National Institute
for Public Health
and the Environment

Report 320015003/2010

W. ter Burg | S.W.P. Wijnhoven | A.G. Schuur

Observations on the methodology for quantitative risk assessment of dermal allergens

rivm

RIVM Report 320015003/2010

Observations on the methodology for quantitative risk assessment of dermal allergens

W. ter Burg
S.W.P. Wijnhoven
A.G. Schuur

Contact:
Gerlienke Schuur
Centre for Substances and Integrated Risk Assessment
gerlienke.schuur@rivm.nl

This investigation has been performed by order and for the account of the Ministry of Health, Welfare and Sport, within the framework of V/320015 (kennisvraag 5.1.11) Prioritering en blootstelling van chemische stoffen in consumentenproducten

rivm

Abstract

Observations on the methodology for quantitative risk assessment of dermal allergens

People can have allergic reactions when sensitised by a chemical substance, meaning that they have dermal complaints after the next contact with the substance. Presently, the threshold of this effect can be determined using a quantitative method. According to RIVM, it is important to implement this method for determining the risk on allergic reactions while using (fragrant) substances. A transparent guidance is required in order to be able to assess this risk. This conclusion is based on the literature review by RIVM, performed on request of the Ministry of Health, Welfare, and Sports. The observations are illustrated with a risk assessment of the fragrance citral, which is found in consumer products such as cosmetics, cleaning agents, and air fresheners.

The quantitative method proposed by the International Fragrance Organization /Research Institute for Fragrance Materials (IRFRA/RIFM) in the basis is satisfactory. Some adaptations are considered to be necessary. The most important issue to be included in the QRA is the fact that people are often exposed to more than one source of allergens (aggregated exposure). An example is the use of several cosmetic products or cleaning agents.The total risk could be underestimated, when the risk is determined per product. The case study with citral underlines the findings from literature. To estimate the exposure of a substance from multiple sources, criteria are needed to determine the relevancy for skin allergy. Relevant parameters for further investigation are the exposed location on the body, and also the duration and repetition of exposure.

Key words: aggregated exposure, allergens, skin sensitisation, quantitative risk assessment

rivm

Rapport in het kort

Opmerkingen bij de methodologie voor een kwantitatieve risicobeoordeling van dermale allergenen

Mensen kunnen allergische reacties krijgen als zij door een stof worden ’gesensibiliseerd’, wat betekent dat ze na een volgend contact huidklachten krijgen. De drempelwaarde voor dit effect kan tegenwoordig met behulp van een kwantitatieve methode worden bepaald. Volgens het RIVM is het nu zaak deze methode te implementeren om risico’s op allergische reacties bij het gebruik van (geur)stoffen te bepalen. Om deze risico’s vast te kunnen stellen is wel een heldere handleiding nodig. Dit blijkt uit literatuuronderzoek van het RIVM, dat in opdracht van VWS is uitgevoerd. Het onderzoek is geïllustreerd met een risicobeoordeling van de geurstof citral, die in producten als cosmetica, schoonmaakmiddelen en luchtverfrissers zit.

Gebleken is dat de kwantitatieve methode die door de International Fragrance Organization/ Research Institute for Fragrance Materials (IFRA/RIFM) is ontwikkeld, in de basis voldoet. Wel zijn enkele aanpassingen nodig. De belangrijkste daarvan is dat bij het vaststellen van risico’s meegenomen moet worden dat mensen vaak niet aan één maar aan meerdere bronnen van allergenen worden blootgesteld (geaggregeerde blootstelling). Een voorbeeld is het gebruik van meerdere cosmeticaproducten of schoonmaakmiddelen. Wanneer het risico per product wordt bepaald, zou het totale risico kunnen worden onderschat. De casestudy naar citral onderstreepte de bevindingen hierover uit de literatuur. Om de blootstelling van een stof uit meerdere bronnen te kunnen schatten, zijn nog wel criteria nodig die relevant zijn voor huidallergie. Denk hierbij aan de plek op het lichaam die aan een stof staat blootgesteld, evenals de duur en herhaling van de blootstelling.

Trefwoorden:
Geaggregeerde blootstelling, allergenen, huid sensibilisatie, kwantitatieve risicobeoordeling

rivm

Contents

Summary

Skin sensitisation is a complex mechanism that can lead to allergic contact dermatitis. The mechanism is characterised by two steps, i.e., the induction of skin sensitisation and an elicitation step. Risk assessment for dermal allergens is aimed at protecting the population against becoming sensitised (induction) from exposure to a consumer product.

Until early 2000, traditional risk assessment for skin sensitisation aimed at hazard identification only and resulted in the classification of chemicals either as a sensitiser or non-sensitiser. Recently, a new method for quantitative risk assessment (QRA) of dermal allergens has been proposed by IFRA/RIFM, in which newly developed systems for hazard characterisation are used. These methods are more focused on delivering dose-response curves and potency information, making the risk assessment quantitative.

The steps in the QRA method are discussed in the current report. The proposed method is promising, although some suggestions for improvement of the proposed method, partly reported earlier, are mentioned here. These suggestions are based on risk assessment issues in general, together with specific points for RA for dermal allergens.

* Aggregate exposure to one substance originating from different sources/ products should be taken into account in the QRA method according to IFRA/RIFM. This method results in a safe concentration of the dermal allergen in a (cosmetic) product. However, exposure to the same chemical from multiple sources (aggregate exposure) is not considered.

* In the hazard assessment according to IFRA/RIFM, a No Expected Sensitisation Induction Level (NESIL) is derived based on a weight of evidence approach, including both human test results and animal data. Human testing has been done, but for new chemicals could be unethical as subjects might become sensitised during those tests and they are not necessarily the most robust data. Animal data could also be useful as a point of departure for the QRA, in that case an interspecies factor of 10 is necessary (see also next bullet).

* In the risk characterisation, Sensitisation Assessment Factors (SAFs) are used. For intraspecies differences the classic factor of 10 is used. More specific for the endpoint sensitisation, AFs for matrix and use are incorporated. The IFRA/RIFM method elegantly made a classification of cosmetic products into several categories. To each category a specific matrix and use AF is assigned. Consideration should be given if these factors relating to exposure need to be taken into account when performing the exposure assessment (and not in AFs). The IFRA/RIFM method does not incorporate an interspecies factor because human test results are included. It is suggested that animal data (LLNA) alone should also be considered as a point of departure. In that case, the default interspecies AF is set to 10.

After the description of the QRA method and mentioning some considerations, an attempt is made to illustrate the suggested improvements in a case study with citral as an example. The case study clearly illustrates that exposure from multiple sources should not be ignored in a risk assessment for dermal allergens. However, for now clear criteria are lacking as to when aggregated exposure should be considered. Furthermore, it was shown that the derivation of the NESIL based on a weight of evidence approach is not transparent.

Overall, it is concluded that more research is necessary to clarify the mechanism of action of skin sensitisation in relation t aggregate exposure. Parameters as location of exposure on the body and the time (repeated) of exposure are important factors in the induction of sensitisation. Ideally, the mechanism of action should drive the way the aggregate exposure assessment is performed; however, to date there is insufficient knowledge on how exactly exposure factors influence the process of sensitisation.
Regarding the QRA methodology, it is recommended that more transparent guidance is given for general risk assessors in the (weight of evidence) derivation of a threshold (the NESIL). Transparent criteria for in- or exclusion and other criteria such as the vehicle to be chosen as most relevant, are needed. It is recommended to develop guidance on how to perform a robust QRA, including an aggregate exposure assessment, for dermal allergens that is widely accepted by scientists and risk assessors.

rivm

Samenvatting

Huidsensibilisatie is een complex mechanisme dat kan leiden tot allergische contactdermatitis. Het mechanisme bestaat uit twee stappen, de inductie van huidsensibilisatie en elicitatie van huidsensibilisatie wat tot uiting komt als contactdermatitis. De risicobeoordeling is gericht op het beschermen van de populatie tegen het gesensibiliseerd raken (de inductie) na blootstelling aan een stof vanuit een consumentenproduct.

De klassieke risicobeoordeling voor huidallergenen was enkel gebaseerd op gevaarsidentificatie die resulteerde in wel of niet classificatie van chemische stoffen als huidallergeen. Recent is er een nieuwe methodiek voor een kwantitatieve risicobeoordeling (QRA) van huidallergenen voorgesteld door IFRA/RIFM. De voorgestelde methodiek ziet er veelbelovend uit. In dit rapport worden de verschillende stappen in de QRA methode besproken, waarbij enkele, deels eerder gepubliceerde suggesties voor verbetering worden benoemd
.
* Blootstelling aan dezelfde chemische stof uit andere bronnen (geaggregeerde blootstelling) zou moeten worden meegenomen in de IFRA/RIFM methode. Deze QRA methode resulteert in een veilige concentratie van een dermaal allergeen in een enkel (cosmetisch) product, maar houdt daarbij geen rekening met mogelijke andere bronnen.

* In de effectbeoordeling volgens de IFRA/RIFM methode wordt een ‘No Expected Sensitisation Induction Level’ (NESIL) afgeleid. Deze is gebaseerd op een ‘weight of evidence’ aanpak, gebruik makend van humane testresultaten. Het inzetten van humane testen wordt echter gezien als onethisch, omdat mensen gesensibiliseerd zouden kunnen raken tijdens zo’n test. Voorgesteld wordt om data vanuit dierstudies (LLNA) ook mee te nemen als een mogelijk startpunt voor de QRA. In dat geval zal een interspecies assessment factor van 10 gebruikt moeten worden.

* In de risicobeoordeling worden ‘Sensitisation Assessment Factors’ gebruikt. Voor intraspecies verschillen wordt de klassieke factor 10 gebruikt. Meer gericht op het eindpunt sensibilisatie zijn ook assessment factoren voor matrix en gebruik ingevoerd. De IFRA/RIFM methode heeft een classificatie van cosmetische producten in verschillende categorieën opgesteld. Aan elke categorie is een specifieke matrix- en gebruiks-AF toebedeeld. Het zou overwogen moeten worden of deze AFs wel nodig zijn, of dat deze factoren niet meegenomen zouden moeten worden in de blootstellingschatting.

Na de beschrijving van de QRA methode, en het benoemen van een aantal overwegingen hierbij, is geprobeerd de voorgestelde verbeteringen te illustreren in een case studie met citral als voorbeeld. In deze casus wordt duidelijk aangetoond dat blootstelling aan een stof vanuit meerdere bronnen niet genegeerd kan worden in een risicobeoordeling van een dermaal allergeen. Een bijkomend probleem is dat er in de literatuur geen duidelijke criteria zijn wanneer een geaggregeerde blootstelling meegenomen zou moeten en mogen worden op basis van wetenschappelijke gronden. Daarnaast is in dit voorbeeld aangetoond dat de afleiding van de NESIL vanuit humane en diergegevens geen transparant proces is.

Geconcludeerd kan worden dat verder onderzoek naar het mechanisme van huidsensibilisatie in samenhang met geaggregeerde blootstelling noodzakelijk is. Parameters zoals de plek van de blootstelling op de huid en de tijdsduur (herhaald) van blootstelling zijn belangrijke factoren bij de inductie van sensibilisatie. In de ideale situatie zou het mechanisme van sensibilisatie de wijze van

aggregeren van de blootstelling moeten aansturen, maar op dit moment zijn er onvoldoende gegevens over hoe blootstellingsfactoren het proces van sensibilisatie beïnvloeden.

Als aanbeveling voor de QRA methodologie is aangegeven dat er een noodzaak is voor een heldere guidance voor de algemene risicobeoordelaar. Punten van aandacht hierbij zijn hoe het beste een 'weight of evidence' afleiding van de drempelwaarde (NESIL) uitgevoerd moet worden, transparante criteria voor het wel of niet meenemen van studies of criteria zoals het vehikel, en aandacht voor een geaggregeerde blootstellingschatting. Aanbevolen wordt dat een guidance wordt ontwikkeld waarin staat hoe een goed onderbouwde QRA voor dermale allergenen uitgevoerd moet worden, inclusief een geaggregeerde blootstellingschatting. Brede acceptatie van de guidance is noodzakelijk, door zowel immunotoxicologen en risicobeoordelaars.

1 Introduction

In general, allergic diseases are among the most important causes of health problems world wide with a prevalence of 15-30% of the population in developed countries (European Allergy White Paper, 1997). Contact dermatitis is relatively common compared to other conditions such as asthma, hay fever and food allergies, with a prevalence of around 4% in the Netherlands. In the future, the prevalence of allergies is expected to increase, partly due to changes in environment and lifestyle (Health Council, 2007). To react to that development, the Dutch Food and Safety Authority assigned the National Institute for Public Health and the Environment (RIVM) to set up a website to collect complaints on allergic reactions by cosmetic products from consumers (CESES; www.cosmeticaklachten.nl in Dutch). Substances in products that come into contact with the skin play an important role as exogenous factors in the triggering of allergenic contact eczemas at work, but also at home. A number of known allergic substances have been identified in a wide range of consumer products (Wijnhoven et al., 2008).
Skin sensitisation risk assessment of new products is critical before introduction into the market place. However, in the past, traditional risk assessment methods for skin sensitisation aimed at hazard identification only and were used for classifying and labelling chemicals either as sensitisers or non-sensitisers. It was assumed that allergic reactions to chemicals followed the mechanistic principle of an all-or-none response that lacks dose-response relationships and thresholds. Since the year 2000 and after a lot of research, it became more and more accepted that skin sensitisation as well as elicitation (see section 2.1.) is only occurring above threshold doses and follows predictable dose-response relationships (Griem et al., 2003). A more quantitative method for hazard characterisation and adequate risk assessment approach can thus be conducted.

In 2008, the RIVM report 'Allergens in consumer products' (Wijnhoven et al., 2008) was written at request of the Food and Consumer Product Safety Authority in The Netherlands to provide more insight in the different aspects that are related to allergies due to the use of consumer products. It contains an inventory on the kind and concentrations of allergic substances in consumer products. Limit values have been set by law for a number of allergenic substances in products, however they are not based on a quantitative risk assessment (QRA). The report from Wijnhoven et al. concluded that it is important to critically look at a quantitative method for risk assessment for dermal sensitisation, which at that time was under development. Later on, Api et al. (2008) proposed a QRA for dermal sensitisation of fragrance ingredients in cosmetic products (further referred to as the IFRA/RIFM QRA method). Although the proposed IFRA/RIFM method seems promising, one of the main criticisms noted by the Scientific Committee on Consumer Products (SCCP, 2008) and also by the Wijnhoven report is that in the QRA proposed, a safe limit is derived for an allergic substance in *one* product. Exposure to the specific substance from several cosmetic or other consumer products (aggregate exposure), however, is not taken into account.

At RIVM, three reports have recently been written on aggregate exposure. Aggregate exposure is defined as the combined exposure to a chemical via several routes (inhalatory, dermal and oral) and via several sources (diet, consumer products, air, dust and so on). The first report titled 'Aggregating human exposure to chemicals. An overview of tools and methodologies' (Delmaar and van Engelen, 2006), was requested within the framework of an EU project INTARESE (Integrated Assessment of Health Risks of Environmental Stressors in Europe). The aim of this report was to examine the suitability of available computer models for evaluating aggregate exposure to consumer products. In a second report (Wolterink et al., 2009), requested by the Food and Consumer Product Safety Authority and the Ministry of Health, Welfare and Sport (VWS), four case studies were performed using an aggregate exposure assessment, to explore the current possibilities and limitations of an aggregate risk

assessment. It was concluded that aggregate exposure is the bottleneck in chemical risk assessment which is often due to the lack of relevant data necessary for exposure assessment. Finally, for the Ministry of Health, Welfare and Sport, an overview was made on available guidance and use of aggregate exposure assessment in different legal frameworks (Schuur et al., 2009). The authors of this report concluded that performing an aggregate exposure assessment should be stimulated in more legal frameworks, for instance by the EU processes of methods and guidance documents. It was recommended that risk assessments need to be performed focusing on a substance, not on a product. And, when an aggregated risk assessment spans different legal frameworks, it should be discussed whether the level of protection, the model choices and the exposure levels are fitted correctly. When there is a risk, it should be discussed which framework will take the lead in addressing the problem.

