Regulatory Toxicology and Pharmacology **31,** 166–181 (2000)
doi:10.1006/rtph.1999.1362, available online at http://www.idealibrary.com on IDEAL®

Criteria for Development of a Database for Safety Evaluation of Fragrance Ingredients

Richard A. Ford,*[,1] Bruce Domeyer,† Otho Easterday,‡[,2] Kurt Maier,§ and John Middleton¶

*Research Institute for Fragrance Materials, Inc., Two University Plaza, Suite 406, Hackensack, New Jersey 07601-6209; †The Procter & Gamble Co., Miami Valley Labs, 11810 East Miami River Road, Ross, Ohio 45061; ‡International Flavors & Fragrances, Inc., 1515 Highway 36, Union Beach, New Jersey 07735-3597; §Haarmann & Reimer GmbH, Muehlenfeldstrasse 1, D-37603 Holzminden, Germany; and ¶Quest International Ltd., Ashford, Kent TN24 OLT, England

Received December 23, 1998

Over 2000 different ingredients are used in the manufacture of fragrances. The majority of these ingredients have been used for many decades. Despite this long history of use, all of these ingredients need continued monitoring to ensure that each ingredient meets acceptable safety standards. As with other large databases of existing chemicals, fulfilling this need requires an organized approach to identify the most important potential hazards. One such approach, specifically considering the dermal route of exposure as the most relevant one for fragrance ingredients, has been developed. This approach provides a rational selection of materials for review and gives guidance for determining the test data that would normally be considered necessary for the elevation of safety under intended conditions of use. As a first step, the process takes into account the following criteria: quantity of use, consumer exposure, and chemical structure. These are then used for the orderly selection of materials for review with higher quantity, higher exposure, and the presence of defined structural alerts all contributing to a higher priority for review. These structural alerts along with certain exposure and volume limits are then used to develop guidelines for determining the quality and quantity of data considered necessary to support an adequate safety evaluation of the chosen materials, taking into account existing data on the substance itself as well as on closely related analogs. This approach can be considered an alternative to testing; therefore, it is designed to be conservative but not so much so as to require excessive effort when not justified.

INTRODUCTION

The Research Institute for Fragrance Materials, Inc. (RIFM), was founded in 1966 as an independent and distinct entity charged with three principal objectives concerning substances used as fragrance ingredients (some with a long history of use): (1) to assure that there are adequate data available to support the safety of these materials under their conditions of use; (2) to review and evaluate standards for testing fragrance ingredients; and (3) to communicate this information to the industry and the scientific community.

To help achieve these objectives, an Expert Panel (hereafter referred to as the Panel), composed of leading scientists external to the fragrance industry and widely recognized for their expertise in their respective disciplines relevant to chemical toxicology, was created to provide scientific advice. Such advice includes rendering decisions as to whether a fragrance ingredient can be considered safe under the conditions of use based on all available data. It is the responsibility of RIFM staff to assure that the available data are sufficient for an adequate safety evaluation by the Panel.

Historically, RIFM has approached this responsibility by assuring that at least a basic battery of tests, acute oral and dermal toxicity, dermal irritation and sensitization, and, where necessary, photoirritation and photosensitization, was available for each fragrance ingredient reviewed (IFRA, 1999). The need for additional data beyond this basic battery was determined on a case-by-case basis after careful review of all existing toxicity data as well as the chemical structure, volume of use, and estimate of exposure from use in consumer products.

Based on industry surveys, an inventory of approximately 1300 fragrance ingredients in use was developed and data for their safety evaluation were gathered by reviewing the published literature, by soliciting the members of RIFM for unpublished data, and, as appropriate, by undertaking a test program to

[1] To whom correspondence should be addressed.
[2] This article is dedicated to the memory of Otho Easterday, who passed away on November 11, 1997.

0273-2300/00 $35.00

fill the gaps. Summaries of these data were published in individual substance monographs, approximately 1000 of which have been published so far (e.g., Ford *et al.*, 1992).

Three recent developments have prompted a review of the RIFM procedures and the preparation of this document. First was the publication of the "Indicative Non-Exhaustive List" (INEL) of fragrance ingredients (European Commission, 1996). This publication almost doubled the number of fragrance ingredients reported by the industry to be in use and therefore the responsibility of RIFM. Second it has been over 25 years since the publication of the first collection of monographs (Opdyke, 1973). Clearly conditions of use, available data, and the basis for safety evaluation have evolved over the years prompting the need for updating of the monographs. Third there have been concerns about the selection process employed by RIFM regarding materials to be reviewed and the criteria used in developing the database for that review.

GENERAL APPROACH

The INEL lists some 2600 substances reported to be in use by the fragrance industry including all of the substances in the original RIFM database. Obviously, a comprehensive review of this number of substances is a massive undertaking and requires a systematic and scientific approach both in determining the order in which they should be reviewed and in development of the necessary database for such a review.

The general approach is based on the same principles utilized by others when faced with assessing the safety of a large number of substances that are currently in use. It is logical that substances with higher volumes of use and higher exposures should be addressed first and, in general, should have the most rigorous safety data to support their use. Additionally, the approach takes advantage of recognized structure–activity relationships in determining both the order of review and the toxicological data needed for that review.

Since both the selection methodology and the criteria for database development depend on volume of use, exposure, and structural considerations, these will be discussed prior to the actual procedures for using these criteria.

Volume of Use

RIFM has conducted a survey of the more than 2600 INEL substances for their annual worldwide volumes used in formulating fragrances. Usage volumes range from a few thousand metric tons per year to less than 1 kg/year, with some materials having no reported use during the year surveyed (RIFM, 1997). Larger volumes of use indicate wider potential for exposure and, therefore, higher priority for review. This quantity also determines, to a certain extent, the data required for an adequate safety review. One practical point is that approximately 60% of fragrance use is in soaps, fabric softeners, cleaners, and detergents (Somogyi *et al.*, 1995). These end products have lower levels of added fragrance and have less skin contact than cosmetics and fine fragrances thereby resulting in lower exposure. Thus, estimating exposure requires considerations beyond simple volume of use.

Exposure

There are two types of exposure relevant to the safety of fragrance ingredients, level of use (skin surface concentration) and total systemic exposure. These vary widely depending on the type of product in which a fragrance ingredient is used and whether that ingredient is used as the dominant note in the fragrance or for a modifying effect at a much lower concentration. Levels of use in consumer products, and the resulting exposures, can range over 5 orders of magnitude or more for any specific fragrance ingredient depending upon its desired contribution to the overall odor profile.

Where the skin is the target organ, the potential for adverse effects is directly related to skin concentration. These adverse effects include dermal irritation, allergic sensitization, photoirritation, and photoallergy. Alcohol–based products, such as perfumes, aftershaves, colognes, and eau de toilettes have the highest concentrations of fragrance ingredients (Boeck and Fergen, 1991). Thus, in order to estimate skin concentration, the International Fragrance Association (IFRA) currently surveys the major international fragrance companies for the top 10 concentrations of specific fragrance ingredients in fragrance mixtures that are used in these alcohol-based products. It is then assumed that the concentrations of these fragrance mixtures may reach a level up to 20% in the final product. This allows estimation of a maximum skin concentration, which is used in determining priorities and testing needs.

