model has helped refine the assessment of fragrance materials and has made a substantial impact on both the improvement of consumer safety of fragrances and the reduction of animal testing (Comiskey et al., 2015, 2017, 2017; Safford et al., 2015a, b, 2017).

The Creme-RIFM Model is built on large volumes of market surveys and scientific data from a wide variety of peer-reviewed and validated sources. The model estimates aggregate exposure to fragrance materials in consumer products. The model uses a probabilistic simulation, sampling from distributions of measured variables (e.g., amount of fragranced product applied or frequency of application) for individuals across a population, to provide a realistic estimate of aggregate exposure to fragrance materials used in a range of common consumer products.

NCS have been surveyed, and exposure data are available for over 900 NCS. The exposure data are updated at a minimum of every 5 years. Data on discrete chemicals reflect the use of the chemical added as such in addition to the contribution from NCS. The exposure to the NCS reflects the use of the NCS.

The innovative Creme RIFM Aggregate Exposure Model is a substantial advance from previous methods and the most comprehensive model of its kind. The model provides more realistic exposure data for RIFM's Safety Assessments for fragrance materials. By using large datasets representing consumer behavior and statistical models, the model provides more representative estimates of consumer exposure, often successfully demonstrating that they are below the level of concern.

3. Threshold of Toxicological Concern (TTC)

While the RIFM Database contains the largest amount of flavor and fragrance data in the world, including more than 75,000 references and 135,000 studies, there is a lack of toxicological data that are up to OECD standards for NCS. When consumer exposure is below a certain level for which there are no adverse health effects (i.e., the TTC), no animal testing is required. Since RIFM has collected concentration exposure survey data for over 900 NCS, for each endpoint, where data on the whole NCS is insufficient or unavailable, the feasibility of Step 2 (exposure-based waiving TTC) is evaluated.

The TTC for systemic exposure is a long-standing, broadly accepted toxicological concept based on measured data (Munro et al., 1996a, 1996b, 2008; Kroes et al., 2004, 2007; Yang et al., 2017). RIFM has continued in recent years to expand upon these datasets and strengthen the TTC approach (Patel et al., 2020). After examining the large body of available exposure data, the 95th percentile chronic exposure for nearly a third of the NCS in the RIFM inventory have chronic systemic exposure (including dermal, inhalation, and oral routes of exposure) that falls below the most conservative TTC limits (Creme RIFM Aggregate Exposure Model, 2021 v3.1.3). When exposure falls below the TTC, there is no appreciable concern for risk. All materials are examined for genotoxicity, and the TTC is only applied when there is no concern for genotoxicity or carcinogenicity.

This concept of TTC can also be applied to evaluating potential skin sensitizers using the dermal sensitization threshold (DST). The DST for both reactive and non-reactive materials are levels below which there is no appreciable risk for the induction of dermal sensitization. These levels are based on data and the probabilistic analysis of potency data for a diverse set of chemical allergens (Safford, 2008; Safford et al., 2011, 2015a, 2015b, 2017; Roberts et al., 2015). Recently, a DST has been calculated for High Potency Chemicals (HPC). The HPC were previously excluded from the DST application. This value presents a useful default approach for unidentified substances in ingredients considering, as a worst-case scenario, that the unidentified compound may be a potent skin sensitizer (Nishijo et al., 2020). Finally, in the case of inhalation exposure, data from the Creme RIFM Aggregate Exposure Model show that >99% of all NCS in the RIFM inventory fall below the most conservative TTC limits (Carthew et al., 2009). In general, whether examining single, discrete ingredients or complex mixtures, exposure to fragrance materials is very low. When exposure-based waiving is not applicable for the evaluation of NCS, it becomes necessary to examine the components of the NCS. Currently, the majority of NCS components are also discrete chemically defined fragrance materials, which allows RIFM to leverage the extensive toxicology data available on those components to assess the safety of NCS.

4. Composition

The fragrance industry, through the International Fragrance Association and International Organization of the Flavor Industry's Complex Ingredient Constituent Compendium (IFRA-IOFI CICC), compiles the composition of the NCS used to define the "typical" composition for a group of similar NCS (IFRA-IOFI unpublished data website). The CICC is the most comprehensive collection of compositional information for NCS used in the fragrance industry. Compositional information reported in the IFRA-IOFI CICC is reported to 0.1% and is the "typical" composition to be used in the RIFM Safety Assessment. This composition was prepared by fragrance industry experts with knowledge of the substance currently on the market and acknowledging the variability inherent in the growth, sourcing, and production. It does not represent a standard specification for use in material production or for use in regulatory compliance. Compositional information available is largely focused on the volatile fraction of aromatic raw materials. However, non-volatile components are identified wherever possible, and material composition is normalized to the total composition, as provided in the CICC, not just the volatile fraction (https://ifrafragrance.org/).

The European Federation of Essential Oils (EFEO)/IFRA (2015) guidelines on identification and sameness of NCS under REACH and CLP are followed in the CICC and, therefore, in the RIFM Safety Assessments on NCS. For example, the following names are used:

- Well-defined mono-constituent substance: Substances in which 1 constituent is present at a concentration of at least 80% (w/w);
- Well-defined multi-constituent substance: Substances consisting of several main constituents present at concentrations generally above or equal to 10% and below 80% (w/w);
- "UVCB substances" (substances of unknown or variable composition, complex reaction products, or biological materials): NCS is where the source is biological and the process is refinement. The name of the UVCB should use a combination of the source and the process, starting with the source.

The typical composition is reported down to 1% in all RIFM NCS Safety Assessments. Substances below 1% but greater than or equal to 0.1% that are components with IFRA Standards or classified by the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) as Carcinogenic, Mutagenic and Reprotoxic (CMRs), Persistent Bioaccumulative Toxic (PBTs), Sensitizers, or Photoirritant and Photoallergenic components are also reported in the Safety Assessment. In addition, any component with an IFRA Standard that is below 0.1% will also be included in the Safety Assessment.

If toxicological testing is required in order to complete a Safety Assessment on an NCS (see Fig. 2, Step 4), the test material composition of NCS samples should represent the agreed-to "typical" composition. Where possible, the test sample will have a minimum of 80% of the total "typical" composition described in the CICC. If 20% of the composition differs from the "typical" composition, then the remaining differences are examined closely for chemical clustering to ensure that the differences represent materials that remain in the same chemical cluster. The test material will strive to be a quality that is currently available on the market.

5. Grouping similar NCS

The purpose of grouping is to maximize review efforts based on similarity in composition for prioritization. The motivation for assessing chemical similarity within all NCS in the RIFM inventory is to maximize the efficiency of the Safety Assessment process. This is accomplished by examining similar NCS at the same time and trying to address data gaps on individual NCS without further testing.

To determine chemical similarity, all NCS are first identified by plant taxonomy: Family, Genus, and Species. Further identifiers, including plant part and extraction processing methodologies, are used to clarify substance identification. Addressing the chemical similarity of complex mixtures begins with 2 distinct questions (see Fig. 1): 1) is the component structural identity the same across NCS ("structural identity" meaning the number of components common versus uncommon), and 2) are compositional percentages of the components the same? The structural similarity of the components is examined in the same manner as for the discrete fragrance ingredients described in the RIFM Criteria for the RIFM Safety Evaluation Process for Fragrance Ingredients (Criteria Document; Api et al., 2015) and read-across rules (Date et al., 2020). At the component level, expert judgment is applied through both known chemical reactivity and toxicological mechanisms.

