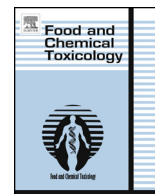




Contents lists available at ScienceDirect

Food and Chemical Toxicology

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

Criteria for the Research Institute for Fragrance Materials, Inc. (RIFM) safety evaluation process for fragrance ingredients

A.M. Api ^{a,*}, D. Belsito ^b, M. Bruze ^c, P. Cadby ^d, P. Calow ^e, M.L. Dagli ^f, W. Dekant ^g, G. Ellis ^h, A.D. Fryer ⁱ, M. Fukayama ^j, P. Griem ^k, C. Hickey ^{a,1}, L. Kromidas ^a, J.F. Lalko ^a, D.C. Liebler ^l, Y. Miyachi ^m, V.T. Politano ^a, K. Renskers ⁿ, G. Ritacco ^a, D. Salvito ^a, T.W. Schultz ^o, I.G. Sipes ^p, B. Smith ^l, D. Vitale ^{a,2}, D.K. Wilcox ^a

^a Research Institute for Fragrance Materials, Inc., 50 Tice Boulevard, Woodcliff Lake, NJ 07677, USA

^b Member RIFM Expert Panel, Columbia University Medical Center, Department of Dermatology, 161 Fort Washington Ave., New York, NY 10032, USA

^c Member RIFM Expert Panel, Malmo University Hospital, Department of Occupational & Environmental Dermatology, Sodra Forstadsgatan 101, Entrance 47, Malmo SE-20502, Sweden

^d Firmenich SA, Case Postale 239, 1211 Geneva 8, Switzerland

^e Member RIFM Expert Panel, University of Nebraska Lincoln, 230 Whittier Research Center, Lincoln, NE 68583-0857, USA

^f Member RIFM 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 RIFM Expert Panel, University of Wuerzburg, Department of Toxicology, Versbacher Str. 9, 97078 Würzburg, Germany

^h Givaudan International SA, 5 chemin de la parfumerie, CH1214 Vernier, Switzerland

ⁱ Member RIFM Expert Panel, Oregon Health Science University, 3181 SW Sam Jackson Park Rd., Portland, OR 97239, USA

^j International Flavors & Fragrances Inc., 800 Rose Lane, Union Beach, NJ 07735, USA

^k Symrise AG, Muehlenfeldstrasse 1, D-37603 Holzminden, Germany

^l Member RIFM 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

^m Member RIFM Expert Panel, Department of Dermatology, Kyoto University Graduate School of Medicine, 54 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

ⁿ Takasago International Corporation, 4 Volvo Drive, Rockleigh, NJ 07647, USA

^o Member RIFM Expert Panel, The University of Tennessee, College of Veterinary Medicine, Department of Comparative Medicine, 2407 River Dr., Knoxville, TN 37996-4500, USA

^p Member RIFM Expert Panel, Department of Pharmacology, College of Medicine, University of Arizona, 1501 North Campbell Avenue, P.O. Box 245050, Tucson, AZ 85724-5050, USA

ARTICLE INFO

Article history:

Received 4 November 2014

Received in revised form 13 November 2014

Accepted 17 November 2014

Available online

Keywords:

Fragrance material

Safety assessment

Human health toxicology

Environmental toxicology

ABSTRACT

The Research Institute for Fragrance Materials, Inc. (RIFM) has been engaged in the generation and evaluation of safety data for fragrance materials since its inception over 45 years ago. Over time, RIFM's approach to gathering data, estimating exposure and assessing safety has evolved as the tools for risk assessment evolved. This publication is designed to update the RIFM safety assessment process, which follows a series of decision trees, reflecting advances in approaches in risk assessment and new and classical toxicological methodologies employed by RIFM over the past ten years. These changes include incorporating 1) new scientific information including a framework for choosing structural analogs, 2) consideration of the Threshold of Toxicological Concern (TTC), 3) the Quantitative Risk Assessment (QRA) for dermal sensitization, 4) the respiratory route of exposure, 5) aggregate exposure assessment methodology, 6) the latest methodology and approaches to risk assessments, 7) the latest alternatives to animal testing methodology and 8) environmental risk assessment. The assessment begins with a thorough analysis of existing data followed by *in silico* analysis, identification of 'read across' analogs, generation of additional data through *in vitro* testing as well as consideration of the TTC approach. If necessary, risk management may be considered.

© 2014 Elsevier Ltd. All rights reserved.

* Corresponding author. Research Institute for Fragrance Materials, Inc., 50 Tice Boulevard, Woodcliff Lake, NJ 07677, USA. Tel.: +1 201 689 8089 ext 103; fax: +1 201 689 8089.

Email address: amapi@rifm.org (A.M. Api).

¹ Present address: Firmenich Inc., Post Office Box 5880, Princeton, NJ 08543, USA.

² Present address: Energizer Personal Care, LLC, 75 Commerce Drive, Allendale, NJ 07401, USA.

1. Introduction

Fragrance materials are used in a wide variety of consumer products including both personal care and household products. Fragrance compounds (also called fragrance mixtures or fragrance oils) are formulations consisting of specific combinations of individual materials or mixtures. Consumer exposure to fragrance materials ranges from skin contact to inhalation. To help ensure the safe use of fragrance materials, the Research Institute for Fragrance Materials (RIFM) was founded. Its mission is to:

1. Engage in research and evaluation of fragrance materials through an independent Expert Panel.
2. Determine safety in use.
3. Gather, analyze, and publish scientific information.
4. Distribute scientific data and safety assessment judgments to RIFM members, industry associations and other interested parties.
5. Maintain an active dialogue with official international agencies.

In 2000, RIFM published its process for assessing the safety of fragrance materials (Ford et al., 2000). This process was further refined and its detailed application to the safety assessment of fragrance materials was documented in a 2003 publication by the RIFM Expert Panel (Bickers et al., 2003). Importantly, many of the fundamental criteria outlined in these documents are still applicable today. The objective of this work is to update the process to:

- incorporate new scientific information that includes a framework for choosing structural analogs and groups (Wu et al., 2010),
- add consideration of the Threshold of Toxicological Concern (TTC) (Kroes et al., 2004, 2007),
- add consideration of the Quantitative Risk Assessment for dermal contact sensitization (QRA) (Api et al., 2008),
- add consideration of the respiratory route of exposure,
- update exposure assessment methodology,
- incorporate the latest methodology and approaches to risk assessments,
- incorporate an intelligent testing strategy which includes appropriate use of alternatives to animal testing methodology, and
- incorporate the current state of environmental risk assessment in support of the International Fragrance Association (IFRA) Standards.

The original criteria document was developed in part in response to regulatory changes. Similarly, with the implementation of REACH (the European regulation on Registration, Evaluation, Authorization and Restriction of Chemicals) (REACH, 2006) substantial additional data are becoming available for chemicals including fragrance materials. While it is theoretically possible to evaluate each and every chemical, there are ethical and practical considerations such as the aim to minimize animal use and testing laboratory capacity that drive the need to use data for one or more compounds to support related chemicals that do not have sufficient data.

The first step in the approach outlined in the original criteria document is to prioritize materials for review by evaluating volume of use, exposure, and chemical structure. Prioritization of assessments is more heavily weighted on direct consumer exposure than on volume of use. Although high volume of use suggests the potential for high human exposure, there are instances where high volume fragrance materials are used in products that result in relatively low human exposure. Conversely there may be lower volume materials that, in part due to their scent characteristics, are used in products with relatively high exposure potential. In addition, other factors to be considered in prioritization of assessments include existing data of concern and/or need for additional information on one

or more toxicological endpoints under review and/or regulatory requirements.

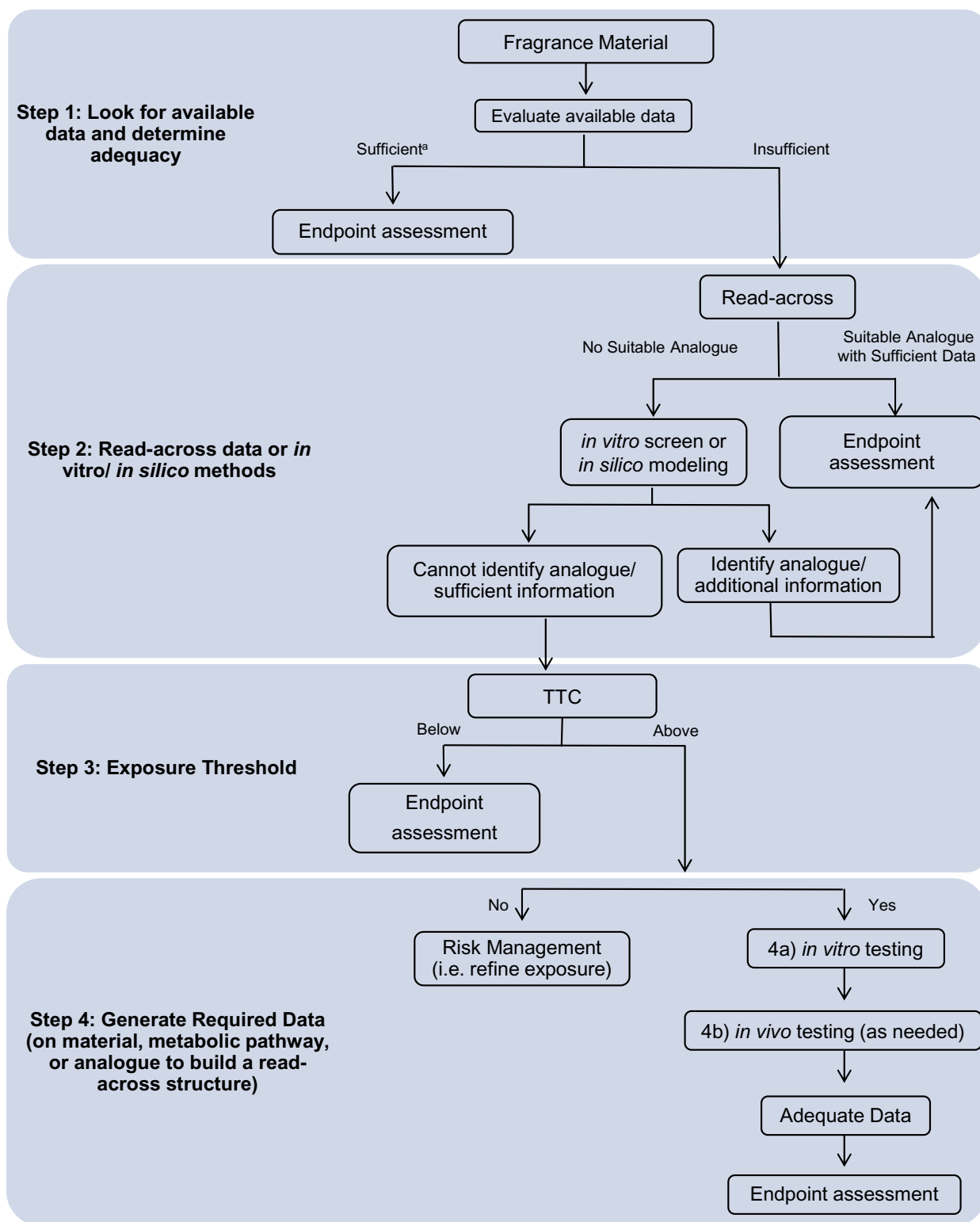
Implementation of REACH in the European Union has in essence resulted in a volume-based “prioritization” of chemicals, including fragrance materials, for review. Dossiers for many of the highest volume fragrance materials (>1000 tons and 100–1000 tons) have been submitted to the European Chemicals Agency (ECHA) for review and dossiers for the lower volume materials are or will be prepared on an ongoing basis. Currently, Robust Study Summaries for most of the materials submitted to ECHA are publicly available (unless accepted by ECHA as Confidential Business Information or CBI) and more will become available as registration of materials in the lower volume bands continues. This further emphasizes the need for careful evaluation since the summaries are available to non-governmental organizations or NGOs and regulatory authorities.

The primary objectives of this update are to outline the steps for a process to develop a complete toxicological profile for a fragrance material, to identify data needs and develop a preliminary exposure assessment to be used in a risk assessment. The exposure and risk assessment of any fragrance material should be an iterative process that incorporates the available hazard data for the key toxicological endpoints coupled with the exposure assessment. Key toxicological endpoints include genotoxicity, repeated dose toxicity, developmental and reproduction toxicity, skin sensitization, phototoxicity, local inhalation effects and environmental considerations. Hazard and exposure evaluations can be developed almost simultaneously since low exposures may permit use of the Threshold of Toxicological Concern (TTC) approach or evidence of a specific toxicity concern may indicate the need for additional data or a decision not to use the material. Another consideration is that since higher volume materials are potentially more data-rich it may be possible to build read-across Structure Activity Relationship (SAR) arguments supporting one or more of the toxicological endpoints for a lower volume material by using data from the high volume material.

Any safety assessment must consider both the human and the environmental impact of a material. As such, the environmental assessment is an integral part of a safety assessment process. In addition, RIFM is responsible for the environmental safety assessment of fragrance materials. RIFM routinely screens for potential impacts to the freshwater aquatic environment since 1999. The processes for assessing human health and environmental safety, while not identical, are complementary in their design following a tiered screening approach to set safety assessment priorities. The published “RIFM Environmental Framework” (Salvito et al., 2002) provides the model used for this effort. It is a conservative model comparing a ‘down the drain’ discharge concentration (through wastewater treatment) with an estimated effect on fish using a large uncertainty factor to avoid false negatives in the use of this screening tool. It is comprised of scenarios for both Europe and North America. While there are no significant changes to the process for environmental safety assessment of fragrance materials, it is presented here for completeness.

A decision-scheme outlining the general steps needed for an overall evaluation to draw conclusions regarding acceptable exposures to a fragrance material is shown in Fig. 1. Similar decision schemes that incorporate endpoint specific considerations are described later. In general, the first step in the process is to gather all available relevant data for the material under consideration. These data should be evaluated for scientific robustness, including whether the study-type and protocol used are adequate and the test material was adequately characterized.

Safety assessments of materials used in fragrances should be carried out by evaluating the available data for relevant toxicological endpoints for local and systemic effects, including (but not limited



^a Acceptability of existing data are evaluated for each endpoint base on established methods (i.e. OECD guidelines), accepted evaluation criteria (i.e. Klimisch et al. 1997). For studies which do not necessarily meet accepted guidelines, a weight of evidence approach may be considered.