The concentration of the fragrance material in a product when combined with the quantity and frequency of use of the product and its wash-off characteristics, as described in more detail in Appendix A, gives an estimate of total skin exposure. The use patterns of a fragrance ingredient in cosmetic products on a product-by-product basis are provided by the cosmetic and fragrance industries (COLIPA, 1997). In estimating total skin exposure, it is assumed that an individual consumer will use various categories of products repeatedly and all of these products will contain the particular fragrance material at the 97.5 percentile level of use. This provides a conservative approach for estimating the applied dermal exposure from different products as developed in Table A-1.

The systemic exposure, then, depends on the percentage of the applied dose of fragrance material that is absorbed percutaneously versus the amount that is washed off or evaporates. For the initial evaluation, it is assumed that 100% of the applied dose is absorbed and is systemically available. However, if dermal absorption and/or evaporation data are available, these can be used to adjust that assumption.

In summary, estimates of the highest skin concentrations are used for setting priorities and for determining the need for testing where the skin is the target organ. Since a consumer rarely, if ever, will encounter the maximum use level in all categories of cosmetic products on a repeated basis, the 97.5 percentile use level, with the assumption of 100% absorption, is used to provide initial conservative estimates of systemic exposure.

Consideration of Chemical Structures

The scientific basis for evaluating potential toxic effects from physicochemical properties of molecules, from specific structural moieties within the molecule, and from data for closely related structures has improved considerably in recent years with advances in the understanding of intoxication and detoxication pathways and with the desire to find alternatives to the use of animals for toxicity testing. Such considerations are particularly valuable for initial evaluation and priority setting. For example, it has been convincingly demonstrated that consideration of broad classes of chemical structures can be very effective in placing substances into clearly separate classes of chronic no observed adverse effects levels (NOAELs) (Cramer *et al.*, 1978; Munro *et al.*, 1996) as well as providing an efficient process for safety evaluation of large groups of structurally related substances (e.g., Adams *et al.*, 1998). Similarly, consideration of molecular structure can serve as an alert for certain toxicological endpoints and can, therefore, be of aid in determining the adequacy of the toxicological database.

Just as volumes and levels of use vary over a wide range, the chemical structures of fragrance materials also vary considerably. Many fragrance ingredients have simple and innocuous structures, such as the fatty acid esters or simple aliphatic alcohols, which would raise few safety concerns even if used in relatively high volumes or at levels resulting in significant exposures. Other fragrance materials have chemical structures that raise much more concern, such as the nitroaromatic substances or the epoxides. For such materials, more extensive databases and higher priorities are expected even when exposures and volumes of use are relatively low.

Components of certain molecular structures that have been associated with specific toxic effects in studies of a wide variety of synthetic and natural chemicals are presented in Appendix B. These structural components, hereinafter referred as "structural alerts," are based on published structure–activity relationships and were developed separately for (a) topical effects, (b) acute and/or systemic effects, and (c) carcinogenic and/or mutagenic effects. Although these structural alerts evolved from a large number of studies it should be obvious that the presence of a structural alert does not necessarily result in the same potential for toxicity. However, where one or more structural alerts exist in a molecule, it increases the priority and attention is drawn to certain endpoints that need to be specially addressed.

For topical effects the structural alerts are those derived by the "Deductive Estimation of Risk from Existing Knowledge" (DEREK) expert system from a large number of irritation and sensitization results (Barratt *et al.*, 1994a,b; Basketter *et al.*, 1994; Barratt and Basketter, 1994). For acute and/or systemic effects, the decision tree approach (Cramer *et al.*, 1978) is used to obtain structural alerts. The foundation for the carcinogenic and/or mutagenic structural alerts is based on the compilations of J. Ashby, R. W. Tennant, and D. Paton (Ashby, 1985, 1994; Tennant and Ashby, 1991; Ashby and Tennant, 1991; Ashby and Paton, 1993).

As new knowledge about structure–activity relationships becomes available, Table B-1 will require modification. In a few cases, as explained in Appendix B, additional structural alerts have already been added based on the knowledge and experience of the authors.

It should be noted that RIFM has developed the numerical scores for the various types of potential toxic effects listed in Table B-1 solely for assisting in determining the order of review as explained below.

Chemical structure can also be used to assist in making the database development and the safety evaluation process more efficient and more comprehensive. Grouping of chemicals into closely related chemical classes provides insight as to the type and quantity of toxicological data needed for evaluation of individual members of the group. For example, the known potential for photosensitization of the nitroaromatic material musk ambrette (Cronin, 1984) prompted more rigorous examination of this endpoint for other nitroaromatic materials. This concept extends beyond fragrance ingredients. For example, the knowledge that several nitroaromatic substances not used in fragrances are genotoxic led to careful examination of this endpoint with structurally similar fragrance materials (Api *et al.*, 1995, 1996).

Judicious consideration of structural relationships can also be used as an alternative for animal testing. For example, simple fatty esters are well known to be readily hydrolyzed *in vivo* in their component acids and alcohols. Where adequate data on the potential for systemic effects of the components exist, it is generally

unnecessary to obtain such data for the parent ester as long as the exposure is comparable or lower. This approach must be used with caution. For example, with sterically hindered esters, it may be necessary to demonstrate ready hydrolysis.

Structural relationships can also be used to increase the confidence of a conclusion of lack of potential for adverse effects for structurally similar materials. For example, testing of linalool and nine linalyl esters for skin sensitization shows no significant potential for sensitization under conditions of use as fragrance materials. This strengthens the conclusion that linalyl esters do not have significant potential for sensitization. According to the "Guidelines for determining an adequate database for safety evaluation" described below, a linalyl ester would not have to be tested for sensitization if use levels in the final product are less than 0.1% and total worldwide quantity of use is less than 1 metric ton per year.

DETERMINATION OF ORDER OF REVIEW

RIFM has developed a method that combines volume of use, chemical structures, and levels of use in an approximately evenly weighted manner to determine the initial priority for review. Because all three of these values may not always be known, default scores are assigned which later may be lowered, as the data become available.

Volume and Level of Use

Priority scores have been developed for different broad categories of levels and volumes of use (see Appendix C). These scores range from zero to 16. A material used at less than 0.05% in a consumer product would receive a zero score for use level while a substance used at greater than 5% would receive a score of 16. Similarly, a substance used as a fragrance ingredient at less than 0.1 metric tons per year worldwide would receive a zero while those used at greater than 1000 metric tons would receive scores of 16. In both cases, intermediate values are given intermediate scores. The default scores of 8 are assigned because not only is it unusual for a fragrance ingredient to be used at greater than 1000 metric tons/year and/or at greater than 5% concentration in a consumer product, but also for any such material, both values would certainly be known making default values unnecessary.

Chemical Structure

The use of chemical structure involves functional group analyses to determine which substances might have the most likely potential for adverse effects. Three separate endpoints have been addressed: dermal toxicity, systemic toxicity, and carcinogenicity/mutagenicity because these endpoints have different levels of seriousness which should be considered when setting priorities. A substance with a structural alert for mutagenicity, for example, is considered before one with only a structural alert for irritation.

RIFM has developed a simple scoring system that assigns chemical structure only slightly less weight (maximum score 12) than that for volume or level of use. This score depends on the severity of the endpoint addressed. For topical effects, the score is 2 if there is a structural alert for this endpoint while the score is 6 for carcinogenic/mutagenic structural alerts. For acute/systemic structural alerts, the score may be 0, 2, or 4 depending on the structural class as defined by Cramer *et al.* (1978). This method is a decision tree approach that classifies materials into one of three classes, low, medium, or high presumed toxicity, based on a series of 33 questions about chemical structure, metabolism, physiological occurrence, etc.