Summarising, two developments have taken place: 1) a promising method has been proposed in the field of QRA for dermal sensitisers and 2) the increasing interest in the use of aggregated exposure assessments in risk assessment. Taking all the information above into account, the question presented itself within the project financed by the Ministry of Health, Welfare and Sport (VWS) to combine these two fields of investigation: i.e., further development of a quantitative risk assessment for dermal sensitisation while taking aggregate exposure assessment into account.

For this reason, the different elements of published IFRA/RIFM QRA method for dermal sensitisers in cosmetic products (Api et al., 2008) were evaluated. As Api et al. themselves note ‘Although it is desirable to use aggregate exposure, there are insufficient data to allow this to occur at this time. This is identified as an area of refinement for a QRA approach’. The SCCP (2008) also commented on the absence of integrating exposure to a substance from several products. This absence of an aggregate exposure assessment as well as the possibility to integrate aggregate exposure into this QRA will be studied, amongst others by using citral, a well-known fragrance material, as a case study.

It should be realised that the current report focuses on dermal sensitisation only. The issue of respiratory sensitisation is even less investigated and needs separate attention. A first inventory on scented products with fragrances able to cause respiratory sensitisation has recently been made and has been presented in another RIVM letter report (Ezendam et al., 2009). In addition, the World Health Organisation - International Programme of Chemical Safety (WHO-IPCS) is currently preparing a Guidance Document on Immunotoxicity for Chemical Risk Assessment, including immunosuppression, immunostimulation, sensitisation and autoimmunity associated with chemical exposure. This document will also include a chapter on guidance for a QRA of dermal allergens.

Outline of the report

The aim of this report is to evaluate the IFRA/RIFM method for a quantitative approach of a risk assessment of a dermal sensitiser and where possible, to suggest improvements. At first, some background information on sensitisation and risk assessment, including aggregate exposure assessment in general, will be given (Chapter 2). This chapter also highlights possible challenges in using aggregate exposure assessment in a QRA for skin sensitisation. Then, the IFRA/RIFM QRA method (Api et al., 2008) will be described and possible shortcomings in the method will be identified (Chapter 3). To obtain more insight, an example substance, citral, will be assessed for its dermal sensitisation effect according to the QRA method, including an aggregate exposure assessment. Citral was chosen because it is a weak to moderate sensitiser, used as fragrance ingredient and present in cosmetics as well as detergents (Chapter 4). The results of the evaluation and outcome of the case study will be discussed in Chapter 5, resulting in suggestions for improvement. Furthermore, recommendations for future implementation of a QRA of dermal allergens are given (Chapter 5).

2 Background chapter

2.1 Allergic contact dermatitis

Allergy is defined as an adverse condition which manifests itself following a hypersensitivity reaction caused by exposure to an exogenous antigen. Allergic contact dermatitis (ACD), caused by dermal allergens, is a type IV or delayed type hypersensitivity reaction, which means that it is an allergic response that is mediated by T cells. From the occupational, consumer and environmental health point of view, ACD and hypersensitivity responses in the respiratory tract represent the most important types of allergy induced by chemicals. A hypersensitivity reaction develops in two phases: a learning phase without symptoms, sensitisation, followed by the immune response effector phase, elicitation. Because of these two phases, hypersensitivity reactions pose some particularly challenging problems for quantitative risk assessment. Induction of sensitisation can go without notice or without signs or symptoms of allergy. After sensitisation, contact with the same allergen, even at lower concentrations required to induce sensitisation, will lead to symptoms of allergic disease. The dose responses for sensitisation and elicitation are different but not entirely independent. As a general rule, the concentration needed for elicitation is inversely related to the concentration at which the induction process was triggered (Friedman, 2007). In practice, it is sometimes difficult to determine the point at which sensitisation ends and elicitation begins, especially after long-term exposure. Development of sensitisation is always a systemic reaction, although allergic reactions may preferably occur at localised sites: in case of ACD at exposed skin areas or even at non-exposed skin areas. Inhalatory exposure might also result in ACD in previously sensitised subjects. On the other hand, oral exposure to dermal sensitisers might result in oral tolerance, thereby somehow protecting subjects from becoming sensitised.

allergic contact dermatitis

step 1. sensitisation
- binding of chemical (= hapten) to proteins or other macromolecules
- internalisation of hapten-modified proteins
- hapten-induced activation of Langerhans cells (LCs), migration and process hapten protein complexes
- presentation of antigens by LCs to naïve specific T-cells
- proliferation of antigen specific T-cells; memory T-cells are formed
- hapten-specific memory T-cells leave the lymph node and enter the circulation

step 2. elicitation
- re-exposure to the chemical
- release of cytokines and chemokines attracting cells to the skin from the circulation
- inflammatory response within 24-48 h, symptoms of ACD

2.2 Risk assessment for sensitisers

Risk assessment for chemically induced hypersensitivity has two components:

- the likelihood that a chemical will induce sensitisation in a previously non-sensitised individual
- the likelihood that a chemical will provoke an allergic reaction in those who are already sensitised.

However, in most cases, the risk assessment focuses on the first step, resulting in a safe situation of not becoming sensitised. Effectively and obviously, this would also protect subjects from allergic reactions unless already sensitised. Hence, in most cases, the lower level of elicitation is hardly taken into account in risk assessment.

The general risk assessment of chemicals consist in hazard characterisation (including hazard identification and dose response assessment) and exposure assessment (including external exposure assessment and toxicokinetics). The base elements of risk assessment are schematically depicted in Figure 2.1. After the assessment of the (external) exposure (the total ingested, dermally applied or inhaled dose of chemicals), information on toxicokinetics of the chemicals (the absorption, distribution, metabolism and excretion) determines the internal dose (the dose of chemicals that reaches the systemic circulation, organs and tissues of man).

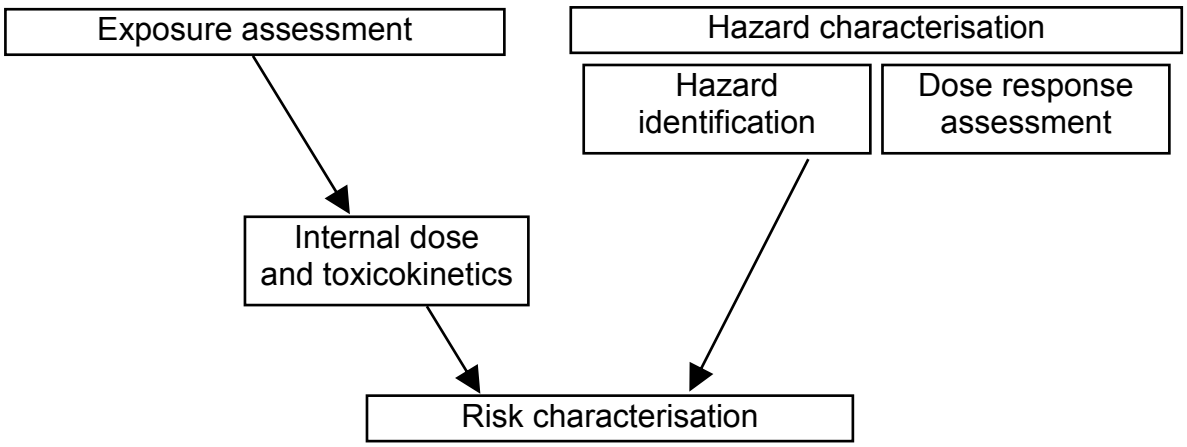

Figure 2.1 Base elements of general risk assessment.

In the past, characterisation of the risks of being exposed to dermal allergens was restricted to an indication of the hazard. In legal frameworks, such as for classification and labelling, this is still common practice. Animal and human testing was aimed at identification of allergenic properties of a chemical, but generally information on potency or amount of substance required for sensitising was lacking. However, during the last ten years, following extensive research, it became clear that skin sensitisation is no different from other toxicological effects in the sense that skin sensitisation as well as allergy elicitation is only occurring above threshold doses and follows predictable dose-response relationships (reviewed in Griem et al., 2003). Newly developed systems for hazard characterisation, such as the Local Lymph Node Assay (LLNA) in mouse (Van Och et al., 2000; Kimber et al., 2001; Griem et al., 2003) and the Human Repeat Insult Patch Test (HRIPT) in humans, are therefore more focused on delivering dose-response curves and potency information. Using these systems, a more

quantitative and adequate risk assessment for sensitisers can be conducted. This will be described more specifically in paragraph 2.5.

2.3 Exposure assessment

The exposure assessment should always be set up in view of the anticipated toxicological effect, in this case dermal sensitisation. Although ACD results from a systemic effect, the mechanism of (dermal) sensitisation is not completely systemic. It is supposed that the antigen specific T-cells, drained by a certain local lymph node, must reach a certain threshold before a subject is actually sensitised (Basketter et al., 2000). This mechanism of action should drive the aggregated exposure assessment. To understand better the difficulties in aggregate exposure assessment of dermal sensitisers, some basic principles regarding the important exposure factors are first listed by taking the mechanism into account. The dose metrics that drive the induction of skin sensitisation should be appreciated. The available evidence suggests that under most normal conditions of exposure, it is the dose per unit area of chemical, rather than the total delivered dose, that has an overriding impact on the effectiveness of sensitisation (Kimber et al., 2008).

2.3.1 Important exposure factors for skin sensitisation

Next to certain chemical characteristics of a dermal allergen, there are a number of exposure factors which are considered to be (possibly) relevant for the induction of skin sensitisation. The dose of a chemical applied per area of skin (µg/cm^2/day) is considered to be the most relevant dose metric. Allergen exposures in the literature are mostly expressed as percentages (weight of allergen per volume of vehicle applied to the skin). This leads to the assumption that in any given test system an equal percentage of a chemical will lead to a similar incidence and/or severity of skin sensitisation. However, the dose applied in combination with the surface area on which the allergen is applied and the layer thickness of the product on the skin are much more important than chemical percentages in a product, since a certain number of LCs are required to be activated in order to initiate the cascade of events that leads to induction of skin sensitisation (discussed in Api et al., 2008). Also important to know is the applied versus the intradermally delivered dose, given that there are several factors that can influence the effective amount of the material delivered to the viable epidermis – evaporation, binding/ sequestration in the skin, integrity of the skin and dynamic processes, such as metabolism. Based on the above, the factors that might be important for the dermal load in the exposure assessment regarding dermal sensitisation are thought to be the location of the exposure, product matrix effects, duration of the exposure, repeated exposure and finally, aggregation of exposures from multiple sources (Basketter et al., 2006; De Jong et al., 2007; Api et al., 2008).

2.3.1.1 Location of exposure

Dermal absorption is a very important step in skin sensitisation. A substance must be able to cross the skin barrier in order to elicit reactions that lead to sensitisation of the skin. The skin is mainly an exterior barrier to all kinds of insults, such as heat, cold, radiation and chemical substances. In short, the skin is built up from epidermis, with the outermost barrier stratum corneum and hair follicles, the dermis and the hypodermis. The stratum corneum is the most important skin layer for protection against hydrophilic and lipophilic chemicals. The thickness of human stratum corneum is site-specific and values observed range from approximately 10 to 50 µm. Human skin may differ significantly between individuals (age-groups and race) and within an individual (site-specific differences). Skin metabolism might also ‘activate’ substances by forming reactive metabolites, which are able to induce skin sensitisation. Metabolism of substances in the skin by enzymes present may also differ site-specifically. The observed skin differences in humans indicate that skin is a complex organ, which is

probably the cause of the large variation in presence and severity of skin effects observed in humans (Marzulli and Maibach, 1987; Scott et al., 1991; Wilkinson et al., 2006).

Next to genetic influences on the skin, personal behaviour may affect the skin barrier, such as the shaving of the skin or sun bathing resulting in sunburns. Certain skin diseases, such as psoriasis, can also affect the skin barrier function. Dermal absorption of substances through damaged skin is thought to be more effective than in the case of healthy, normal skin.

Dealing with the skin differences between and within humans in an exposure assessment is hardly possible. A higher dermal absorption due to a weaker skin barrier would lead to a higher concentration of the chemical in the skin in contrast to a skin site with a better barrier to the insult. Consequently, a higher concentration in the skin leads potentially to more severe skin sensitisation. In the QRA for dermal allergens the dermal load per unit skin is used as the preferred dose metric, which in fact is a surrogate for the concentration in the skin. However, the same dermal load per unit skin could result in a range of concentrations in the skin dependent on the location of the skin. In addition, this concentration in the skin is not determined during toxicity testing. Based on the above, the concentration in the skin cannot be used as point of departure (PoD) in the QRA. Thus, in agreement with the proposed method by IFRA/RIFM, it is recommended to use the dermal load per unit skin as PoD, and to apply an assessment factor to account for skin differences.

It is noted that layer thickness of a product on the skin influences the possibility of the chemical contacting the skin. For example, if 100 mg of the same product is spread on 10 cm^2 or 100 cm^2 skin, the same amount would result in concentrations of 10 mg/cm^2 and 1 mg/cm^2, respectively. Although the concentration in the latter case is lower, theoretically a relatively higher amount of the chemical can come into contact with the skin because of the thinner layer thickness. In the QRA for dermal sensitisers, the dermal load per unit skin is used. This implicitly assumes that all chemicals within the product put on the skin come into contact with the skin, regardless of the layer thickness. Thus, an overestimation is made of the amount of chemicals coming into contact with the skin.

The location of exposure also plays a role in the accumulation of a chemical after repeated exposure to a certain chemical from the use of a single product repeatedly or from the use of multiple products. This issue will be discussed in section 2.4 under ‘aggregate exposure’.

2.3.1.2 Matrix effects

The matrix of the product containing the dermal allergen can influence the actual exposure to the chemical (Api et al., 2008). The matrix may for instance contain irritants or skin penetration enhancers, which may lead to a higher dermal absorption of the chemical and thus to a higher concentration in the skin. On the other hand, the chemical of interest may be bound to ingredients that for example decrease dermal absorption. In animal testing, the chemical of interest is administered using a vehicle, such as ethanol or certain oils, which may also influence the exposure (Api et al., 2008; Lalko and Api, 2008). Accounting for differences between matrices on the one hand (exposure estimation) and vehicles on the other (hazard setting) is difficult. Therefore, an assessment factor for matrix effects is proposed to be included in the QRA to cover these differences (see 3.4.3.3).

2.3.1.3 Area size of exposure

It is the dose per unit area of a chemical that determines the level of sensitisation (Friedman, 1990; 2007; Kimber et al., 2008). The exception to this rule is when the area of the application site drops below a certain critical level, as has been observed with 2,4-dinitrochlorobenzene (DNCB) and with other contact allergens (Kligman, 1966). The mechanistic basis for the importance of dose per unit area

is the activation of lymph nodes draining the site of contact with contact allergens. There is a need for a certain critical number of antigen-presenting cells reaching the draining lymph nodes. This is to provide a signal of sufficient magnitude and intensity to overcome the threshold required for the triggering of a primary immune response. Translating this into the dose metric found to be relevant for the acquisition of skin sensitisation; there is a certain critical level of antigen (= chemical allergen) and a critical level of LCs available for transport required in the area of skin from which there is lymphatic drainage to a single lymph node, or to a series of lymph nodes in the same area. With sufficient numbers of LCs carrying sufficient allergen reaching the lymph node, the threshold for effective stimulation is reached and exceeded. Consequently, as the critical mass increases further, then so will the strength of induced immune responses (and the extent to which skin sensitisation is acquired). The consequence is that if the same amount of chemical is distributed over a larger area of skin, where lymphatic drainage is served by several lymph nodes not in dose proximity, the critical mass of available allergen per lymph node will be reduced and the level of immune activation will be diminished accordingly (Kimber et al., 2008).