Fig. 1. Overall process flow chart.

to): genotoxicity/carcinogenicity, reproductive and developmental toxicity, repeated dose toxicity, skin sensitization, respiratory toxicity, phototoxicity and environmental effects. These data are put into context with the expected exposure from various fragranced

products via the dermal route (including both leave on and rinse off applications) and from inhalation exposure (Bickers et al., 2003). Oral exposure is also relevant for fragrance materials that are used in oral care products, such as toothpaste and mouthwash.

As mentioned above, the approach that RIFM utilizes for the human health safety assessment of fragrance materials has been previously described (Bickers et al., 2003; Ford et al., 2000). The fundamentals of this published review process remain intact, such as the reliance on the RIFM Expert Panel. This present publication documents the changes in the process to assure the safe use of fragrance materials as significant advances take place in the science used in safety evaluations. One of the primary areas of advancement in safety evaluations is the increased applicability and acceptance of *in vitro* and *in silico* methods to evaluate toxicological endpoints. This publication includes RIFM's incorporation of these new methodologies into its evaluation process, as well as a greater focus on individual fragrance materials than in past assessments that were developed and published as summaries of structurally similar fragrance materials (Belsito et al., 2011). This new process does not abandon the application of the principle of read-across to support materials but begins with the material under evaluation. The identification and selection of structural analogs will be determined taking into account physical-chemical properties and metabolism as described by Wu et al. (2010).

Fragrance materials assessed by RIFM have multiple uses and those vary significantly among fragranced products. It is also necessary to consider all toxicological endpoints during the safety assessment process. This document outlines: 1) how each toxicological endpoint will be considered in the safety assessment; 2) where and when data or testing are required; and 3) how the data and/or testing results will be incorporated into the safety assessment. This guidance will be used for safety assessments on fragrance materials performed by RIFM.

2. Evaluating the toxicological profile of a fragrance material: how the endpoint assessment process is structured

This evaluation process employs a tiered testing strategy. As shown in Fig. 1, the process provides an overall profile for the material of interest and operates for each of the different endpoints through a series of steps. This tiered testing strategy is followed for each key toxicological endpoints – genotoxicity, repeated dose toxicity, developmental and reproduction toxicity, skin sensitization, phototoxicity, local inhalation effects and environmental considerations:

(i) STEP 1:

Q. Is sufficient information available on the endpoint to complete a risk assessment and is it of sufficient quality?

YES: → Complete an endpoint assessment

NO: → Proceed with Step 2

(ii) **STEP 2:** Examine metabolism, chemistry, physico-chemical properties, and toxicological properties for read-across to analogs that have been identified and judged appropriate for each endpoint. *In vitro* screening assays and *in silico* tools/modeling (i.e., structural similarities, reactive groups, etc.) can also be used to support identification of analogs or, where appropriate, as the basis for the preparation of the endpoint assessment.

Q. Can one or more appropriate analogs be identified or is sufficient information provided from the *in vitro* screening assays and/or *in silico* models?

YES: → Complete an endpoint assessment

NO: → Proceed with Step 3

NOTE: It may be possible, based on existing data, to skip Step 2 and proceed directly to Step 3 to complete an Endpoint Assessment.

(iii) **STEP 3:** Consider the application of the Threshold of Toxicological Concern (TTC) concept (Kroes et al., 2007).

Q. Is the exposure below the critical threshold?

YES: → Complete an endpoint assessment

NO: → Proceed with Step 4

(iv) **STEP 4:** Generate the data required to complete an endpoint assessment or proceed to risk management. Risk management may include a refinement of exposure (e.g., use level) to support an adequate margin of exposure. Data generation may include data on the material itself, metabolic pathways, or other materials to build an appropriate read-across structure.

NOTE: For Step 4 studies generating additional needed information will be carried out in the following order:

Step 4 a) *In vitro* studies

Step 4 b) *In vivo* studies

Every possible *in vitro* approach to generate the necessary data will be carried out before consideration of *in vivo* studies.

3. Important considerations for the RIFM safety assessment process

3.1. Exposure

Exposure is an essential part of the safety assessment process and is required in order to conduct a safety assessment. Fragrances are used in a wide variety of products including decorative cosmetics, fine fragrances, shampoos, toilet soaps, and other toiletries as well as in other consumer products such as household cleaners, detergents, oral and air care products. RIFM has access to two types of exposure data on fragrance. The first is volume of use data, which is provided by the IFRA approximately every two years through a comprehensive survey of IFRA and RIFM member companies that manufacture fragrances. The second method is an aggregate exposure model using deterministic and probabilistic exposure data to describe real life consumer exposure to a specific fragrance material (Comiskey et al., 2014; Safford et al., 2014a). This model will address exposure from all routes, including that from inhaled products.

As appropriate, more detailed information on inhalation exposure can be calculated separately using the 2-Box Air Dispersion Model (Petry et al., 2013). This model is an indoor environment air model that characterizes the dispersion of a single chemical inside two connected, enclosed zones and determines air exposure concentrations from both far field (applied to products released to the air but not specifically toward the body) and near field (applied to products that are intentionally sprayed toward the body) analyses. The 2-Box Air Dispersion Model is a unique hybrid of the Dutch National Institute for Public Health and the Environment, RIVM Consumer Exposure (ConsExpo) model and the U.S. Environmental Protection Agency (EPA) Multi-Chamber Chemical Exposure Model (MCCEM). Model output includes a temporal profile of chemical concentrations in the two zones.

3.2. Threshold of toxicological concern (TTC)

The TTC approach is based on the concept that reasonable assurance of safety can be given, even in the absence of chemical-specific toxicity data, provided the exposure is sufficiently low, i.e. that an exposure level can be defined below which there is no significant risk to human health. The TTC is based on the Threshold of Regulation, FDA's priority-based assessments of food additives (Hattan and Rulis, 1986), which was expanded to include consideration of the chemical structure in conjunction with toxicity data (Kroes et al., 2004; Munro et al., 1996). These analyses originally focused on systemic exposure following oral administration. More recently, the TTC approach was extended to consider systemic exposure following topical application of cosmetic products including

the use of default skin penetration values (Blackburn et al., 2005; Kroes et al., 2007). In 2012, a joint opinion from the European Scientific Committees (Scientific Committee on Consumer Safety (SCCS), Scientific Committee on Health and Environmental Risks (SCHER), Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR)) considered the TTC approach, in general, scientifically acceptable for human health risk assessment of systemic toxic effects caused by chemicals present at very low levels, as based on sound exposure information (SCCP, 2012).

The TTC concept has also been applied to evaluating potential skin sensitizers. The dermal sensitization threshold (DST) establishes a level below which there is no appreciable risk for the induction of sensitization, and is based on a probabilistic analysis of potency data for a diverse range of known chemical allergens (Safford, 2008; Safford et al., 2011). There has also been the suggestion that TTC can be applied to inhalation exposure and risk assessment (Carthew et al., 2009; Drew and Frangos, 2007; Escher et al., 2010; Kroes et al., 2007). With respect to inhalation exposure another important consideration is the potential for site of contact (local) effects in all parts of the respiratory tract.

3.3. Read-across

Read-across is an important technique utilized by RIFM to estimate missing data for a single or limited number of chemicals using an analog approach. In principle, read-across utilizes common endpoint information, including physicochemical properties and toxicity, for one (or more) chemical(s) to make a prediction on the same endpoint for another chemical. It may be performed in a qualitative or quantitative manner. This process can help to avoid the need to carry out specific tests on every substance for every endpoint. The criteria for providing sufficient information via read-across will be specific to each analogous set of chemicals, and may be specific to each endpoint. Analogous sets of chemicals are selected based on structural, reactivity, metabolic and physicochemical similarities (Blackburn et al., 2011; Wu et al., 2010). Structural analogs can be found by expert review of other chemical substances in combination with the OECD QSAR Toolbox (OECD, 2014), or other computational models as appropriate.

As described by the European Chemicals Agency (ECHA, 2010b), within an analogous set of chemicals, specific endpoint assessments can be achieved by performing read-across in the following ways to fill data gaps: 1) one-to-one (one analog used to make an estimation for a single chemical), 2) many-to-one (two or more analogs used to make an estimation for a single chemical), 3) one-to-many (one analog used to make estimations for two or more chemicals), or 4) many-to-many (two or more analogs used to make estimations for two or more chemicals). There is generally more confidence in a read-across when data from more than one source chemical are considered (e.g. many-to-one) or more than one *in silico* tool agrees. Furthermore, when considering data points from multiple source chemicals, a clear trend may be more apparent in which a conservative value for the target chemical may be identified, or a weight of evidence approach may be used to increase the confidence in filling a data gap. To address the uncertainty surrounding read-across, assessment factors may be considered (e.g. if there is higher uncertainty surrounding the read-across, then the more conservative approach for extrapolating data is to increase the assessment factor when calculating the safe exposure levels (i.e. margin of exposure)). Alternatively, defining a range of values rather than a specific data point may be sufficient.

3.4. Adequacy of data

Step 1 in all endpoint assessments examines the acceptability of any existing data. Focus will always be on studies carried out ac-

cording to established methods and latest guidelines (OECD, 2013) and that are considered reliable based on accepted evaluation criteria such as Klimisch scores (Klimisch et al., 1997). The quality of the test material for all studies evaluated should be defined and correspond to fragrance materials in commerce. For studies which do not necessarily meet accepted guidelines, a weight of evidence approach may be considered.

4. Human health endpoints

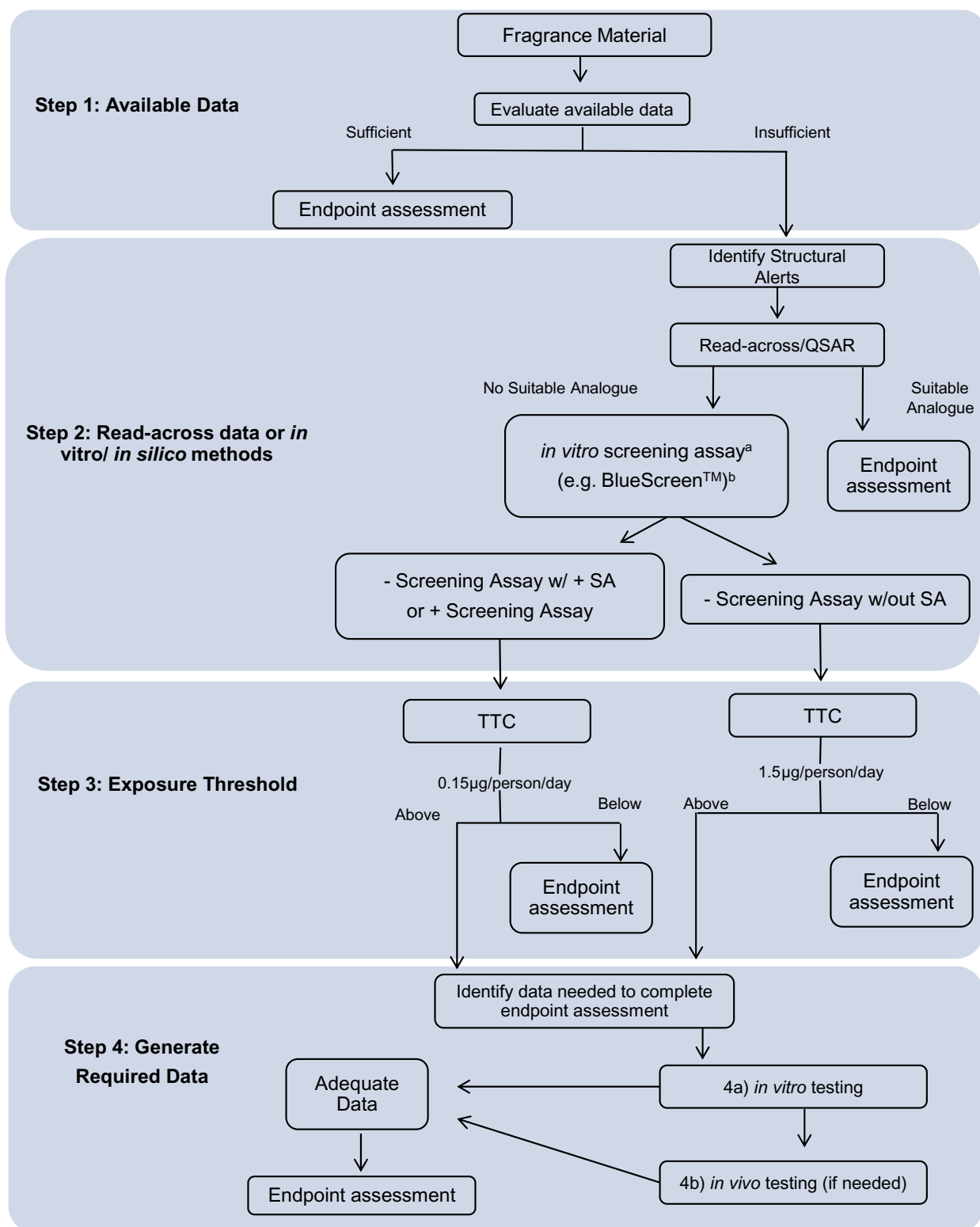
4.1. Genotoxicity

The safety assessment of potential genotoxic hazard for fragrance materials is carried out through a series of steps according to the flow chart shown in Fig. 2.

Step 1 examines the acceptability of any existing data on the fragrance material 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 understood that no single assay can be utilized to predict genotoxic effects to humans; rather a combination of tests which 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 which cover one of the key endpoints: gene mutation or clastogenicity. For fragrance materials that are judged by **Step 1** to have sufficient data, an endpoint assessment is conducted without further testing needs. Concerning genotoxicity it has to be taken into account that the current battery of *in vitro* tests (Ames, and *in vitro* clastogenicity) has a very low specificity; it has been shown that up to 90% of non-carcinogens are positive in at least one of these tests (Kirkland et al., 2005). Therefore, results of *in vivo* genotoxicity tests should be given a greater weight of evidence than respective *in vitro* results.