To obtain the total structural alert priority score, the score for each endpoint is simply added. Thus, the highest score based on structure would be 12 (2 + 4 + 6) while the lowest score, no structural alerts, would be zero. If the structure is unknown, a default score of 12 is assigned. Appendix C contains more details and provides examples for applying the numerical scores to fragrance ingredients to determine review priorities.

It is fully recognized that this approach is empirical and may oversimplify the situation by ignoring, for example, that sensitization clearly has a systemic component, as does carcinogenicity. However, as a practical approach to priority setting, such simplification is useful.

Calculation of Order of Review

The total priority score is obtained by simply adding the three scores for volume, use level, and structure. A substance used at greater than 1000 metric tons/year (score = 16) *and* at a level of greater than 5% in consumer products (score = 16) *and* having a structure with maximum scores for all 3 structural endpoints (score = 12) would receive a maximum priority score of 44 (16 + 16 + 12). Conversely, a substance with no structural alerts *and* which is used in consumer products at no greater than 0.05% *and* the total worldwide volume of use is less than 100 kg/year would have a score of zero. For a material for which the volume and level of use is unknown (default scores = 8) but which contains no structural alerts the priority score would be 16, but this priority would likely change as those unknown figures become available. Again, it should be emphasized that, for example, a score of zero does not establish zero adverse effects, but rather gives some degree of confidence that an adverse effect is not likely to occur and that a lower priority for safety evaluation is justified. Conversely, a high score of 44 does not mean that the material presents a risk, only that its

use and structural characteristics are such that a high priority for review is called for.

Obviously, this approach for determining the order of review does not substitute for a safety evaluation; that requires careful consideration of all available data on a case-by-case basis. However, such an approach must be sensitive enough not to assign a low priority to a material that may have significant potential for adverse effects under conditions of use. The fact that the overall priority is set using not only chemical structure but also exposure and volumes of use minimizes this possibility. On the other hand, the method should not result in high priorities for materials of relatively low potential for adverse effects, thereby wasting resources and diverting attention from materials of more concern.

CRITERIA FOR DEVELOPMENT OF AN ADEQUATE DATABASE

Structure, volume, and level of use can also help in determining the type and amount of data that are considered necessary for an adequate safety evaluation. It is tempting when discussing the extent of data needed for a safety evaluation to say that all potential endpoints must be measured by specific testing of each and every material undergoing evaluation. This is simply unrealistic. To require a full and comprehensive database on each and every one of the 2600+ fragrance materials currently in use is not only impractical and unnecessary, it is a waste of resources not to mention an unacceptable use of animals. This is especially true when it is recognized that approximately 2/3 of these materials are used worldwide in volumes of less than 1 metric ton per year and for over 1000 substances, worldwide usage does not exceed 100 kg/year (RIFM, 1997).

On the other hand, each of these substances must be fully evaluated as to their potential for adverse effects under their conditions of use. The challenge is to accomplish this in an efficient but comprehensive manner so as to assure protection of the consumer while considering resources.

As a practical approach RIFM has developed volume and use level cutoffs for certain endpoints that preclude certain types of testing as long as adequate data exist on closely related analogues. This approach has been successfully used before (e.g., FDA, 1982, 1993); Rulis, 1989; Munro *et al.*, 1998). However, because the skin is the primary route of exposure for fragrance materials, it was necessary to develop a method that is applicable to this route of exposure.

This approach may be best illustrated by example. At present, no simple saturated fatty ester has been shown to have a significant potential for skin sensitization. If such a material is used in fragrances at levels such that the final skin concentration does not exceed 0.1% in a consumer product *and* the worldwide volume of use is less than 1 metric ton per year *and* there are adequate sensitization test data on close analogues showing no effects, it is considered unnecessary to conduct such a test on this material. On the other hand, since it is well known that quinones or substances metabolically converted to quinones can be skin sensitizers, it would normally be necessary to conduct sensitization tests on such materials regardless of use level or volume.

The limits used in the database criteria, with the exception of those used in consideration of subchronic testing, are based on expert judgement and are considered as practical limits to the need for toxicological testing of some materials used in small quantities and/or at low levels. In the consideration of subchronic testing, the thresholds are based on a very large collection of chronic toxicity NOAELs (Munro *et al.*, 1996).

This approach builds upon methods developed and implemented over the past 30 years based on a practical approach to evaluation and testing. As science progresses, some of the principles will change and the methods described herein will need to evolve. This is especially true in the areas of structure–activity relationships (SAR), mechanism-based approaches, and *in vitro* testing, all of which are expected to increasingly serve as alternatives to animal testing and to be increasingly helpful in designing necessary animal experiments. Of course, when the definitive safety evaluation of the material is undertaken by experts, such as the RIFM Expert Panel, there is always the option of requesting additional test data. The method presented here, however, should, in most cases, assure that adequate data are available prior to such a review.

Guidelines for Specific Toxicological Tests

A safety evaluation considers a broad range of safety-related information concerning acute toxicity, dermal absorption, dermal irritation, mucous membrane irritation, skin sensitization, subchronic toxicity, toxic mechanisms, mutagenicity, photoirritation, photoallergy, toxicokinetics, developmental and reproductive toxicity, and carcinogenicity. (Specific target organ endpoints such as neurotoxicity and immunotoxicity are included under subchronic toxicity.) All of these endpoints must be addressed in the development of the database for each fragrance ingredient but that does not mean that specific testing for each endpoint is necessarily needed.

The rationale for defining the need for specific studies is described here. In several cases, this rationale involves limits below which testing is not required as long as there are no concerns raised from data on closely related substances.

Nevertheless, in each case where experimental studies are determined not to be necessary, a scientific

justification must be provided. Such scientific justification may include, but is not necessarily limited to, an assessment that there are sufficient existing data on the material or relevant analogues, determination that the exposure is trivial, consideration of likely detoxication pathways, or evidence that the substance is a natural component of the body (or that it metabolizes to such).

Low exposure is generally not in itself sufficient to conclude that there is no requirement for toxicity testing. However, the need for toxicity testing is mitigated when low exposure is accompanied by available information on chemicals that are closely related structurally. While it is possible that a material could have such a low exposure as to justify no testing regardless of chemical structure (Rulis, 1989), such low exposures are not currently known for fragrance ingredients.

In general, the methodology preferred by RIFM is mentioned. It is beyond the scope of this publication to provide specific details of the protocols; however, sample protocols can be obtained on request from RIFM. All studies should be conducted under Good Laboratory Practice Guidelines and, if available, according to OECD Guidelines.

Acute Toxicity

The nature of use and resulting exposure to fragrance ingredients is such that acute toxicity is rarely an issue; nevertheless, an estimate of acute toxicity can be useful. Reasonable estimates are made by analogy to other substances for which comprehensive data are available. Only if there are insufficient data on analogous substances, or if required by regulations, are animal studies conducted to determine oral and dermal acute toxicity. If such a test is needed, a limit test at 2 g/kg orally with rats or dermally with rabbits is the first step. In the event of 50% or fewer deaths, this is considered adequate. Higher mortality calls for progressively lower doses until an estimate of the LD_{50} is determined.

Percutaneous Absorption

RIFM initially assumes that a fragrance ingredient is 100% absorbed through the skin; i.e., everything remaining in contact with skin surface would be absorbed into the general circulation of the body. In cases where this conservative assumption is not considered adequate, absorption studies are necessary. *In vitro* studies using viable excised skin from animals or human subjects are used to estimate the amount of the fragrance ingredient capable of penetrating the skin. These methods are often sufficient for the safety evaluation. The method commonly employed by RIFM uses rat and human skin in a flow through *in vitro* cell (e.g., Bronaugh, 1995). *In vivo* studies in animals and humans, sometimes using radiolabeled material, are done when more refined safety evaluations are needed. Such studies have been conducted by RIFM for several fragrance materials under simulated exposure conditions. The methodology and a discussion of the ethics of such studies are being published elsewhere (Ford *et al.*, 1999).