Considering each component and its percentage in the NCS as discrete variables, NCS are grouped. The grouping method is multi-tiered. The top tier of grouping considers plant taxonomy, family, and genus. In the second tier, the agglomerative hierarchical clustering method is applied to group all NCS under a single genus. Hierarchical clustering is a method of clustering that seeks to build a hierarchy. Agglomerative clustering is a bottom-up approach in which the starting

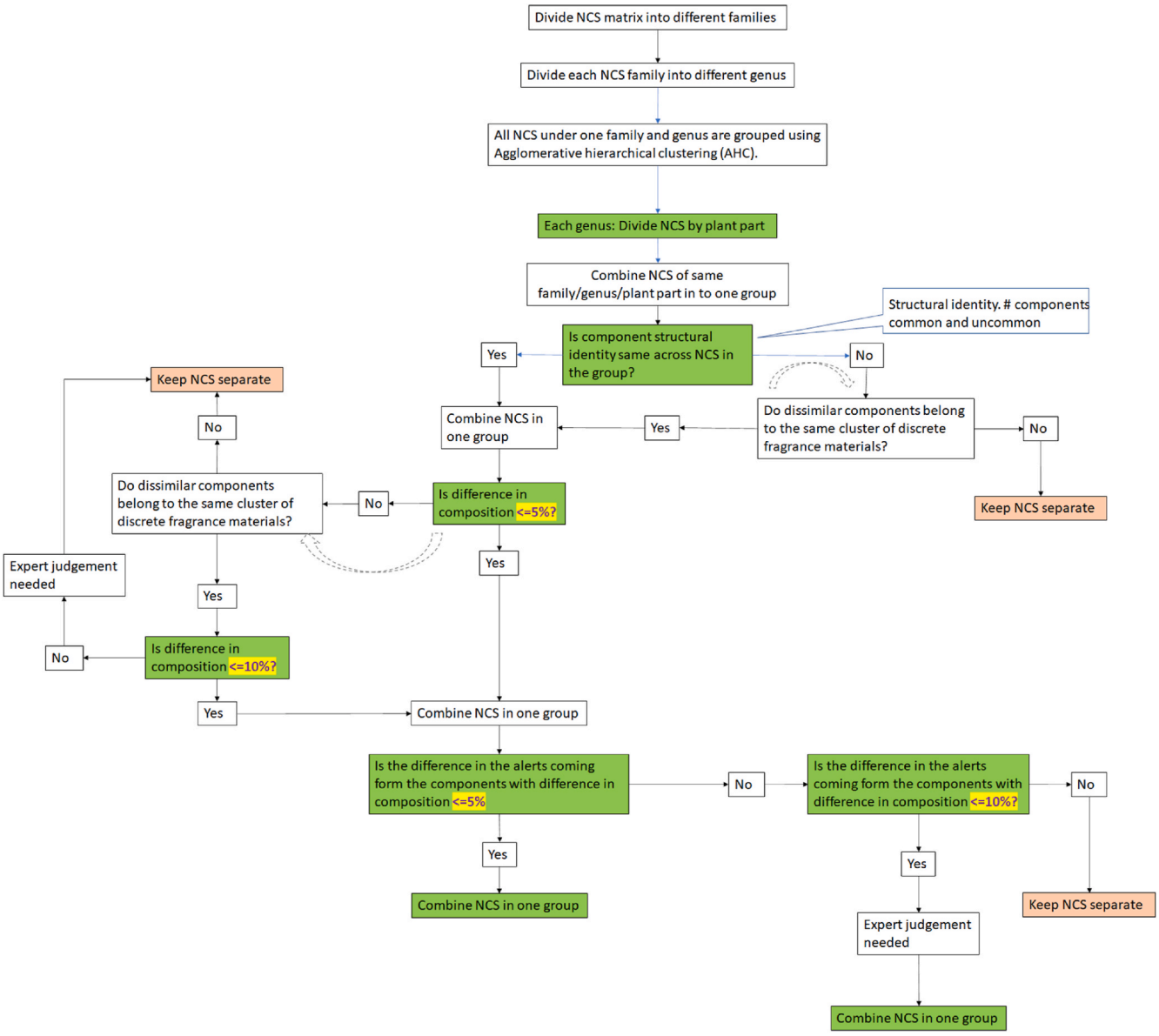

Fig. 1. Determining NCS similarity for prioritization.

point is a single cluster, and pairs of clusters are merged as one moves up the hierarchy (Nielsen, 2016). In the third tier, the source plant part, the process of extraction, and the final composition of NCS are considered. The resulting subgroups of NCS from this tiered process are then subjected to further refinement in a decision tree. This refinement is carried out according to structural similarity and the percentage of compositional difference. A generally accepted literature criterion for reporting the synthesis of a compound is a minimum of 95% purity (Portoghese, 2009). Based on this, the resulting NCS groups that have components with <5% compositional difference were placed in the same cluster. Any NCS failing this condition is allocated into a separate group.

6. Read-across justification for specific components

Where feasible under Step 3, RIFM will also use a read-across approach between components of the NCS that are considered discrete chemicals. The read-across analogs for the NCS components are identified using the RIFM fragrance materials chemical inventory clustering and read-across search criteria (Date et al., 2020). These criteria follow the strategy for structuring and reporting a read-across prediction of toxicity as described in Schultz et al. (2015) and are consistent with the guidance provided by OECD within Integrated Approaches for Testing and Assessment (OECD, 2017) and the European Chemical Agency (ECHA) read-across assessment framework (ECHA, 2017).

It may be possible that if sufficient toxicological data are available, RIFM will consider utilizing NCS-to-NCS read-across. There must be a strong similarity between all the dimensions of an NCS; only then will there be a possibility to perform NCS-NCS read-across.

7. Human health and environmental stepwise evaluation strategies

7.1. Genotoxicity

The first step in examining the genotoxicity of a material is determining the acceptability of any existing data on the whole NCS being evaluated. Where a number of tests have been undertaken that do not necessarily meet accepted guidelines, it may still be decided by RIFM that there is sufficient weight-of-evidence on which to base a conclusion as to the genotoxicity/non-genotoxicity of the substance. It is well