In cases where there are no data or existing test data are considered insufficient to evaluate the genotoxic potential of a fragrance material, **Step 2** is applied. **Step 2** provides detailed examination of the fragrance material utilizing *in silico* tools, *in vitro* assays, and read-across. First, the presence of known structural alerts (i.e. molecular substructures or reactive groups) associated with genotoxic properties is evaluated. All fragrance materials are assessed against the structural alerts identified by the methods of Ashby and Tennant (1991) and Benigni and Bossa (2008). In some cases metabolites of the fragrance material in question may be determined. In such cases the literature is mined to identify any genotoxicity data on these metabolites. Known impurities should also be reviewed for genotoxicity. Alternatively, a wide range of *in silico* tools may be used to assess the genotoxic potential of a material (e.g., DEREK, MultiCASE, Oncologic, TOPKAT, TIMES and OECD toolbox). Some of these are based on more extensive lists of structural alerts than the Ashby and Tennant, or Benigni and Bossa structural alerts, and others are derived from expert knowledge, data from bioassays and other models that use molecular descriptors to form rules for prediction. Several of the available models are based on the potential for a chemical to react with DNA, and therefore they have been shown to best correlate with Ames test data (Benigni et al., 2010). It is important to note that there are few models which have been designed to predict *in vivo* genotoxicity, or identify other genotoxic mechanisms than DNA reactivity. Furthermore, the output from *in silico* tools is not always straightforward and therefore may require expert judgment.

A range of key principles has been identified to assess the adequacy of model predictions (Worth et al., 2010). In particular, one key criterion which must be demonstrated for any evaluation



^a In the absence of a screening assay, a TTC level of 0.15 µg/person/day can be applied.

^b BlueScreen™ test (Gentronix, Manchester, UK) is a screening tool currently being used as a high throughput screening assay to provide further weight of evidence as to genotoxicity potential.

Fig. 2. Genotoxicity flow chart.

performed is that the substance is within the applicability domain of the respective model and the prediction is reliable for the class of chemical being assessed. The current status of available software models has recently been reviewed (Serafimova et al., 2010),

and the applicability of selected models in predicting the genotoxic potential of specific substances has also been evaluated (Worth et al., 2010). While the use of *in silico* tools can indicate a potential genotoxic hazard it is important to note that 'no alert' does not mean

a material does not possess a genotoxic hazard as the chemical may fall outside the applicability domain of the model. The structural alerts will be used to prioritize the evaluation of fragrance materials using other methodologies.

The identification of structural alerts, together with the evaluation of potential transformation, assists in the evaluation of potential analogs for read-across and QSAR. Identification of suitable analog(s) may provide sufficient data for the fragrance material of interest, in which case an endpoint assessment is conducted with the information available. In addition to structural alerts, potential chemical reactivity and transformation should be evaluated based on knowledge and evaluation of the material's structure as outlined by Wu et al. (2010).

For fragrance materials where suitable analogs are not identified, and/or do not provide sufficient data, a high throughput screening assay may be utilized to provide further weight of evidence as to the genotoxic potential of the material, provide anchoring data for read-across, and identify fragrance materials for further evaluation. For example, BlueScreen™ test (Gentronix, Manchester, UK) is a screening tool currently being used in this capacity. This assay provides information on multiple genotoxic mechanisms and can be performed with and without metabolic activation (Billinton et al., 2008; Birrell et al., 2010; Hastwell et al., 2006, 2009; Hughes et al., 2012; Knight et al., 2009). Following this evaluation step the material moves into Step 3 which involves evaluating the exposure threshold. The total systemic exposure to a fragrance ingredient is determined using an aggregated exposure model explained by Comiskey et al., 2014 and Safford et al., 2014.

Fragrance materials identified to be negative in the screening assay and which have no structural alerts move into **Step 3** using a default TTC value for non-genotoxic materials of 1.5 µg/person/day. This level corresponds to the threshold of regulation derived by the US Food and Drug Administration (Rulis, 1986, 1989, 1992) to be applied to substances that do not contain a structural alert for genotoxicity/carcinogenicity, but intended to protect against all types of toxicity including carcinogenicity. Therefore, an endpoint assessment can be conducted on a fragrance material which has no structural alert, and a negative result in the screening assay, and has been identified as having a consumer exposure threshold below 1.5 µg/person/day or 0.025 µg/kg/day.

For fragrance materials which exhibit: 1) a structural alert, 2) no structural alert, but a positive screening assay Bluescreen™; or 3) structural alert and no screening data, the material is assessed using a default TTC value for potentially genotoxic materials of 0.15 µg/person/day. The default TTC value for potentially genotoxic materials of 0.15 µg/person/day 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. If the substance were to be tested and shown to be a true genotoxic carcinogen this would result in an expected Margin of Exposure (MOE) of $\geq 100,000$ (Benford et al., 2010; Cheeseman et al., 1999; EFSA, 2012). An additional level of conservatism is built into our system by the exclusion of structures that are closely related to recognized highly potent genotoxic carcinogens (i.e. aflatoxin-like, azoxy- or N-nitroso-compounds), which would also pass directly to **Step 4** (Kroes et al., 2004). For these substances the upper bound lifetime risk for cancer is estimated to be greater than one in a million even at an exposure of 0.15 µg/person/day.

An illustration of the conservatism built into the TTC values was provided by Munro (1990). Using linear extrapolation of TD_{50} (carcinogenic dose to 50% of the test animals) results from a subset of the data in the CPDB, Munro demonstrated a more conservative estimate of 10^{-6} risk level than would be estimated by other models. The probability of exceeding the risk level depends on the likelihood that a chemical is actually carcinogenic and its potency. When

assuming that as much as 10% of all chemicals are genotoxic carcinogens, the probability of any untested chemical being a carcinogen with a virtual safe dose (VSD) below 1.5 µg/person per day was determined to be 4%; the corresponding percentage for the lower value of 0.15 µg/person per day was 1% (Fung et al., 1995). These estimates also make a worst-case assumption that any untested substance that is a carcinogen would have a similar potency to the 15% most potent carcinogens in the CPDB, which is unlikely for fragrance ingredients given the structural dissimilarities between them and these carcinogens. Therefore, there is very low probability that untested substances below the TTC value of 0.15 µg/person/day pose any appreciable cancer risk to humans. Data also show that the TTC value of 0.15 µg/person/day likely also covers heritable effects (EFSA, 2012).

If sufficient data cannot be obtained to make a valid assessment, additional evaluation is required in **Step 4**. The next step in this process is an *in vitro* analysis following OECD or ICH guidelines as appropriate (Step 4a). The recommended assays can be adapted based on the assessment but are typically:

- 1) An *in vitro* gene mutagenicity assay (i.e. Ames (OECD, 1997))
- 2) An *in vitro* cytogenicity assay (i.e. *in vitro* Micronucleus Test (OECD, 2010b))

The recommendation of the above *in vitro* tests is based on the analysis of Kirkland et al. (2005, 2011).

Positive results in one or both of these tests would trigger further testing, and even when the *in vitro* tests are negative, further testing may be required (i.e. where there is a pattern of positive *in vivo* genotoxicity data for the material's chemical class). The next step (4b) may include a number of additional types of evaluation, including *in vivo* analysis, to generate sufficient information to complete the endpoint assessment.

The process for the tiered assessment of the endpoint of genotoxicity is summarized in Fig. 2.

4.2. Repeated dose toxicity

A systemic effect is an effect that is observed distant from the initial site of contact after the agent is absorbed. This would include non-genotoxic carcinogenicity. The risk of repeated dose toxicity of all chemically-defined fragrance materials is assessed through a series of steps according to the flow chart in Fig. 3.

Step 1 examines the adequacy of any existing test data carried out according to established methods. In addition to data from subchronic toxicity studies, other data for consideration include short-term repeated dose, acute, reproduction, metabolism, and skin absorption studies. This includes the type of effect e.g. adaptive versus adverse (Lewis et al., 2002; Williams and Iatropoulos, 2002) and the slope of the dose–response curve. All available data should be evaluated to determine the NOAEL for repeated dose toxicity. In cases where only a Lowest-Observable-Adverse-Effect-Level (LOAEL) is available, the data will be evaluated to see whether a NOAEL can be extrapolated. The European Chemicals Agency supports an extrapolation factor of 3 (minimum, majority of cases) to 10 (maximum, exceptional cases) to determine a from a LOAEL (ECHA, 2010a). In determining the point of departure for the endpoint assessment, route-to-route differences also need to be taken into account.

An endpoint assessment for repeated dose toxicity will be generated for substances with data that are judged by **Step 1** to be sufficient. A MOE is calculated based on the ratio of the NOAEL from the repeated dose study to the aggregate exposure level (see Section 3A). MOE values below 100 have been used by regulatory agencies as flags for further evaluation (Faustman and Omenn, 2008). Historically, the 100-fold factor was introduced in the United States

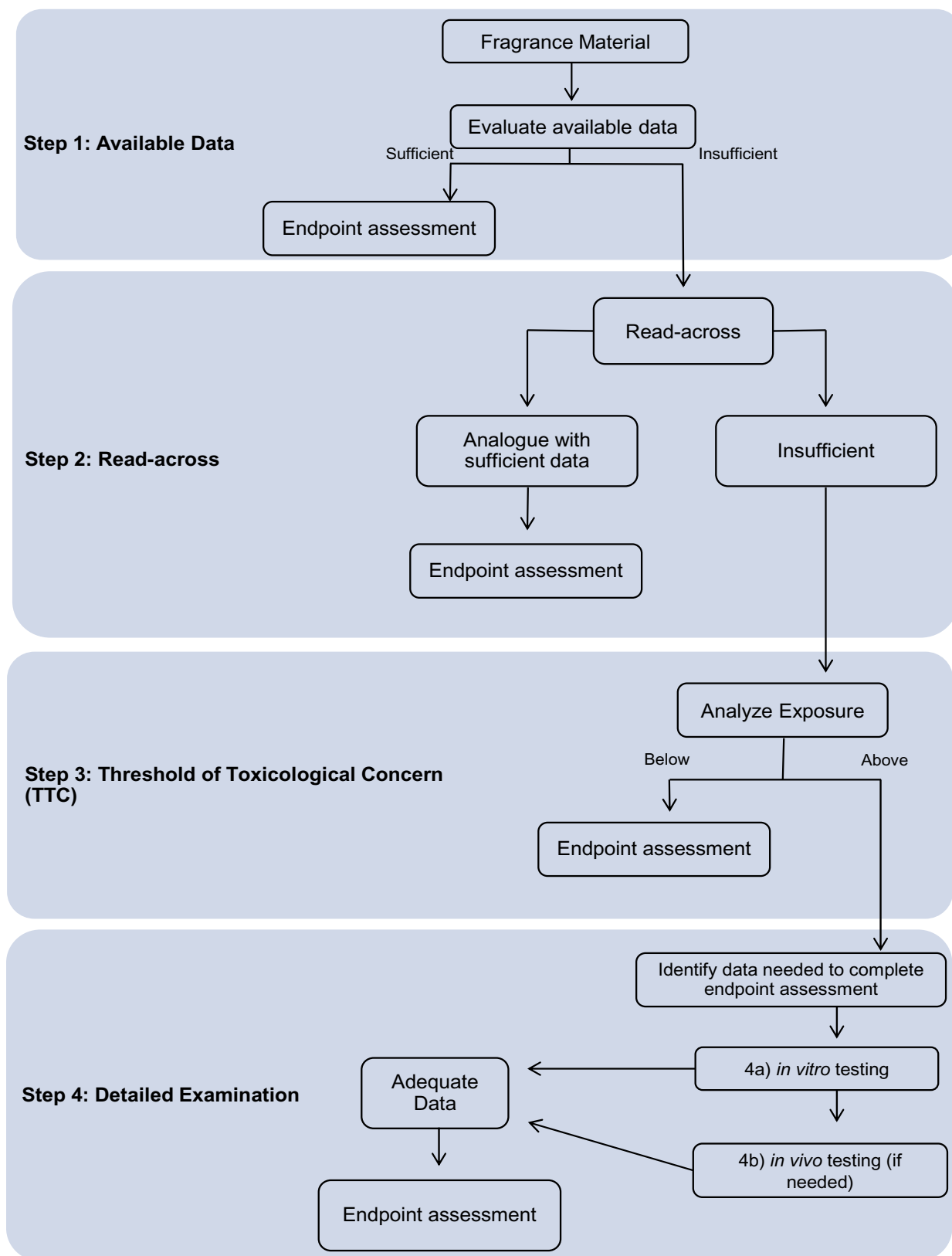


Fig. 3. Repeated dose, developmental, and reproductive toxicity flow chart.

in the mid-1950's in response to legislative guidelines for food additives. The 100-fold factor was comprised of a factor of 10 to reflect the hypothesized increased sensitivity of a relative to laboratory test animals and an additional factor of 10 to take into account inter-

individual variability. This fundamental approach has been adopted into guidelines and recommendations by several international agencies and governmental bodies (ECETOC, 1995). If the MOE is insufficient, the endpoint assessment can be refined using

data-derived assessment factors rather than the defaults. In some instances, absorption and/or metabolism data to supplement the existing data may be generated to complete the assessment. In these instances the additional data may enable refinement of the exposure assessment from Step 1.

In cases where existing test data are considered insufficient or there are no test data to characterize the repeated dose effects of the material, **Step 2** is applied. For certain materials read-across analysis may be appropriate. Structural analogs may be identified by expert review of similar fragrance materials, OECD QSAR Toolbox (OECD), EPA Analog Identification Method, or other appropriate computational models. In this case, all available data on other fragrance materials in the same structural class and, additionally, any non-fragrance materials that are identified to be structurally related that have data should be reviewed. Data could include, but are not limited to, subchronic toxicity studies (normally 90-day repeated dose studies in rats), screening studies addressing repeated dose toxicity, knowledge of available metabolic pathways, and skin absorption. These data will be closely reviewed to decide if a read-across NOAEL can be derived for the purposes of completing an endpoint assessment, which could include the use of an additional extrapolation factor.

When an endpoint assessment cannot be completed in Step 2, the material will move to **Step 3**. **Step 3** uses an exposure based threshold such as the Threshold of Toxicological Concern (TTC). The fragrance material is assigned to a Cramer class, provided it has no genotoxicity structural alerts or those alerts have been addressed by testing or read-across (Cramer et al., 1978), based on its structure, and the specific class (I, II or III) is used to assign a level for the TTC. Cramer Classes I, II, and III have limits of 1800, 540, and 90 µg/person/day, which correspond to exposures of 30, 9, and 1.5 µg/kg bw/day, respectively (Kroes et al., 2007). If the daily aggregate exposure is below the respective TTC value, the repeated dose toxicity assessment endpoint is complete and the fragrance material is considered safe for this endpoint at the current use levels.