2.3.1.4 Duration of exposure

There is little information on the impact of duration of exposure on the induction of contact allergy. Most of the studies were performed in allergic subjects, looking at elicitation responses. In subjects allergic to *p*-paraphenylenediamine (PPD), it has been shown that there is a clear relationship between the exposure time and allergic responses to 1.0 % PPD: exposure for 5 minutes: 16% responded, for 30 min: 38% responded, for 120 min: 69% responded. Exposure to lower concentrations of PPD showed similar relationships, albeit with lower responses. There was no response when exposure was 1 or 2 min. Similarly, there was a direct relationship between the concentration applied and the number of patients that responded (Hextall et al., 2002). Thus, both exposure time and concentration have been shown to have a direct relationship with elicitation of an allergic reaction.

Whether this relationship also exists in non-patients (meaning not already sensitised) was researched by Basketter et al. (2006), by varying the exposure duration to 1.0% PPD. They have shown that reduction of exposure duration from 48 hr to 5 min decreased the incidence of sensitisation from 54% to 3%. Based on this latter result it can be assumed that the exposure duration does play a role in skin sensitisation. However, in an extended clinical study it was observed that infrequent but longer duration and higher concentration exposure to PPD was significantly less likely to induce sensitisation than more frequent, short duration and lower concentration exposure. This was shown in small groups of individuals using different concentrations of PPD, different durations and different frequencies.
Based on the findings above, it appears that concentration, exposure duration, and the time between exposures together have an effect on skin sensitisation. Supposedly, there would be a minimal duration of exposure required for a chemical to cross the skin barrier and induce a reaction. Furthermore, it may require several exposure events to reach the threshold for induction. Likely, the required minimal exposure duration is chemical dependent and is logically also dependent on the effectiveness of the skin barrier at a specific location.

Contact sensitisers can rapidly be taken up in the epidermis, as has been shown in the murine LLNA. Epidermal cytokines were activated already 30 minutes after application of DNCB.

In the exposure assessment for QRA this would mean that in general, there is no need to take exposure duration into account, because even after five minutes 3% of the subjects were sensitised. Furthermore, migration to the draining lymph node occurred within 4 hours after application and continued for up to 72 hours (Cumberbatch et al., 2005) showing that the skin sensitisation process may continue after cessation of the exposure. Exposure duration should be taken into account when it is demonstrated that

the dermal absorption of a certain chemical is very slow. The default assumption that even a short-term exposure could induce skin sensitisation would be protective and is therefore recommended in the QRA for dermal sensitisers.

2.3.1.5 Frequency of exposure

Repeated exposure to a dermal allergen may be important for inducing skin sensitisation. It has been shown that frequent, short duration exposure to lower concentrations is more likely to induce sensitisation than infrequent but longer duration to higher concentrations of PPD (Basketter et al., 2006). In mice, it has been shown that PPD accumulates in the skin (White et al., 2007) and repeated application of low doses of PPD might accumulate to levels above the threshold. This has been shown in subjects allergic to PPD, where repeated exposure to low concentrations of PPD gave similar results as a single exposure to a higher concentration (White et al., 2007). That accumulation of chemicals in the skin after repeated exposure to concentrations below the threshold might lead to (induction of) sensitisation has been demonstrated for formaldehyde donors by De Jong et al. Formaldehyde donors (chemicals that are highly reactive with proteins and may thus persist in the skin) were administered at concentrations that induce a sensitisation index of two after a single exposure, which is below the threshold for hazard identification (a sensitisation index of three is required). Results show that repeated and prolonged exposure to this concentration can still induce sensitisation in mice (De Jong et al., 2007). However, for isoeugenol it was shown that repeated exposure to low concentrations in a repeated open application test (ROAT) did not elicit allergic reactions in subjects already sensitised to isoeugenol (Johansen et al., 1996). Currently, we do not know whether or not isoeugenol can accumulate in the skin and thus, no explanation can be given for this observation.

Kimber et al. (2008) remarked that when repeated exposures at the same site are closely consecutive (such that the first dose has not cleared before the second exposure), then the true dose per unit area is more difficult to ascertain and calculation of the total dose may represent a more pragmatic solution. Although single sequential exposures to separate lymph nodes are still able to induce a cutaneous immune response, the acquisition of skin sensitisation will be more vigorous when repeated applications are focused through a single lymph node.

The frequent short duration exposure is characteristic for the use of many products by consumers, e.g., the use of deodorant or perfume. Repeated exposure to a product and the chemical of interest is therefore considered very important for risk assessment. Because the same chemical can be present in several products and thus repeated exposure may also result from the use of several products, the risk assessment should not be limited to a single product.

2.4 Aggregate exposure

Human exposure to almost all allergens results from many different sources. Sometimes, exposure may occur only via one product (e.g., a cosmetic product) but in many cases, other cosmetic products and product groups, such as detergents or clothing, are also involved. Establishing an exposure estimate is therefore complex and requires an aggregate exposure assessment. Aggregate exposure is defined as the total exposure that arises from multiple sources via different pathways and routes. The level of detail at which aggregation might be done, should be dictated by the scope and purpose of the assessment (Delmaar and Van Engelen, 2006). For the performance of a realistic quantitative aggregate exposure assessment, many data are needed.

Exposure to multiple products containing the same chemical of interest, the so-called aggregated exposure assessment, should be estimated in the QRA for dermal allergens. This requires information on consumer behaviour and detailed use information of the specified products. For instance, the repeated exposure to the chemical should be more or less in the same skin area for inducing skin sensitisation. In a paper of Kligman (1966), the induction of skin sensitisation with four different products containing the same contact allergen was studied in human volunteers. In this study it was shown that it is far more effective to expose a person to four sequential 48 hour exposures to different, but adjacent sites on one extremity than to give four sequential 48 hour exposures of exactly the same dose per unit area to each of the four extremities (arms and legs). The conclusion drawn by Kligman was that '*bombardment of the same lymph node is superior to stimulation of four different nodal systems'*. Thus, the location of use is crucial in aggregating the exposure. Next to that, it must be clear whether or not a chemical will accumulate in the skin, instead of being metabolised. In the latter situations, time obviously plays a vital role.

To achieve such an aggregated exposure assessment, a high-end user of products containing the chemical of interest should be identified. The next step is to identify which products would lead to an exposure at the same skin area within a certain time period. The exposure to the individual products containing the chemical should ultimately be summed up. The basic assumption underlying this approach is that the chemical is able to accumulate in the skin and past exposure would still contribute to the induction of skin sensitisation.

2.5 Effect assessment

2.5.1 Hazard identification

Hazard identification of sensitising chemicals has been comprehensively discussed in a WHO document (1999). It is focused on responses of the immune system and not on the general screening of changes in all body systems. There is a certain relationship between dose (exposure) and effect. It is assumed nowadays that there is a certain threshold for the induction of skin sensitisation. This relation, however, only becomes apparent after a second challenge, due to the mechanism of induction and elicitation within skin sensitisation, which makes testing somewhat more complicated. The elicitation response usually occurs at lower concentrations than are required for the induction step of skin sensitisation.

2.5.2 Hazard characterisation

Predictive methods such as the Buehler Assay or Guinea Pig Maximisation Test (GPMT) do not incorporate a dose-response analysis, and activity is measured as frequency of responses rather than as severity of responses. Therefore, there is a need for information on potency defined as the quantity of chemical necessary to induce sensitisation (or elicitation). This is very important for the risk assessment aiming to prevent sensitisation as well as elicitation (excerpted from WHO, 1999). In the following sections, human and animal tests to identify and characterise skin sensitisers are described.

2.5.2.1 Human data - induction

HRIPT tests done in classical design using several different induction doses are appropriate for determination of a dose-response curve, no-observed effect level (NOEL) and lowest observed effect level (LOEL). In case of testing only one dose, problems exist when a high percentage of subjects are sensitised and no LOEL can be identified, since there is no information on a level at which no

sensitisation occurs. These data are only to be used when the sensitisation rate is below 50%. To extrapolate to a suitable LOEL, a factor of 3 has been suggested to apply to doses producing sensitisation rates between 10 and 25% and a factor of 10 for sensitisation rates between 25 and 50% (Griem et al., 2003).

In human hazard identification tests used in earlier times, only one high dose was normally tested. This resulted in many subjects being sensitised without identifying a LOEL and NOEL. Today, for ethical reasons, human assays are no longer used for identification of skin sensitisation hazard; this is at first determined in animal assays. The HRIPT is sometimes employed as a confirmatory test for lack of sensitisation at 'the' NOEL in animal models or derived from QSARs.

Sometimes, epidemiological data are also available that provide hazard and exposure assessment information. Data include information from occupational and non-occupational cohorts, the general population or dermatology clinic patients and may comprise patch testing and/or questionnaire data. No proof of sensitising hazard can be concluded from negative epidemiological data, but the prevalence of ACD in an exposed (sub)population may indicate a sensitising hazard and provide dose-response information if the dose is adequately assessed and can be expressed in terms of skin area dose. At best, a NOEL, LOEL or bench mark dose (BMD) can be derived from these data.

2.5.2.2 Human data - elicitation

Elicitation is more often tested in humans; however, an elicitation threshold has been experimentally determined only for a small number of chemicals. This is because the determination of a NOEL/ LOEL for elicitation is not usually the aim of the patch test study. Determination of a dose-relationship for elicitation is possible after different experimental set-ups: in clinical patch tests and the repeated open application test (ROAT).

As PoD for risk assessment, different options have been proposed: the patch test minimum elicitation threshold (MET_{10}, inducing a threshold response in 10% of the subjects tested), a NOEL or BMD from ROAT or from a use test (Weaver et al., 1985; Sosted et al., 2006; Zachariae et al., 2006).
When elicitation thresholds are determined in newly sensitised subjects, a relationship between sensitisation dose and elicitation threshold is visible; the higher the sensitisation dose, the lower the elicitation threshold (Friedmann et al., 1983). Previously sensitised subjects may become more sensitive to a challenge after repeated exposure, where they may respond to decreasing levels of exposure. Although it has not formally been shown that a 'minimum threshold' is finally approached over time, the thresholds in well-established allergic individuals seem more reliable than those determined after experimental sensitisation.

2.5.2.3 Animal data - induction

The LLNA was originally used for qualitatively identifying sensitising chemicals when the stimulation index is ≥3. However, since three different doses are used, the study will provide a dose-response curve for induction of sensitisation. The sensitising potency is expressed as the EC3 value, the effective concentration of a chemical leading to a threefold increase in proliferation of lymph node cells compared to non-exposed controls.
When comparing human NOELs and BMDs with LLNA thresholds, the average ratio of both values is close to 1, indicating a direct comparison between mice and humans (EC3 of 10 μmol/cm^2 in mice corresponds to a NOEL/ BMD of 10 μmol/cm^2 in humans). EC3 is therefore proposed as a surrogate NOEL in QRA (Api et al., 2008; Basketter et al., 2000, 2005a,b; Gerberick et al., 2001a,b; Griem et al., 2003; Schneider and Akkan, 2004). The REACH guidance (ECHA, 2008) considers the EC3 value as a LOAEL for induction, although it is commented in a footnote that some papers suggest that the LLNA

EC3 value is close enough to the human NOAEL and therefore can be used as a surrogate for the NOAEL.
Tests in guinea pigs (guinea pig maximisation test (GPMT), Buehler test), have been used to identify possible sensitisation hazards. However, they provide poor information with regard to sensitisation potency. Although the protocols of these tests have recently been modified in order to generate more useful potency data (Andersen et al., 1995; Van Och et al., 2001; Yamano et al., 2001), no validation of the protocols has taken place yet.
In case no QRA can be performed and no NOEL can be derived, there are some semi-quantitative approaches using potency categories (ECETOC, 2003; Felter et al., 2002; Gerberick et al., 2001a). All data from LLNA tests are put together in a weight of evidence (WoE) approach to assign a substance to one of several potency categories. As a starting point for risk assessment, the lower boundary of the potency category in which the substance is grouped is used. Different systems of categorisation exist but the disadvantage of all these systems is that 'artificial' tenfold steps are introduced in a continuum of sensitisation potencies. Figure 2.2 summarises the different potency categories as they have been described in literature until now.

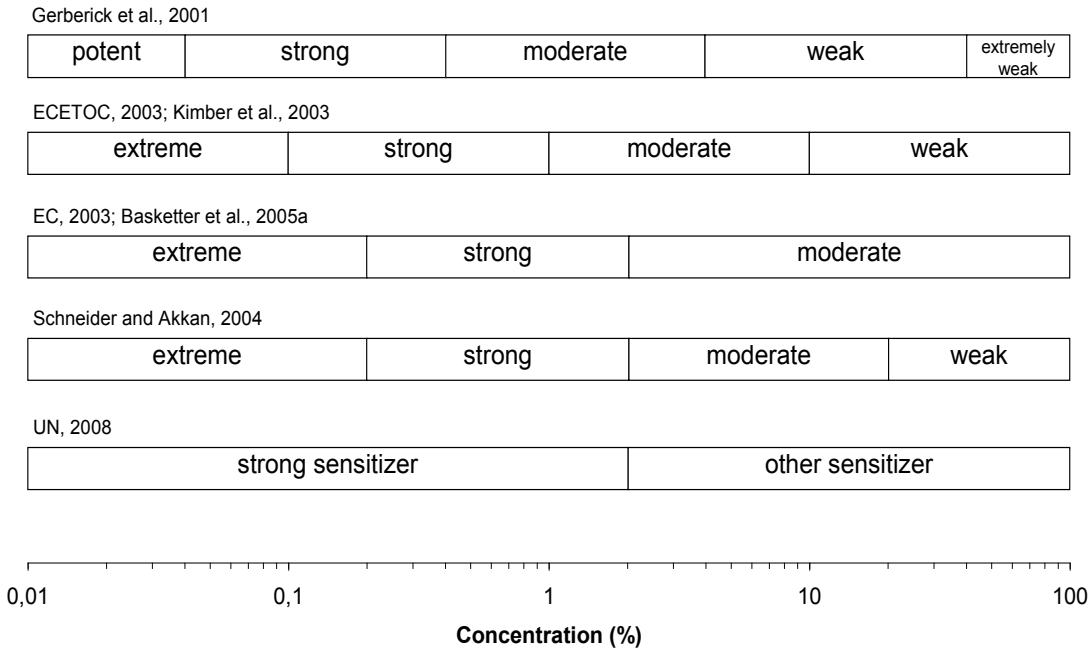

Figure 2.2 Overview of potency categories for skin sensitisers based on LLNA EC3 values.

Recently, a two-category system has been introduced for classification and labelling of sensitisers under the Globally Harmonised System (GHS): category 1A, strong sensitisers (EC3 ≤ 2%) and category 1B, other sensitisers (EC3 > 2%) (UN, 2008).

2.5.2.4 Animal data - elicitation

Human data concerning elicitation are considered to be more reliable than animal data because the data are based on human subjects with a well-established skin allergy (see 2.5.2.2). For this reason, the elicitation process is generally not considered in animal studies.