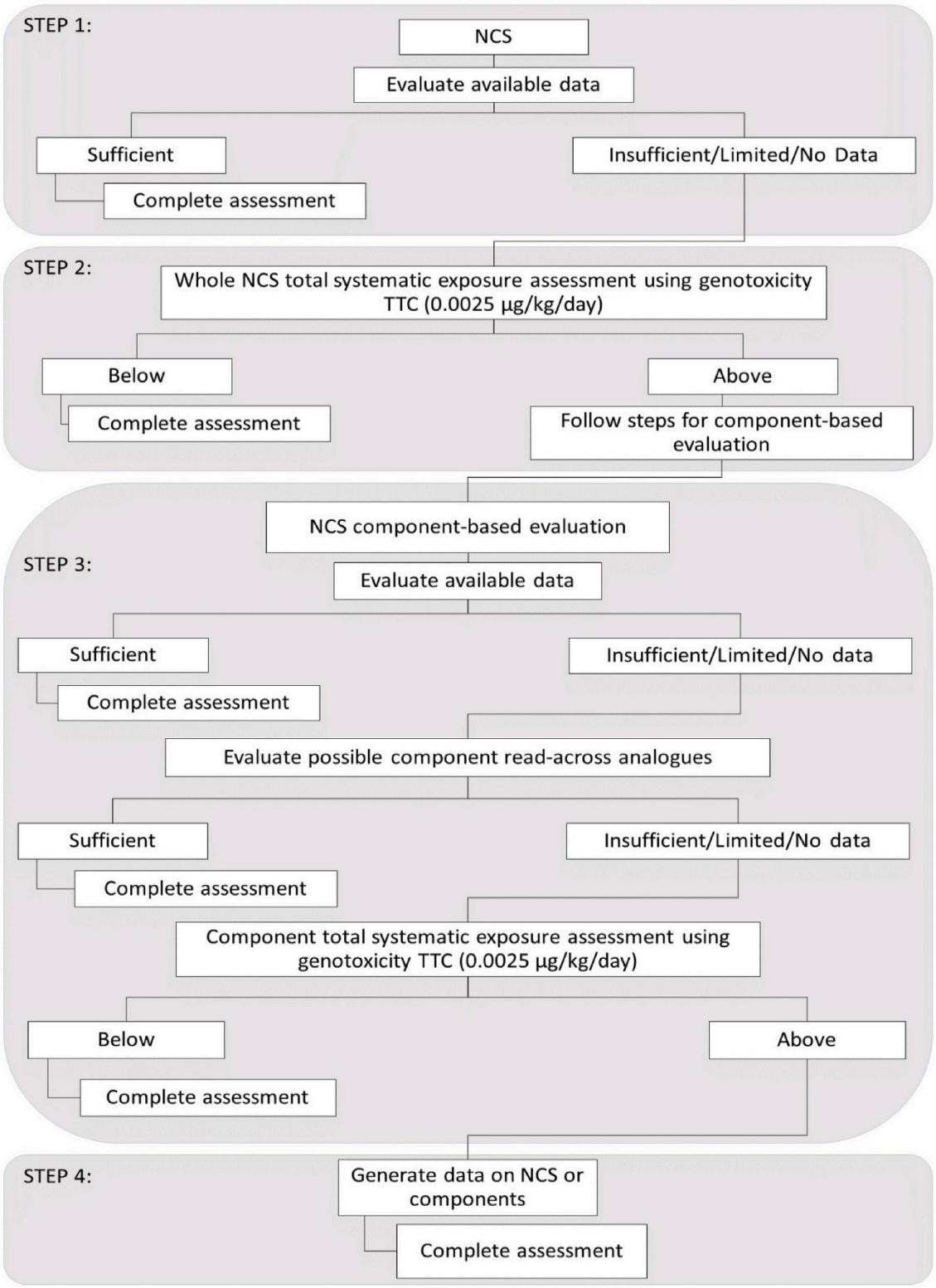

Fig. 2. Genotoxicity material evaluation.

understood that no single assay can be utilized to predict genotoxic effects to humans; rather, a combination of tests that address different genetic endpoints must be considered. For example, a sufficient data set should address both gene mutation (i.e., an Ames or HPRT test) and cytogenetic (clastogenic) potential. An insufficient data set for a material is considered to be no data at all or only data that cover one of the key endpoints: gene mutation or clastogenicity. For fragrance materials that are determined to have sufficient data, an endpoint assessment is conducted without further testing needs (see Fig. 2, Step 1). If there are no data and/or a partial dataset available for the whole NCS, then the exposure must be considered (see Fig. 2, Step 2).

In the absence of any genotoxicity data on the whole NCS, a TTC

level of 0.15 μg/person/day (0.0025 μg/kg/day) can be applied. This default TTC value for potentially genotoxic materials of 0.15 μg/person/day, which has been established by Kroes et al. (2004), is derived from the extensive Carcinogenic Potency Database (CPDB) of Gold and co-workers (Gold et al., 1984, 1989, 1997) and is based on linear extrapolation down to a 1 in a million (10−6) risk. In case there are no data on the whole NCS, and the exposure is above the TTC level of 0.15 μg/person/day, then component-based evaluation is conducted (see Fig. 2, Step 3).

In a component-based evaluation, target data are evaluated for all the components identified in the NCS in a similar manner as in an individual fragrance Safety Assessment described in detail in Api et al. (2015). First, target data are evaluated for the components. If not sufficient, read-across analog data are considered. If there are no data on the target or read-across for the identified components, exposure-based assessment for the components is considered. If exposure is below the TTC level of 0.15 μg/person/day (0.0025 μg/kg/day), then the Safety Assessment is concluded. If the exposure is above the TTC and no data are available, then testing may be required (see Fig. 2, Step 4). When there are no data available for the whole NCS and/or its components and exposure is above the TTC, testing of either the whole NCS or a component is required. The testing strategy may vary on a case-by-case basis, based on expert opinions.

7.2. Repeat Dose Toxicity, Developmental Toxicity, and Fertility

If available data on the whole NCS are determined to be sufficient, a no observable effect level (NOAEL) will be derived, and the Margin of Exposure (MoE) will be calculated. The Safety Assessment is considered complete and acceptable if the MoE >100. Unless skin absorption data on the whole NCS are available, skin absorption will be considered to be 100% (see Fig. 3, Step 1).

For instances where there are insufficient data available on the whole NCS, and the genotoxicity concern has been excluded, the whole NCS exposure will be compared to the Cramer Class III limit (1.5 g/kg/day; Kroes et al., 2007) (see Fig. 3, Step 2). However, if >95% of the NCS components are identified in the same Cramer Class, then that whole NCS is classified in the same Cramer Class, as long as the remaining 5% derived exposure does not exceed the Cramer Class III limit. If exposure to the whole NCS is above the appropriate TTC Cramer Class limit, the NCS will be evaluated on a component basis in a multistep approach following the RIFM Criteria Document (Api et al., 2015). Briefly, data available on the component will be evaluated thoroughly, and if considered sufficient, a NOAEL will be determined (see Fig. 3, Step 3a). For components with insufficient data, an appropriate read-across analog will be selected (Date et al., 2020), followed by data evaluation (see Fig. 3, Step 3b). If data are determined to be insufficient, the derived exposure of the component will be compared to its respective Cramer Class exposure threshold. If the derived exposure to each component continues to be above the TTC threshold, the exposure for each component will be refined using *in silico* or *in vitro* skin absorption data (Shen et al., 2014). Derived exposure for each component is calculated by multiplying the total systemic exposure (in μg/kg/day) of the whole NCS and the percentage of each component in the whole NCS.

If all of the above-described approaches are deemed inadequate to support the Repeat Dose and/or Developmental and Reproductive Toxicity safety evaluation for the whole NCS, RIFM will consider the need for testing or recommend risk management measures (see Fig. 3, Step 4).

7.3. Skin Sensitization

Whole NCS data will first be evaluated, and if it is determined to be sufficient, it will be concluded as a non-sensitizer or quantitative risk assessment (QRA) will be conducted using a weight-of-evidence No Expected Sensitization Induction Level (NESIL) (see Fig. 4, Step 1). In the case of incomplete or limited data availability, a reactive DST of 64 μg/cm^2 will first be applied to the reported NCS exposure (Safford, 2008; Safford et al., 2011, 2015a,b; Roberts et al., 2015) (see Fig. 4, Step 2).

If there are insufficient data and the exposure of the whole NCS is above the reactive DST, the NCS is assessed based on its components (see Fig. 4, Step 3). In this step, all components that are present in the NCS are assessed individually. When existing data are insufficient on a component, a search for a read-across analog is conducted. If sufficient data on the component or its read-across analog support that the component is not a sensitizer, the component is considered safe in the context of the NCS. If sufficient data on the component or its read-across analog indicate the component is a sensitizer, the NESIL of the component is used for QRA. For the QRA of the component, the maximum acceptable concentrations in finished products are calculated based on the component NESIL. The current derived exposure of the component is then compared with the calculated Maximum Acceptable Concentrations for each product category. The derived exposure is calculated by multiplying the current dermal exposure of the NCS for each product category by the typical percentage of the component in the NCS. The component is considered safe under the current use level in the context of the NCS if the derived exposure is below the maximum acceptable concentrations. If the derived exposure is above the maximum acceptable concentration, further testing may be considered. If additional testing is not possible or appropriate, then risk management measures may be considered. When insufficient data are available for a component, and no appropriate read-across analog can be found, the reactivity of the component and its potential metabolites and autoxidation products with skin proteins are assessed utilizing the existing data, information from structural analysis, and *in silico* tools. Depending on the reactivity of the component and its potential metabolites and autoxidation products, the derived exposure of the component is benchmarked utilizing the non-reactive DST of 900 μg/cm^2 or reactive DST of 64 μg/cm^2 (Safford, 2008; Safford et al., 2011, 2015a,b; Roberts et al., 2015). For components with current derived exposures above the respective DST, follow-up *in vitro* testing may be needed (see Fig. 4, Step 4). If there is enough existing data to move to a human study to support the current use level, then a Confirmation of No Induction in Humans test (CNIH) will be conducted (Na et al., 2020).