If the exposure exceeds the TTC the fragrance material will move to **Step 4** where the data needed to complete the endpoint assessment are generated. The battery of tests could include in a tiered approach: Step 4a *in vitro* metabolism and/or *in vitro* skin absorption, Step 4b *in vivo* toxicokinetics, and/or *in vivo* 90-day subchronic toxicity.

Systemic or repeated dose toxicity has been considered as a far too heterogeneous and complex endpoint to be encoded in a single predictive model (Adler et al., 2011). The status of possible alternatives to animal testing, such as *in vitro* assays, will continue to be monitored by RIFM. Currently, the European Centre for the Validation of Alternative Methods (ECVAM) has not scientifically validated any alternative models for repeated dose toxicity (ECVAM, 2012). In the United States, the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) have at present no validated methods for subchronic repeated dose toxicity (ICCVAM, 2012). Potential alternative approaches (both *in vitro* and *in vivo*) for assessing repeated dose toxicity, with a focus on cosmetic use, were thoroughly reviewed by Adler et al. (2011). Although no full replacement alternatives for *in vivo* studies for repeated dose toxicity are currently available, alternative methods and the use of integrated testing strategies have been utilized to refine and reduce the use of laboratory animals. RIFM will continue to investigate the potential of SARs (e.g. OECD QSAR Toolbox, DEREK, MultiCASE, ToxCast, TOPKAT, TIMES etc.) for their scientific attributes concerning repeated dose endpoints and consider embracing methodologies as they are scientifically supported.

The process for the tiered assessment of the endpoint of repeated dose toxicity is summarized in Fig. 3.

4.3. Developmental and reproductive toxicity

Developmental toxicity refers to effects on growth and developmental retardation, malformations, and functional deficits in the fetuses, neonates, and maturing offspring. Reproductive toxicity refers to, but is not confined to, effects such as reduced fertility, effects on gonads, oogenesis, spermatogenesis, and general disturbances to the reproductive cycle. The risk of both developmental and reproductive toxicity of all chemically-defined fragrance materials is assessed through a series of steps according to the flow chart shown in Fig. 3.

Step 1 examines the adequacy of any existing data carried out according to established methods and considered reliable. Data from developmental and reproductive studies, supporting data such as subchronic, metabolism, and skin absorption studies, are considered. For example, a very large margin of exposure from a subchronic test could significantly reduce concern regarding developmental or reproductive effects (Janer et al., 2007). Studies may include a rodent 1- or 2-generation reproductive toxicity study and a developmental toxicity study following the most recent guidelines (OECD, 1983, 2001). The type of effect e.g. adaptive versus adverse (Lewis et al., 2002; Williams and Iatropoulos, 2002) and the slope of the dose-response curve should also be considered. All available data should be evaluated to determine the NOAELs for both developmental and reproductive toxicity endpoints, including data on the reproductive organs generated in a subchronic study. In cases where only a LOAEL is available, the data will be evaluated to see whether a NOAEL can be extrapolated. The European Chemicals Agency supports an extrapolation factor of 3 (minimum, majority of cases) to 10 (maximum, exceptional cases) to determine a NOAEL from a LOAEL (ECHA, 2010a). In determining the point of departure for the endpoint assessment, route-to-route differences also need to be taken into account. An endpoint assessment for both the reproductive and developmental toxicity endpoints will be generated for substances with data that are judged by **Step 1** to be sufficient.

As with Repeated Dose endpoint analysis, in some instances additional information to complete the endpoint assessment may be obtained by generating absorption and/or metabolism data to supplement the existing exposure data in combination with existing data relevant to the repeated dose endpoint. In these instances the additional data may enable refinement of the MOE calculation in Step 1.

In cases where existing data are considered insufficient, **Step 2** is applied. In **Step 2**, read-across may be used for those fragrance materials for which there are sufficient developmental and reproductive toxicity data on close structural analogs. All available data on other fragrance materials in the same structural class will be reviewed and, additionally, any non-fragrance materials that are identified to be structurally related to have data. Data could include, but are not limited to, developmental or reproductive toxicity studies, reproductive toxicity screening studies, subchronic/chronic studies with information on the reproductive organs, sperm analysis, knowledge of available metabolic pathways, and comparison of skin absorption. These data will be closely reviewed to decide if a read-across NOAEL can be adopted for the purposes of setting safe use levels. If a NOAEL is identified from read-across data, the MOE will be derived from the NOAEL and the aggregate systemic exposure data and an endpoint assessment will be generated.

In the case where there are no sufficient data on the fragrance material itself or on structurally-related materials, **Step 3** will be applied. **Step 3** uses the TTC concept. It has been reported that the TTC for general toxicity can also be applied to both the developmental and reproductive toxicity endpoints (Laufersweiler et al., 2012). Piersma et al. concluded that all endpoints in reproductive toxicology have shown thresholds of adversity, thus there is evidence for the presence of dose levels with no appreciable increase in risk (Piersma et al., 2011). For reproductive toxicity, the same TTC values have been suggested as for general toxicity, given that NOAELs of reproductive toxicity studies tend to be similar or higher than

those observed in general toxicity studies (Kroes et al., 2007). Cramer Classes I, II, and III have limits of 1800, 540, and 90 $\mu\text{g}/\text{person}/\text{day}$, which correspond to exposures of 30, 9, and 1.5 $\mu\text{g}/\text{kg bw}/\text{day}$, respectively (Kroes et al., 2007). If the systemic exposure is below the respective TTC, an endpoint assessment for both the reproductive and developmental toxicity endpoints will be generated for the fragrance material.

If the exposure exceeds the applicable TTC value, the fragrance material will move to **Step 4**. Additional testing may be recommended as determined on a case-by-case basis. The evaluation guidance recommended by the January 2011 RIFM Reproduction Workshop will be followed, which involves a tiered approach to testing, which includes obtaining information on skin absorption. In the absence of any data, testing will begin with consideration of metabolism. This will include Step 4a, an *in vitro* comparative metabolism study in rat, rabbit and human hepatocytes. Step 4b is an *in vivo* oral and/or dermal toxicokinetic study in rats. If the *in vitro* metabolism studies show that the rat is not a relevant model for human risk assessment, testing in a more appropriate species will be considered. Dosages and route of administration will be carefully selected from the toxicokinetic data. The goal of the metabolism and toxicokinetic work is to determine relevant (bioavailable) doses in the appropriate species. Doses that produce extreme toxicity or overwhelm the metabolic pathways should not be considered relevant for further evaluation. After these studies are completed, a reproduction/developmental screening test (OECD 421) will be conducted in rats or the designated appropriate species with dietary administration (encapsulating if there are concerns for palatability). Gavage administration will not be used due to bolus dose concerns; however, dermal administration may be considered if deemed appropriate. If the reproduction/developmental screening test shows no indication of developmental or reproductive risks and an acceptable margin of exposure exists, then based on this pilot study an endpoint assessment is possible and no further testing is needed on this fragrance material. If effects are observed, the reproduction/developmental screening test will serve as the dosage-range finder for an enhanced 1-generation reproduction toxicity study (OECD 415) with dietary administration conducted in rats. A RIFM enhanced OECD 415 1-generation reproductive toxicity study includes estrous cycling, sperm analyses, and ovarian follicle analysis in parental generation rats. The pups, individually identified, are evaluated for sexual maturation endpoints (anogenital distance (all pups, PND 1 and 22), presence of nipples (all pups, ~PND 12), and preputial separation and vaginal patency (1 male and 1 female from each litter, out to ~PND60)). The enhanced 1-generation reproduction toxicity study will identify the NOAELs used in the endpoint assessment. In each step of the process, the data will be carefully evaluated and compared to any existing data on the fragrance material and any read-across materials. Other testing deemed necessary to complete the safety assessment could include *in silico* or *in vitro* dermal absorption, subchronic toxicity, or developmental toxicity.

4.4. Skin sensitization

The risk assessment for skin sensitization of all chemically-defined substances is carried out through a series of steps according to the flow chart shown in Fig. 4.

Step 1 examines the acceptability of any existing experimental data. In keeping with the criteria described by Api et al. (2008), this evaluation considers all studies that have given non-equivocal results and that have been carried out according to established and reliable methods. These methods may include all well run Local Lymph Node Assays (OECD, 2010a), Guinea Pig Maximization Tests (OECD, 1992), Buehler tests (OECD, 1992) and Open/Closed Epicutaneous Tests in Guinea Pigs. Human Repeated Insult Patch Tests (HRIPT) will also be taken into account if carried out according to accepted

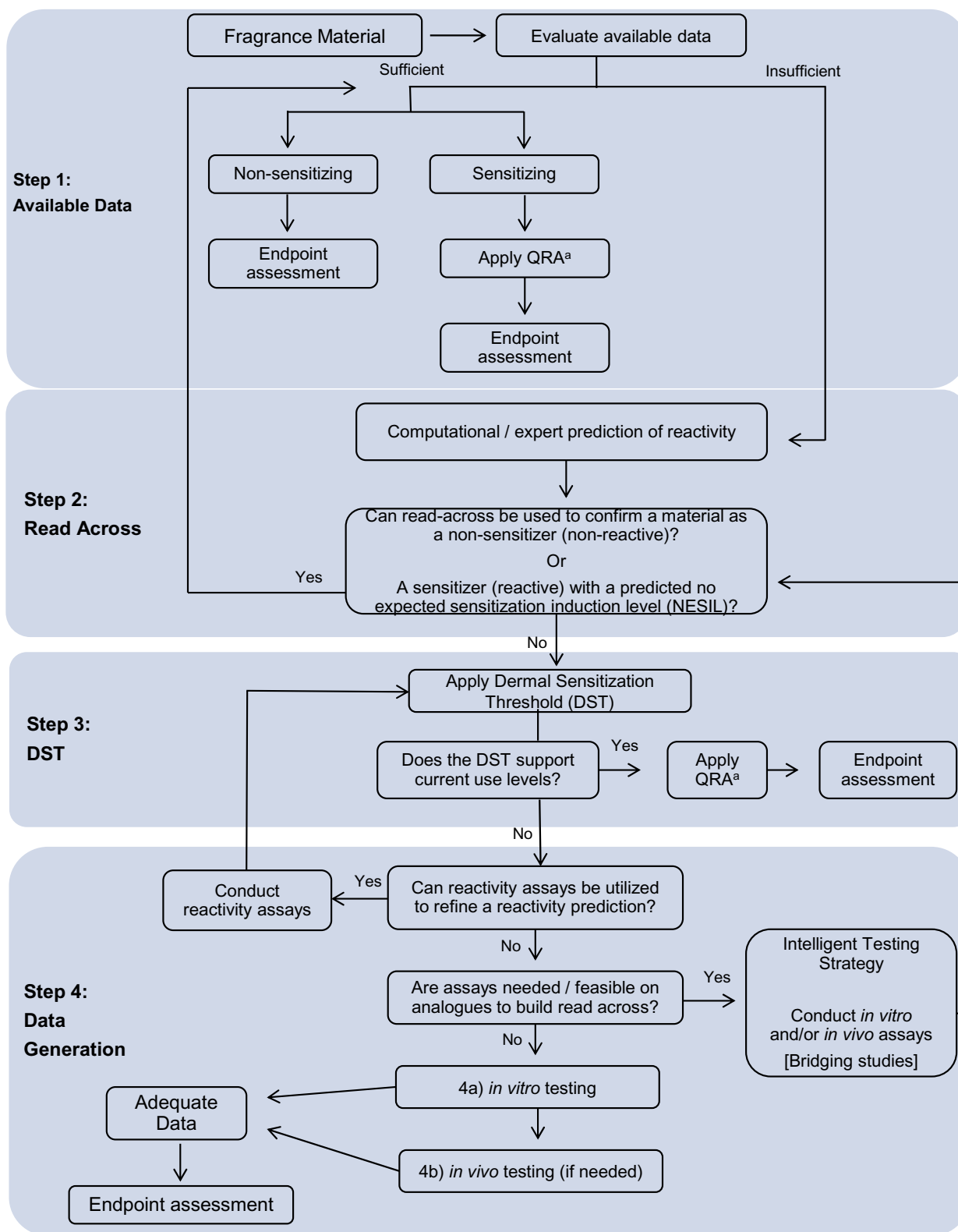
methods (Politano and Api, 2008). The HRIPT is used as a confirmatory study to substantiate a predicted no expected sensitization induction level (NESIL), and is not utilized to establish a sensitization hazard or to define the sensitizing potency of a material. In cases where historical HRIPT data are available, it may be possible to identify a lowest observed effect level (LOEL) or maximum tested no effect level (MT-NOEL). In the absence of data to support a higher use level, the MT-NOEL will form the basis of a NESIL for use in a dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. Where a number of tests have been undertaken but do not meet all of these requirements, it may be concluded based on a weight of evidence that a material is either sensitizing or non-sensitizing. Substances that are judged by **Step 1** to have sufficient data are referred for an endpoint assessment and the supporting evidence documented.

In cases where existing test data are considered to be insufficient to conclude that the substance is not a sensitizer or to derive an adequate NESIL, **Step 2** is applied. This step utilizes structure activity models and read across to make a determination on the sensitizing potential of a fragrance material. A key step in the development of skin sensitization is the covalent interaction of an electrophilic chemical with nucleophilic amino acid residues within proteins (Lepoittevin and Cribier, 1998; Smith et al., 2001). The classification of chemicals as either 'non-reactive' or 'reactive' provides an initial step in hazard identification and aids the development of read across. Structure activity models, such as those provided within the OECD Toolbox (Enoch et al., 2008), to classify substances as either reactive or non-reactive may be utilized. Materials that are initially classified as "non-reactive" are further evaluated for their potential to form reactive species through metabolic or abiotic mechanisms. These initial screening results should be reviewed, and further evaluation conducted, by an independent expert chemist¹ to confirm the classification (Aptula and Roberts, 2006b). In **Step 2**, read-across is utilized for those substances for which there is sufficient sensitization test data on other close structural analogs. The validity of this read-across approach may be supported by *in vitro* studies such as the direct peptide reactivity assay (Gerberick et al., 2004, 2007), keratinocyte assays (Natsch et al., 2007, 2010, 2011) or other comparable *in vitro* assays (Ade et al., 2006; Aleksic et al., 2009; Aptula et al., 2006a; Ashikaga et al., 2006; Chipinda et al., 2010; Gerberick et al., 2007, 2009; Goebel et al., 2012; Natsch and Gfeller, 2008; Rousset et al., 2002; Sakaguchi et al., 2006; Troutman et al., 2011). However, for the purpose of potency assessment, a quantitative relationship has to be demonstrated between the parameter measured and the known potency for the mechanistic domain being assessed.