Methods are under development for predicting skin absorption based on structure and physicochemical properties (e.g., Schaefer and Redelmeier, 1996). These methods have not been fully validated and therefore should be used with caution. In some cases assumptions of less than 100% absorption can be considered for fragrance ingredients predicted to have low absorption based on data for close chemical analogs.

Dermal Irritation

Studies of dermal irritation are conducted in animals and/or humans as the preliminary test for skin sensitization and the need, therefore, is based on the same guidelines as for this test (see below). Such studies are normally conducted by using a single occluded patch under the same conditions as used in the skin sensitization test.

Mucous Membrane Irritation

Concentrations of individual fragrance ingredients rarely exceed 1 to 3% in consumer products; therefore, an examination of mucous membrane irritation is not generally conducted.

Skin Sensitization

Potential for skin sensitization should be tested for all fragrance ingredients for which any *one* of the following three conditions apply: (1) structural alerts for topical effects, (2) an estimate of the maximum concentration in hydroalcoholic products that is equal to or exceeds 0.1%, or (3) a worldwide use for compounding of fragrance mixtures that is equal to or exceeds 1 metric ton per year.

Even if a decision is made that testing is unnecessary for fragrance ingredients below these limits and having no structural alerts for topical effects, a careful comparison should be made to other structurally related substances that have demonstrated no evidence of potential for induction of skin sensitization.

RIFM considers the definitive test for potential to induce dermal sensitization as the human repeat insult patch test (HRIPT) using occluded patches (Draize *et al.*, 1944; Draize, 1959). Tests for skin sensitization in humans are usually preceded by tests for dermal irritation and are conducted at nonirritating doses, usually initially determined as a multiple of the highest reported use level in a consumer product.

Such testing normally must be preceded by animal testing in a validated model such as the Buehler (1965)

guinea pig test with induction at the minimally irritating dose and challenge at the maximum nonirritating dose. In some cases, results from such animal tests may be considered adequate if sufficient data are available on closely related analogs.

Photoirritation

Testing for photoirritation is normally required for all fragrance ingredients having significant ultraviolet absorbance in the range of 290–400 nm if they have a maximum use concentration of 0.1% or higher and an annual volume of use greater than 1 metric ton. Wavelengths less than 290 nm do not reach the earth's surface since they are absorbed, predominantly by ozone, in the stratosphere; the 290- to 400-nm range includes wavelengths that elicit most of the known chemical phototoxic and photoallergic reactions (Kornhauser *et al.*, 1987).

The judgment as to whether there is significant ultraviolet absorption may be based on established rules from electronic structures (Silverstein *et al.*, 1981; Wingrove and Caret, 1981). When there is uncertainty, an ultraviolet spectrum is obtained.

While testing for photoirritation on humans is considered to be the definitive test, such tests are necessarily preceded by appropriate *in vitro* or animal tests which in themselves may provide adequate data. Testing for photoirritation is not necessary if lack of photoirritation potential has been demonstrated with appropriately validated *in vitro* tests such as the 3T3 NRU test (Spielmann *et al.*, 1998).

With groups of closely related materials, testing on representative substances in the group may be sufficient. The test methodology is essentially the same as for irritation except the patch site is irradiated either immediately after application of test material or after patch removal.

Photoallergy

True photoallergy is a rare phenomenon in comparison to photoirritation. Nevertheless because of its severity, testing for photoallergenicity is required on all fragrance ingredients as described above for photoirritation, except that the limit for exposure is lowered to a concentration of 0.01%, unless closely related substances have been tested and shown not to have photoallergic potential. Fragrance ingredients below these exposure thresholds are evaluated carefully relative to other substances in their chemical class as relevant to a decision not to test. The test methodology is essentially the same as for sensitization except the patch site is irradiated either immediately after application of test material or after patch removal.

Subchronic Toxicity

The need for studies of subchronic toxicity (normally 28- or 90-day repeated-dose studies in rats) is based on a consideration of all available data. This includes systemic exposure, knowledge of available metabolic pathways, and data on other substances in the same structural class. Such studies are usually conducted using the dermal route of exposure.

In assessing the need for subchronic testing a decision may be made by reference to the publications of Cramer *et al.* (1978) and Munro *et al.* (1996). Cramer *et al.* (1978) have provided a decision tree approach for estimating toxic risk based on structure. Their method places chemicals into one of three Classes, I, II, or III, reflecting a presumption of low, moderate, or serious toxicity, respectively. This method is used here for assistance in determining the extent of testing. For example, while a classification of III does not necessarily mean that the chemical poses a risk, such a chemical, if it has more than a trivial exposure, should be evaluated for potential subchronic toxicity.

Munro *et al.* (1996) examined over 600 chemicals with over 2900 chronic or reproductive effect no observed effect levels (NOELs) on the basis of each chemical's classification by the Cramer *et al.* (1978) decision tree approach. (These NOELs included neurotoxicity, reproductive toxicity, and other target organ effects.) The resulting cumulative frequency distributions of the most sensitive endpoints provide distinct, nonoverlaping, sigmoid curves for each of the three classes. The lower fifth percentile NOEL was determined for each class from the distribution curve as 3000, 910, or 150 μg/kg body wt/day for Class I, II, or III, respectively. The fifth percentile NOEL was considered conservative based on the strict criteria for selecting the NOELs, the extensive testing for a variety of endpoints, and the presumption that there is no genotoxic carcinogenicity for the selected chemical. This is an endpoint that must be evaluated separately (see below).

The NOELs should be reduced to allow for a safety factor for humans and for recognition that dermal exposure usually results in a lower systemic exposure than does oral exposure for the same dose. It is considered that a division by five is reasonable to apply to the lower fifth percentile NOELs for animal oral exposure to obtain a human dermal exposure threshold (HDET). This provides HDET values of 600, 182, and 30 μg/kg body wt/day for Class I, II, or III, respectively. These values are compared with the total systemic exposure.

Exposure values above the HDET generally require testing of the material itself or a structurally closely related analogue. However, even in these cases, consideration of metabolism may preclude the need for such testing. Values below the HDET generally preclude testing when data are available on structurally related analogs.

Before deciding not to perform a subchronic toxicity test for a substance, consideration must be given to existing data for structurally related materials, metabolism to innocuous metabolites, possible intoxication and detoxication mechanisms, and the systemic and mutagenic/carcinogenic structural alerts, all consistent with total weight of evidence.

Mutagenicity

Testing a fragrance ingredient for mutagenicity (genotoxicity) is considered necessary if the annual worldwide quantity of usage is more than 0.1 metric tons and the material possesses any of the alert structures for this endpoint (Appendix B) or if mutagenicity potential has been indicated by other tests on the material or with structurally related materials. An *in vitro* point mutation assay (for example, Ames battery) and an *in vitro* mammalian cell chromosomal aberration test (e.g., mouse micronucleus assay) are normally considered adequate unless positive results arise from these tests or significant evidence that raises questions about this endpoint arises from other studies. In such cases *in vivo* screening tests should be considered. For fragrance ingredients with structural alerts and worldwide quantities of use less than 0.1 metric tons/year, a careful analysis of exposure, the nature of any structural alerts, and data on structurally related materials should be used in deciding the need for testing.