7.4. Photoirritation and Photoallergenicity

Data on the whole NCS will be evaluated first (see Fig. 5, Step 1). If the available data are sufficient, or if there is UV Absorbance on the NCS that shows the NCS does not exhibit significant absorbance, the photoirritation and photoallergy section of the NCS may be completed. If data are insufficient, then exposure-based waiving of further testing will be attempted (see Fig. 5, Step 2). If dermal exposure for the NCS in all product categories except for IFRA Category 12 (Products not intended for direct skin contact, minimal or insignificant transfer to the skin) is below an exposure level below which it is unlikely that any type of phototoxic potential exists, then the NCS can be completed for photoirritation based on exposure. This threshold level was defined by Api et al. (2015) as 5 ppm. If dermal exposure for the whole NCS exceeds 5 ppm, then a component-based evaluation will be made. A component-based evaluation may be made on the basis of UV/Vis absorbance (no absorbance or a Molar Extinction Coefficient [MEC] below the benchmark of concern for photosafety, 1000 L/mol/cm (Henry et al., 2009)) or available study data (see Fig. 5, Step 3a). Available data (UV absorbance data or study data) from read-across analogs may also be used in place of data for the target constituent (see Fig. 5, Step 3b). If data for a component are insufficient, or if there is a component of concern (MEC>1000 L/mol/cm, or study data indicative of photoirritant or photoallergenic effects), the derived dermal exposure for that component will be compared to an exposure level below which it is unlikely that any type of phototoxic potential exists. If

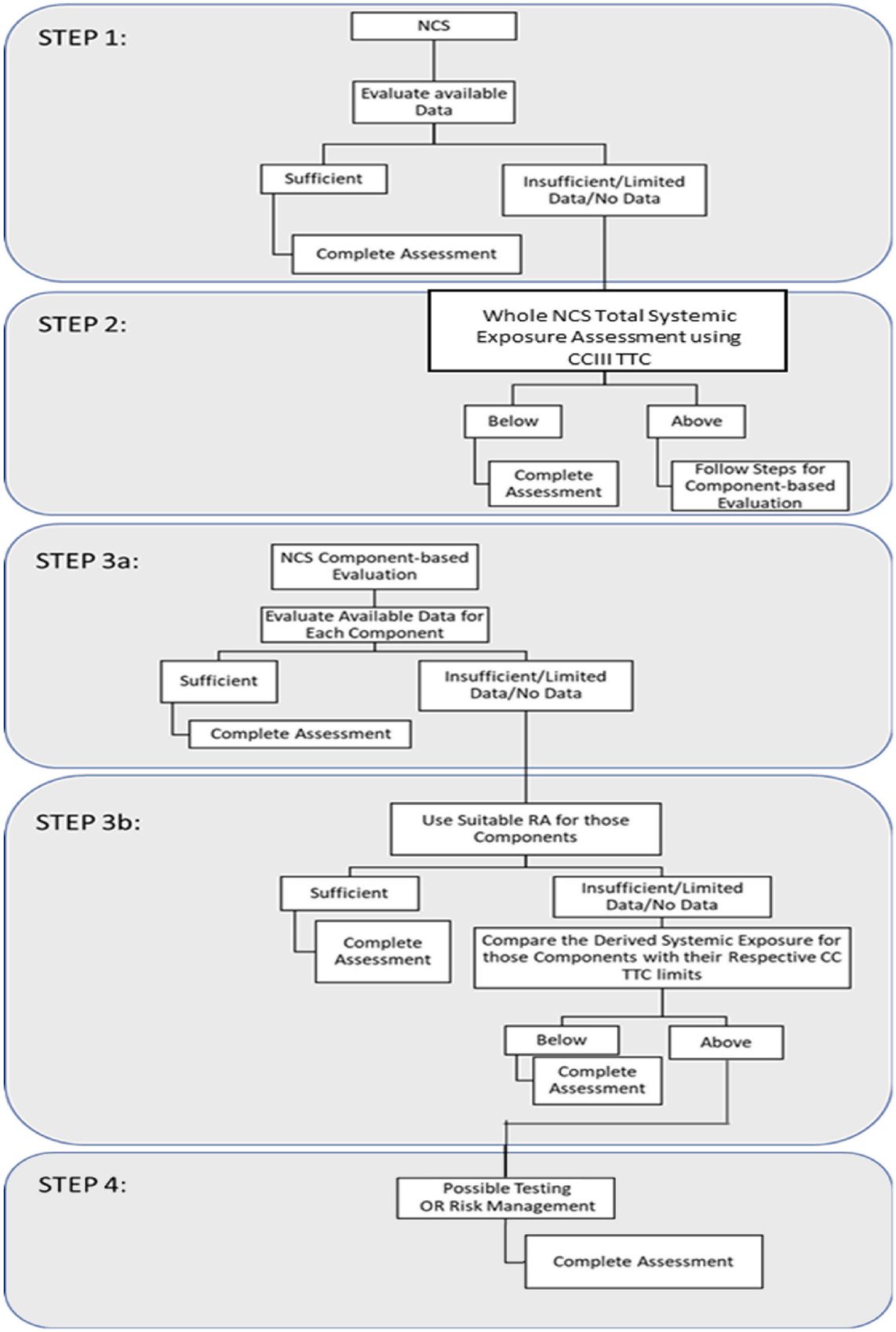

Fig. 3. Repeat Dose Toxicity, Developmental Toxicity, and Fertility material evaluation.