2.5.2.5 In vitro data

For animal welfare reasons and complying with requirements imposed by the chemical legislation in Europe, there is an increasing emphasis on development of in vitro approaches as alternative for in vivo testing methods. These in vitro methods have to address the various elements of the immune response to skin-sensitising chemicals;

- skin penetration (bioavailability)
- quantitative measurement of chemical reactivity with glutathione
- peptide or proteins with and without metabolic activation of chemical reactivity (Gerberick et al., 2007; Natsch et al., 2007)
- measurement of chemical activation of keratinocytes (Van Och et al., 2003; Coquette et al., 2003) and dendritic cells (Aeby et al., 2007; Sakaguchi et al., 2006)
- response of T-cells against haptenated peptides

Since none of the elements alone is representative for skin sensitisation potential, all the results have to be integrated again, afterwards. One method for integration of these tests is to use the data to assign a potency category to the substance (Jowsey et al., 2006; Natsch et al., 2009), for example, as shown in section 2.5.2.3.

2.5.2.6 General sensitisation threshold approach

In toxicological risk assessment the threshold for toxicological concern (TTC) for oral risk assessment has evolved as a useful concept. This concept acknowledges that a human exposure threshold can be determined below which there is no appreciable risk to human health even when the toxicological profile of the substance is not known. Safford (2008) proposed a method to apply this concept in the QRA of dermal allergens, using the IFRA/RIFM method to derive input parameters (see chapter 3). Since the QRA of dermal allergens is still under discussion, it goes beyond the scope of this report to describe the TTC concept according to Safford (2008) in more detail. Developments in the future should be followed, as this concept may be useful when there is very low skin exposure to chemicals with insufficient data on sensitisation hazard and/or potency.

3 Quantitative risk assessment of dermal allergens

3.1 Introduction

As already mentioned in the previous chapter, characterisation of the risks of being exposed to dermal allergens was until recently restricted to an indication of the hazard. Since then, several attempts have been made to develop a method for a QRA for skin sensitisers. One such recent method is the IFRA/RIFM QRA method proposed by Api et al. (2008). This method solely deals with consumer (dermal) exposure to fragrance ingredients. In brief, the QRA is as schematically described in the figure below (Figure 3.1), indicating which parts are required for the QRA. In the present report, the IFRA/RIFM method will be discussed stepwise, by giving attention in more detail to the way the method deals with exposure assessment, toxicology and ultimately, the risk assessment. Suggestions for improvement of the QRA for dermal allergens will be made, where considered appropriate.

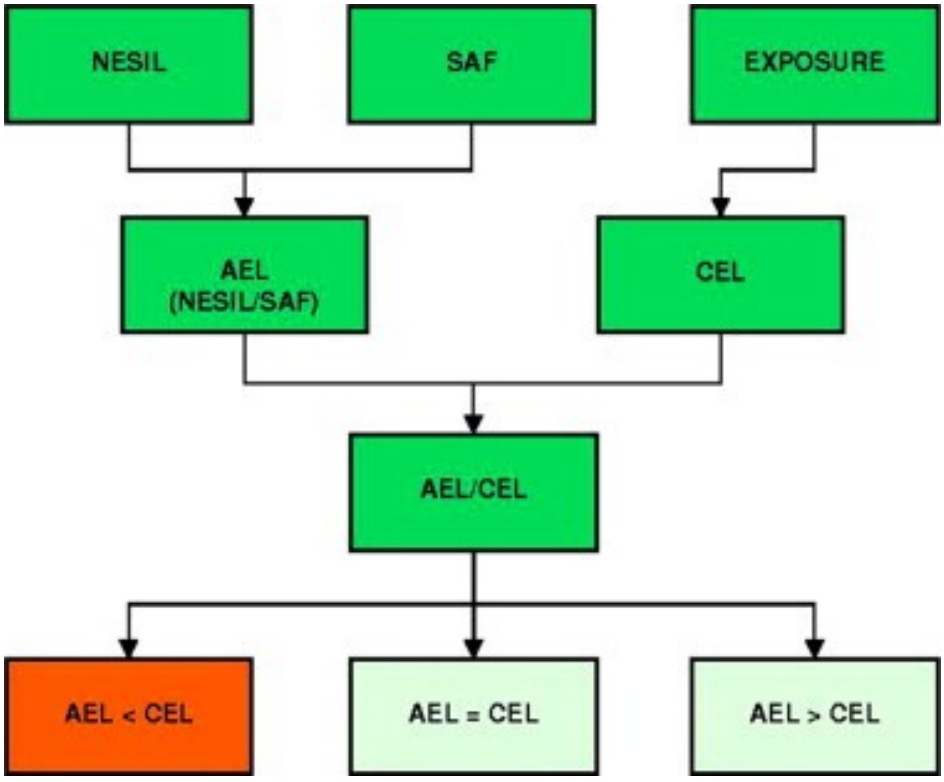

Figure 3.1. Key steps of the IFRA/RIFM QRA process.
NESIL (no expected sensitisation induction level) is the predicted dose threshold for the induction of skin sensitisation in humans. SAF (sensitisation assessment factors) represent uncertainties associated with inter-individual variability, matrix differences and exposure considerations. AEL (acceptable exposure levels) are calculated by dividing the NESIL by the product of the three SAFs. The AEL is then compared to the CEL (consumer exposure level), which results in acceptable risk if the AEL is equal to or larger than the CEL. If the CEL exceeds the AEL, re-evaluation of the risk management is required (according to Loveless et al., 2009).

3.2 Exposure assessment

3.2.1 Exposure assessment in the QRA for dermal allergens

The exposure assessment is a crucial component in QRA for dermal allergens. The exposure estimate must match the exposure that is required for the specific effect, i.e., skin sensitisation. Based on experiences in the field of skin sensitisation, it is widely accepted that the relevant dose metric for QRA for dermal allergens is the dermal load per cm^2 skin (mg/cm^2 or mg/cm^2/day (see section 2.3.1)). To derive such exposure estimates, specific information on the use of the product and the concentration of the allergen of interest is required.

In general, an exposure assessment can be performed in many ways. Measurements of the dermal load after application or unintended exposure are preferred, but this type of data is very scarce. In consumer exposure assessments, it is often opted for modelling the exposure, because actual measurements are frequently lacking or not fit for the specific purpose. Where possible, measurement data are used as input for modelling. Consumer surveys or questionnaires can also be used to derive vital input for exposure assessments.

3.2.2 Exposure assessment according to IFRA/RIFM

In the proposed IFRA/RIFM QRA for dermal allergens not much attention has been paid to exposure assessment as a part of the QRA. The exposure assessment in Api et al. (2008) and Api and Vey (2008) starts with determination of the Consumer Exposure Level (CEL) for a certain product or product category containing a fragrance material. The exposure estimates (CELs) posed as defaults are mainly based on unpublished data gathered by industry. As a rule, the most conservative values (i.e., the highest amount used per day measured and smallest skin site area reported) were used to derive the CELs. Exceptions are made when certain advantages in a study would give preference to a certain source, for instance when measurements approach reality more accurately. Api et al. (2008) derived CELs for a number of products and underlying data were shown, whereas Api and Vey (2008) described the CELs derived per product category (see Table 3.1 below).

The product category CEL is determined in the following way: first product categories are determined in which product types with similar exposure profiles are placed together, such as hand cream or deodorants/antiperspirants.
Secondly, for each product within that product category, the product specific CEL (g used per cm^2 per day) is determined based on the literature (such as the fact sheet developed by RIVM on consumer products) and measurement data. Api and colleagues acknowledge that the mechanism of action of skin sensitisation and the exposure to elicit such an effect is complex. There are many factors which may play an important role in the exposure assessment (see 2.3.1) as indicated by both Api et al. (2008) and others. However, the IFRA/RIFM method accounts for these factors, such as use considerations, specific skin sites, repeated exposures to a product on one day, by applying SAFs (sensitisation assessment factors) to the hazard to cover the uncertainties in the exposure. Thus, rather than adjusting the exposure estimate, an assessment factor that is considered sufficiently protective is applied to the NESIL to derive the Acceptable Exposure Level (AEL).

Finally, the product(s) that ‘scores’ the highest CEL and the highest SAF may serve as a surrogate for exposure for the entire product category. Hence, a high CEL and high SAF for a product or product category would provide a worst case exposure surrogate for exposure within this product category.

Table 3.1 Product type consumer exposure levels that drive the IFRA/RIFM QRA category (adopted from IFRA/RIFM QRA information booklet (IFRA and RIFM, 2008))

IFRA QRA category	Category consumer exposure[1] ($mg/cm^2/day$)	Product type that drives the category consumer exposure level
Category 1	11.7	Lip products
Category 2	9.1	Deodorants/antiperspirants
Category 3	2.2	Hydroalcoholics for shaved skin
Category 4	2.2	Hydroalcoholics for unshaved skin
Category 5	4.2	Hand cream
Category 6	1.4	Mouthwash
Category 7	4.4	Intimate wipes
Category 8	1.0	Hair styling aids
Category 9	0.2	Rinse-off hair conditioners
Category 10	0.1	Hard surface cleaners
Category 11	0.00033	Candles

The category consumer exposure level ($mg/cm^2/day$) is driven by the product type in that category with the highest consumer exposure level and the highest Sensitisation Assessment Factor (SAF). Categories 3 and 4 have the same category consumer exposure but are different categories because of a different SAF for the use of products on shaved skin.

Considerations

- It is not always transparent how CELs have been derived for certain products, because some CELs were based using unpublished industry data. The quality of the underlying data therefore cannot be evaluated.
- Although the aim of the IFRA/RIFM approach is to provide a conservative risk assessment, there are no indications whether the approach is indeed worst-case and if so, to what magnitude. For example, the category CEL might be more realistic for one product than for another product in that category. At this moment, there might be too little information available for the individual products to justify such a category approach.
- Api et al. (2008) stated that there are uncertainties concerning the exposure profile and subsequent exposure assessment. Issues such as the location and duration of the exposure, frequency of exposure and aggregate exposure have been addressed, but rather than taking them into account in the exposure assessment, these issues are accounted for in the effect assessment (incorporated in SAFs).

Based on these considerations, it is suggested to set up the exposure assessment as realistically as possible. The preferred approach is to consider the exposure data that are specific for the product of interest. The exposure assessment should include information on the location and duration of exposure, as well as the frequency of the exposure of the product of interest. Any uncertainty related to the exposure assessment should preferably, where possible, be accounted for in the exposure assessment (such as exposure to multiple sources).

Alternatively, when product specific data are lacking, the product category approach proposed by IFRA/RIFM might be a pragmatic one. The above-mentioned considerations should be taken into account, especially the consideration that uncertainties concerning the exposure profile should be considered in the exposure assessment. Care must be taken to ascertain that the category consumer exposure estimate sufficiently protects the consumer.

3.3 Effect assessment

The other important aspect of risk assessment of dermal sensitisers is effect assessment, consisting in hazard identification and hazard characterisation. In Chapter 2, an outline is given of the development in test methods over the last 10 years, resulting in more quantitative results that make risk assessment of dermal sensitisers much more adequate.

3.3.1 Derivation of point of departure according to IFRA/ RIFM

As already mentioned in Chapter 2, different values can be used as PoD in risk assessment for skin sensitisation (induction and elicitation). According to the IFRA/RIFM QRA method described by Api et al. (2008), the PoD for risk assessment of fragrance ingredients is the No-Expected Sensitising Induction Level (NESIL). The NESIL, introduced by Api and colleagues, is a benchmark dose that is derived using a weight of evidence (WoE) approach, including animal and human data, and is expressed as a dose per unit area (e.g., µg/ cm^2). The different human and animal test methods underlying the data have already been described in detail in Chapter 2. It appears that Api and colleagues have assigned the LLNA as the preferred animal test method. In addition, IFRA/ RIFM are recommending the use of the RIFM standard HRIPT protocol, as described by Politano and Api (2008) for the generation of confirmatory human data for use in the QRA. Api et al. recommend using a WoE approach for determining the NESIL for fragrance ingredients because the historical data that are used to determine the sensitisation potential of a material may be of variable quality and robustness. To this end, WoE guidelines have been developed by IFRA/RIFM. These guidelines are developed specifically for fragrance ingredients and are intended to be applied only to fragrance ingredients. They may also address some unusual situations for which discrepancies between data generated in the various tests need to be resolved.

In the following text box, the eight different guidelines are briefly summarised.

Guideline #1
The quantity of chemical per unit area of the skin (e.g., µg/cm^2) is considered as the most appropriate dose metric for skin sensitisation.
Guideline #2
A NOEL from a well-performed HRIPT (> 100 subjects, published methodology) will be given precedence over NOELs form other repeated exposure clinical tests conducted in humans. It is important to evaluate the robustness of the studies.
Guideline #3
Where LOEL data exist from other human tests lower than the NOEL derived from a HRIPT, this will be considered (unless there are reasons to disregard the LOEL data).
Guideline #4
In the absence of a NOEL from a HRIPT, a NOEL from a different predictive human test (HMT) can be used to set the NESIL, provided that it is supported by an EC3 value from a well-conducted LLNA.
Guideline #5
Adjuvant tests in animals and non-adjuvant tests in guinea pigs shall not be used as primary sources for derivation of NESILs in this context. They may be used to contribute information to determine potency classification and be incorporated in a WoE approach.
Guideline #6
When only LLNA data are available, then a confirmatory HRIPT should be considered. A cautious approach will be used for selection of the dose level of fragrance ingredient in the conduct of any such HRIPTs. Exceptionally, the weighted average EC3 value can be used to define the NESIL.

Guideline #7
A NOEL from a well-run HRIPT will (even if higher) has precedence over all other NOELs. When there is a significant discrepancy between a HRIPT NOEL and a LLNA EC3 value, further consideration in setting the NESIL will be required. A LLNA EC3 level that exceeds a NOEL determined by a HRIPT will not be used to define the NESIL. If the HRIPT NOEL is the lowest NOEL available, it shall take precedence in deriving the NESIL. Additional sources of data should be considered.
Guideline #8
Data from diagnostic patch test studies cannot be used directly in a weight of evidence approach for the determination of NESILs for the induction of contact allergy to fragrance ingredients. These studies can be useful to help determine the need for additional data. The absence of relevant positive reactions following testing in dermatology clinics may provide support to current exposures to the fragrance ingredient.

Considerations

- The HRIPT data takes precedence over other data including other predictive human tests in the guidance rather than including all data in the WoE approach. However, experience with HRIPT is scarce outside industry and the test is not part of any official guidelines and very little method description exists.
- The HRIPT protocol is recommended by IFRA/RIFM to generate confirmatory human data in case only animal data (LLNA) are available. Confirmatory tests where subjects may become sensitised are considered unethical.
- The WoE guidelines for derivation of a NESIL as suggested by Api et al. have no general validity and scientific support is lacking (SCCP, 2008).

The WoE approach can be used to derive the NESIL, where preferably both human data, e.g., HRIPT or other predictive human test and animal data, e.g., the LLNA, are considered. However, if human data is not available it is not recommended to perform a HRIPT for confirmation purposes. Instead, it is suggested to use animal data from the LLNA (if available) as the PoD for deriving the NESIL. In Scheepmaker (ed. 2006) a fact sheet was dedicated to the use of LLNA data in risk assessment; more specifically on how to derive a NOAEL based on LLNA data both when only one study is available or more (see Figure 3.2). In case of the latter however, the fact sheet does not give guidance on how to determine which LLNA studies should be included or discarded.