Developmental and Reproductive Toxicity

These studies may be necessary if significant dermal absorption occurs and data from other toxicity tests indicate an effect on reproductive organs. They may also be considered necessary if closely related substances have previously shown effects on reproduction or fetal development. In particular indicators for hormonal-related changes such as may be observed in subchronic toxicity studies provide a basis for the need for these studies.

Testing a representative member of a chemical class may be warranted if significant dermal absorption occurs and no relevant data exist on any member of the chemical class or closely related analogs.

Carcinogenicity

Because of the complexity of this endpoint and the different mechanisms that might be involved in carcinogenicity (Clayson and Kitchen, 1998), RIFM would not normally undertake a carcinogenicity bioassay prior to review by the RIFM Expert Panel where their expertise would be depended upon. Indirect evidence relevant to this endpoint can be gained from several of the above-described studies.

For example, chemical structure and known intoxication pathways can serve as an alert for electrophilicity, a chemical property often associated with potential for carcinogenicity. Clearly data from genotoxicity testing are relevant to this endpoint. Indeed, most of the structural alerts for mutagenicity/carcinogenicity in Table B-1 are related to this property. Evidence for enhanced cellular proliferation and other epigenetic mechanisms can often be obtained from subchronic studies.

In most cases, however, indications of potential human carcinogenicity will result in the withdrawal of the substance from use as a fragrance ingredient unless there is indication for a nongenotoxic mechanism, which then requires further evaluation.

Toxicokinetics and Metabolism

Toxicokinetic and/or metabolic data are required when it is necessary to resolve issues raised by the results of other testing or by consideration of structurally related substances. Such data are often used, however, to justify the use of data on analogs.

For example, if a fragrance ingredient is rapidly converted to a metabolite for which the toxicity is well known, the data on the metabolite are used in the evaluation of the precursor. Examples of this are acids, alcohols, and their esters. It is usually assumed that systemic testing of an ester is not necessary where adequate data exist on the hydrolyzed acid and alcohol or vice versa. However, when the ester is sterically hindered, it may be necessary to demonstrate rapid hydrolysis under physiological conditions.

Other examples are secondary alcohols and the corresponding ketones, functional groups that are well recognized to be metabolically interchangeable; acetals and ketals and their corresponding aldehydes, ketones, and alcohols; and primary alcohols and their corresponding carboxylic acids. While transformation and interconversion are generally assumed in these cases, it may be necessary in special cases to conduct studies to support that assumption.

Use of Human Data

Relevant data developed from safety tests on human subjects (e.g., data on the testing of cosmetic products from which fragrance ingredient information can be derived) are analyzed to determine whether conclusions regarding individual fragrance ingredients can be drawn.

Testing of individual fragrance ingredients on human subjects is sometimes necessary. Human data appropriately obtained usually lead to more consequential safety evaluations; therefore, when conclusions stemming from human data contradict those arising from animal tests, the former takes precedence. When adequate-in-use experience is documented, this information should be utilized as part of the safety assessment.

TABLE A-1
Calculation of Dermal Exposure to a Specific Fragrance Ingredient in Cosmetic Products

Type of cosmetic product	Grams applied	Applications per day	Retention factor	Mixture/ product	Ingredient/ mixture	Ingredient/ product	Ingredient mg/day	Ingredient mg/kg/day
Body lotion	8.00	0.71	1.000	0.004	X	0.004X	22.720X	0.378X
Face cream	0.80	2.00	1.000	0.003	X	0.003X	4.800X	0.0800X
Eau de toilette	0.75	1.00	1.000	0.080	X	0.080X	60.000X	1.0000X
Fragrance cream	5.00	0.29	1.000	0.040	X	0.040X	58.000X	0.9667X
Antiperspirant	0.50	1.00	1.000	0.010	X	0.010X	5.000X	0.0833X
Shampoo	8.00	1.00	0.010	0.005	X	0.005X	0.400X	0.0067X
Bath products	17.00	0.29	0.001	0.020	X	0.020X	0.099X	0.0016X
Shower gel	5.00	1.07	0.010	0.012	X	0.012X	0.642X	0.0107X
Toilet soap	0.80	6.00	0.010	0.015	X	0.015X	0.720X	0.0120X
Hair spray	5.00	2.00	0.010	0.005	X	0.005X	0.500X	0.0083X
								Total = 2.55X

Note. X is the fractional amount of fragrance ingredient/fragrance mixture.

Documentation of the Database

The complete database for each fragrance material and any justification of decisions that a test is not necessary must be fully documented. The data may be presented in special formats or dossiers tailored to meet specific needs of the particular group of experts that will conduct the safety evaluation. It is the intent of RIFM to publish the documentation along with the safety evaluation by the RIFM panel in the open and peer-reviewed literature.

CONCLUSIONS

The above-described methods and criteria for determining order of review and for establishing an adequate database for safety evaluation of fragrance ingredients are designed to be practical but comprehensive and scientifically sound. They are published for public review and comment. These methods are currently being implemented and the first priority fragrance ingredients and their structurally related materials have been through the process and have been reviewed by the Panel. Data on two other groups are currently being compiled.

It is fully recognized that RIFM relies on the Panel for final decisions on the adequacy of the complied databases. It is also recognized that these databases must be maintained and updated to assure that the most recent methodology and technology are used in an efficient and practical manner.

APPENDIX A

Basis for Determining Estimated Dermal Exposure for Users of Cosmetic Products Containing a Specific Fragrance Ingredient

Human exposure to fragrance materials results primarily from the use of cosmetic products. The determinant factors for this exposure are quantities of cosmetic used, frequency of use, and concentration of the fragrance material in these products.

The quantity used and frequency of application for a range of cosmetic products are presented in Table A-1. Usage is expressed as a daily exposure although it is based on cosmetic products likely to be used in a weekly period. Thus, it has been estimated that a body lotion may be used 5 days per week (i.e., 0.71 times per day), a fragranced cream or a bath product each 2 days per week (i.e., 0.29 times per day), and a shower gel 15 days per 2-week period (i.e., 1.07 times per day). Usage figures have been established at typical levels (COLIPA, 1997).

This analysis regarding usage involves several assumptions. For example, it is assumed that a body lotion (5 days per week) and a fragrance cream (i.e., a body lotion containing a higher level of fragrance) (2 days per week) will not both be used on the same day. The use of face cream includes the use of makeup and foundation. The use of antiperspirants includes the use of deodorants. The use of eau de toilette includes the use of all alcohol-based products (i.e., perfumes, aftershaves, colognes). These products are not all used on one occasion; when one of these products is used at a lower concentration, it is assumed that a larger amount is applied, thus equating to the values used for eau de cologne. Retention factors for the skin are conservative estimates from known use of products, taking into account wash-off characteristics.

The concentration of the fragrance mixture in a cosmetic product type has been determined by the cosmetic industry (COLIPA, 1997).

The concentration of a fragrance ingredient in a fragrance mixture is based on data obtained by the fragrance industry from the knowledge of commercialized formulations containing the fragrance ingredient. The 97.5 percentile concentration of a specific fragrance

ingredient in fragrance mixtures is currently provided by a group of perfumes from several leading fragrance companies (IFRA, 1999).