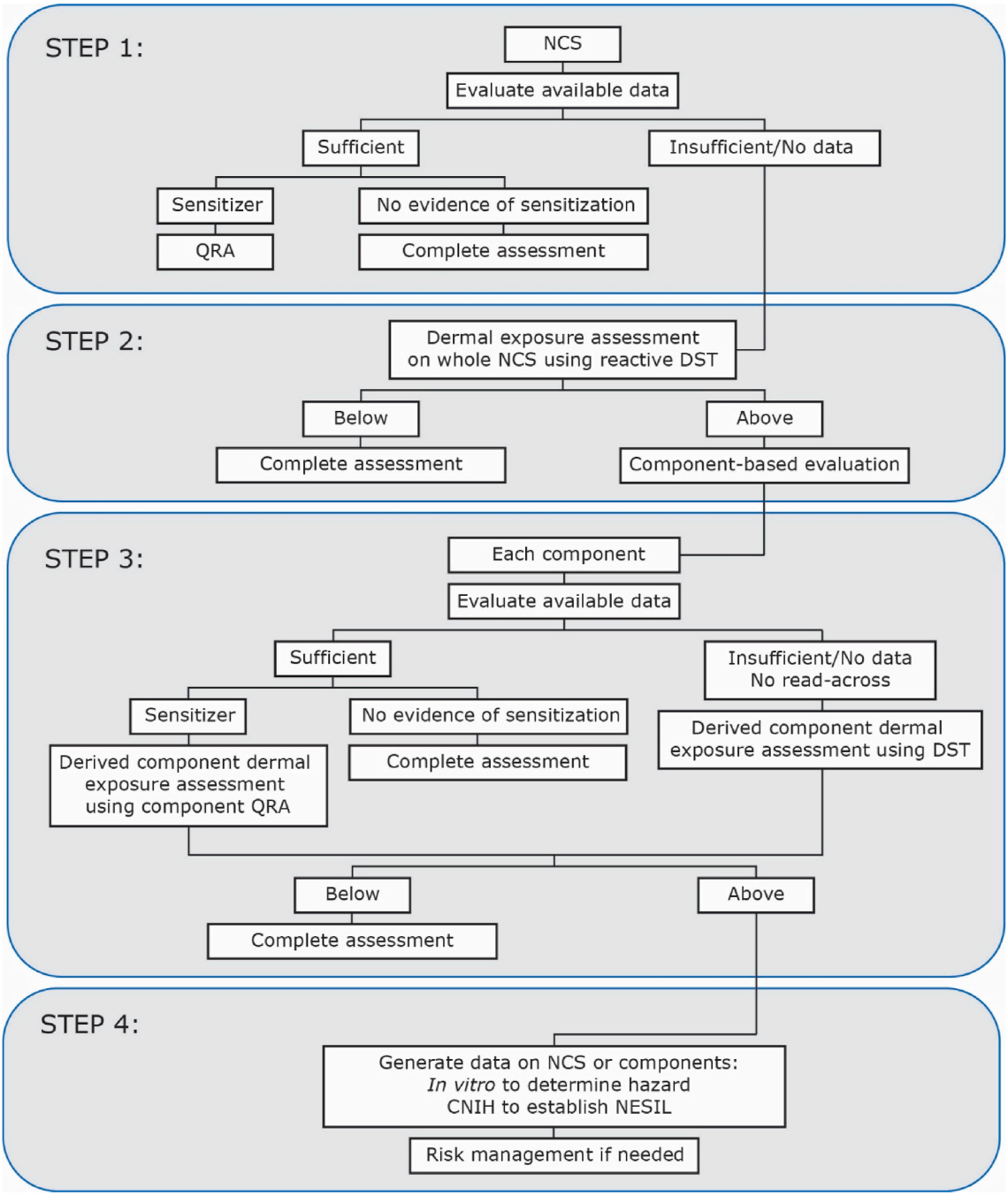

Fig. 4. Skin Sensitization material evaluation.

none of these options are viable, then testing of the NCS will be conducted (see Fig. 5, Step 4). A tiered testing approach will be utilized, starting with UV/Vis absorbance (OECD 101), followed by the 3T3-Neutral Red Uptake phototoxicity assay (OECD 432) if needed, and finally, if required, a reconstructed human epidermis phototoxicity assay in conjunction with a human photoirritation test for confirmation of a no-effect level in humans for photoirritation. Apart from UV/Vis absorbance, the testing strategy does not address photoallergenicity as there are currently no recognized non-animal test methods to address photoallergenicity. When this is the case, it will be clearly stated in the Safety Assessment.

7.5. Local Tespiratory Toxicity

NCS data will be evaluated first, and if determined to be sufficient, a No Observed Adverse Effect Concentration (NOAEC) will be derived, and an MoE will be calculated (see Fig. 6, Step 1). If there is no data available on the whole NCS (or the data is deemed insufficient), exposures will be compared to the more restrictive inhalation TTC limit (Cramer Class III) for local effects defined by Carthew et al. (470 g/day; Carthew et al., 2009) (see Fig. 6, Step 2). However, if >95% of the NCS components are identified in the same Cramer Class, then that whole NCS is classified in the same Cramer Class, as long as the remaining 5% derived exposure does not exceed the Cramer Class III limit. If exposure of the NCS is above this inhalation TTC limit, then the substance will be evaluated on a component basis: 1) examine data on each target

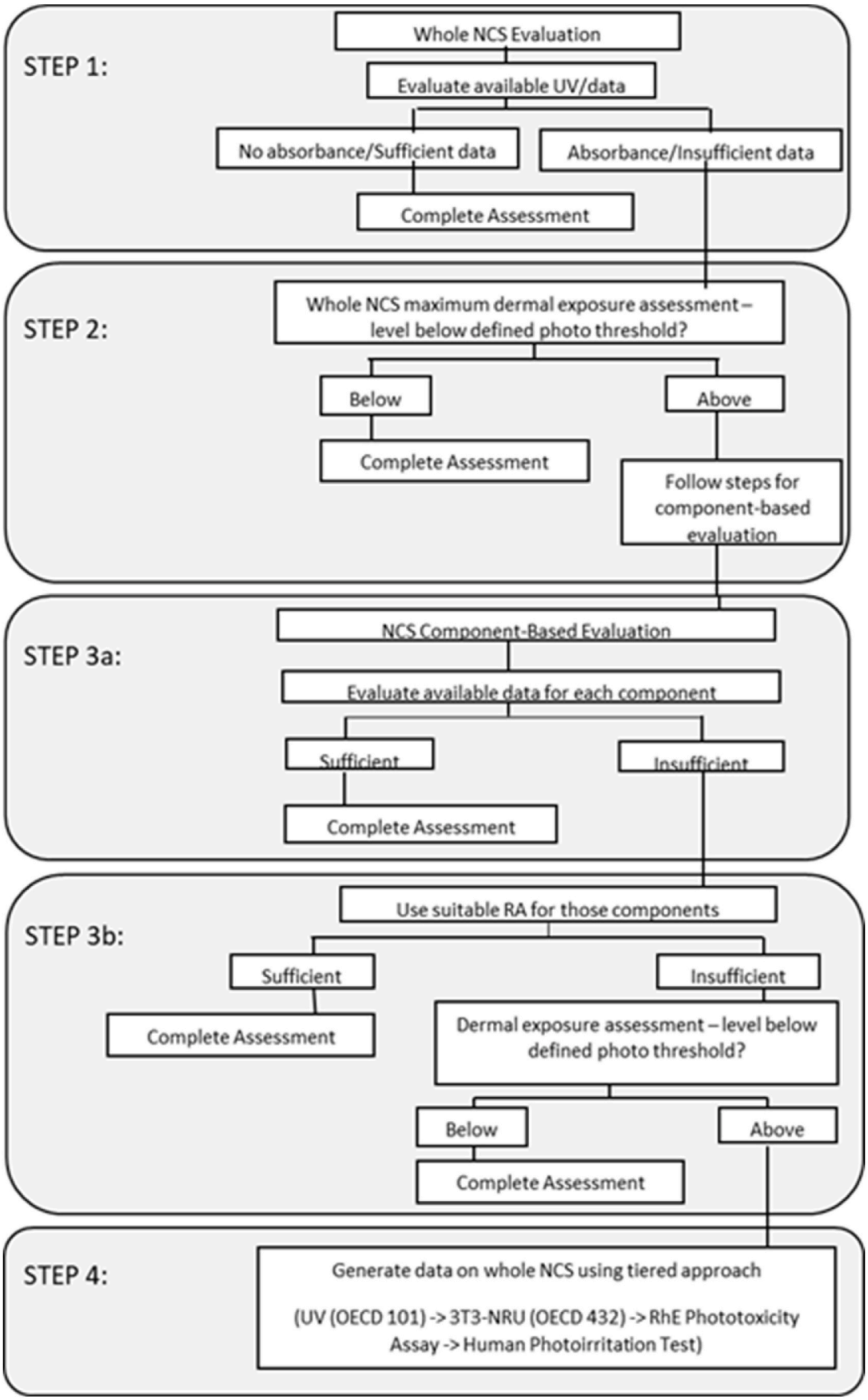

Fig. 5. Photoirritation & Photoallergenicity material evaluation & testing strategy.