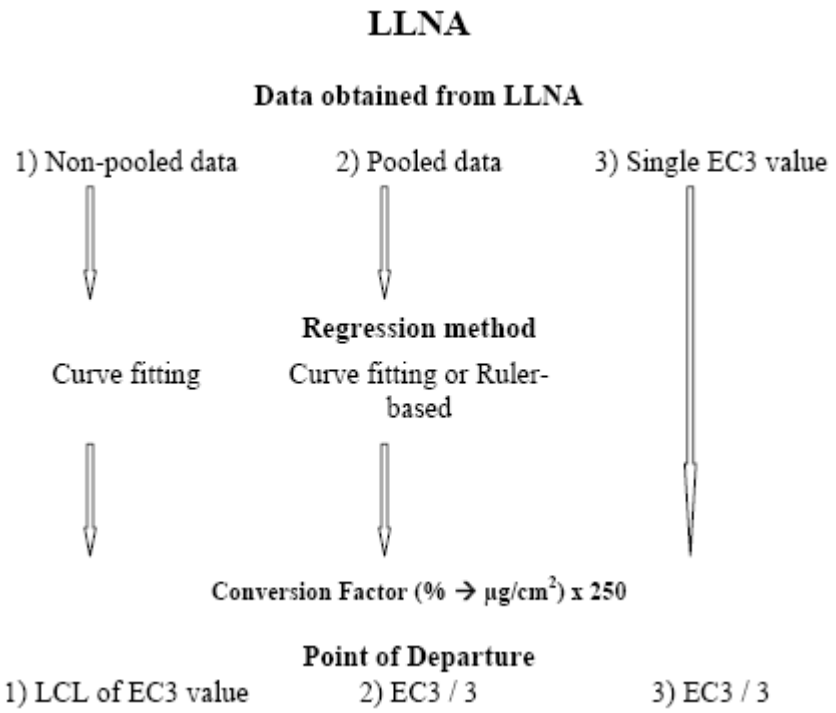

Figure 3.2 Deriving the NESIL from LLNA data (figure adopted partly from Scheepmaker ed., 2006).

3.4 Risk assessment method according to IFRA/RIFM

The total picture of the risk assessment according to the IFRA/RIFM method is shown in Figure 3.1. In short, the NESIL (see section 3.3) is divided by a total uncertainty factor (based on the combined SAFs) in order to derive the acceptable exposure level (AEL), which is then compared to the actual CEL (see 3.2). In case the AEL is equal to or greater than the CEL, the product is considered safe. The IFRA/RIFM method uses the ratio between the AEL and CEL to determine a safe concentration for a certain fragrance material in the product of interest (see text box below).

IFRA/RIFM risk assessment for fragrance ingredients in a product:

WoE NESIL/ SAF = AEL

CEL product (category)

AEL/ CEL = factor X

X/ 100 = x%

When concentration of fragrance in product = < x%, then the product is SAFE

3.4.1 Sensitisation Assessment Factors

In the IFRA/RIFM method, Sensitisation Assessment Factors (SAFs) are assessment factors that are applied to the NESIL in determining the AEL. According to Api et al., the relevant AFs for sensitisation are inter-individual (or intraspecies) variations, matrix effects and user considerations. The overall SAF is determined by multiplying the separate SAFs, which in theory may range from 10 to 1000 (see Figure 3.3). However, Api et al. (2008) state that it is unlikely that the overall SAF would exceed 300 for fragrance products.

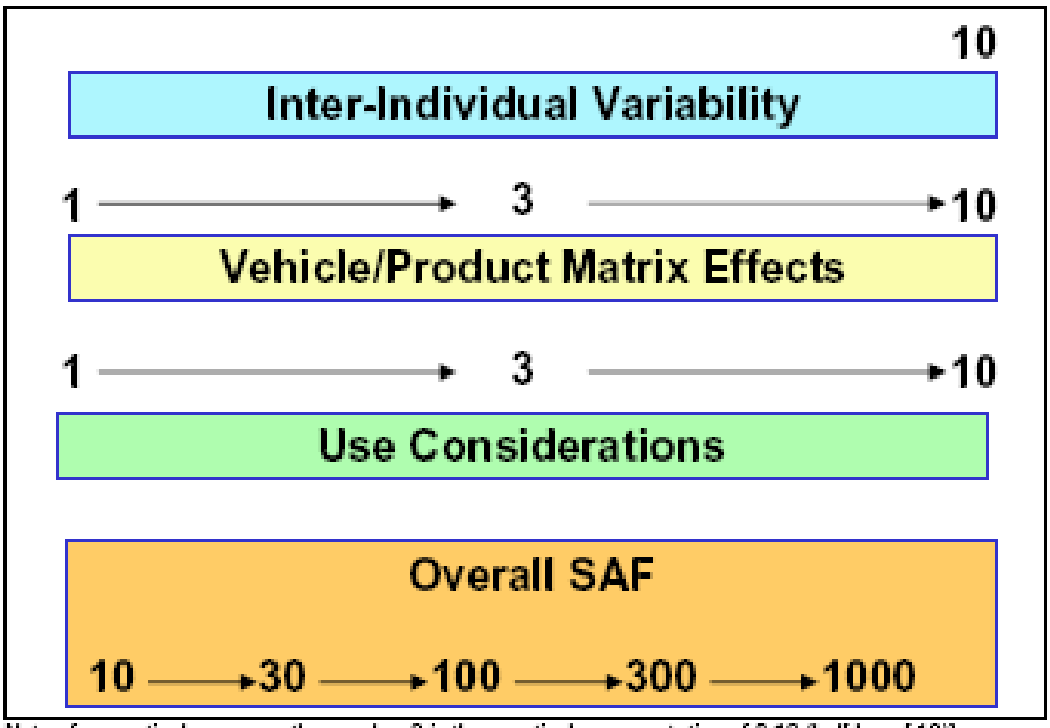

Figure 3.3: Generation of the overall SAF by multiplying the SAFs for the sub-factors (adopted from Api et al. (2008)).

3.4.1.1 Interspecies variation

Historically, for non-cancer risk assessments, a default tenfold assessment factor (AF) has been used to extrapolate from laboratory animal species to humans. The approach for contact sensitisers when using mouse data is somewhat different. Is has been shown that the mouse EC3 data in general closely correlate with the NOEL from human sensitisation tests designed to confirm lack of induction. Furthermore, the NESIL is based on a weight of evidence approach including human data. Therefore, it has been decided that for the QRA of skin sensitising fragrance ingredients, no separate AF for interspecies variation was necessary (Api et al., 2008).

Consideration

The above-mentioned argument is only valid in the case of available and suitable human data. When only mouse data are available, then an interspecies AF factor should be considered. In the fact sheet on LLNA strategy (in Scheepmaker ed., 2006), it was argued for interspecies differences to use a factor of three instead of the classical factor of ten. With regard to skin penetration, rodents tend to display a considerably higher skin penetration for most chemicals than humans (Scheepmaker ed., 2006). The

use of murine data is therefore considered conservative, since three- to tenfold higher penetration has been reported. Furthermore, the biological process that takes place for the immune system to respond to sensitisers is considered a practically similar process across mammalian species. In the case only animal data is available for deriving the NESIL, in Scheepmaker (ed. 2006) it is suggested to follow the scheme as drawn up in Figure 3.2. Essentially, this scheme takes into account an interspecies factor of three.

The REACH guidance for the endpoint sensitisation (ECHA, 2008) mentions that EC3 data generally correlate well with human skin sensitisation thresholds derived from historical predictive testing; however there are cases where this correlation is poor and the two values may differ by tenfold or more. In view of this variation, the REACH guidance recommended default AF of ten for interspecies variation should be used, unless there is evidence (e.g., from a close analogue of the substance in question) of good correlation between the EC3 and human NOAEL/LOAEL, in which case the interspecies AF could be lowered. In cases where there is good agreement between the LLNA EC3 value and the NOAEL/LOAEL derived from good quality historical human predictive tests, the REACH guidance recommends to use the lowest threshold value without applying an interspecies AF.

In the case only animal data is available for deriving the NESIL, it is suggested to account for the interspecies variation as described in Scheepmaker (ed., 2006, see Figure 3.2).

3.4.1.2 Inter-individual variations

This assessment factor accounts for possible intraspecies differences and variability. It covers genetic differences, sensitive subpopulations, inherent barrier function, age, gender and ethnicity, which might lead to subjects being more susceptible to becoming sensitised. Basically, all factors can influence the strength of the healthy skin barrier function due to genetic effects or illnesses, or is inherently different due to age, gender or ethnicity. This factor is always set at ten in the IFRA/RIFM method.

Consideration
The use of a factor of ten is in accordance with general principles of toxicology (SCCP, 2008). The REACH guidance on the endpoint sensitisation (ECHA, 2008) refers to the general AF for intraspecies variation, thus a factor of ten. Felter et al. (2002) concluded in their paper on a scientific basis for sensitisation uncertainty factors that for the intraspecies AF a tenfold default would be adequate to protect the majority of the population. Age, sex and race do not appear to be major influences, while disease states, genetic differences and associated differences in metabolic activity and general integrity of the skin are more likely to have a significant effect on inter-individual sensitivity. However, as previously noted by Wijnhoven et al. (2008); more information is needed on how to deal with susceptible individuals, as it is anticipated that the factor of ten may not always cover inter-individual variations. Until more information becomes available, a factor of ten is considered appropriate.

3.4.1.3 Matrix effects

The SAF for matrix effects covers the differences between the matrix/vehicle used during animal testing and that of the product in which the chemical is used. In some cases, products may contain irritants or penetration enhancers, which may serve as promoters of skin sensitisation or help penetrate the allergen through the stratum corneum. The larger the difference between experimental and real-life scenarios, the greater the matrix SAF will be. The proposed range in the IFRA/RIFM method for this SAF is one to ten. In general, the products have been assigned a factor of three, where only one product was assigned a factor of ten (depilatory) by Api et al. (2008).

Consideration
When the actual human exposure involves exposure to sensitisers in a different or more complex matrix compared to the test situation, which might increase the potential for induction of sensitisation (e.g., matrix with irritant or/and penetration enhancing properties), we agree that the application of an additional AF of one- to tenfold should be considered. If human exposure is expected in a matrix with no penetration enhancers or irritants, an AF of three might be sufficient, or if the matrix is very similar to the matrix used to determine the NOAEL/EC3 and is not expected to increase the potential for induction of sensitisation, the AF may be reduced to one. This is in accordance with the guidance under REACH (ECHA, 2008). In practice, this will lead to the same AF values that Api proposes. However, we consider it appropriate to review on a case-by-case basis (not on a product category basis) whether the test conditions are comparable to the product conditions.

3.4.1.4 Use considerations

The exposure to a skin sensitiser in the experimental situation is defined and controlled, in contrast to the real-life situation. Therefore, a SAF to cover the differences in exposure between the experimental and real-life situation was introduced, i.e., a SAF for user considerations. According to Api et al. (2008), exposures in the latter situation are generally less exaggerated, more variable and controlled by the consumer. The types of user considerations are the site of contact and related skin integrity, occlusion and time and frequency related to the use of the product.

The site of contact and the related skin integrity are directly related to the use of the product. Whereas in human patch tests the back or the arm is generally exposed, consumer products can be applied all over the body. The differences in skin integrity can be substantial; the plant foot arch being 12-fold less permeable to ^{14}C-labelled hydrocortisone, while the forehead was 3-fold more permeable relative to the back area (Marzulli and Maibach, 1987). The skin integrity can also be affected by the product itself or from another product used at the same location. Adults may suffer from dermatitis, while infants may have diaper rash. Sun-burned skin and shaved skin are also known to be relative more permeable than healthy skin. The use of products at the affected locations may result in increased dermal absorption of the skin sensitiser.

Considerations
Differences in exposure profiles between experimental and real-life settings are covered by applying an AF for use consideration in the IFRA/RIFM method. Some differences, such as the duration and frequency of exposure, also covered by this AF, should in our opinion be included in the exposure assessment. The importance of the duration and frequency of the exposure with respect to the risk of becoming sensitised is discussed in section 2.3. It should be noted that the experimental settings are aimed to reveal sensitising properties of a chemical and exposure is relatively intense. On the other hand, studies have shown that a single exposure may already be sufficient to sensitise a subject. More research is needed before a general or perhaps a category-fit statement can be made on the relevance of duration and frequency of the exposure in skin sensitisation. In the fact sheet on LLNA strategy (Scheepmaker ed, 2006), a correction factor for exposure duration of ten is mentioned. This was suggested because of the accumulating effect of formaldehyde found by De Jong et al. (2007). Animals exposed to levels below the EC3 value were still sensitised after repeated exposure. This supports the need to take repeated exposure into account in the risk assessment.

Other differences between experimental and real-life settings, such as occlusion of the skin or reduced skin integrity caused by product use, should indeed be covered by this AF. In the experimental situation, occasionally, the skin is occluded by a bandage to ascertain dermal exposure, for instance in

the case of volatile compounds. In the real-life situation, skin can be occluded by clothes or products that form some kind of film.

Therefore, on a case-by-case basis an AF (1 – 10) should be considered to account for specific exposure condition considerations (that are not considered in the exposure assessment). It is important to consider that repeated exposure may lead to the induction of skin sensitisation at exposures lower than the experimentally derived induction threshold.

3.4.2 Determining the risk

After the (Weight of Evidence) NESIL and the SAF have been determined for the chemical substance and the product in which the substance is contained, the AEL can be derived. The NESIL is divided by the overall SAFs to obtain the AEL. The next step in the IFRA/RIFM method is to compare the AEL with the CEL for that product or product category. The risk is considered acceptable when the AEL is greater than or equal to the CEL. If the ratio is below one, the risk is considered unacceptable.

Consideration

The general principle that has been applied, i.e., the ratio between the AEL and CEL, to determine the risk is approved by both SCCS and RIVM; however, as mentioned before, risk assessment should not be aimed at determining safe levels of fragrance materials in products. Instead, risk assessment should be aimed at identifying possible risks of a specific substance and its use in products.

4 Example

4.1 Introduction to citral

To stipulate the need for performing an aggregate quantitative risk assessment of a dermal allergen and to gain more insight into possible pitfalls when performing such risk assessments, an example was carried out with citral as sensitiser. Citral (CAS 5392-40-5) is a mixture of two acyclic monoterpenoids, neral and geranial, which can be regarded as branched chain aliphatic unsaturated aldehydes (cis- and trans-3,7-dimethyl-2,6-octadien-1-al). Citral is therefore also known as geranial (or citral a) and neral (or citral b). It is common in lemongrass, lemon and citrus fruits. Next to limonene, citral is one of the characteristic odours related to the citrus family.

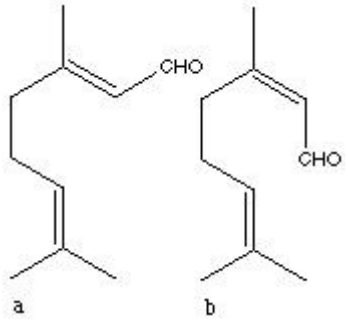

Figure 4.1. Structure of citral a (geranial) and b (neral) (from www.food-info.net).

Citral occurs widely in varying component isomer ratios in many natural products including citrus oils such as lemon oil and orange oils, in lemongrass oils, Litsea cubeba oil, black pepper oil, verbena oil, melissa oil, ginger oil, etc. Most people are exposed to citral in their daily lives when citrus fruits are peeled and cut by hand. Citral is also regularly found in the diet as a natural or synthetic flavouring component of some spices and in fruit-based or fruit-flavoured soft drinks (Lalko and Api, 2008).

4.1.1 Presence in consumer products

Citral is found in several consumer products, as shown in the inventory in the report of Wijnhoven et al. (2008), which is the source of the information given here. Citral was found by the Danish EPA in 26.1% of the cosmetic products containing a fragrance substance (in 23 out of 88 products), with a concentration ranging from 38.8 to 553.9 ppm (equals mg/l). In children's cosmetics, it was present in 8.2% (17 out of 208) products in a concentration range of 4.0 to 73 ppm. In the Netherlands, the Food and Consumer Product Safety Authority (VWA) also determined citral in children's cosmetics reporting its presence in 8.7% (2 out of 23) products in a concentration ranging from 109 to 168 ppm. Citral is also used in cleaning products and detergents. The Danish EPA investigated 43 cleaning products, of which 7 products (16.2%) contained citral in concentrations up to 0.0501 m/m%. Similarly, the VWA measured fragrances in 52 cleaning products and found citral in 1 product in a concentration of 8 ppm. In air fresheners, the Danish EPA determined citral in 36.8% (7 out of 19) of the products tested in a concentration range of 200 to 26,000 ppm. BEUC measured emissions from 74 different air fresheners. Citral could be detected in 2 (2.6%) air fresheners with a concentration of

2.0 – 48 µg/m^3. Furthermore, citral was found in a toy (27 ppm, Danish EPA). This overview shows that citral is present in a number of consumer products in the low to very high concentration range.