Substances without a suitable read across and for which no NESIL could be derived are passed to **Step 3** and evaluated utilizing the *Dermal Sensitization Threshold* (DST). The DST applies the concept of the TTC to the evaluation of dermal sensitization. The DST establishes a level below which there is no appreciable risk for the induction of sensitization, and is based on a probabilistic analysis of potency data for a diverse range of known chemical allergens (Keller et al., 2009; Safford, 2008; Safford et al., 2011). For non-reactive substances a DST of 900 $\mu\text{g}/\text{cm}^2$ is applied and assumed to represent a worst case estimate of the NESIL (Safford, 2008; Safford et al., 2011).² Substances that are classified as "reactive" are attributed a lower DST of 64 $\mu\text{g}/\text{cm}^2$ (Safford et al., 2014b). This NESIL is then utilized in a QRA (Api et al., 2008) and will result in the

¹ An expert chemist should be qualified by scientific/academic training in organic chemistry and have ongoing and demonstrated practical experience.

² Based on analysis of the results of Local Lymph Node Assays for substances considered to be "non-reactive" these authors predict that there is a 99.74% probability that an untested chemical, when classified as "non-reactive" will have a LLNA EC3 of 900 $\mu\text{g}/\text{cm}^2$ or higher.



^aQRA = Dermal Sensitization Quantitative Risk Assessment for fragrance ingredients (Api et al., 2008)

Fig. 4. Skin sensitization flow chart.

maximum limits being imposed on the current IFRA categories of consumer products as indicated in Table 1.

In cases where the use of a substance does not exceed the limits given in Table 1 (Step 3) it will move to an endpoint assessment. On the other hand, if it is known that a substance exceeds these levels, the substance must be further examined (Step 4).

Where it is not possible to derive a reliable NESIL from available data/read-across or the application of the DST exceeds the maximum limits, additional testing is required (Step 4). In cases where the application of the DST is exceeded, a suitable chemical reactivity assay may be conducted (such as those described above) to refine the prediction of (non-)reactivity followed by a reevalua-

Table 1

Maximum use levels in consumer products based on the Dermal Sensitization Thresholds for fragrance materials judged to be “non-reactive” or “reactive”.

Category of consumer product ^a	“Non-reactive” materials (NESIL ^b = 900 µg/cm ²)	“Reactive” materials (NESIL ^b = 64 µg/cm ²) ^c
1	0.025%	0.0018%
2	0.033%	0.0023%
3	0.14%	0.010%
4	0.41%	0.029%
5	0.21%	0.015%
6	0.65%	0.046%
7	0.068%	0.0048%
8	0.91%	0.065%
9	4.5%	0.32%
10	7.5% ^d	0.53%
11	No specific limit ^e	No specific limit ^e

^a See Api et al. (2008) for a definition of each product category and description of the QRA limits calculations.

^b NESIL is the predicted no expected sensitization induction level.

^c Safford et al., 2014b.

^d The maximum pragmatic level allowed in the category is defined as 2.5% (Api and Vey, 2008).

^e Category 11 includes all non-skin contact or incidental skin contact products. Due to negligible skin contact the concentration of a fragrance ingredient should not exceed the usual concentration of the fragrance compound in the finished product.

tion utilizing the DST. In other cases, further testing may involve the use of appropriate *in vitro* methods and data (such as those described above). Consideration is given to testing suitable structural analogs along with the material of interest to support and develop the use of read-across (Step 4a). As with all other endpoints, these bridging studies should be conducted utilizing an intelligent testing strategy designed to eliminate or reduce the need for *in vivo* testing. Following the exhaustion of all other approaches, then Step 4b should be followed and a Local Lymph Node Assay (OECD, 2010a) may be considered. The NOEL for induction of dermal sensitization, identified through the evaluation process above, may be confirmed in a HRIPT (Politano and Api, 2008). From these data a NESIL may be derived by a weight of evidence approach (Api et al., 2008) for subsequent use in the QRA.

The process for the tiered assessment of the endpoint of sensitization is summarized in Fig. 4.

4.5. Photoirritation/photoallergenicity

The assessment of the risk for phototoxicity (including photo irritation and photoallergenicity) of all chemically-defined substances is carried out through a series of steps (Fig. 5).

In Step 1, if there are already sufficient data on the material itself to address phototoxicity, then the next step is an endpoint assessment. If there are insufficient data, Step 2 is applied utilizing either read across to structurally related material(s) or a benchmark evaluation of the UV/VIS absorbance. Similar to the previous step, if there are sufficient data on read across material(s) to address phototoxicity, then an endpoint assessment is conducted. However, where insufficient data exist the UV/VIS absorbance spectrum is evaluated to eliminate materials that have no significant absorbance in the range of 290–760 nm. This is considered a valid screen because a photo-biological response is dependent on photo-activation of the test material and changes in absorbance following interaction with the skin are rare (Lovell and Sanders, 1992). However, it has been debated as to what should be considered “no significant absorption”. In its introduction to the Technical Guidance for the 3T3 NRU test (OECD, 2004), it is stated that:

“...if the molar extinction/absorption coefficient [MEC] is less than 10 litre \times mol⁻¹ \times cm⁻¹ the chemical is unlikely to be photo-reactive.”

More recently other authors (Henry et al., 2009) have pointed out that this choice was:

“based on an early OECD guidance note on UV measurement technique and does not reference any specific relationship with photo-safety issues. This MEC value is very low. In spectrophotometric practice it represents only a slight increase in absorption over the instrument baseline measurements and ultimately represents the practical limit of detections for most drug-like molecules.”

Henry et al. (2009) have studied the molar extinction coefficients of 35 phototoxic substances and have concluded that:

“All the compounds had one or more peak maxima at or above 290 nm, with MEC’s typically greater than 3000 L mol⁻¹ cm⁻¹” and “Molecules with an MEC less than 1000 L mol⁻¹ cm⁻¹ deemed less of a photo-safety risk since this low level of light absorption is unlikely to prove harmful.”

This threshold of 1000 L \times mol⁻¹ \times cm⁻¹ has been, in principle, agreed to by the European Medicines Agency in their explanation on the “Note for guidance on photosafety testing” published 24 June 2010. As a consequence, 1000 L \times mol⁻¹ \times cm⁻¹ is recommended as the maximum value for the MEC at any wavelength above 290 nm in Step 2 to exclude substances from further evaluation.

Step 3 represents an exposure level below which it is unlikely that any type of phototoxic potential exists. Clearly, setting such a limit is complicated by the differences in the mechanism(s) of action between phototoxic substances, the photo-toxicological endpoint and the conditions under which they express these effects (e.g. length of exposure to the substance and vehicle used, spectral output of the light source, the overall “dose” of UV radiation as well as the duration and timing of exposure to UV radiation and the number and frequency at which these exposures to both substance and UV radiation are repeated (Man et al., 2004; Ortel and Gange, 1990; Taylor et al., 2002). It is possible that the strongest evidence is derived from the more potent phototoxicants to which humans are exposed, such as furocoumarins.

A stringent limit of 1 ppm of furocoumarins is applied to products used explicitly for situations with substantial light exposure such as sun screen products. This limit is in line with the European Cosmetics Regulation (European Union, 2009), which has restricted furocoumarins in “sun protection and in bronzing products” to that level. A risk assessment demonstrating the safety of 5 mg/kg of furocoumarins in all other leave-on cosmetic products and 50 mg/kg for rinse-off products was prepared by industry (Colipa/EFFA, 2005). Industry has proposed applying these limits to the content of any combination of 7 furocoumarins serving as markers: 5-Methoxypsoralen (bergapten), bergamottin, byacangelicol, epoxybergamottin, isopimpinellin, oxypeucedanin, and 8-Methoxypsoralen (xanthotoxin). These were selected based on their relevance with regard to the potential amount present in various citrus oils used by the industry as well as potency considerations. On the basis of the above, i.e. the broad photo-toxicological effects and the relatively high potency of furocoumarins (Colipa/EFFA, 2005), it is proposed that substances with a use level below those shown in Table 2, are considered as having a negligible phototoxic potential.

Substances that have a maximum molar absorption coefficient greater than 1000 L mol⁻¹ cm⁻¹ at any wavelength above 290 nm and which exceed the respective limits in Table 2 must pass through Step 4a and/or b. This step comprises a tiered application of *in vitro* and *in vivo* testing, the latter being used only as a last resort. In step 4a, the 3T3 NRU Assay according to the OECD TG 432 (OECD, 2004) may be conducted. Like any *in vitro* cell-based model, there are limitations of the 3T3 NRU assay including methodological issues (e.g., solubility of test material) and interpretation of outcomes particularly those associated with testing complex mixtures. As well,

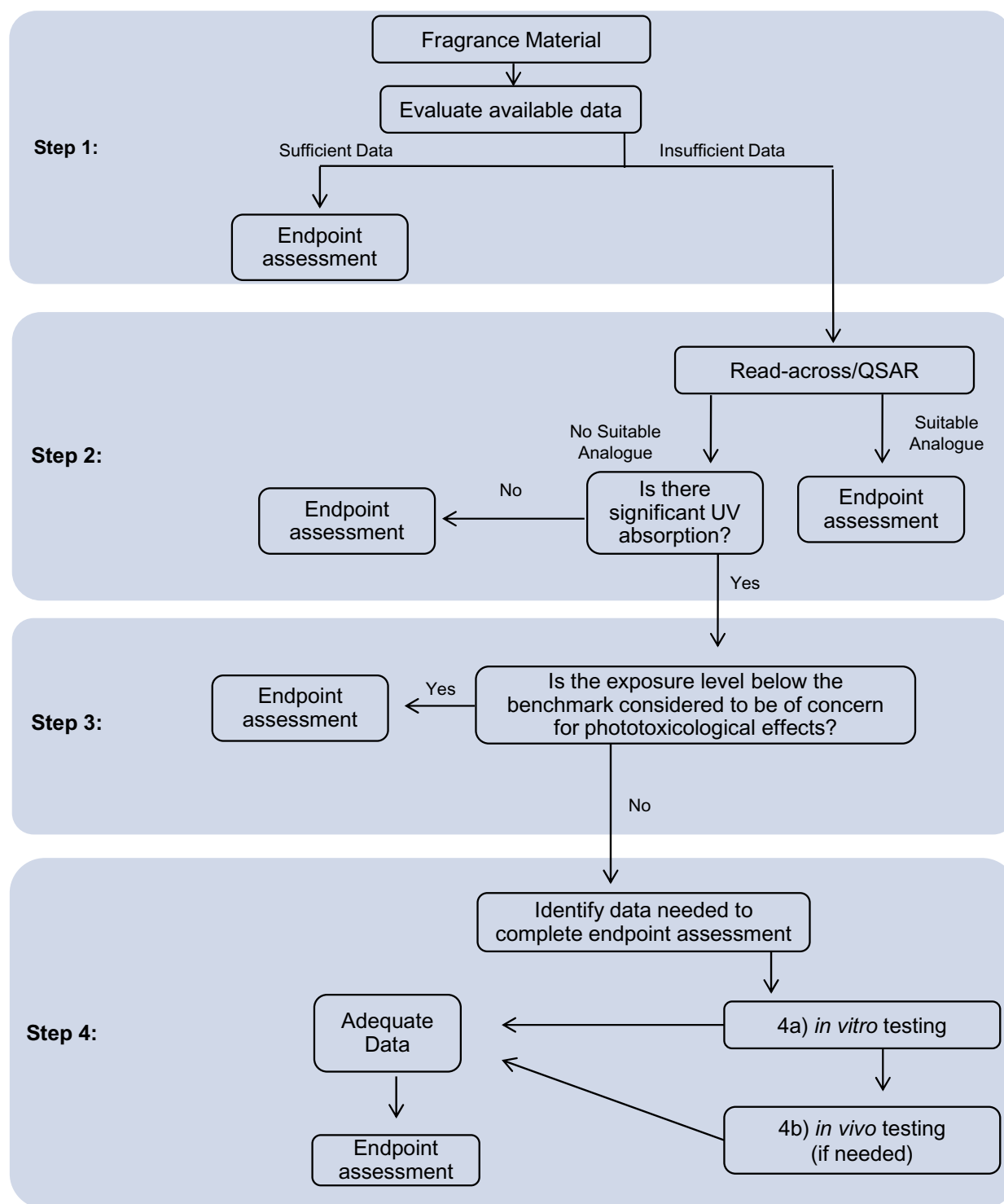


Fig. 5. Flow chart for phototoxicity.

the 3T3 NRU assay is mainly a hazard identification assay and provides little to no information of potency of the test material. Despite these limitations, it is accepted that the absence of any phototoxic response in the 3T3 NRU assay is sufficient to classify the substance as having negligible photo-toxicological potential (Ceridono et al., 2012). As an adjunct to the 3T3 NRU test, *in vitro* assays utilizing reconstructed human skin equivalents may be considered such as the EST (epidermal skin test) 1000 assay (Jones et al., 2001; Liebsch et al., 1995). Assays utilizing reconstructed human skin provide an

advantage over the 3T3 NRU, allowing for topical administration in a relevant vehicle and in skin cells of human origin.

Finally, if the 3T3 NRU, or comparable assay, indicates phototoxic potential, it may be possible to proceed cautiously to an exploratory test in **Step 4b** (Barratt and Brown, 1985; Lovell, 1993). A human or guinea pig study may be conducted (Ichikawa et al., 1981; Kaidbey and Kligman, 1981). A full review of available animal models and human clinical testing for photoirritation and photoallergenicity can be found in Nash (2009).