Total dermal exposure to a consumer for a specific fragrance ingredient is determined by adding figures for the different product types expressed as milligram per kilogram body weight per day based on a 60-kg adult. In view of the stated assumptions, this value for dermal exposure must be regarded as conservative; it is extremely unlikely that a consumer will consistently use a number of different cosmetic products all of which are perfumed with a fragrance mixture containing the 97.5 percentile concentration of that specific fragrance ingredient.

Table A-1 is used for estimating the total dermal exposure that could be available for percutaneous absorption and systemic toxicity simply by substituting the 97.5 percentile level of use in fragrance mixtures for X. For decisions about testing for dermal effects, an alternate method using the top ten concentrations in hydroalcoholic products currently is used.

Where appropriate, a refinement of the use level data is obtained by collecting through IFRA the distributions of the concentrations of use of the fragrance ingredient in marketed fragrances for the use categories listed in Table A-1. This allows for the determination of an upper percentile value for X in each category. Where IFRA guidelines restrict the use level or levels of a substance, the restricted level(s) should be used as X in Table A-1.

APPENDIX B

The Use of Structural Alerts, i.e., Structural Moieties That Elicit Alerts for Potential Toxic Effects

The scientific basis for evaluating potential toxic effects from structural moieties within the molecule has improved markedly over past years. Table B-1 summarizes the current views of the authors for application to the three basic types of potential toxic effects. It should be clearly understood that the table is dynamic and must be modified as new information, or a questionable correlation, comes to the attention of the users. The literature to support the Table B-1 is Barratt *et al.* (1994a,b), Basketter *et al.* (1994), Barratt and Basketter (1994), Cramer *et al.* (1978), Ashby (1985, 1994), Tennant and Ashby (1991), Ashby and Tennant (1991), and Ashby and Paton (1993). In addition, the authors have, based on their own experience, added a few alerts that were not included in the original publications as described below.

It must be understood that the structural alerts are derived from studies of a large number of natural and synthetic substances and must not be construed to mean that any other substance that may contain the same or a similar structural feature necessarily has the same toxic potential. They are used here to aid in the setting of priorities and to be used as part of the criteria for determination of an adequate toxicological database.

All of the structural alerts that are provided are not necessarily found in fragrance ingredients. They have been provided for completeness and possible applications for the future. The assigned scores only have meaning for the purpose of setting priorities. They were determined with an attempt to make the highest possible structural alert score, 12 (2 for topical alerts plus 4 for acute/systemic alerts plus 6 for carcinogenic/mutagenic alerts), approximately equal to the highest score for use volume or use levels, 16.

A detailed discussion of the reasoning and possible mechanisms associated with each structural alert is well beyond the scope of this document. In some cases, such as number 29, the mechanism is well understood. (The alkoxy groups and the double bond act as a stabilizer for the incipient carbonium ion, the electrophile that is DNA reactive, and results from metabolic oxidation of the methylene group followed by formation of the sulfate). Some are a result of caution such as 15 which assumes that all nitriles may act as cyanide donors even though many are known not to. Some are based on knowledge of detoxication mechanisms such as 7 which is based on the knowledge that the only route available for tertiary alcohols is conjugation and excretion, and since such alcohols are at least somewhat sterically hindered, this process can be slowed. Most, however, are based on results from tests of substances containing the particular functional moiety.

Some general comments can be made, however.

For topical effects a score of 2 was given to each of the structural alerts for skin sensitization taken from the cited literature and summarized in a structural alert table in Barratt *et al.* (1994a). In three cases, the authors have added topical structural alerts that were not included in Barratt *et al.* (1994a) based on personal knowledge and experience: alert 2 based upon the known sensitizing potential of methyl octine carbonate and methyl heptine carbonate; alert 16 based on the known photosensitization potential of musk ambrette; and alert 31 based on the known irritancy potential of some phenols.

For acute/systemic effects, scores of 0, 2, or 4 were given to structures consistent with structure Classes I, II, or III, respectively, based on the classification scheme of Cramer *et al.* (1978) and supported by the database presented in Munro *et al.* (1996). In four cases, the authors have added acute/systemic structural alerts that were not included in Cramer *et al.* (1978): alert 1 for conjugated dienes based on the known toxicity of butadiene and the propensity of these structures to form reactive epoxides; alert 7 for tertiary alcohols and their esters for the reasons given in the general discussion above; alert 30 based on the known

TABLE B-1
Structural Alerts, i.e., Structural Moieties That Elicit Alerts for Potential Toxic Effects

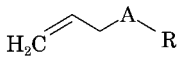
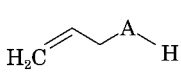
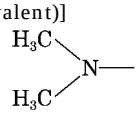
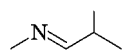
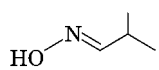
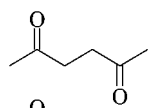
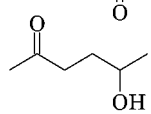
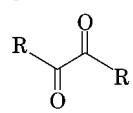
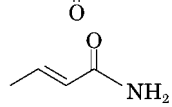
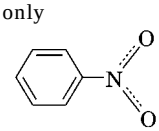
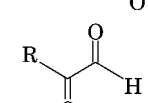
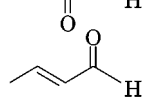
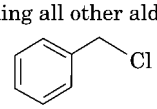
No.	Structural alert	Topical effects	Acute/systemic effects	Carcinogenic/mutagenic effects
1	 Cannot be part of a benzene ring	0	2	6
2	—C≡C—	2	2	0
3	Cyclopropyl and/or cyclobutyl derivatives	0	2	0
4	Cycloalkanones and cycloalkenones with $C_n > 4$	2	2	0
5	Cycloalkanols and cycloalkenols with $C_n > 4$	0	2	0
6	  A = O, S, N	2	2	6
7	*Tertiary* alcohols and their esters	0	2	0
8	R-D [where D = any atom other than C, H, O, S (divalent), N (trivalent)]	2	4	0
9		0	4	6
10	 	0	4	0
11	  Including ketals of the above	0	4	0
12		2	4	0
13		2	2	6
14	Heterocycles except lactones with a ring size >4	0	4	0
15	R—CN R = aliphatic C only	0	4	0
16		2	4	6
17	  Including all other aldehydes	2	0	6 (restricted)
18		0	4	6

TABLE B-1—*Continued*

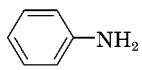
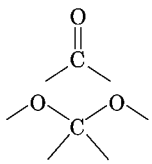
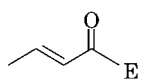
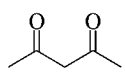
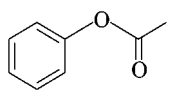
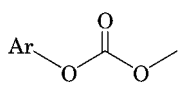
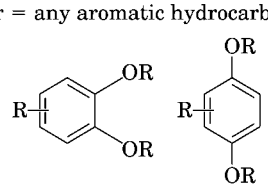
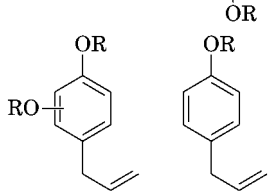
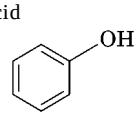
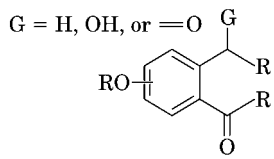
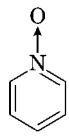
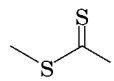
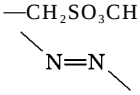
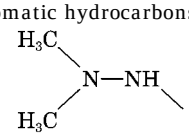
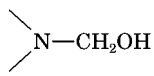
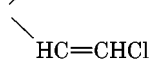
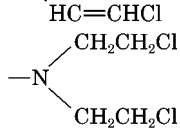
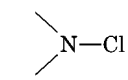
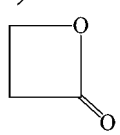
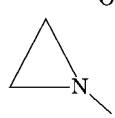
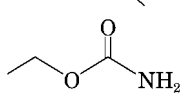
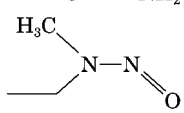
No.	Structural alert	Topical effects	Acute/systemic effects	Carcinogenic/mutagenic effects
19	 Including only ring substituted anilines	2	2	6
20	 including arene oxides	2	4	6
21		2	0	0
22	including any carbon structures on open bonds	2	4 unless known to be hydrolyzed; if so treat as hydrolyzed	0
23	Enols, enol ethers, enol esters, acid anhydrides, isocyanates, isothiocyanates, β-lactams, quinones, disulfides, 1,2-diamines, quaternary ammonium cations, precursors of α, β-unsaturated aldehydes, thiazoles, and thiazolines	2	0	0
24	 Double bond not part of an aromatic ring E = C, N, O (but not OH)	2	0	0
25		2	0	0
26		2	0	0
27	 Ar = any aromatic hydrocarbon	2	0	0
28		2	0	0
29		0	4	6
30	Ethers and esters of ethylene glycol or hydroxyacetic acid	0	2	0
31		2	0	0