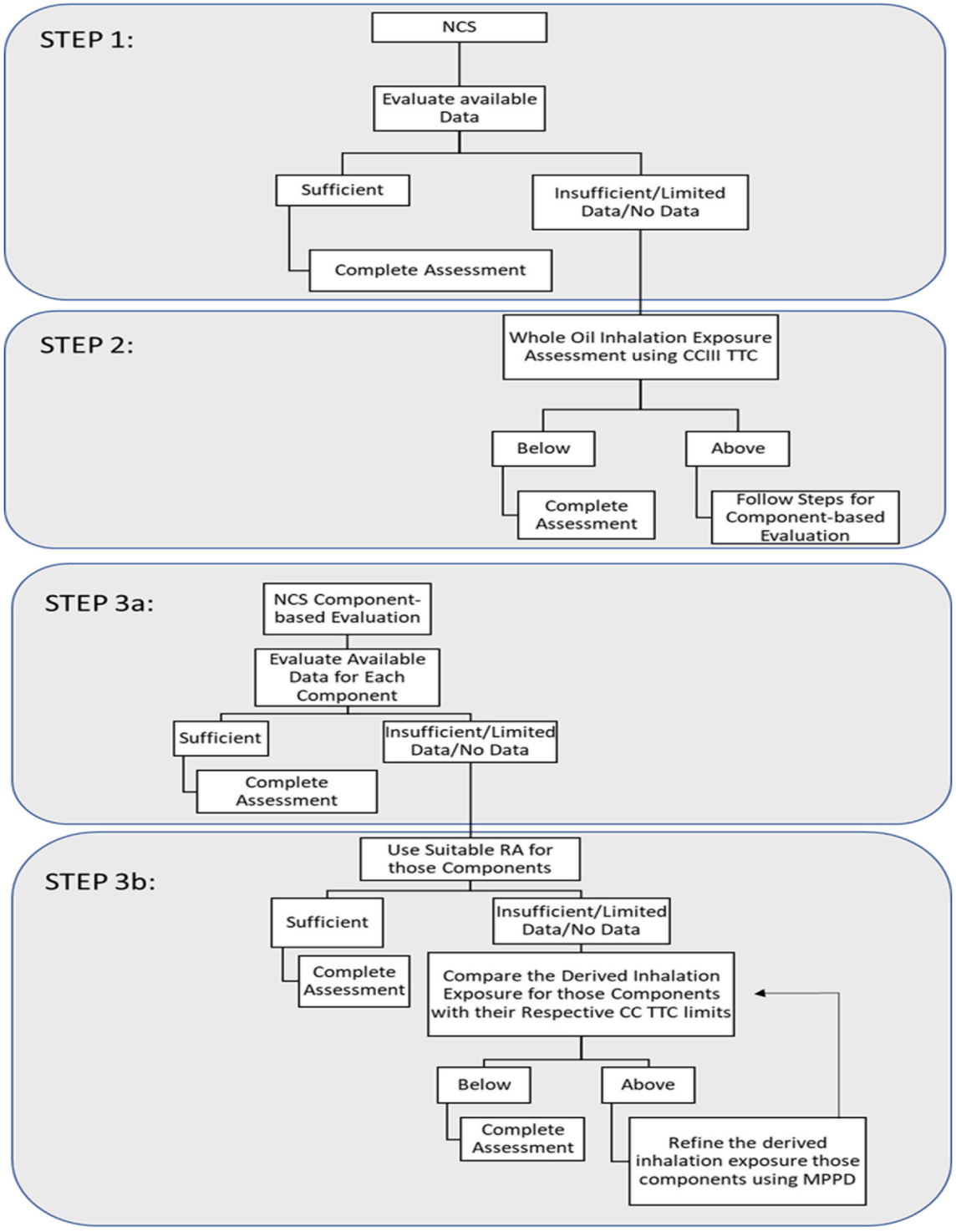

Fig. 6. Local Respiratory Toxicity material evaluation.

component or on read-across analogs; 2) evaluate each component against respective Cramer Class TTC (if unavailable, then default to Cramer Class III TTC limit); or 3) refine each component exposure using the multiple-path particle dosimetry model (MPPD) and re-assess against the inhalation TTC (see Fig. 6, Steps 3a-3b). Components will be evaluated using a derived exposure value: inhalation exposure of the NCS x component % = derived exposure (mg/day). Fig. 6 outlines the criteria for local respiratory toxicity evaluation.

The MPPD model is an *in silico* tool used to simulate inhalation exposure of a fragrance ingredient by considering its physical-chemical properties and refining the deposition fraction for each region within the respiratory tract.

An analysis of all available fragrance ingredient aggregate chronic inhalation exposure data was conducted (Creme RIFM Aggregate Exposure Model, Version 3.1.3, April 2021). The analysis showed that >99% of NCS chronic aggregate inhalation exposure falls below the most conservative TTC limits. Therefore, under current conditions of use, exposure-based waiving is sufficient for the evaluation of local respiratory toxicity.

8. Environmental Risk and Hazard

Guidance for the assessment of NCS materials to comply with REACH requirements has been previously developed as a joint effort between IFRA, the IFRA Environmental Task Force, EFEO, and RIFM (See EFEO/IFRA Guidelines on the Environmental Assessment of NCS, http://www.efeo.eu/fileadmin/user_upload/REACH/EFEO_IFRA/Dokument_26_May_2016/eco _tox_essential_oil_guidance_en.pdf).

This guidance outlines the use and applicability of component, block, and whole NCS approaches to address both risk and PBT assessment.

8.1. Approach

(1) The "known constituents' approach": This approach can be used when a substance is known to contain specific constituents at relevant concentrations that are suspected of having (v)P, (v)B and T properties.
(2) The "block approach" (fraction profiling): The substance is divided into fractions/blocks of structurally similar constituents, or which follow a regular, predictable pattern of structures.
(3) The "whole NCS" approach: The NCS is considered to be a single chemical substance for the purpose of the assessment and testing (see Fig. 7).

Choice of approach (see Fig. 7) will depend on several factors such as:

(1) Knowledge of constituents and/or fractions in the whole NCS.
(2) Differences in properties amongst them.
(3) The ability to characterize these.
(4) Technical limitations in testing and feasibility to generate new data will influence the choice of the approach. In some cases, the strategy will require a stepwise approach starting with one approach and improving the assessment by using or combining other approaches to different constituents or groups of constituents.

8.1.1. The known constituents' approach

(1) Can apply when a substance is well characterized and/or is known to contain specific constituents that are relevant for classification and for the PBT/vPvB assessment when they are suspected, based on screening-level information, to represent the worst case of the (v)P, (v)B and T properties.
(2) Can also be used if the specific constituents can be isolated or separately manufactured for testing or if there is existing available data for the individual constituents.

8.1.1.1. Risk assessment. In a "constituent approach," each constituent would be assessed individually by assessing the hazard of constituent x, the exposure of constituent x and the risk of constituent x. The risk assessment for the substance is based on individual constituents.

8.1.1.2. Hazard assessment (PNEC determination). Predictive No-Effect Concentration (PNECs) are determined for an individual constituent in the same manner as for discrete compounds. The use of QSAR or measured data for the constituents are used, and appropriate assessment factors as described in Salvito et al. (2002) are applied to determine the PNEC. PEC is determined by the amount of the component in the evaluated NCS. These are then compared against their respective PEC (Predicted environmental concentration) to determine the overall RCR (risk characterization ratio).

8.1.2. The "block approach" (or "fraction profiling")

(1) Constituents that are structurally similar or that follow a regular, predictable pattern of structures are grouped into fractions that are normally considered as if they were single constituents. The assessment and/or testing are conducted on the fraction itself, not on individual (or surrogate) constituents.
(2) The substance is divided into fractions containing constituents that are expected to have the same degradation behavior:

8.1.2.1. Risk assessment. A constituent (or related structure) can be chosen to represent each block and data required to complete the risk assessment collected (e.g., adsorption properties such as log K_{ow}, log K_{oc}). For example, an essential oil composed of sesquiterpene alcohols and sesquiterpene hydrocarbons may be regarded as 2 blocks of constituents based on their water solubility and adsorption properties.