4.1.2 Setting the example

Citral was selected as case study because the fragrance material is commonly used not only in cosmetic products but also in detergents (Wijnhoven et al., 2008). This makes an aggregate exposure assessment relevant. Furthermore, when tested in the LLNA, citral is found to be in the weak to moderate potency range of skin sensitisers, with an EC3 value of 5.6 (Api et al., 2008) or 5.7 (Loveless et al., 2009). In 2008, IFRA/RIFM performed a QRA for citral according to their method.

In this example the RIFM/ IFRA method of performing a QRA for skin sensitisers is compared to an adapted version in order to include an aggregate exposure assessment to several products. Therefore, dermal exposure to citral resulting from the use of three selected consumer products, i.e., shower gel, deodorant spray and cleaning product, is taken into account in the QRA.
Inhalation or ingestion of citral is not considered in this example, although these routes might have an influence on dermal sensitisation.

4.2 Health effects

In the past, citral has been tested extensively in guinea pigs, mice and humans for its skin sensitising abilities. In all species, citral has tested positive for skin sensitisation. Citral was found to be sensitising in the guinea pig at 1% in vaseline (the highest non-irritating concentration) (OECD SIDS, 2001). A survey of sensitisation data from tests on materials containing citral was conducted under the auspices of the Soap and Detergent Association (SDA) (OECD SIDS, 2001). This survey was restricted to skin patch tests on human subjects conducted in the USA by member companies of SDA and by perfume suppliers. None of the personal care or household products containing citral induced hypersensitivity attributed to citral in 10,660 patch tests and there were no confirmed reactions to citral in 2,098 patch tests on fragrance blends containing the substance. A total of 22 induced sensitised cases occurred in 174 tests conducted at 1 to 5% pure citral in ethanol, but no inductions occurred at the 0.5% citral concentration level in 82 test subjects.

4.2.1 Deriving the NESIL for citral

Lalko and Api (2008) reviewed literature on skin sensitisation tests with citral to derive a threshold for induction of sensitisation. Test results were reported for the guinea pig (GPMT, Buehler Assay, Draize, Maguire, FCAT, OET, CET and SIAT tests), mouse (LLNA) and human (HRIPT, HMT). Where the results in the guinea pig and human are predominantly used to identify a skin sensitiser, the LLNA test can provide input for both identification and dose-response relationship of a skin sensitiser. To derive a threshold for skin sensitisation, Lalko and Api used the LLNA data, which are supported by human data. The EC3 value obtained in the LLNA is considered the threshold at which the animal has become sensitised. The weighted mean of 11 different EC3 values, ranging from 300 to 3250 µg/cm^2, was calculated to be 1414 µg/cm^2. In order to derive the weighted mean, weights (that were not shown by authors) were assigned to the type of vehicle used in the LLNA tests. Human data were also used to provide a weight of evidence for the threshold level. Based on findings in the human HRIPT, a NOEL of 1400 µg/cm^2 and a LOEL of 3876 µg/cm^2 was obtained. Taking the animal and human data together, the NOEL of 1400 µg/cm^2 from the HRIPT was finally set as the threshold for induction for skin sensitisation, also referred to as the NESIL.
Currently, no other NESILs have been derived for citral. It is noted, however, that the derivation of the NESIL by Lalko and Api is not completely transparent as to how the weighted mean of the EC3 values

was derived. But, more importantly, it is unclear why a weighted mean was derived at all. The reported range of obtained EC3 values by Lalko and Api (2008) is wide and sometimes unexplainable.
Two EC3 values obtained using the same vehicle and reported by the same reference differed by more than a factor of 4 (i.e., 375 and 1700 µg/cm^2). The question raised is whether or not one should use a weighted mean or simply take the lowest reported EC3 value, derived from a well-performed study. Logically, the choice for taking a different EC3 value would alter the entire risk assessment. Basketter et al. (2008) reported for citral an EC3 value of 3300 µg/cm^2 and a human NOEL of 775 µg/cm^2 . Using the same method as Lalko and Api (2008), a much lower NESIL would have been derived.
At this moment, it is noted that more guidance is needed on how to derive a NESIL when only animal data are available, or different data are available that all can be used. This, however, goes beyond the scope of this example. For comparison purposes and simplicity, the derived threshold for induction of 1400 µg/cm^2 citral will be used as the NESIL in this example.

4.3 Quantitative risk assessment for citral

4.3.1 QRA using the IFRA/RIFM method

Citral has previously been used as a case substance in an example calculation of the QRA according to the IFRA/RIFM method. This method uses categorical exposure estimates to determine the 'surrogate' exposure for a product within that category. In the paper of Api and Vey (2008), where the example calculation was made, the derivation of a categorical consumer exposure estimate was performed for several product categories. The origin of the underlying data for such estimates has been described in more detail in Api et al., 2008 (see also section 3.2.2).

The exposure assessment using the IFRA/RIFM method is relatively simple. The estimate of exposure for a specific product is based on the *categorical exposure estimate* to which that product belongs. The products deodorant, shower gel and cleaning agent belong to categories 2, 9 and 10, respectively (see Api and Vey, 2008), with corresponding product exposure estimates of 9.1, 0.2, and 0.1 mg/cm^2/day (see Table 3.1). Note that the categorical exposure estimate for (all) deodorants was based on data of a deodorant roller. Other data however show that deodorant sprays may give higher exposures (Api and Vey, 2008). In the Cosmetics Fact Sheet (Bremmer et al., 2006) developed by RIVM, it has been shown that the amount of deodorant product on the skin after spraying is much higher than after using a roller, i.e., 3.4 g versus 0.5 g. In comparison, the IFRA/RIFM method takes 0.9 g into account, assuming a surface area of 100 cm^2 (9.1 mg/cm^2/day x 100, Api et al., 2008).

In the exposure assessment according to the IFRA/RIFM method in this example, product exposures were taken from the paper of Lalko and Api (2008). In this paper, product exposure values are established for several product categories, which implicitly include exposure factors such as dilution or retention factors.

Subsequently, *weight fractions* of citral for deodorant and cleaning agent were obtained from Wijnhoven et al. (2008), where the upper limits taken were 0.06 and 0.05%, respectively. The weight fraction of citral in shower gel is based on the upper limit of perfume content, as described in the fact sheet Cosmetics (Bremmer et al. 2006), assuming that the perfume is solely based on citral. Ideally, product specific information on the weight fraction is used but at the moment, information is lacking for shower gel.

Combining the modelled exposure estimate and the established NESIL in the QRA according to IFRA/RIFM, it shows that both shower gel and cleaning agent can be considered safe. Exposure to deodorant, however, results in an unacceptable risk (see Table 4.1). Following the approach by IFRA/RIFM, a safe concentration can be derived assuming that the CEL equals the AEL. The concentration of citral in deodorant that would still be considered safe is 0.052% (see Table 4.1, third column). Hence, if the concentration of citral in deodorants was restricted to a limit of 0.052%, then all three products (separately) would be considered safe in the QRA. The aggregate exposure to citral from these products is not considered in the IFRA/RIFM method.

Table 4.1. Determination of the risk in the separate products by the IFRA/RIFM method

citral	**Deodorant spray**	**Deodorant spray**	**Shower gel**	**Hard surface cleaner**
WoE NESIL (µg/cm^2)	1400	1400	1400	1400
SAF	300	300	100	100
AEL (mg/cm^2/day)	0.0047	0.0047	0.014	0.014
Category product exposure (mg/cm^2/day)	9.1	9.1	0.2	0.1
Concentration of citral in the product (%)	0.06	0.052 (calculated)[a]	0.5	0.05
CEL (mg/cm^2/day)	0.0055	0.0047	0.001	0.00005
Risk assessment (AEL/CEL)	0.86 (**risk**)	1 (**safe**)	14 (**safe**)	280 (**safe**)

[a] The concentration of citral in deodorant is a calculated safe concentration, i.e., when CEL is equal to AEL.

4.3.2 QRA with aggregate exposure assessment

One of the main messages of this report is to apply an aggregate exposure assessment in a QRA for dermal sensitisers. Therefore, several possible consumer product sources containing citral were selected, of which the aggregate exposure from the combination of use within a certain period of time is considered to be relevant for becoming sensitised. The combination of the use of a shower gel, deodorant and cleaning agent was chosen. For reasons of simplicity, it was decided neither to perform population estimations of the exposure nor to apply probabilistic techniques to scale the variability and uncertainty in the assessment. It is acknowledged that these aspects could be considered in a full-blown QRA, but is beyond the scope of this report.

The following exposure scenario was chosen: an individual is cleaning the kitchen, takes a shower afterwards and subsequently uses a deodorant spray on one day. It is anticipated that these exposures would lead to an aggregated exposure. The exposure estimates for the individual products are therefore added, which is in contrast to what is done in the IFRA/ RIFM method, where the safety of using the products is considered separately (see 4.3.1). Exposure estimates resulting from the category approach following the IFRA/RIFM method might result in an overestimation when used in an aggregate assessment. The exposure level within each category is a worst-case estimate based on one sentinel product, which might be different than the product of interest. Aggregated exposure based on sentinel products may lead to an irrelevant and highly unrealistic exposure estimate, because the substance

exposure from the different products is not drained to the same lymph node and the exposure estimate may be too worst case, since the product exposure estimates were already worst-case estimates.

4.3.2.1 Exposure parameters

The exposure parameters have been obtained from the ConsExpo fact sheets on cosmetics (i.e., deodorant spray and shower gel) and cleaning products (i.e., cleaning agent) (developed by RIVM, Bremmer et al. 2006; Prud'homme de Lodder et al. 2006). These exposure parameters were based on a realistic worst-case exposure scenario set for the specific product. This information is used to determine the amount of product contacting the skin and the surface area of skin contact. Weight fractions considered to be safe according to the IFRA/RIFM method were taken forward: 0.052, 0.5 and 0.05%, respectively for deodorant, shower gel and cleaning agent. The fact sheets include corrections for dilution or retention of the product, which are taken forward when relevant in this example.

Table 4.2 Exposure estimates based on input from fact sheets (suggested method)

	Suggested method this report
Deodorant	
Weight fraction citral product (%)	0.052
Amount on skin (g)	3.4
Surface area of contacted skin (cm^2)	100
Exposure estimate (mg/cm^2)	*0.0177*
Shower gel	
Weight fraction citral product (%)	0.3-**0.5** (based on classic soap or soap gel)/3 due to dilution
Amount on skin (g)	26.1
Surface area of contacted skin (cm^2)	17500
Exposure estimate (mg/cm^2)	*0.0025*
Cleaning agent	
Weight fraction citral product (%)	0.05/80 due to dilution
Amount on skin (g)	19
Surface area of contacted skin (cm^2)	1900
Exposure estimate (mg/cm^2)	*0.000063*
Aggregate exposure estimate (mg/cm^2)	**0.020**

4.3.2.2 Risk characterisation

The risk is determined for the use of several citral containing products combined on one day. The consumer exposure level derived in Table 4.2 (i.e., 0.020 mg/cm^2) is compared to the AEL for citral in the selected products. The same PoD, i.e., the NESIL of 1400 $\mu g/cm^2$, will be used to derive the AEL. However, since aggregate exposure of the products is considered, the SAF might be built up differently than it would be following the IFRA/RIFM method. The overall SAF is built up as follows: no assessment factor for interspecies is included in the QRA because the NESIL is based on human data. The assessment factor for intraspecies variability is set to ten as default. The assessment factor for matrix effects is set to three, as the products are unlikely to contain skin penetration enhancers or irritants. Furthermore, the shower gel and cleaning agent are diluted with water. The assessment factor

for use is chosen to be ten in this case. This is based on the use SAF of ten for deodorant. The rationale for ten is that the exposure area under the arm can be shaven and is easily irritated. As the contribution of the exposure of the deodorant covers 88.5% of the total, even for the aggregate risk assessment, the overall use SAF is chosen to be ten. Thus, a total SAF of 10 x 3 x 10 = 300 was derived. This results in an AEL of 0.0047 mg/cm^2 (1.4 mg/cm^2 / 300). The ratio between the AEL and CEL is then 0.24, which is considered to be an unacceptable risk (see Table 4.3). Obviously, the exposure to citral resulting from the deodorant use is the main contributor to the risk. Based on deodorant alone, the AEL:CEL ratio would have been 0.27 and thus also a risk. The reason for this difference with the IFRA/RIFM method, where deodorant was considered safe with the same concentration of citral, is the much higher calculated exposure using input from the Cosmetics fact sheet (0.0177 vs. 0.0047 mg/cm^2).

Table 4.3. Determination of the risk with aggregate exposure to citral from the selected products deodorant, shower gel and cleaning agent

citral	**Aggregate exposure (using exposure estimates based on input from fact sheets)**	**Aggregate exposure (using exposure estimates resulting from IFRA/RIFM method)**
WoE NESIL (µg/cm^2)	1400	1400
SAF	300	300
AEL (mg/cm^2/day)	0.0047	0.0047
Aggregated CEL (mg/cm^2/day)	0.020	0.00575
Risk assessment (AEL/CEL)	0.24 (**risk**)	0.82 (**risk**)

4.4 Concluding remarks

The example shows that exposure to citral from multiple products contributes to the total exposure and thus to the risk as well. Where all products separately could be considered safe in the IFRA/RIFM QRA method, it is not considered safe when the aggregate exposure estimate is considered. This has been shown using the aggregate exposure estimates based on input using defaults from the ConsExpo fact sheets, as well as using the aggregate exposure estimates from the IFRA/RIFM method. The much higher exposure estimate for deodorant in the suggested method is mainly responsible for the risk as determined here. In this example, the contribution from the (only) two other products in comparison to the exposure from deodorant is rather low but should not be disregarded. If other products, known to contain citral and commonly used, are included, the total exposure could increase even further.

The SCCP (2008b) published an opinion on a QRA performed by IFRA/RIFM in 2008 for citral, farnesol and phenylacetaldehyde. For citral, the ‘safe’ limits based on the dermal sensitisation QRA would be, e.g., 0.05% citral in deodorants and 0.6 % in hydroalcoholic products for unshaved skin. In liquid soaps 7% citral can be used according to the dermal sensitisation QRA model, 8.2% can be used in shampoos and 100% citral in baby diapers and hand dishwashing. These ‘safe’ limits calculated from the dermal sensitisation IFRA/RIFM QRA methodology for these wash-off products were changed into maximum pragmatic concentrations of 5% for liquid soaps and shampoos and 2.5% for baby diapers and hand dishwashing.