Table 2

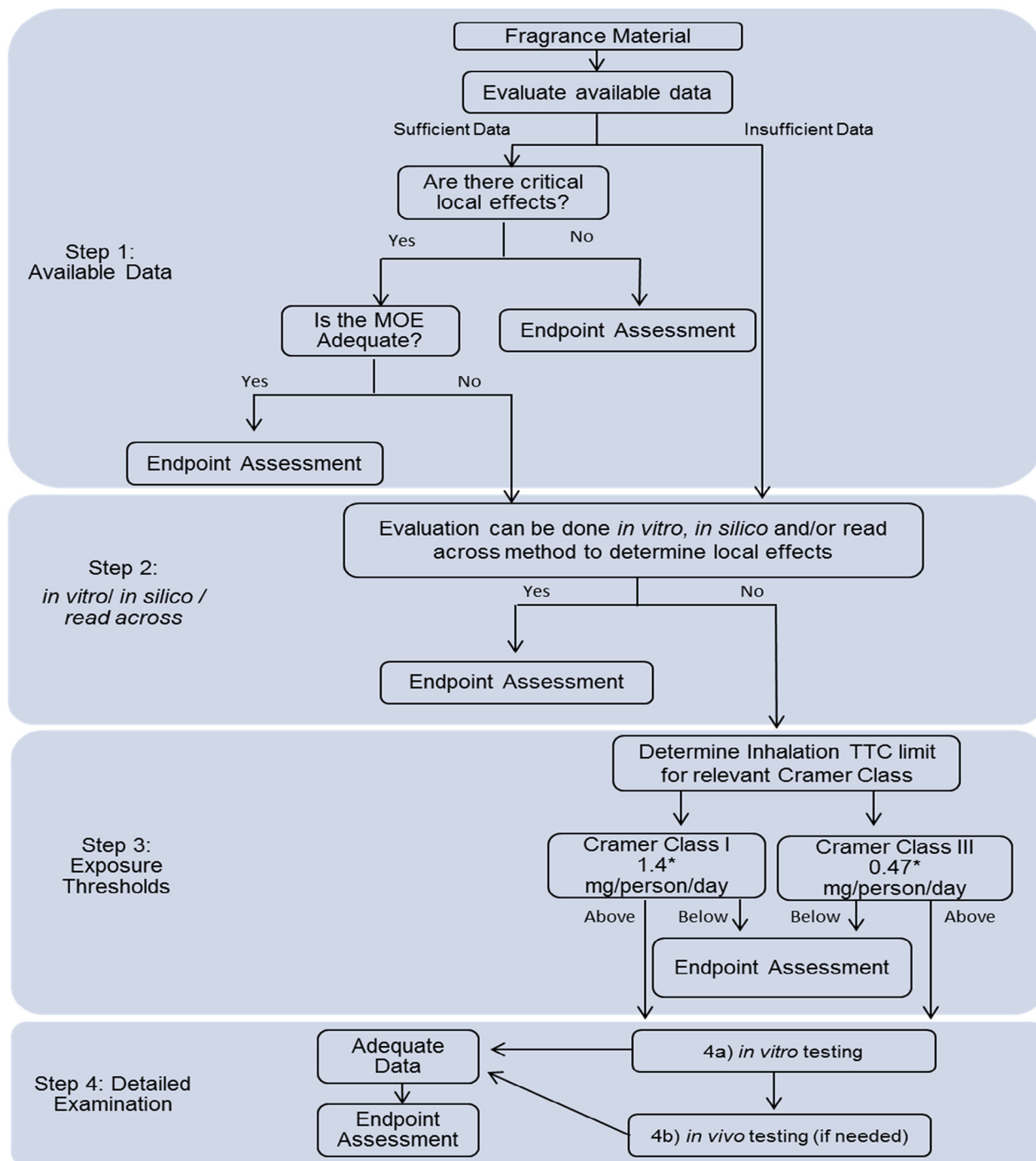
Maximum default limits of furocoumarins in consumer products based on a formal risk assessment (Colipa/EFFA, 2005).

Category of consumer product	Maximum limit
Sun-care products	0.0001%
Other leave-on cosmetics	0.0005%
Rinse-off cosmetics	0.005%
Household products rinsed from the skin	0.005%
Incidental contact products ^a	No specific limit

^a Includes all non-skin contact or incidental skin contact products. Due to negligible skin contact the concentration should not exceed the usual concentration of the fragrance compound in the finished product.

4.6. Respiratory (local)

The risk assessment for local respiratory tract effects caused by inhalation of chemically-defined substances is carried out through a series of steps according to the flow chart in Fig. 6. The focus of the assessment is local effects at the site of contact (target organs: nose, larynx, pharynx, tracheobronchial tree (trachea, bronchi, and bronchioles), and pulmonary region (respiratory bronchioles, alveolar ducts, alveoli)) (Elberling et al., 2006; Kleno and Wolkoff, 2004; Millqvist et al., 1999; Walker et al., 2001a, 2001b). Systemic effects



*TTC limits are based on Local Inhalation Effects from Carthew et al. 2009 (uses human lung weight of 650 g). Cramer Class II materials default to Cramer Class III.

Fig. 6. Respiratory toxicology local effects flow chart.

from exposure to fragrance materials by all routes (i.e., inhalation, oral, and/or dermal) are covered in Section 4.2 of this document regarding Repeated Dose Toxicity.

Prior to Step 1, the first step after gathering all available toxicological data is to develop an estimate of potential exposure. **Step 1** is then used to evaluate the potential for induction of effects in the respiratory tract based on available inhalation toxicity data. If the material under consideration induces local effects the no effect concentrations (NOAEC) can be compared to the predicted exposure to determine the MOE. If a low observable adverse effect concentration (LOAEC) is available instead of a NOAEC, a NOAEC can be derived by including an additional extrapolation factor. If the MOE is relatively low (less than 100) or if exposure exceeds the NOAEC/LOAEC, the risk assessment would need to be refined by proceeding to **Step 2**.

Using **Step 2** a structure–activity analysis can be conducted to determine if there are appropriate structural analogs that have relevant data that can be used as surrogates for the risk assessment for local effects. Alternatively, if appropriate *in vitro* methods and/or *in silico* (i.e. computational) tools become available, the fragrance material can be evaluated and the results compared with available read-across data (Abraham et al., 1998; Enoch et al., 2009, 2010; Kimber et al., 2001; Selgrade et al., 2012; Veith et al., 2009). If no appropriate analogs are identified or if no *in vitro* or *in silico* evaluations can be done, the material moves to **Step 3**.

Step 3 is applied when the existing inhalation toxicity data are insufficient to complete the endpoint assessment. This step is based on the application of the Threshold of Toxicological Concern (TTC). The TTC principles, which were originally developed with a focus on systemic exposure following oral administration, have been extended to consider systemic exposure following topical application of cosmetic products (Blackburn et al., 2005; Kroes et al., 2007). More recently there have been analyses proposing that the TTC can be applied to inhalation exposure and risk assessment (Carthew et al., 2009; Drew and Frangos, 2007; Escher et al., 2010; Kroes et al., 2007). A key consideration with respect to inhalation is the development of separate Cramer Class TTC values for site of contact (local) effects. **Step 3** uses the TTC for local effects upon inhalation exposure for a material belonging to either Cramer Class I (1.4 mg/day) or III (0.47 mg/day) assuming a human lung weight of 650 g (Carthew et al., 2009). In order to derive these values, a group of 92 chemicals used primarily in consumer products was evaluated for both systemic and site of contact effects (Carthew et al., 2009). The authors established NOAECs for site contact effects and assigned a Cramer class for each chemical. Genotoxic carcinogens, *in vivo* mutagens, heavy metals, dioxins, PCBs, organophosphates and polymers were excluded from the analysis. Most of the chemicals evaluated belong to Cramer Classes I or III. For conservative purposes if a material is determined to be of Cramer Class II, it is assigned the Cramer Class III value. The data set used to establish these respiratory TTC values is relatively small compared to the much larger dataset used to establish the TTC values for repeated dose toxicity. If the calculated exposure is less than the TTC recommended limit, the fragrance material can be used without additional testing. If the calculated exposure is greater than the recommended TTC limit, one needs to move to **Step 4**.

In **Step 4**, additional studies are identified, using available *in vitro* models first (**Step 4a**) and ultimately *in vivo* analysis (**Step 4b**), to generate sufficient data to complete the endpoint assessment.

The process for the tiered assessment of the endpoint of respiratory effects is summarized in Fig. 6.

5. Environmental endpoint assessment

The most significant route of exposure to the environment for fragrance ingredients is down the drain discharge to freshwater. This

exposure is driven by the annual volume of use (VoU). As noted above, RIFM has routinely screened for potential impacts to the freshwater aquatic environment since 1999. The published “RIFM Environmental Framework” (Salvito et al., 2002) provides the model used for this effort. It is a conservative model comparing a ‘down the drain’ discharge concentration (through wastewater treatment) with an estimated effect on fish using a large uncertainty factor to avoid false negatives in the use of this screening tool. It is comprised of scenarios for both Europe and North America. These scenarios take into account the differences in wastewater treatment and per capita water use. In order for a material to be considered ‘safe for use’ the calculated (or measured) exposure must be less than the calculated (or measured) effects on aquatic organisms (i.e., the Risk Screening Criteria (risk quotient (RQ) or Predicted Environmental Concentration/Predicted No-Effect Concentration (PEC/PNEC ratio)) is <1).

The RIFM Environmental Framework follows a three tiered approach. In Tier 1 aquatic risk is estimated using only physical-chemical properties (log Kow and Mol Wt (Molecular Weight) and regional VoU for two exposure scenarios: North America and Europe. In Tier 2, ECOSAR is used to refine the PNEC for all materials whose Tier 1 PEC/PNEC is >1. In Tier 3 existing or *de novo* experimental data or other validated QSARs may be used to refine the risk assessment further for any materials whose PEC/PNEC is still >1 (see Fig. 7).

For completeness RIFM determines if a material is persistent, bioaccumulative, and toxic to the environment (PBT), or very persistent and very bioaccumulative (vPvB) and, presently, are the same criteria as are used in the EU for REACH (ECHA, 2012).

The models in EPISUITE (Estimation Programs Interface (EPI) Suite™ (version 4.1; Software Copyright 2000–2011); US Environmental Protection Agency’s Office of Pollution Prevention and Toxics and Syracuse Research Corporation (<http://www.epa.gov/opptintr/exposure/pubs/episuite.htm>)) are used for *in silico* determinations of P, B and T values, useful both in hazard and risk assessment. The relevant models in EPISUITE are BIOWIN (for predicting wastewater biodegradation), BCFBAF (for predicting fish bioaccumulation) and ECOSAR (for predicting toxicity to freshwater aquatic organisms). However, hazard screening is focused on materials that are potential P and B materials; therefore, ECOSAR (ECOLOGICAL Structure Activity Relationship Class Program for Microsoft Windows (version 1.11; Software 2012); US Environmental Protection Agency’s Office of Pollution Prevention and Toxics and Syracuse Research Corporation (<http://www.epa.gov/oppt/newchems/tools/21ecosar.htm>)) information is more relevant for risk than hazard screening. Within BIOWIN the modules BIOWIN 2 (non-linear model), BIOWIN 6 (MITI non-linear model), and BIOWIN 3 (ultimate biodegradation timeframe) are used as follows (i.e., Persistence screening criteria):

- A material is potentially P if either BIOWIN 2 < 0.5 or BIOWIN 6 < 0.5 and BIOWIN 3 < 2.2,
- Furthermore, a material is considered borderline potentially P if either BIOWIN 2 < 0.5 or BIOWIN 6 < 0.5 and BIOWIN 3 < 2.7.

A material would be considered potentially bioaccumulative if the model BCFBAF predicts fish bioconcentration greater than 2000 L/kg; and very bioaccumulative if this value is greater than 5000 L/kg (i.e., bioaccumulation screening criteria).

From the generated screening list, RIFM then identifies available data or generates new information to provide, based on a weight of evidence, the necessary assessment to confirm or refute that a material is a PBT (see Fig. 7).

Environmental Assessment

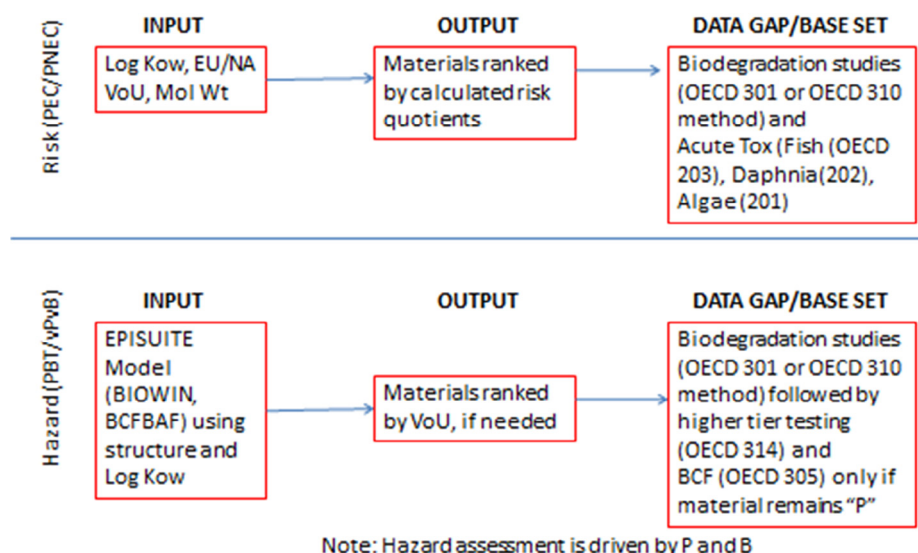


Fig. 7. Environmental data assessment.

6. Evaluation of complex ingredients

Some fragrance ingredients can be complex mixtures (e.g. essential oils, plant extracts, reaction products, and mixtures of positional and geometric isomers) and are often referred to as UVCB substances (substances of unknown or variable composition, complex reaction products or biological origin) under REACH (REACH, 2006) and other international regulations (e.g., Canada's Domestic Substance List or DSL).

The assessment of an UVCB begins with a search for data on the UVCB itself and a determination of the components of the UVCB.

If sufficient data are available on the UVCB proper then an assessment will be completed following the process flow outlined in Fig. 1.

When constituents are known (>95%) and sufficient data are not available on the complex material proper, the assessment may be conducted using methods to assess complex mixtures (Smith et al., 2001; Toxic Equivalent Approach OECD, ECHA; Ellis, 2010; Price et al., 2009). When an individual component does not have data to address a specific toxicological endpoint identification of an appropriate analog(s) will be undertaken.