8.1.2.2. Hazard assessment (PNEC determination). For blocks of constituents of substances of similar structure and physical-chemical properties, QSARs can be applied using a worst-case scenario (i.e., highest log K_{ow}), or if data are available on members of the block, the lowest value for an aquatic toxicity endpoint (NOEC, EC50, LC50) can be utilized, and appropriate assessment factors are applied to determine the PNEC. These are then compared against their respective PECs to

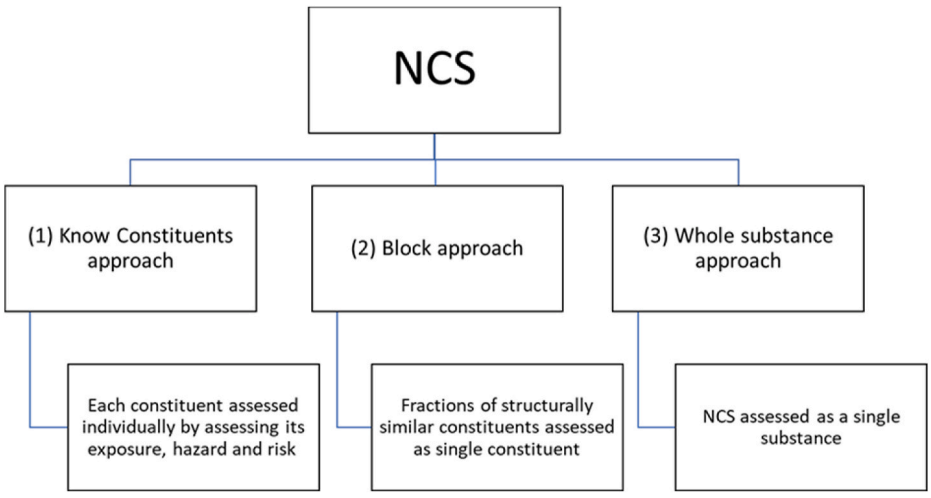

Fig. 7. Environmental risk & hazard evaluation.

determine the overall RCR.

8.1.3. The "whole NCS" approach

(1) When all the constituents are expected to have very similar properties, standard test methods may be applicable.
(2) The whole NCS can be considered as a single chemical substance for the purpose of the assessment and testing.
(3) When the whole NCS is composed of constituents with dissimilar properties, the use of the whole NCS approach may still be applicable:

8.1.3.1. Risk assessment. The NCS will be assessed as a whole NCS.

8.1.3.2. Hazard assessment (PNEC determination). A whole NCS approach to PNEC assessment may be appropriate as well.

This could provide confirmatory or complementary data to either the constituent or block approach (or both). Furthermore, where feasible, a PNEC derived from a Water Accommodated Fraction (WAF) could present a more environmentally realistic PNEC.

9. Summary

During the period from 2014 to 2020, approaches to the safety assessment of complex substances were limited to the general guidance provided within the Criteria Document (Api et al., 2015). This publication is designed to update RIFM's approach to evaluate NCS. As with all RIFM Safety Assessments, they will be re-evaluated on a 5-year basis in order to include any new data and/or new toxicological methods of assessment. This includes a review of the exposure data, especially in cases where threshold waiving measures or use of were employed.

The RIFM approach to evaluating NCS employs a series of decision trees, reflecting advances in risk assessment approaches of mixtures and toxicological methodologies, in a tiered approach for each endpoint using a 4-step process with testing only as a last resort: 1) evaluate available data on NCS; 2) verify whether TTC can be applied; 3) verify whether the NCS risk assessment can be achieved on a component basis; and 4) determine whether data must be generated. The process includes:

- A source of extensive complex natural substance composition (the IFRA_IOFI (CICC) https://ifrafragrance.org);
- An outline of the effects of botany, plant taxonomy, and processing for substance identification of NCS;
- a decision tree for determining the chemical similarity of NCS for prioritization; and
- A stepwise evaluation process for human health and environmental toxicity regarding NCS using the information on the NCS and its components.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. RIFM staff are employees of the Research Institute for Fragrance Materials, Inc. (RIFM). The Expert Panel receives a small honorarium for time spent reviewing the subject work.

References

Api, A.M., Belsito, D., Bruze, M., Cadby, P., Calow, P., Dagli, M.L., Dekant, W., Ellis, G., Fryer, A.D., Fukayama, M., Griem, P., Hickey, C., Kromidas, L., Lalko, J.F., Liebler, D.C., Miyachi, Y., Politano, V.T., Renskers, K., Ritacco, G., Salvito, D., Schultz, T.W., Sipes, I.G., Smith, B., Vitale, D., Wilcox, D.K., 2015. Criteria for the Research Institute for fragrance materials, Inc. (RIFM) safety evaluation process for fragrance ingredients. Food Chem. Toxicol. : Int. J. Publ. Br.Indust. Biol. Res. Assoc. 82 (Suppl. l), S1–S19.

Bernal, J., Nozal, M.J., Bernal, J.L., Ares, A.L., 2020. Determination of carvacrol and thymol in honey by using a simple and efficient headspace-gas chromatography-mass spectrometry method. Food Anal. Methods 13, 2138–2146.

Canter, P.H., Thomas, H., Ernst, E., 2005. Bringing medicinal plants into cultivation: opportunities and challenges for biotechnology. Trends Biotechnol. 23, 180–185.

Carthew, P., Clapp, C., Gutsell, S., 2009. Exposure based waiving: the application of the toxicological threshold of concern (TTC) to inhalation exposure for aerosol ingredients in consumer products. Food Chem. Toxicol. : Int. J. Publ. Br.Indust. Biol. Res. Assoc. 47, 1287–1295.

Comiskey, D., Api, A.M., Barratt, C., Daly, E.J., Ellis, G., McNamara, C., O'Mahony, C., Robison, S.H., Safford, B., Smith, B., Tozer, S., 2015. Novel database for exposure to fragrance ingredients in cosmetics and personal care products. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 72, 660–672.

Comiskey, D., Api, A.M., Barrett, C., Ellis, G., McNamara, C., O'Mahony, C., Robison, S. H., Rose, J., Safford, B., Smith, B., Tozer, S., 2017. Integrating habits and practices data for soaps, cosmetics and air care products into an existing aggregate exposure model. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 88, 144–156.

Date, M.S., O'Brien, D., Botelho, D.J., Schultz, T.W., Liebler, D.C., Penning, T.M., Salvito, D.T., 2020. Clustering a chemical inventory for safety assessment of fragrance ingredients: identifying read-across analogs to address data gaps. Chem. Res. Toxicol. 33, 1709–1718.

Dudareva, N., Klempien, A., Muhlemann, J.K., Kaplan, I., 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. New Phytol. 198, 16–32.

European Chemical Agency (ECHA), 2017. Read-across assessment framework (RAAF). Retrieved from. https://echa.europa.eu/documents/10162/13628/raaf_en.pdf.

EFEO/IFRA Guidelines on Substance Identification and Sameness of Natural Complex Substances (NCS) under REACH and CLP, 2015. Retrieved from. https://ifrafragrance.org/docs/default-source/guidelines/23628_gd_2015_08_14_guidelines_on_substance_identification_and_sameness_of_natural_complex_substances_(ncs)_under_reach_and_clp.pdf?sfvrsn=516b038a_0.

Geng, Q., Zhan, P., Tian, H., Wang, P., Chen, H., 2019. Gradual optimization of headspace solid-phase microextraction conditions of volatiles in pepper chicken soup combined with gas chromatography-mass spectrometry and principal component analysis. Int. J. Anal. Chem. https://doi.org/10.1155/2019/8963191, 2019.