SCCP commented on several points of the methodology:

- The data provided show that the application of the dermal sensitizstion QRA approach would allow increased exposures to allergens, already known to cause allergic contact dermatitis in consumers. This is because the separate products, according to the QRA, may be allowed to contain a higher concentration before it is considered unsafe. There is no confidence that the levels of skin sensitisers identified by the dermal sensitisation QRA are safe for the consumer.
- It is of concern that the model operates with multiple product categories without considering risk from aggregated exposures.
- Occupational exposures are not considered although they have been identified as an important area of development of the dermal sensitisation QRA.
- QRA model is based primarily on data from experimental sensitisation tests in humans, e.g., HRIPT. There is a lack of in-depth method description and the experience with this test, its validity, sensitivity and reliability is sparse outside the industry. Such experimental sensitisation tests in humans are considered unethical.
- Epidemiological and experimental data, providing information on sensitisation/elicitation reactions in consumers by fragrance ingredients in marketed products, are not integrated in the dermal sensitisation QRA model.
- It is of concern that scientific consensus has not been achieved concerning the choice of safety factors.
- Identification of safe levels of exposure to existing substances known to cause allergic contact dermatitis in the consumer should be based on clinical data and/or elicitation low effect levels. Currently these are the only methods, which have proven efficient in reducing/preventing existing problems of sensitisation/allergic contact dermatitis in the consumer.

The first three comments by the SCCP would be covered if an aggregate exposure assessment was considered. Taking the results in the example and the comments by the SCCP into consideration, we strongly recommend including aggregate exposure assessments in the QRA for dermal sensitisers.

5 Discussion

In the present document, recent developments have been reported as well as suggestions for improvement of the current approaches with regard to the risk assessment of dermal sensitisers. First the IFRA/RIFM method (Api et al., 2008) for a quantitative risk assessment (QRA) for dermal allergens has been described and commented. One of the shortcomings of the current approaches was the lack of consideration of aggregate exposure. Aggregate exposure assessment has become more and more important in the field of exposure assessment and therefore, should be incorporated in the QRA on dermal sensitisation as the main suggestion for improvement. To be able to put the suggestions for improvement into perspective, an example QRA of a fragrance (citral) in cosmetics and detergents has been performed. Although the IFRA/RIFM QRA method was developed for risk assessment of fragrance ingredients in cosmetic products, the method is considered to be applicable in other fields as well. In that case, the aggregated exposure becomes even more relevant in the QRA.

At this moment, the WHO-IPCS is preparing a Guidance Document on Immunotoxicity for Chemical Risk Assessment, including immunosuppression, immunostimulation, sensitisation and autoimmunity associated with chemical exposure. It is expected that the guidance concerning dermal sensitisation will be based to a large extent on the proposed method by IFRA/RIFM but will cover a broader field than the use of fragrance ingredients in cosmetic products. The release of the guidance document is due in early 2011 and was therefore not taken into account in this report.

5.1 Considerations on the IFRA/RIFM method

As already mentioned in Chapter 3, we have some thoughts on the existing methodologies for quantitative risk assessment of dermal sensitisers. These considerations are summarised below.

5.1.1 Exposure assessment

With respect to exposure in the IFRA/RIFM method, the so-called Consumer Exposure Level (CEL) is used to determine the exposure to a chemical from a certain product. For each product category set up by IFRA/RIFM, a CEL has been derived. The CEL, however, is not necessarily based on the product of interest, but can be based on a sentinel product that is supposed to represent a similar exposure profile with the same or higher exposure level. As previously noted in Chapter 3, the basis for such categorisation might be (too) weak, as there is a lack of exposure data. Furthermore, the derivation of the product CELs is not always transparent, as was also previously noted by the SCCP (2008). Although in the current report we did not look in detail into the choices made in the so-called product amount, the citral example already shows that there might be some discussion possible on certain choices for every product category. For instance, we noticed that there are large differences in the product amount of deodorants between rollers and sprays. In the product category approach, the amount of the roller has been used, being much lower than the spray.

Unless it is not possible due to lack of data, preference is given to a product-specific approach rather than using a category approach. In this case, the exposure assessment would be more tailor-made and more transparent, considering the choices for exposure parameters made. Furthermore, the exposure estimate is directly related to the product of interest and can therefore better be judged on its merits by a risk assessor, for instance, whether or not the estimation is a best-case estimate or a realistic worst-case estimate. Obviously, reliable information on the use pattern of consumer products is vital in such an approach. Recently, detailed information on the use of cosmetics has become available (Loretz et

al., 2005; 2006; 2007) and RIVM has developed fact sheets (see www.consexpo.nl) incorporating product use information and defaults for exposure assessment. The product-specific approach also enables an aggregate exposure assessment, which in several frameworks is required, e.g., according to the Biocide Directive (98/8/EC), aggregation of exposure should be taken into account. Since dermal sensitisation depends on exposure to a chemical and not the use of a single product, multiple sources should be considered within a QRA for dermal sensitisation. Discussion on aggregate exposure within the QRA for dermal sensitisers is described in section 5.2.

5.1.2 Derivation of point of departure (NESIL)

In the IFRA/RIFM method, a weight of evidence approach is stated for determination of the NESIL. However, no specific guidance for this is given, making expert knowledge of the possible points of departure and weighing factors necessary. In principle, the IFRA/RIFM method favours the HRIPT test and the LLNA test to be used as primary sources in the QRA, as these studies provide quantitative data. Other studies may only serve as supporting evidence or to determine potency classification. However, it is not transparent which criteria are used for the in- or exclusion of tests for the WoE approach. The IFRA/RIFM method does give some criteria for the acceptance of a human patch test and a specific protocol is proposed (Politano and Api, 2008), but it is not clear if these criteria are generally accepted in the field of sensitisation experts. In addition, guidance is needed to judge skin sensitisation tests on their quality and to possibly correct for that in the WoE approach.

This need for guidance was illustrated in the citral example. Citral, a relatively well-known fragrance, has been tested numerous times in the murine LLNA and in human HRIPT and HMT studies. In the citral example in this report, the NESIL derived by Lalko and Api (2008) was used, which was based on a weighted mean of EC3 values and a NOEL in the HRIPT. It is noticed that the calculation of the weighted mean is based on characteristics of the different vehicles (Lalko and Api, 2008), but it is not specified what vehicle is given more or less weight in the calculation. Next to that, the citral case indicated that it is difficult to judge the quality of PoDs reported in several human HRIPT or HMT studies found in the literature.

When human data are not available, it is recommended by IFRA that a confirmatory HRIPT test (according to the protocol of Politano and Api) is carried out. This approach has ethical drawbacks, since healthy people can become sensitised by such a confirmatory test. If human data is lacking, it is suggested in this report to use the mouse data from the LLNA, where the EC3 value is considered the threshold for sensitisation. Furthermore, this value can be considered as a point of departure (PoD) to derive the NESIL. An interspecies factor to extrapolate from animal to human is then required (see 5.1.3). As shown in the example, EC3 values derived for citral displayed a large range. In case more data is available, a weight of evidence approach can be applied to obtain the proper PoD, but if there is only one LLNA study, the EC3 value is subject to uncertainty. Then, an additional factor of three, on top of the interspecies factor, might be applied if only one animal test is available (fact sheet LLNA in Scheepmaker ed., 2006). Whether or not these uncertainty factors sufficiently cover the uncertainty should, however, be investigated further.

Overall, more specific guidance for the risk assessor to be used in practice is definitely needed and might be added to the upcoming WHO/IPCS guidance. In our opinion, there is a need for more clear guidance for general risk assessors in the derivation of a threshold (the NESIL). Criteria which help the WoE approach could be very useful.

5.1.3 Sensitisation assessment factors

There is overall agreement on the use of ten for the *intraspecies AF* for inter-individual differences. The IFRA/RIFM method, the fact sheet for LLNA strategy, and the REACH guidance, all propose the factor of ten. Only the SCCP (2008) commented that differences up to 100 were found in sensitisation studies. Their conclusion was based on the Kligman (1966) study, where 5/24 volunteers were sensitised by exposure to 0.1% PPD and 24/24 volunteers at exposure to 10% PPD. However, this observation does not prove a 100-fold difference in susceptibility amongst individuals, since it is unknown which concentrations between 0.1% and 10% PPD would already lead to 100% response in the subjects.

IFRA/RIFM state that no *interspecies AF* is needed, since the NESIL is based on human data. However, this only holds in the case of available and suitable human data. In case of animal data, a factor of three might be sufficient, since in general the EC3 values seem to correlate well with the human data and species sensitivity is not expected to vary significantly. This is also recommended in the fact sheet on LLNA strategy, because a three- to tenfold higher penetration of chemicals through the murine skin have been reported and similar processes are considered across mammalian species. However, from a table in Wijnhoven et al. (2008) based on Basketter et al. (2008), for some substances a factor of 4-12 could be found, in which humans were more sensitive to the substance. Although the factor of 4-12 is on species differences, it also includes other uncertainties, such as the reliability of the EC3 value derived from the LLNA test. If there is evidence that humans are more sensitive than the mouse with respect to sensitisation, because of a higher dermal absorption for instance, then a higher interspecies factor than three can be considered.

With respect to skin sensitisation specific AFs, an *AF for matrix effects* is used in the IFRA/RIFM method. The REACH guidance also mentions an AF for vehicle or matrix effects, which could be between one and ten, depending on the situation.

As mentioned previously, an attempt should be made to include uncertainties of exposure in the exposure assessment and not to cover them by using SAFs, as is the case for possible aggregate exposure or maybe other use considerations. However, some differences between the exposure situation in the toxicity test and the actual human exposure situation might be expressed in an *AF for use considerations,* because these differences are too difficult to take into account in the exposure assessment. This is also defined in the REACH guidance (ECHA, 2008). Parameters that might be taken into account can be occlusion, skin integrity, location on the body of the exposure, exposure frequency or duration differences and so on.

5.2 Aggregated exposure in QRA for dermal sensitisers

Currently, aggregate exposure is not taken into account in the IFRA/RIFM method, although it is acknowledged that exposure from multiple sources may play an important role in the QRA (Api et al., 2008). Aggregate exposure has been taken into account in other frameworks, where the chronic exposures from several products were ‘simply’ summed up. The toxicological endpoints to which the chronic exposures were compared were generally systemic effects seen after chronic exposure in laboratory species. Dermal skin sensitisation, however, proved to be a much more complex toxicological mechanism.

When performing an aggregate exposure assessment, the starting point of setting the boundaries for the exposure assessment should be the process of skin sensitisation. Kligman (1966) concluded that

frequent exposure at the same location increases the risk of sensitisation. In general, the relevant exposure site of the body is determined by the closest present draining lymph nodes. Aggregation is only allowed when the same substance is applied at sites close to each other and drained by the same lymph nodes (so not feet and head). The second important point is the time frame of exposure. What is the maximum time span relevant for dermal sensitisation to allow repeated exposures to be added and if so, how? Data in the literature do not really answer this question, but do suggest that answers might be chemical specific or might be dependent on the dermal absorption rate. In the case study of citral presented here, products were chosen that could be used in a small period of time of one to two hours, and applied on one extremity (e.g., hands or armpits). However, the exact influence of the location and duration of the exposures with respect to aggregated exposure are still unknown and need to be further investigated. Obviously, if these factors are demonstrated to be vital, then more in-depth exposure scenarios with reliable data are required.

The dose metric used for dermal sensitisation in risk assessment is the unit of $\mu g/cm^2$. This seems to be the most logical choice, as it is the outcome of the LLNA or other sensitisation tests. It is also no problem to calculate the exposure estimate in this unit using consumer exposure models. However, in the IFRA/RIFM method, the exposure is calculated as $\mu g/cm^2$ per day. When performing an aggregate risk assessment for citral in this report, the problem arose that apart from the location of the exposure to the different products, it might also be important when or within what time frame the exposures occur. Hence, it might not be relevant whether or not there is daily exposure. Especially, when aggregate exposure is considered estimated over a short period of the day, the extension 'per day' would not make any sense. In addition, the use of the dose metric of mg/cm^2/day by Api and colleagues and others is not completely correct. This dose metric implies that the exposure estimate resembles the time weighted average over 24 hrs, but it appears that they assumed that the dermal load was constant over time (that is, no fluctuations occur during the day). For skin sensitisation this only makes sense in case the same product is used continuously at the same location. It is therefore preferred to use mg/cm^2 to indicate exposure, where the extension 'per event' can be added if properly described.

It is acknowledged that using product categories is pragmatic and for instance, might aid registrants of chemicals under REACH with their exposure estimate and QRA for dermal sensitisers. Within REACH, aggregate exposure should also be taken into account if the aggregate exposure may lead to a higher risk. We mentioned earlier in the citral case study (Chapter 4) that the present categorisation by IFRA/RIFM does not allow performing an aggregate exposure assessment because aggregating the exposure of sentinel products might be irrelevant and too worst case. Instead, an alternative categorisation can be developed where the categories are predominantly based on exposure profile and certain combinations of categories are preset, based on aggregation possibilities. A prerequisite should be that exposure to the different products leads to drainage to the same lymph node. Obviously, any suggested product category approach should be widely accepted before it can be used in the QRA.

5.3 General remarks for the risk assessment

First of all, it is noted that the suggested improvements for risk assessment of dermal sensitisers does not include protection for allergic responses in previously sensitised subjects. To date, the knowledge on dermal sensitisation does not allow deriving reliably 'safe' levels for elicitation. Nevertheless, by protecting subjects from becoming sensitised indirectly, the subjects are also protected from allergic responses in the future. In the IFRA/RIFM method, the QRA is (only) performed for the endpoint of dermal sensitisation, not taking into account elicitation. It should be discussed by policy makers what

they want to prevent, only the prevention of new subjects becoming sensitised or also the prevention of elicitation in already sensitised subjects.

In many legal frameworks, a risk assessment for a substance in a product is performed in a preventive way, using the actual concentrations present in the product. In the case of the IFRA/RIFM method, the result of the QRA is a safe level per product. Thus, this is done beforehand, resulting in a maximum allowable level of a specific substance causing dermal sensitisation in a specified product.
The worked example in this report showed that ignoring aggregate exposure might lead to a risk when separate QRAs for products might end up safe. Allowing safe levels to be determined that way may result in a possible total exposure of consumers causing dermal sensitisation. A possible solution for this problem might be that for the product categories defined by Api et al. (2008), some kind of 'filling up' of the total NESIL is established. So, for example (just for the sake of the argument), deodorants might use up 60% of the NESIL, while soap and shampoo only 10%.

5.4 Conclusions and recommendations

Clarify the mechanism of action
Exposure from multiple sources cannot be ignored in the QRA for dermal sensitisation. Exact mechanism and relevant factors leading to skin sensitisation, especially with regard to aggregate exposure, are not yet completely understood. Fundamental research is needed to further clarify the mechanism of action of skin sensitisation and the influence of certain exposure parameters on that process, such as location and duration of exposure and time between exposures. This might answer the question if it is relevant to add exposure estimate for, for instance, deodorant to the exposure estimate of a hand cream.

Set up guidance for aggregate exposure assessment in relation to skin sensitisation
Criteria for applying to an aggregate exposure assessment for the endpoint sensitisation should be formulated. Relevance of performing an aggregate assessment should be discussed first, focusing on the product types, taking into account the time (duration and repetition) and location of use. When performing an aggregate risk assessment, a critical look should be given to the product types in which the specific chemical is or might be used. Further, there should be a high probability that products will be used within a certain small period of time (for instance, four hours), while being relevant to add exposures together. It is recommended to describe criteria for aggregation of exposure for product categories (or better, all product types) to be able to assess when aggregation is relevant.

Set up guidance for derivation of the NESIL
There is a need for more clear and more specific guidance for general risk assessors in the derivation of a threshold for skin sensitisation (the NESIL). The guidance should at least include: criteria for in- or exclusion of a study and criteria used in the weight of evidence derivation of a NESIL. The issues concerning the sensitisation assessment factors can also be addressed in the guidance.

Find consensus in several frameworks
Obviously, scientific consensus and wide acceptation is needed to broadly implement a QRA method for skin sensitisation. A broader discussion is needed on this QRA method, where different groups, institutions and countries should be involved. It is very important to create support for a quantitative risk assessment method for dermal sensitisation within legal frameworks and other groups. The forthcoming guidance from WHO/IPCS might be a starting point.

rivm

Acknowledgements

The valuable input of Janine Ezendam, Henk van Loveren, Jacqueline van Engelen and Marja Pronk is highly acknowledged.