When the data are insufficient on either the UVCB proper or on one or more individual constituents the components of the UVCB may be analyzed using a combination of analog identification for 'read across' and the TTC (Threshold of Toxicological Concern) approach (Kroes et al., 2005; Price et al., 2009).

The evaluation may be completed by identifying the appropriate endpoint analysis needed and outlining a testing strategy to develop adequate information. The testing plan will be developed on a case by case basis. It may include testing on a blend of commercially available material, testing using variants (e.g., a subset of terpene chemicals that identify a subset of the UVCB) whose constituents "bracket" the available variants or a series of tests and analysis on individual materials, blocks or complete material depending on the endpoint of interest.

7. Summary

During the period from 2000 to 2003 approaches to safety assessment of fragrances and estimations of patterns of use in

consumer products were published (Bickers et al., 2003; Cadby et al., 2002; Ford et al., 2000). This publication is designed to update RIFM's risk assessment process reflecting advances in approaches in risk assessment over the past ten years and include:

- incorporate new scientific information that includes a framework for choosing structural analogs and groups (Wu et al., 2010),
- add consideration of the Threshold of Toxicological Concern (TTC) (Kroes et al., 2004, 2007),
- add consideration of the Quantitative Risk Assessment for dermal contact sensitization (QRA) (Api et al., 2008),
- add considerations for the respiratory route of exposure,
- update exposure assessment methodology,
- incorporate the latest methodology and approaches to risk assessments,
- incorporate the latest alternatives to animal testing methodology, and
- incorporate the current state of environmental risk assessment in support of the IFRA Standards.

Conflict of interest

The authors declare that there are no conflicts of interest.

Transparency document

The Transparency document associated with this article can be found in the online version.

References

- Abraham, M.H., Kumarsingh, R., Cometto-Muniz, J.E., Cain, W.S., 1998. A quantitative structure-activity relationship (QSAR) for a Draize eye irritation database. *Toxicol. In Vitro* 12, 201–207.
- Ade, N., Martinozzi-Teissier, S., Pallardy, M., Rousset, F., 2006. Activation of U937 cells by contact sensitizers: CD86 expression is independent of apoptosis. *J. Immunotoxicol.* 3, 189–197.
- Adler, S., Basketter, D., Creton, S., Pelkonen, O., Benthem, J.V., Zang, V., et al., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects- 2010. *Arch. Toxicol.* 85, 367–485.

- Aleksic, M., Thain, E., Roger, D., Saib, O., Davies, M., Li, J., et al., 2009. Reactivity profiling: covalent modification of single nucleophile peptides for skin sensitization risk assessment. *Toxicol. Sci.* 108, 401–411.
- Api, A.M., Vey, M., 2008. Implementation of the dermal sensitization Quantitative Risk Assessment (QRA) for fragrance ingredients. *Regul. Toxicol. Pharmacol.* 52, 53–61.
- Api, A.M., Basketter, D.A., Cadby, P.A., Cano, M.F., Ellis, G., Gerberick, G.F., et al., 2008. Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. *Regul. Toxicol. Pharmacol.* 52, 3–23.
- Aptula, A.O., Roberts, D.W., 2006b. Mechanistic applicability domains for nonanimal-based prediction of toxicological end points: general principles and application to reactive toxicity. *Chem. Res. Toxicol.* 19, 1097–1105.
- Aptula, A.O., Patlewica, G., Roberts, D.W., Schultz, T.W., 2006a. Non-enzymatic glutathione reactivity and in vitro toxicity: a non-animal approach to skin sensitization. *Toxicol. In Vitro* 20, 239–247.
- Ashby, J., Tennant, R.W., 1991. Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutat. Res.* 257, 229–306.
- Ashikaga, T., Yoshida, Y., Hirota, M., Yoneyama, K., Itagaki, H., Sakaguchi, H., et al., 2006. Development of an in vitro skin sensitization test using human cell lines: the human cell line activation test (h-CLAT) I. Optimization of the h-CLAT protocol. *Toxicol. In Vitro* 20, 767–773.
- Barratt, M.D., Brown, K.R., 1985. Photochemical binding of photoallergens to human serum albumin: a simple in vitro method for screening potential photoallergens. *Toxicol. Lett.* 24, 1–6.
- Belsito, D., Bickers, D., Bruze, M., Calow, P., Dagli, M.L., Fryer, A.D., et al., 2011. A toxicological and dermatological assessment of macrocyclic lactone and lactide derivatives when used as fragrance ingredients. *Food Chem. Toxicol.* 49 (Suppl. 2), S219–S241.
- Benford, D., Bolger, P.M., Carthew, P., Coulet, M., Dinovi, M., Leblanc, J.C., et al., 2010. Application of the Margin of Exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. *Food Chem. Toxicol.* 48 (Suppl. 1), S2–S24.
- Benigni, R., Bossa, C., 2008. Structure alerts for carcinogenicity, and the Salmonella assay system: a novel insight through the chemical relational databases technology. *Mutat. Res.* 659, 248–261.
- Benigni, R., Bossa, C., Tcheremenskaia, O., Giuliani, A., 2010. Alternatives to the carcinogenicity bioassay: in silico methods, and the in vitro and in vivo mutagenicity assays. *Expert Opin. Drug Metab. Toxicol.* 6, 809–819.
- Bickers, D.R., Calow, P., Greim, H.A., Hanifin, J.M., Rogers, A.E., Saurat, J.H., et al., 2003. The safety assessment of fragrance materials. *Regul. Toxicol. Pharmacol.* 37, 218–273.
- Billinton, N., Hastwell, P.W., Beerens, D., Birrell, L., Ellis, P., Maskell, S., et al., 2008. Interlaboratory assessment of the GreenScreen HC GADD45a-GFP genotoxicity screening assay: an enabling study for independent validation as an alternative method. *Mutat. Res.* 653, 23–33.
- Birrell, L., Cahill, P., Hughes, C., Tate, M., Walmsley, R.M., 2010. GADD45a-GFP GreenScreen HC assay results for the ECVAM recommended lists of genotoxic and non-genotoxic chemicals for assessment of new genotoxicity tests. *Mutat. Res.* 695, 87–95.
- Blackburn, K., Stickney, J.A., Carlson-Lynch, H.L., McGinnis, P.M., Chappell, L., Felter, S.P., 2005. Application of the threshold of toxicological concern approach to ingredients in personal and household care products. *Regul. Toxicol. Pharmacol.* 43, 249–259.
- Blackburn, K., Bjerke, D., Daston, G., Felter, S., Mahony, C., Naciff, J., et al., 2011. Case studies to test: a framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. *Regul. Toxicol. Pharmacol.* 60, 120–135.
- Cadby, P.A., Troy, W.R., Vey, M.G., 2002. Consumer exposure to fragrance ingredients: providing estimates for safety evaluation. *Regul. Toxicol. Pharmacol.* 36, 246–252.
- 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.* 47, 1287–1295.
- Ceridono, M., Tellner, P., Bauer, D., Barroso, J., Alepee, N., Corvi, R., et al., 2012. The 3T3 Neuronal red uptake phototoxicity test: practical experience and implications for phototoxicity testing – the report of an ECVAM-EFPIA workshop. *Reg. Tox. Pharmacol.* 63, 480–488.
- Cheeseman, M.A., Machuga, E.J., Bailey, A.B., 1999. A tiered approach to threshold of regulation. *Food Chem. Toxicol.* 37, 387–412.
- Chipinda, I., Ajibola, R.O., Morakinyo, M.K., Ruwona, T.B., Simoyi, R.H., Siegel, P.D., 2010. Rapid and simple kinetics screening assay for electrophilic dermal sensitizers using nitrobenzenethiol. *Chem. Res. Toxicol.* 23, 918–925.
- Colipa/EFFA 2005. Formal Risk Assessment of Furocoumarins in finished cosmetic products. Furocoumarins (FCs) in Fragrances – Submission of Formal Risk Assessment. Submitted by the European Flavour & Fragrance Association to the European Commission Cosmetics and Medical Devices, DG Enterprise on December 5, 2005.
- Comiskey, D., Api, A.M., Clapp, C., Daly, E.J., Ellis, G., McNamara, C., et al., 2014. Novel database for exposure to fragrance ingredients in cosmetics and personal care products. *Reg. Tox. Pharmacol.* accepted for publication.
- Cramer, G.M., Ford, R.A., Hall, R.L., 1978. Estimation of toxic hazard – a decision tree approach. *Food Cosmet. Toxicol.* 16, 255–276.
- Drew, R., Frangos, J., 2007. The concentration of no toxicological concern (CoNTC): a risk assessment screening tool for air toxics. *J. Toxicol. Environ. Health A* 70, 1584–1593.
- ECETOC, 1995. Assessment Factors in Human Health Risk Assessment. Technical Report No. 68. ISSN-0773-8072-68.
- ECHA 2010a. Guidance on information requirements and chemical safety assessment. Chapter R.8: characterisation of dose [concentration]–response for human health. In: European Chemicals Agency (ed.) Version 2.
- ECHA 2010b. *Practical guide 6: How to report read-across and categories* [Online]. <http://echa.europa.eu/doc/publications/practical_guides/pg_report_readacross_cat.pdf>. (accessed 6.28.11).
- ECHA, 2012 (<http://echa.europa.eu/>) Guidance on information requirements and chemical safety assessment Chapter R.11: PBT Assessment, November 2012 ver 1.1.
- ECVAM. 2012. *Scientifically Validated Methods* [Online]. European Centre for the Validation of Alternative Methods.
- EFSA 2012. European Food Safety Authority. Outcome of the public consultation on the draft scientific opinion on Threshold of Toxicological Concern (TTC). Supporting Publications 2012:EN-293. [93 pp.].
- Elberling, J., Dirksen, A., Johansen, J.D., Mosbech, H., 2006. The capsaicin cough reflex in eczema patients with respiratory symptoms elicited by perfume. *Contact Dermatitis* 54, 158–164.
- Ellis, G., 2010. Special issue 40th ISEO: Toxicological challenges for essential oils in REACH. *Flavour and Fragrance Journal* 25 (3), 138–144.
- Enoch, S.J., Madden, J.C., Cronin, M.T., 2008. Identification of mechanisms of toxic action for skin sensitisation using a SMARTS pattern based approach. *SAR QSAR Environ. Res.* 19, 555–578.
- Enoch, S.J., Roberts, D.W., et al., 2009. Electrophilic reaction chemistry of low molecular weight respiratory sensitizers. *Chem. Res. Toxicol.* 22, 1447–1453.
- Enoch, S.J., Roberts, D.W., Cronin, M.T., 2010. Mechanistic category formation for the prediction of respiratory sensitization. *Chem. Res. Toxicol.* 23, 1547–1555.
- Escher, S.E., Tluczkiewicz, I., Batke, M., Bitsch, A., Melber, C., Kroese, E.D., et al., 2010. Evaluation of inhalation TTC values with the database RepDose. *Regul. Toxicol. Pharmacol.* 58, 259–274.
- European Union 2009. Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products (recast). Official Journal of the European Communities. L342/59–209, December 22, 2009.
- Faustman, E.M., Omenn, S., 2008. Risk Assessment. In: Klaassen, C.D. (Ed.), *Casarett and Doull's Toxicology The Basic Science of Poisons*. New York.
- Ford, R.A., Domeyer, B., Easterday, O., Maier, K., Middleton, J., 2000. Criteria for development of a database for safety evaluation of fragrance ingredients. *Regul. Toxicol. Pharmacol.* 31, 166–181.
- Fung, V.A., Barrett, J.C., Huff, J., 1995. The carcinogenesis bioassay in perspective: application in identifying human cancer hazards. *Environ. Health Perspect.* 103, 680–683.
- Gerberick, G.F., Vassallo, J.D., Bailey, R.E., Chaney, J.G., Morrall, S.W., Lepoittevin, J.P., 2004. Development of a peptide reactivity assay for screening contact allergens. *Toxicol. Sci.* 81, 332–343.
- Gerberick, G.F., Vassallo, J.D., Foertsch, L.M., Price, B.B., Chaney, J.G., Lepoittevin, J.P., 2007. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. *Toxicol. Sci.* 97, 417–427.
- Gerberick, G.F., Troutman, J.A., Foertsch, L.M., Vassallo, J.D., Quijano, M., Dobson, R.L., et al., 2009. Investigation of peptide reactivity of pro-hapten skin sensitizers using a peroxidase-peroxide oxidation system. *Toxicol. Sci.* 112, 164–174.
- Goebel, C., Aeby, P., Ade, N., Alepee, N., Aptula, A., Araki, D., et al., 2012. Guiding principles for the implementation of non-animal safety assessment approaches for cosmetics: skin sensitisation. *Regul. Toxicol. Pharmacol.* 63, 40–52.
- 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.
- Hastwell, P.W., Chai, L.L., Roberts, K.J., Webster, T.W., Harvey, J.S., Rees, R.W., et al., 2006. High-specificity and high-sensitivity genotoxicity assessment in a human cell line: validation of the GreenScreen HC GADD45a-GFP genotoxicity assay. *Mutat. Res.* 607, 160–175.
- Hastwell, P.W., Webster, T.W., Tate, M., Billinton, N., Lynch, A.M., Harvey, J.S., et al., 2009. Analysis of 75 marketed pharmaceuticals using the GADD45a-GFP 'GreenScreen HC' genotoxicity assay. *Mutagenesis* 24, 455–463.
- Hattan, D.G., Rulis, A.M., 1986. FDA's priority-based assessment of food additives. III. Specific toxicity parameters. *Regul. Toxicol. Pharmacol.* 6, 181–191.
- 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* 96, 57–62.
- Hughes, C., Rabinowitz, A., Tate, M., Birrell, L., Allsup, J., Billinton, N., et al., 2012. Development of a high-throughput gaussia luciferase reporter assay for the activation of the GADD45a gene by mutagens, promutagens, clastogens, and aeneugens. *J. Biomol. Screen.* 17, 1302–1315.
- Ichikawa, H., Armstrong, R.B., Harber, L.C., 1981. Photoallergic contact dermatitis in guinea pigs: improved induction technique using Freund's complete adjuvant. *J. Invest. Dermatol.* 76, 498–501.