TABLE B-1—*Continued*

No.	Structural alert	Topical effects	Acute/systemic effects	Carcinogenic/mutagenic effects
32		0	4	0
33	 All aromatic ring *N*-oxides	0	4	6
34		2	4	0
35		2	4	0
36	Aliphatic 1^0 or 3^0 amines	0	2	0
37	$—CH_2SO_3CH_3$	0	0	6
38		0	4	6
39	Polycyclic aromatic hydrocarbons	0	4	6
40		0	0	6
41		0	0	6
42		2	4	6
43		2	4	6
44		0	4	6
45		0	4	6
46		0	4	6
47		0	0	6
48		0	4	6
49	$—CX_n$ where X = halogen	2	4	6
50	Default (when no structure is available)	2	4	6
51	When no structural alert is present	0	0	0

Note. See instructions for use in the introduction to Appendix B.

potential for reproductive effects of glycol ethers; and alert 34 as a default because this structure does not seem to have been considered in Cramer *et al.* (1978).

For carcinogenic or mutagenic effects a score of 6 was given for structures that may be DNA-reactive and is based on the alerts given in the publications by Ashby, Tennant, and Paton (Ashby, 1985, 1994; Tennant and Ashby, 1991; Ashby and Tennant, 1991; Ashby and Paton, 1993). In addition, the authors have supplemented the list of alerts: alert 1 based on the same reasoning as described in the previous paragraph; alert 28 based on the known carcinogenicity of safrole and related materials; and alert 39 based on the known carcinogenicity of several polycyclic aromatic compounds.

Alert 17 is labeled "restricted" for carcinogenic or mutagenic effects even though Ashby, Tennant, and Paton list it as a structural alert for mutagenicity. Their system is primarily aimed at predicting the potential for mutagenicity in *in vitro* systems, whereas *in vivo,* aldehydes are known to be rapidly detoxicated by oxidation to carboxylic acids. Therefore, the presence of an aldehyde group in a structure will not automatically require consideration of a high alert for carcinogenicity/mutagenicity. Rather the high alert will be restricted pending evaluation of the entire molecule with consideration of metabolic and other factors that may alter the toxic potential.

In Table B-1, the following conventions have been used: wherever there is a bond that is open ended, it is intended that the bond must connect with another carbon atom; "R-" refers to any carbon with one free bond for attachment; and "Ar-" refers to any aromatic carbon.

APPENDIX C

Setting Priorities for Safety Review Using Structure, Volume, and Level of Use

The stepwise approach to setting priorities for safety evaluation is based on the sum of scores for structural alerts, scores for quantity of use, and scores for concentration levels in the final product.

TABLE C-1

Quantity of use (metric tons per year)	Score
<0.1	0
0.1–<1	1
1–<10	2
10–<100	4
100–<1000	8
≥1000	16
Default[a]	8

[a] Use default when the quantity of use is unknown.

TABLE C-2

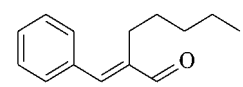
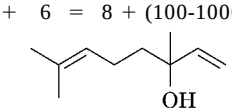
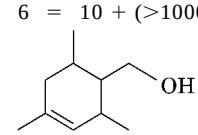
Concentration level in final product (%)	Score
<0.05	0
0.05–<0.1	1
0.1–<0.5	2
0.5–<1.0	4
1.0–<5.0	8
≥5.0	16
Default[a] (use 3%)	8

[a] Use default when concentration level is unknown.

For structural alert scores see Appendix B for topical effects, acute/systemic effects, and carcinogenicity/mutagenicity effects. Note that while the scores for each type of alert are added, multiple alerts for the same type are not. Thus, if there are structural alerts for both topical effects and carcinogenicity/mutagenicity effects, the score would be 8 (2 + 6). However, if there were two structural alerts for topical effects, the score would only be 2. If there are two or more alerts for acute/systemic effects, use the highest score only.

The volume of use score is based on the annual worldwide volume of use in fragrance mixtures and does not include material used as chemical precursors (Table C-1).

TABLE C-3
Examples of Determining Overall Priority Scores for Fragrance Ingredients

Rule	Alert score T	A/S.	C/M	
				
	α-Hexylcinnamaldehyde			
17	2	0	6	
24	2	0	0	(Volume) Score (Level) Score Total
Overall	2 +	0 +	6 =	8 + (100-1000 tn) 8 + (1.2%) 8 = 24
				
	Linalool			
6	2	2	6	
7	0	2	0	(Volume) Score (Level) Score Total
Overall	2 +	2 +	6 =	10 + (>1000 tn) 16 + (2.0%) 8 = 34
				
	2,4,6-Trimethyl-3-cyclohexenyl methanol			
51	0	0	0	(Volume) Score (Level) Score Total
Overall	0 +	0 +	0 =	0 + (0.1-1tn) 1 + (0.4%) 2 = 3

The level of use score (Table C-2) is typically based on the highest value reported for alcohol-based products, as discussed in appendix A. Where IFRA guidelines restrict the use level of a substance, the restricted level should be used as the maximum level.

Examples are given in Table C-3.

ACKNOWLEDGMENTS

The authors are indebted to all the many contributors, but acknowledge a few: Emil Pfitzer for extensive critique and editorial improvements and Normal Baker, Peter Cadby, Wing Fung, Mark Fukayama, Friedrich Grundschober, Charles Letemplier, Dan Mooney, Kenneth Schrankel, William Troy, and Hans van Bergen for careful review of criteria applications.

REFERENCES

Adams, T. B., Greer, D. B., Doull, J., Munro, I. C., Newberne, P., Portoghese, P. S., Smith, R. L., Wagner, B. M., Weil, C. S., Woods, L. A., and Ford, R. A. (1998). The FEMA GRAS assessment of lactones used as flavour ingredients. *Food Chem. Toxicol.* **36**(4), 249–278.

Api, A. M., Ford, R. A., and San, R. H. C. (1995). An evaluation of musk xylene in a battery of genotoxicity tests. *Food Chem. Toxicol.* **33**(12), 1039–1045.

Api, A. M., Pfitzer, E. A., and San, R. H. C. (1996). An evaluation of musk ketone in a battery of genotoxicity tests. *Food Chem. Toxicol.* **34**(7), 633–638.