References

Aeby P, Basketter D, Diembeck W, Eschrich D, Gerberick F, Itagaki H, Kimber I, Le Varlet B, Manou I, Paye M, Rousset F, Rowland J and Sakaguchi H (2006). Dendritic cells research projects. ALTEX. 23 Suppl: 226–228.

Anderson C, Hehr A, Robbins R, Hasan R, Athar M, Mukhtar H and Elmets CA (1995). Metabolic requirements for induction of contact hypersensitivity to immunotoxic polyaromatic hydrocarbons. J. Immunol. 155, 3530–3537.

Api AM, Basketter DA, Cadby PA, Cano M-F, Ellis G, Gerberick GF, Griem P, McNamee PM, Ryan CA and Safford B (2008). Dermal Sensitization Quantitative Risk Assessment (QRA) for Fragrance Ingredients. Regul. Toxicol. Pharmacol. 52, 3–23.

Api AM and Vey M (2008). Implementation of the dermal sensitization Quantitative Risk Assessment (QRA) for fragrance ingredients. Regul. Toxicol. Pharmacol. 52, 53-61.

Basketter DA, Blaikie L, Dearman RJ, Kimber I, Ryan CA, Gerberick GF, Harvey P, Evans P, White IR and Rycroft RJG (2000). Use of the local lymph node assay for the estimation of relative contact allergenic potency. Contact Dermatitis, 42(6), 344–348.

Basketter DA, Andersen KE Liden C, Van Loveren H, Boman A, Kimber I, Alanko K and Berggren E (2005a). Evaluation of the skin sensitizing potency of chemicals by using the existing methods and considerations of relevance for elicitation. Contact Dermatitis, 52(1), 39–43.

Basketter DA, Clapp C, Jefferies D, Safford B, Ryan CA, Gerberick F, Dearman RJ and Kimber I (2005b). Predictive identification of human skin sensitization thresholds. Contact Dermatitis 53, 260–267.

Basketter DA, Jefferies D, Safford BJ, Gilmour NJ, Jowsey IR, McFadden J, Chansinghakul W, Duangdeeden I and Kullavanijaya P (2006). The impact of exposure variables on the induction of skin sensitization. Contact Dermatitis 55, 178–185.

Basketter DA, Gerberick F and Kimber I (2008). Appendix A: LLNA/EC3 Validation – Submission ICCVAM draft background review document use of the murine local lymph node assay (LLNA) to determine skin sensitization potency categories.

Bremmer HJ, Prud'homme de Lodder LCH and Van Engelen JGM (2006). Cosmetics Fact Sheet. RIVM report 320104001.

Coquette A, Berna N, Vandenbosch A, Rosdy M, De Wever B and Poumay Y (2003). Analysis of IL-1and IL-8 expression and release in in vitro reconstructed human epidermis for the prediction of in vivo skin irritation and/or sensitization. Toxicol. In Vitro 17, 311–321.

Cumberbatch M, Scott RC, Basketter DA, Scholes EW, Hilton J, Dearman RJ and Kimber I (1993). Influence of sodium lauryl sulphate on 2,4-dinitrochlorobenzene-induced lymph node activation. Toxicology 77, 181–191.

De Jong WH, De Klerk A, Beek MT, Veenman C and Van Loveren H (2007). Effect of Prolonged Repeated Exposure to Formaldehyde Donors with Doses Below the EC3 Value on Draining Lymph Node Responses. Journal of Immunotoxicology 4, 239–246.

Delmaar JE and van Engelen JGM (2006). Aggregating human exposure to chemicals. An overview of tools and methodologies. RIVM report 630700001.

ECETOC (European Centre for Ecotoxicology and Toxicology of Chemicals) (2003). Contact Sensitization: Classification According to Potency. Technical Report No. 87.

ECHA (European Chemicals Agency) (2008). Guidance on information requirements and chemical safety assessment Chapter R.8: Characterization of dose [concentration]-response for human health. May 2008.

European Allergy White Paper (1997). Allergic diseases as a public health problem in Europe. The UCB Institute of Allergy.

Ezendam J, te Biesebeek JD and Wijnhoven SWP (2009). Exposure to fragrance allergens in scented consumer products. RIVM Letter report 340301002.

Felter SP, Robinson MK, Basketter DA and Gerberick GF (2002). A review of the scientific basis for uncertainty factors for use in quantitative risk assessment for the induction of allergic contact dermatitis. Contact Dermatitis, 47(5), 257–266.

Friedmann PS, Moss C, Shuster S and Simpson JM (1983). Quantitative relationships between sensitizing dose of DNCB and reactivity in normal subjects. Clin. Exp. Immunol. 53, 709–715.

Friedmann PS (1990). The immunology of allergic contact dermatitis: the DNCB story. Advances in Dermatology 5, 175-195; discussion 196.

Friedmann PS (2007). The relationships between exposure dose and response in induction and elicitation of contact hypersensitivity in humans. British Journal of Dermatology 157, 1093–1102.

Gerberick GF, Robinson MK, Felter SP, White IR and Basketter DA (2001a). Understanding fragrance allergy using an exposure based risk assessment approach. Contact Dermatitis 45, 333–340.

Gerberick GF, Robinson MK, Ryan CA, Dearman RJ, Kimber I, Basketter DA, Wright Z and Marks JG (2001b). Contact allergenic potency: Correlation of human and local lymph node assay data. Am. J. Contact Dermatitis 12, 156–161.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG and Lepoittevin JP (2007). Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol. Sci. 97, 417–27

Griem P, Goebel C and Scheffler H (2003). Proposal for a risk assessment methodology for skin sensitization potency data. Regul. Toxicol. Pharmacol. 38, 269–290.

Health Council of the Netherlands. Asthma, allergy and environmental factors. The Hague: Health Council of the Netherlands, 2007; publication no. 2007/15E.

Hextall JM, Alagaratnam NJ, Glendinning AK, Holloway DB, Blaikie L, Basketter DA and McFadden JP (2002). Dose-time relationships for elicitation of contact allergy to para-phenylenediamine. Contact Dermatitis 47, 96–99.

IFRA and RIFM (2008). Information Booklet Version 3.4 (revised July 2008) at http://www.rifm.org/pub/publications.asp (accessed November 2009).

Johansen JD, Andersen KE and Menne T (1996). Quantitative aspects of isoeugenol contact allergy assessed by use and patch tests. Contact Dermatitis 34, 414–418.

Jowsey IR, Basketter DA, Westmoreland C and Kimber I (2006). A future approach to measuring relative skin sensitising potency: a proposal. J. Appl. Toxicol. 26, 341–350.

Kimber I, Basketter DA, Berthold K, Butler M, Garrigue JL, Lea L, Newsome C, Roggeband R, Steiling W, Stropp G, Waterman S and Wiemann C (2001). Skin sensitization testing in potency and risk assessment. Toxicol. Sci. 59, 198–208.

Kimber I, Basketter DA, Butler M, Gamer A, Garrigue JL, Gerberick GF, Newsome C, Steiling W and Vohr HW (2003). Classification of contact allergens according to potency: proposals. Food Chem. Toxicol. 41, 1799–1809.

Kimber I, Dearman RJ, Basketter DA, Ryan CA, Gerberick GF, McNamee PM, Lalko J, Api AM (2008). Dose metrics in the acquisition of skin sensitization: Thresholds and importance of dose per unit area. Regul. Toxicol. Pharmacol. 52, 39–45.

Kligman AM (1966). The identification of contact allergens by human assay. II. Factors influencing the induction and measurement of allergic contact dermatitis. Journal of Investigative Dermatology, 47, 375–392.

Lalko J and Api AM (2008). Citral: Identifying a threshold for induction of dermal sensitization. Regul. Toxicol. Pharmacol. 52, 62–73.

Loretz LJ, Api AM, Barraj LM, Burdick J, Dressler WE, Gettings SD, Han Hsu H, Pan YH, Re TA, Renskers KJ, Rothenstein A, Scrafford CG, Sewall C. (2005). Exposure data for cosmetic products: lipstick, body lotion, and face cream. Food Chem. Toxicol. 43(2):279–91.

Loretz L, Api AM, Barraj L, Burdick J, Davis de A, Dressler W, Gilberti E, Jarrett G, Mann S, Laurie Pan YH, Re T, Renskers K, Scrafford C, Vater S. (2006). Exposure data for personal care products: hairspray, spray perfume, liquid foundation, shampoo, body wash, and solid antiperspirant. Food Chem. Toxicol. 44(12):2008–18.

Loretz LJ, Api AM, Babcock L, Barraj LM, Burdick J, Cater KC, Jarrett G, Mann S, Pan YH, Re TA, Renskers KJ, Scrafford CG. (2008). Exposure data for cosmetic products: facial cleanser, hair conditioner and eye shadow. Food Chem. Toxicol. 46(5):1516–24.

Loveless SE, Api AM, Crevel RWR, Debruyne E, Gamer A, Jowsey IR, Kern P, Kimber I, Lea L, Lloyd P, Mehmood Z, Steiling W, Veenstra G, Woolhiserm M, Hennes C (2009). Potency values from the local lymph node assay: Application to classification, labelling and risk assessment. Regul. Toxicol. Pharmacol., in press.

Marzulli FN and Maibach HI (1987). Dermatotoxicology. Third edn Washington: Hemisphere Publishing Corporation.

Natsch A, Gfeller H, Rothaupt M and 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 and Ellis G (2009). Filling the concept with data: integrating data from different in vitro and in silico assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. Toxicol. Sci. 107, 106–121.

OECD SIDS (2001). OECD SIDS Initial Assessment Report on Citral (CAS N°:5392-40-5). 13th SIAM (Switzerland, November 6–9, 2001).

Politano VT and Api AM (2008). The Research Institute for Fragrance Materials' human repeated insult patch test protocol. Regul. Toxicol. Pharmacol. 52, 35-38.

Prud'homme de Lodder LCH, Bremmer HJ, Van Engelen JGM (2006). Cleaning Products Fact Sheet. To assess the risks for the consumer. RIVM Report 320104003.

QRA Expert Group (2006). Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. Technical dossier.

Safford RJ (2008). The dermal sensitization threshold – A TTC approach for allergic contact dermatitis. Regul. Toxicol. Pharmacol. 51, 195–200.

Sakaguchi H, Ashikaga T, Miyazawa M, Yoshida Y, Ito Y, Yoneyama K, Hirota M, Itakagi H, Toyoda H and Suzuki H (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.

SCCP (Scientific Committee on Consumer Products) (2008). Opinion on Dermal Sensitization Quantitative Risk Assessment (Citral, Farnesol and Phenylacetaldehyde)
Scientific Committee on Consumer Products .
http://ec.europa.eu/health/ph_risk/committees/04_sccp/sccp_opinions_en.htm.

SCCP (Scientific Committee on Consumer Products) (2008b). Opinion on Dermal Sensitization Quantitative Risk Assessment (Citral, Farnesol and Phenylacetaldehyde). adopted at 16th plenary of 24 June 2008. http://ec.europa.eu/health/ph_risk/committees/04_sccp/sccp_opinions_en.htm.

Scheepmaker JWA (ed.) (2006). Fact sheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment. Part VI. RIVM Report 601516014.

Schneider K and Akkan Z (2004). Quantitative relationship between the local lymph node assay and human skin sensitization assays. Regul. Toxicol. Pharmacol. 39, 245–255.

Schuur AG, Delmaar JE and van Engelen JGM (2009). Geaggregeerde blootstelling. Gebruik in verschillende kaders en toekomstige mogelijkheden. RIVM briefrapport 320015002 (in Dutch).

Scott RC, Corrigan MA, Smith F and Mason H (1991). The influence of skin structure on permeability: an intersite and interspecies comparison with hydrophilic penetrants. J. Invest. Dermatol. 96(6):921-5.

Sosted H, Menne T and Johansen JD (2006). Patch test dose–response study of p-phenylenediamine: Thresholds and anatomical regional differences. Contact Dermatitis 54, 145–149.

UN (United Nations) (2008). Updating of the second revised edition of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) – Health hazards – Revision of Chapter 3.4 with respect to strong versus weak sensitizers. Document No. ST/SG/AC.10/C.4/2008/18, dated 15 August 2008.

Van Och, FMM, Slob, W, De Jong, WH, Vandebriel, RJ, Van Loveren, H. (2000) A quantitative method for assessing the sensitizing potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of uncertainty margins. Toxicol. 146, 49–59.

Van Och FMM, Vandebriel RJ, Prinsen MK, De Jong WH, Slob W and van Loveren H (2001). Comparison of dose–responses of contact allergens using the guinea pig maximization test and the local lymph node assay. Toxicology 167, 207–215.

Van Och FM, Vandebriel RJ, De Jong WH and van Loveren H (2003) Effect of prolonged exposure to low antigen concentration for sensitization. Toxicology 184, 23–30.

Weaver JE, Cardin CW and Maibach HI (1985). Dose–response and diagnostic patch testing with sensitized humans. Contact Dermatitis 12, 141–145.

White JML, Basketter DA, Pease CK, Sanders DA and McFadden JP (2007). Intermittent exposure to low-concentration paraphenylenediamine can be equivalent to single, higher-dose exposure. Contact Dermatitis 56, 262–265.

WHO (World Health Organization) (1999). Principles and Methods for Assessing Allergic Hypersensitization Associated with Exposure to Chemicals. WHO/ILO/UNEP International Programme on Chemical Safety (IPCS), Environmental Health Criteria 212.

Wijnhoven SWP, Ezendam J, Schuur AG, van Loveren H, and van Engelen JGM (2008). Allergens in consumer products. RIVM Report 320025001.

Wilkinson SC, Maas WJ, Nielsen JB, Greaves LC, van de Sandt JJ and Williams FM (2006). Interactions of skin thickness and physicochemical properties of test compounds in percutaneous penetration studies. Int. Arch. Occup. Environ. Health 79 (5), 1–9.

Wolterink G, Van de Ven B, Ter Burg W and Verkaik-Kloosterman J (2009). Aggregate exposure to chemicals. RIVM Report 320108002.

Yamano T, Shimizu M and Noda T (2001). Relative elicitation potencies of seven chemical allergens in the guinea pig maximization test. J. Health Sci. 47, 123–128.

Zachariae C, Lerbaek A, McNamee PM, Gray JE, Wooder M and Menne T (2006). An evaluation of dose/unit area and time as key factors influencing the elicitation capacity of methylchloroisothiazolinone/methylisothiazolinone (MCI/MI) in MCI/MI-allergic patients. Contact Dermatitis 55, 160–166.

rivm

List of abbreviations

ACD	allergic contact dermatitis
AEL	acceptable exposure level
BMD	bench mark dose
CEL	consumer exposure level
EA	exposure assessment
ES	exposure scenarios
GPMT	guinea pig maximisation test
HI	hazard identification
HMT	human maximisation test
HRIPT	human repeat insult patch test
LC	Langerhans cells
LOAEL	lowest observed adverse effect level
LLNA	local lymph node assay
MET	minimum elicitation threshold
NESIL	no-expected sensitising induction level
NOAEL	no observed adverse effect level
NOEL	no effect level
PoD	point of departure
QEA	quantitative exposure assessment
QRA	quantitative risk assessment
QSAR	quantitative structure-activity relationship
RA	(qualitative) risk assessment
RMM	risk management measure
ROAT	repeated open application test
SAF	sensitisation assessment factor
TTC	threshold of toxicological concern
WoE	weight of evidence

RIVM

National Institute
for Public Health
and the Environment

P.O. Box 1
3720 BA Bilthoven
The Netherlands
www.rivm.nl