- ICCVAM. 2012. *Test Methods Reviewed or Under Consideration by ICCVAM by Toxicity Endpoint* [Online]. Interagency Coordinating Committee on the Validation of Alternative Methods.
- Janer, G., Hakker, B.C., Piersma, A.H., Vermeire, T., Slob, W., 2007. A retrospective analysis of the added value of the rat two-generation reproductive toxicity study versus the rat subchronic toxicity study. *Reprod. Toxicol.* 24, 103–113.
- Jones, P.A., Lovell, W.W., King, A.V., Earl, L.K., 2001. In vitro testing for phototoxic potential using the epiderm 3-D reconstructed human skin model. *Toxicology Methods* 11, 1–19.
- Kaidbey, K.H., Kligman, A.M., 1981. Photosensitization by coumarin derivatives. Structure-activity relationships. *JAMA Dermatol. (formerly Arch. Dermatol.)* 117, 258–263.
- Keller, D., Krauledat, M., Scheel, J., 2009. Feasibility study to support a threshold of sensitization concern concept in risk assessment based on human data. *Arch. Toxicol.* 83, 1049–1060.
- Kimber, I., Pichowski, J.S., et al., 2001. Alternative approaches to the identification and characterization of chemical allergens. *Toxicol. In Vitro* 15, 307–312.
- Kirkland, D., Aardema, M., Henderson, L., Muller, L., 2005. Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat. Res.* 584, 1–256.
- Kirkland, D., Reeve, L., Gatehouse, D., Vanparys, P., 2011. A core in vitro genotoxicity battery comprising the Ames test plus the in vitro micronucleus test is sufficient to detect rodent carcinogens and in vivo genotoxins. *Mutat. Res.* 721, 27–73.
- Kleno, J., Wolkoff, P., 2004. Changes in eye blink frequency as a measure of trigeminal stimulation by exposure to limonene oxidation products, isoprene oxidation products and nitrate radicals. *Int. Arch. Occup. Environ. Health* 77, 235–243.
- Klimisch, H.J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul. Toxicol. Pharmacol.* 25, 1–5.
- Knight, A.W., Birrell, L., Walmsley, R.M., 2009. Development and validation of a higher throughput screening approach to genotoxicity testing using the GADD45a-GFP GreenScreen HC assay. *J. Biomol. Screen.* 14, 16–30.
- Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., et al., 2004. Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food Chem. Toxicol.* 42, 65–83.
- Kroes, R., Kleiner, J., Renwick, A., 2005. The threshold of toxicological concern concept in risk assessment. *Toxicol. Sci.* 86 (2), 226–230.
- Kroes, R., Renwick, A.G., Feron, V., Galli, C.L., Gibney, M., Greim, H., et al., 2007. Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food Chem. Toxicol.* 45, 2533–2562.
- Laufersweiler, M.C., Gadagbui, B., Baskerville-Abraham, I.M., Maier, A., Willis, A., 2012. Correlation of chemical structure with reproductive and developmental toxicity as it relates to the use of the threshold of toxicological concern. *Regul. Toxicol. Pharmacol.* 62 (1), 160–182.
- Lepoittevin, J.P., Cribier, B., 1998. [Physiopathology of contact eczema]. *Ann. Dermatol. Venerol.* 125, 775–782.
- Lewis, R.W., Billington, R., Debryune, E., Gamer, A., Lang, B., Carpanini, F., 2002. Recognition of adverse and nonadverse effects in toxicity studies. *Toxicol. Pathol.* 30, 66–74.
- Liebsch, M., Doring, B., Donnelly, T.A., Logemann, P., Rhein, S.L.A., Spielmann, H., 1995. Application of the human dermal model Skin 2 ZK 1350 to phototoxicity and skin corrosivity testing. *Toxicol. In Vitro* 9, 557–562.
- Lovell, W.W., 1993. A scheme for in vitro screening of substances for photoallergic potential. *Toxicol. In Vitro* 7, 95–102.
- Lovell, W.W., Sanders, D.J., 1992. Phototoxicity testing in guinea-pigs. *Food Chem. Toxicol.* 30, 155–160.
- Man, I., Dawe, R.S., Ferguson, J., Ibbotson, S.H., 2004. The optimal time to determine the minimal phototoxic dose in skin photosensitized by topical 8-methoxypsoralen. *Br. J. Dermatol.* 151, 179–182.
- Millqvist, E., Bengtsson, U., Lowhagen, O., 1999. Provocations with perfume in the eyes induce airway symptoms in patients with sensory hyperreactivity. *Allergy* 54, 495–499.
- Munro, I.C., 1990. Safety assessment procedures for indirect food additives: an overview. Report of a workshop. *Regul. Toxicol. Pharmacol.* 12, 2–12.
- Munro, I.C., Ford, R.A., Kennepohl, E., Sprenger, J.G., 1996. Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. *Food Chem. Toxicol.* 34, 829–867.
- Nash, J.F., 2009. Phototoxicology. In: Ballantyne, B., Marrs, T.C., Swersen, T. (Eds.), *General and Applied Toxicology*, vol. 3. 3rd ed. Wiley, New York, pp. s1585–s1600.
- Natsch, A., Gfeller, H., 2008. LC-MS-based characterization of the peptide reactivity of chemicals to improve the in vitro prediction of the skin sensitization potential. *Toxicol. Sci.* 106, 464–478.
- Natsch, A., Gfeller, H., Rothaupt, M., Ellis, G., 2007. Utility and limitations of a peptide reactivity assay to predict fragrance allergens in vitro. *Toxicol. In Vitro* 21, 1220–1226.
- Natsch, A., Emter, R., Ellis, G., Adamson, G., 2010. Development of a high-throughput keratinocyte-based standard assay to detect skin sensitizers based on ARE-dependent gene activity. *Toxicologist* 114.
- Natsch, A., Bauch, C., Foertsch, L., Gerberick, F., Norman, K., 2011. The intra- and inter-laboratory reproducibility and predictivity of the keratinoSens assay to predict skin sensitizers in vitro: results of a ring-study in five laboratories. *Toxicol. In Vitro* 25, 733–744.
- OECD. 1983. Organisation for economic co-operation and development. Test Guideline No. 415. OECD guideline for testing of chemicals. One-generation reproduction toxicity study. <http://www.oecd-ilibrary.org/environment/test-no-415-one-generation-reproduction-toxicity-study_9789264070844-en> (accessed 07.11.14.).
- OECD. 1992. Organisation for economic co-operation and development. Test Guideline No. 406. OECD guideline for testing of chemicals. Skin sensitisation. <http://www.oecd-ilibrary.org/environment/test-no-406-skin-sensitisation_9789264070660-en> (accessed 07.11.14.).
- OECD. 1997. Organisation for economic co-operation and development. Test Guideline No. 471. OECD guideline for testing of chemicals. Bacterial reverse mutation test. <<http://www.oecd-ilibrary.org/docserver/download/9747101e.pdf?expires=1358971663&id=id&accname=guest&checksum=5CD866085D25D1F575EC6910488D1112>> (accessed 07.11.14.).
- OECD. 2001. Organisation for economic co-operation and development. Test Guideline No. 414. OECD guideline for testing of chemicals. Prenatal development toxicity study. <http://www.oecd-ilibrary.org/environment/test-no-414-prenatal-development-toxicity-study_9789264070820-en> (accessed 07.11.14.).
- OECD. 2004. Organisation for economic co-operation and development. Test Guideline No. 432. OECD guideline for the testing of chemicals. In vitro 3T3 NRU phototoxicity test. <http://www.oecd-ilibrary.org/environment/test-no-432-in-vitro-3t3-nru-phototoxicity-test_9789264071162-en> (accessed 07.11.14.).
- OECD. 2010a. OECD test guideline for testing of chemicals. 429: skin sensitisation: local lymph node assay. [Online]. <http://www.oecd-ilibrary.org/environment/test-no-429-skin-sensitisation_9789264071100-en> (accessed 07.11.14.).
- OECD. 2010b. Organisation for economic co-operation and development. Test Guideline No. 487. OECD guideline for testing of chemicals. In vitro mammalian cell micronucleus test. <http://www.oecd-ilibrary.org/environment/test-no-487-in-vitro-mammalian-cell-micronucleus-test_9789264091016-en> (accessed 07.11.14.).
- OECD. 2013. OECDiLibrary [Online]. <<http://www.oecd-ilibrary.org/2013>> (accessed 07.11.14.).
- OECD. 2014. *OECD QSAR Toolbox* [Online].
- Ortel, B., Gange, R.W., 1990. An action spectrum for the elicitation of erythema in skin persistently sensitized by photobound 8-methoxypsoralen. *J. Invest. Dermatol.* 94, 781–785.
- Petry, T., Vitale, D., Joachim, F.J., Mascarenhas, R., Schneider, S., Smith, B., et al., 2013. Human health assessment of scented candle emissions. *The Toxicologist* 132 (1), 333.
- Piersma, A.H., Hernandez, L.G., Van Benthem, J., Muller, J.J., Van Leeuwen, F.X., Vermeire, T.G., et al., 2011. Reproductive toxicants have a threshold of adversity. *Crit. Rev. Toxicol.* 41, 545–554.
- Politano, V.T., Api, A.M., 2008. The Research Institute for Fragrance Materials' human repeated insult patch test protocol. *Regul. Toxicol. Pharmacol.* 52, 35–38.
- Price, P.S., Hollnagel, H.M., Zabik, J.M., 2009. Characterizing the noncancer toxicity of mixtures using concepts from the TTC and quantitative models of uncertainty in mixture toxicity. *Risk Anal.* 29 (11), 1534–1548.
- REACH 2006. Regulation (EC) No. 1907/2006 (REACH) *Official Journal of the European Union*.
- Rousset, F., Verda, D., Garrigue, J.L., Mariani, M., Leclaire, J., 2002. In vitro prediction of contact sensitivity with human cell lines. *Contact Dermatitis* 46, 6 (s4), 6.
- Rulis, A.M., 1986. De Minimis and the Threshold of Regulation. Lewis Publishers, Chelsea, MI.
- Rulis, A.M., 1989. Establishing a Threshold of Concern. Plenum Press, New York, NY.
- Rulis, A.M., 1992. Threshold of Regulation: Options for Handling Minimal Risk Situations. American Chemical Society, Washington, DC.
- Safford, R.J., 2008. The dermal sensitisation threshold- a TTC approach for allergic contact dermatitis. *Regul. Toxicol. Pharmacol.* 51, 195–200.
- 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.* 60, 218–224.
- Safford, B., Api, A.M., Clapp, C., Comiskey, D., Daly, E.J., Ellis, G., et al., 2014a. Use of an aggregate exposure model to estimate consumer exposure to fragrance ingredients in personal care and cosmetic products. *Reg. Toxicol. Pharmacol.* accepted for publication.
- Safford, R.J., Roberts, D.W., Lalko, J., Api, A.M., 2014b. Extension of the Dermal Sensitization Threshold (DST) approach to incorporate chemicals classified as reactive. Manuscript submitted to *Regulatory Toxicology and Pharmacology*.
- Sakaguchi, H., Ashikaga, T., Miyazawa, M., Yoshida, Y., Ito, Y., Yoneyama, K., et al., 2006. Development of an in vitro skin sensitization test using human cell lines; human Cell Line Activation Test (h-CLAT). II. An inter-laboratory study of the h-CLAT. *Toxicol. In Vitro* 20, 774–784.
- 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.
- SCCP. 2012. Scientific Committee on Consumer Safety (SCCS), Scientific Committee on Health and Environmental Risks (SCHER), Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR): OPINION ON Use of the Threshold of Toxicological Concern (TTC) Approach for Human Safety Assessment of Chemical Substances with focus on Cosmetics and Consumer Products. The SCCS/SCHER/SCENIHR adopted this opinion by written procedure on 8 June 2012 after public consultation (SCCP/1171/08). <http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/scs_o_092.pdf>.
- Selgrade, M.K., Sullivan, K.S., et al., 2012. Decision trees for evaluating skin and respiratory sensitizing potential of chemicals in accordance with European regulations. *Regul. Toxicol. Pharmacol.* 63, 371–380.

- Serafimova, R., Fuart Gatnik, M., Worth, A. 2010. Review of QSAR Models and Software Tools for predicting Genotoxicity and Carcinogenicity. *JRC Technical Report*. Publications Office of the European Union, Luxembourg.
- Smith, C.K., Cheung, C., Elahi, E.N., Hotchkiss, S.A., 2001. High-performance liquid chromatography method for the quantification of non-radiolabelled cinnamic compounds in analytes derived from human skin absorption and metabolism experiments. *J. Chromatogr. B Biomed Sci. Appl.* 758, 249–264.
- Taylor, D.K., Anstey, A.V., Coleman, A.J., Diffey, B.L., Farr, P.M., Ferguson, J., et al., 2002. Guidelines for dosimetry and calibration in ultraviolet radiation therapy: a report of a British Photodermatology Group workshop. *Br. J. Dermatol.* 146, 755–763.
- Troutman, J.A., Foertsch, L.M., Kern, P.S., Dai, H.J., Quijano, M., Dobson, R.L., et al., 2011. The incorporation of lysine into the peroxidase peptide reactivity assay for skin sensitization assessments. *Toxicol. Sci.* 122, 422–436.
- Veith, G.D., Petkova, E.P., et al., 2009. A baseline inhalation toxicity model for narcosis in mammals. *SAR QSAR Environ. Res.* 20, 567–578.
- Walker, J.C., Kendal-Reed, M., et al., 2001a. Human breathing and eye blink rate responses to airborne chemicals. *Environ. Health Perspect.* 109, 507–512.
- Walker, J.C., Kendal-Reed, M., et al., 2001b. Human responses to propionic acid. II. Quantification of breathing responses and their relationship to perception. *Chem. Senses* 26, 351–358.
- Williams, G.M., Iatropoulos, M.J., 2002. Alteration of liver cell function and proliferation: differentiation between adaptation and toxicity. *Toxicol. Pathol.* 30, 41–53.
- Worth, A., Lapenna, S., Lo Piparo, E., Mostrag-Szlichtyng, A., Serafimova, R. 2010. *RE: the applicability of software tools for genotoxicity and carcinogenicity prediction: case studies relevant to the assessment of pesticides*.
- Wu, S., Blackburn, K., Amburgey, J., Jaworska, J., Federle, T., 2010. A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. *Regul. Toxicol. Pharmacol.* 56, 67–81.