Ashby, J. (1985). Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. *Environ. Mutagen.* **7,** 919–921.

Ashby, J., and Tennant, R. W. (1991). Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutat. Res.* **257,** 229–306.

Ashby, J., and Paton, D. (1993). The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. *Mutat. Res.* **286,** 3–74.

Ashby, J. (1994). Two million rodent carcinogens? The role of SAR and QSAR in their detection. *Mutat. Res.* **305,** 3–12.

Barratt, M. D., Basketter, D. A., Chamberlain, M., Payne, M. P., Adams, G. D., and Langowski, J. J. (1994a). Development of an expert system rulebase for identifying contact allergens. *Toxicol. in Vitro* **8**(4), 837–839.

Barratt, M. D., Basketter, D. A., Chamberlain, M., Admans, G. D., and Langowski, J. J. (1994b). An expert system rulebase for identifying contact allergens. *Toxicol. in Vitro* **8**(5), 1053–1060.

Barratt, M. D., and Basketter, D. A. (1994). Structure–activity relationships for skin sensitization: An expert system. In *Alternative Methods in Toxicology* (A. Rougier, A. M. Goldberg, and H. I. Maibach, Ed.), Vol. 10, pp. 293–301. Mary Ann Liebert, New York.

Basketter, D. A., Bremmer, J. N., Kammuller, M. E., Kawabata, T., Kimber, I., Loveless, S. E., Magda, S., Pal, T. H. M., Stringer, D. A., and Vohr, H.-W. (1994). The identification of chemicals with sensitizing or immunosuppressive properties in routine toxicology. *Food Chem. Toxicol.* **32**(3), 239–296.

Boeck, A., and Fergen, H-U. (1991). Part VI. Production of Perfumes. Chapter 15. Compounding. In *Perfumes: Art, Science and Technology* (P. Müller and D. Lamparsky, Eds.), p. 422. Elsevier, New York.

Bronaugh, R. L. (1995). Methods for in vitro percutaneous absorption. *Toxicol. Methods* **5**(4), 265–273.

Buehler, E. V. (1965). Delayed contact hypersensitivity in the guinea pig. *Arch. Dermatol.* **91,** 171–177.

Clayson, D. B., and Kitchen, K. T. (1998). Interspecies differences in response to chemical carcinogens. In *Carcinogenicity: Testing, Predicting and Interpreting Chemical Effects* (K. T. Kitchen, Ed.), pp. 837–880. Dekker, New York.

COLIPA (1997). The European Cosmetics Toiletry and Perfumery Association, Personal communication. Brussels, Belgium.

Cramer, G. M., Ford, R. A., and Hall, R. L. (1978). Estimation of toxic hazard—A decision tree approach. *Food Cosmet. Toxicol* **16,** 255–276.

Cronin, E. (1984). Photosensitivity to musk ambrette. *Contact Dermatitis* **11,** 88–92.

Draize, J. H., Woodard, G., and Calvery, H. D. (1944). Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *J. Pharmacol. Exp. Ther.* **82**(2), 377–390.

Draize, J. H. (1959). *Dermal Toxicity in Appraisal of the Safety of Chemicals in Foods, Drugs and Cosmetics*, p. 46. The Association of Food and Drug Officials of the United States, Texas State Department of Health, Austin, TX.

European Commission (1996). Commission Decision of 8 May 1996 establishing an inventory and a common nomenclature of ingredients employed in cosmetic products. Section II. Perfume and Aromatic Raw Materials. Directive 96/335/EC. *Official J. Eur. Communities* **39,** 526–679.

FDA (1982). *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food*, Redbook. United States Food and Drug Administration, Bureau of Foods, Washington, DC.

FDA (1993). *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food*, Redbook II (Draft). United States Food and Drug Administration, Bureau of Foods, Washington, DC.

Ford, R. A., Api, A. M., and Letizia, C. S. (1992). Monographs on fragrance raw materials: Special issue VIII. *Food Chem. Toxicol.* **30**(S), 1S–138S.

Ford, R. A., Hawkins, D. R., Schwarzenbach, R., and Api, A. M. (1999). The systemic exposure to the polycyclic musks, AHTN and HHCB, under conditions of use as fragrance ingredients: Evidence of lack of absorption from a skin reservoir. *Toxicol. Lett.* **111**(1–2), 133–142.

IFRA (1999). International Fragrance Association, Code of Practice. Geneva, Switzerland.

Kornhauser, A., Wamer, W., and Giles, A., Jr. (1987). Light-induced dermal toxicity: Effects on the cellular and molecular level. In *Dermatotoxicology*, (F. N. Marzuli and H. I. Maibach, Eds.), third ed., Chapter 17, pp. 378–379. Hemisphere, Washington, DC.

Munro, I., Ford, R. A., Kennepohl, E., and Sprenger, J. G. (1996). Correlation of structural class with no-observed-effect levels: A proposal for establishing a threshold of concern. *Food Chem. Toxicol.* **34,** 829–867.

Munro, I. C., Shubik, P., and Hall, R. (1998). Principles for the safety evaluation of flavouring substances. *Food Chem. Toxicol.* **36,** 529–540.

Opdyke, D. L. (1973). Monographs on fragrance raw materials. *Food Chem. Toxicol.* **11,** 95–115.

RIFM (1997). Survey of volume of use of substances used as ingredients in fragrances in Japan, Europe and the United States in 1996.

Rulis, A. M. (1989). Establishing a threshold of regulation. In *Risk Assessment in Setting National Priorities* (J. J. Bonin and D. E. Stevenson, Eds.), pp. 271–278. Plenum, New York.

Schaefer, H., and Redelmeier, T. E. (1996). *Skin Barrier: Principles of Percutaneous Absorption.* pp. 254–255. Karger, Basel, Switzerland.

Silverstein, R. M., Bassler, G. C., and Morrill, T. C. (1981). *Spectro-*

metric Identification of Organic Compounds fourth ed., pp. 305–331. Wiley, New York.

Somogyi, L. P., Rhomberg, B., and Takei, N. (1995). *Flavors and Fragrances.* Chemicals and Energy Practice, SRI Consulting, Menlo Park, CA.

Spielman, H., Balls, M., Dupuis, J., Pape, W. J. W., Pechovitch, G., De Silva, O., Holzhtter, H. G., Clothier, R., Desolle, P., Gerberick, F., Liebsch, M., Lovell, W. W., Maurer, T., Pfannenbecker, U., Potthast, J. M., Csato, M., Sladowski, D., Steiling, W., and Branton, P. (1998). EU/COLIPA In vitro phototoxicity validation study, results of Phase II (blind trial). 1. The 3T3 NRU phototoxicity test. *Toxicol. in Vitro* **12**(3), 305–327.

Tennant, R. W., and Ashby, J. (1991). Classification according to chemical structure, mutagenicity to *Salmonella* and level of carcinogenicity of a further 39 chemicals tested for carcinogenicity by the U.S. National Toxicology Program. *Mutat. Res.* **257,** 209–227.

Vuilleumier, C., Flament, I., and Sauvegrain, P. (1995). Headspace analysis study of the evaporation of perfume ingredients applied to the skin. *Int. J. Cosmet. Sci.* **17,** 61–67.

Wingrove, A. S., and Caret, R. L. (1981). Organic chemistry. *Spectroscopy I—Spectroscopic Methods: Infrared and Ultraviolet Spectroscopy,* Chapter 12, pp. 502–521. Harper & Row, New York.