Gold, L.S., et al., 1984. A carcinogenic potency database of the standardized results of animal bioassays. Environ. Health Perspect. 58, 9–319.

Gold, L.S., Slone, T.H., Bernstein, L., 1989. Summary of carcinogenic potency and positivity for 492 rodent carcinogens in the carcinogenic potency database. Environ. Health Perspect. 79, 259–272.

Gold, L.S., Slone, T.H., Manley, N.B., Garfinkel, G.B., Rohrbach, L., Ames, B.N., 1997. Carcinogenic potency database. In: Handbook of Carcinogenic Potency and Genotoxicity Databases. CRC Press, Boca Raton, FL, pp. 1–606.

Góra, J., Lis, A., Kula, J., Staniszewska, M., Wolszyn, A., 2002. Chemical composition variability of essential oils in the ontogenesis of some plants. Flavour Fragrance J. 17, 445–451.

Henry, B., Foti, C., Alsante, K., 2009. Can light absorption and photostability data be used to assess the photosafety risks in patients for a new drug molecule? J. Photochem. Photobiol. B Biol. 96 (1), 57–62.

ISO 9235, 2013. Aromatic Natural Raw Materials — Vocabulary.

Johnson, C.B., Kanantzis, A., Skoula, M., Mitteregger, U., Novak, J., 2004. Seasonal, populational and ontogenic variation in the volatile oil content and composition of individuals of Origanum vulgare subsp. Hirtum, assessed by GC headspace analysis and by SPME sampling of individual oil glands. Phytochem. Anal. 15, 286–292.

Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J.G., Wurtzen, G., European branch of the International Life Sciences, I, 2004. Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. Food Chem. Toxicol. : Int. J. Publ. Br.Indust. Biol. Res. Assoc. 42, 65–83.

Kroes, R., Renwick, A.G., Feron, V., Galli, C.L., Gibney, M., Greim, H., Guy, R.H., Lhuguenot, J.C., Sandt, J.J.M.v.d., 2007. Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. Food Chem. Toxicol. 45 (12), 2533–2562.

Munro, I.C., Ford, R.A., Kennepohl, E., Sprenger, J.G., 1996a. Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. Food Chem. Toxicol. : Int. J. Publ. Br.Indust. Biol. Res. Assoc. 34, 829–867.

Munro, I.C., Ford, R.A., Kennepohl, E., Sprenger, J.G., 1996b. Thresholds of toxicological concern based on structure-activity relationships. Drug Metabol. Rev. 28 (1/2), 209–217.

Munro, I.C., Renwick, A.G., Danielewska-Nikiel, B., 2008. The threshold of toxicological concern (TTC) in risk assessment. Toxicol. Lett. 180 (2), 151–156.

Na, M., Ritacco, G., O'Brien, D., Lavelle, M., Api, A., Basketter, D., 2020. Fragrance skin sensitization evaluation and human testing, dermatitis. https://doi.org/10.1097/DER.0000000000000684. November 16, 2020. Volume Publish Ahead of Print Issue. Retrieved from.

Nielsen, F., 2016. Hierarchical clustering. In: Introduction to HPC with MPI for Data Science. Undergraduate Topics in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-21903-5_8. Retrieved from.

Nishijo, T., Api, A.M., Gerberick, G.F., Miyazawa, M., Roberts, D.W., Safford, R.J., Sakaguchi, H., 2020. Application of the dermal sensitization threshold concept to chemicals classified as high potency category for skin sensitization assessment of

ingredients for consumer products. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 117, 104732.
Novak, J., Bahoo, L., Mitterogger, U., Franz, C., 2006. Composition of individual essential oil glands of savory (Satureja hortensis L., Lamiaceae) from Syria. Flavour Fragrance J. 21, 731–734.
OECD, 2017. OECD Series on Testing and Assessment, Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA). Retrieved from. https://doi.org/10.1787/9789264274815-en.
Patel, A., Joshi, K., Rose, J., Laufersweiler, M., Felter, S.P., Api, A.M., 2020. Bolstering the existing database supporting the non-cancer Threshold of Toxicological Concern values with toxicity data on fragrance-related materials. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 116, 104718.
Portoghese, P.S., 2009. Revision of purity criteria for tested compounds. J. Med. Chem. 52, 1-1.
RIFM (The Research Institute for Fragrance Materials), 2021. Private Correspondence with RIFM Natural Complex Substances (NCS) Core Team.
Roberts, D.W., Api, A.M., Safford, R.J., Lalko, J.F., 2015. Principles for identification of high potency category chemicals for which the dermal sensitisation threshold (DST) approach should not be applied. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 72, 683–693.
Safford, R.J., 2008. The Dermal Sensitisation Threshold- a TTC approach for allergic contact dermatitis. Regul. Toxicol. Pharmacol. : RTP 51, 195–200.
Safford, B., Api, A.M., Barratt, C., Comiskey, D., Daly, E.J., Ellis, G., McNamara, C., O'Mahony, C., Robison, S., Smith, B., Thomas, R., Tozer, S., 2015a. Use of an aggregate exposure model to estimate consumer exposure to fragrance ingredients in personal care and cosmetic products. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 72, 673–682.
Safford, R.J., Api, A.M., Roberts, D.W., Lalko, J.F., 2015b. Extension of the Dermal Sensitisation Threshold (DST) approach to incorporate chemicals classified as reactive. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 72, 694–701.
Safford, B., Api, A.M., Barratt, C., Comiskey, D., Ellis, G., McNamara, C., O'Mahony, C., Robison, S., Rose, J., Smith, B., Tozer, S., 2017. Application of the expanded Creme RIFM consumer exposure model to fragrance ingredients in cosmetic, personal care and air care products. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 86, 148–156.
Safford, R.J., Aptula, A.O., Gilmour, N., 2011. Refinement of the Dermal Sensitisation Threshold (DST) approach using a larger dataset and incorporating mechanistic chemistry domains. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 60, 218–224.
Salgueiro, L., Martins, A.P., Correia, H., 2010. Raw materials: the importance of quality and safety. A review. Flavour Fragrance J. 25, 253–271.
Salvito, D.T., Senna, R.J., Federle, T.W., 2002. A framework for prioritizing fragrance materials for aquatic risk assessment. Environ. Toxicol. Chem. 21, 1301–1308.
Schultz, T.W., Amcoff, P., Berggren, E., Gautier, F., Klaric, M., Knight, D.J., Mahony, C., Schwarz, M., White, A., Cronin, M.T., 2015. A strategy for structuring and reporting a read-across prediction of toxicity. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 72, 586–601.
Shen, J., Kromidas, L., Schultz, T., Bhatia, S., 2014. An in silico skin absorption model for fragrance materials. Food Chem. Toxicol. : Int. J. Publ. Br.Indust. Biol. Res. Assoc. 74, 164–176.
Vallat, A., Gu, H., Dorn, S., 2005. How rainfall, relative humidity and temperature influence volatile emissions from apple trees in situ. Phytochemistry 66, 1540–1550.
Waimer, F., Lang, F., Stumpf, H., 2007. The European Pharmacopoeia monograph extracts – implementation of the concepts for 'quantified' and 'standardised' extracts in industrial practice. Pharmeuropa Scientific Notes 1, 15–20.
Yang, C., Barlow, S.M., Muldoon Jacobs, K.L., Vitcheva, V., Boobis, A.R., Felter, S.P., Arvidson, K.B., Keller, D., Cronin, M.T.D., Enoch, S., Worth, A., Hollnagel, H.M., 2017. Thresholds of Toxicological Concern for cosmetics-related substances: new database, thresholds, and enrichment of chemical space. Food Chem. Toxicol. : Int. J. Publ. Br.Indust. Biol. Res. Assoc. 109, 170–193.