



Research Article

SpheraCosmolife: A New Tool for the Risk Assessment of Cosmetic Products

Gianluca Selvestrel¹, Federica Robino², Diego Baderna¹, Serena Manganelli^{1,3}, David Asturiol⁴, Alberto Manganaro⁵, Matteo Zanotti Russo², Giovanna Lavado¹, Cosimo Toma¹, Alessandra Roncaglioni¹ and Emilio Benfenati¹

¹Laboratory of Environmental Chemistry and Toxicology, Environmental Health Department, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milano, Italy; ²Angel Consulting s.a.s., Milano, Italy; ³current affiliation: Chemical Food Safety Group, Nestlé Research, Lausanne, Switzerland; ⁴European Commission, Joint Research Centre (JRC), Ispra, Italy; ⁵Kode s.r.l., Pisa, Italy

Abstract

A new, freely available software for cosmetic products has been designed that considers the regulatory framework for cosmetics. The software allows an overall toxicological evaluation of cosmetic ingredients without the need for additional testing and, depending on the product type, it applies defined exposure scenarios to derive risk for consumers. It takes regulatory thresholds into account and uses either experimental values, if available, or predictions. Based on the experimental or predicted no observed adverse effect level (NOAEL), the software can define a point of departure (POD), which is used to calculate the margin of safety (MoS) of the query chemicals. The software also provides other toxicological properties, such as mutagenicity, skin sensitization, and the threshold of toxicological concern (TTC) to provide an overall evaluation of the potential chemical hazard. Predictions are calculated using *in silico* models implemented within the VEGA software. The full list of ingredients of a cosmetic product can be processed at the same time, at the effective concentrations in the product as given by the user. SpheraCosmolife is designed as a support tool for safety assessors of cosmetic products and can be used to prioritize the cosmetic ingredients or formulations according to their potential risk to consumers. The major novelty of the tool is that it wraps a series of models (some of them new) into a single, user-friendly software system.

1 Introduction

The European regulation on cosmetics represented a paradigm shift in Europe for the safety assessment of cosmetics, which transitioned from the classical toxicological approach based on animal testing towards a completely novel strategy, where the use of animals for toxicity testing is banned (EC, 2009). The European strategy has been followed by an increasing number of countries in the world¹. However, the regulation does not provide details on which alternative methods are to be used. Several ambitious projects have addressed sophisticated alternative testing strategies, such as the European initiatives SEURAT-1² (Berggren et al., 2017) and EU-ToxRisk³.

The outcomes of a workshop organized by the Scientific Committee on Consumer Safety (SCCS) have now been published, indicating some areas of research still to be addressed (Rogiers et al., 2020).

Using *in silico* tools is particularly appealing because they can generate safety data for cosmetic ingredients without testing (Gellatly and Sewell, 2019; Taylor and Rego Alvarez, 2020). There are many tens of thousands of cosmetic ingredients, and this number poses an incredible challenge, not only for *in vivo* methods, but also for other test methods. The possibilities that *in silico* models offer in this respect were demonstrated by calculating various properties for about 20,000 botanical ingredients of cosmetics (Raitano et al., 2019).

¹ <https://www.cosmeticsdesign-europe.com/Article/2019/03/05/Global-ban-on-animal-testing-where-are-we-in-2019>

² <http://www.seurat-1.eu/>

³ <https://www.eu-toxrisk.eu/>

Received October 22, 2020; Accepted April 29, 2021;
Epub May 3, 2021; © The Authors, 2021.

ALTEX 38(4), 565-579. doi:10.14573/altex.2010221

Correspondence: Gianluca Selvestrel, PhD
Laboratory of Environmental Chemistry and Toxicology
Environmental Health Department
Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milano, Italy
(gianluca.selvestrel@marionegri.it)

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.


Tab. 1: Risk assessment steps defined by regulators and their translation into SpheraCosmolife

Regulatory procedure	Technical guidance/Regulation references	SpheraCosmolife	Comparison consideration
HAZARD IDENTIFICATION: carried out to identify the intrinsic toxicological properties of the substance	SCCS NoG* (relevant toxicological studies on cosmetic ingredients) / Annex I. 8 of Reg. 1223/2009 (toxicological profile of the substances)	Experimental values are retrieved from the internal database. In case experimental data are missing, <i>in silico</i> models (expert and statistical-based tools) to predict toxicological properties are provided.	Novelty of computational approaches to refine the assessment. Other toxicological endpoints will be added in future versions.
EXPOSURE ASSESSMENT: based on the declared functions and uses of a substance as cosmetic ingredient, the amount present in the respective cosmetic product categories, and the frequency of use	SCCS NoG* (Exposure assessment) / Annex I. 6 and 7 of Reg. 1223/2009 (exposure to the cosmetic product, exposure to the substances)	Application of the official equations and implementation of predefined exposure scenarios within the software. Introduction of a refined approach based on new models for skin permeation to estimate the internal exposure.	Novelty of computational approaches to refine the assessment. Automatic creation of exposure scenarios.
DOSE-RESPONSE ASSESSMENT: based on calculation of the POD	SCCS NoG*	Automatic calculation of the POD (NOAEL) using a QSAR model. Experimental values are also in the internal database.	Novelty of computational approaches to refine the assessment. New models for NOAEL and LOAEL will be implemented in future versions. Organ-specific toxicity will be indicated.
RISK CHARACTERIZATION: focused on systemic toxicity based on the MoS	SCCS NoG* (General principles for the calculation of the margin of safety and threshold of toxicological concern)	MoS is automatically calculated using the previous results. A decision tree is implemented to provide a TTC assessment.	Automatic calculation of the MoS. A decision tree approach allows estimation of the TTC class and compares the TTC thresholds with the exposure values, incorporating dermal bioavailability into the use of TTC.

*ref: SCCS, 2018. LOAEL, lowest observable adverse effect level; MoS, margin of safety; NOAEL, no observable adverse effect level; NoG, notes of guidance; POD, point of departure; QSAR, quantitative structure-activity relationship; SCCS, Scientific Committee on Consumer Safety; TTC, threshold of toxicological concern

However, apart from hazard identification, risk assessment of cosmetic ingredients requires both hazard characterization and estimation of exposure. In fact, the assessment of the risk posed by one or more components of the cosmetic product depends on the concentration of the ingredients, and the exposure must be examined, starting from skin permeation by the substance of interest. Furthermore, the assessor needs to use different software systems to obtain the various values for the different toxicological endpoints and exposure values.

Within the EC-funded LIFE project VERMEER⁴, a strategy was drawn up to integrate several tools into a single computer software system to support safety assessors of cosmetic products. To comply with the cosmetics regulation, we planned to reproduce the procedure employed for cosmetics as closely as possible, with specific reference to the lists of already regulated ingredients, the thresholds that need to be respected, and the specific

endpoints that must be evaluated. As a result, we built the software system SpheraCosmolife presented here.

2 Methods

The novel software system combines an expert system approach, which refers to the sequence of steps followed by the assessors, with machine-learning and statistical models to provide predictions. Thus, the novel system has a sound theoretical basis derived from the procedure defined by the regulators, which is currently performed manually in most cases. To demonstrate the safety of a cosmetic product prior to placing it on the market, the responsible person must ensure that it has undergone a safety assessment. Annex I of Regulation (EC) 1223/2009 defines the aspects to be evaluated within the cosmetic product safety report (EC, 2009). More-

⁴ <https://www.life-vermeer.eu>

over, as indicated in Article 11 of Regulation 1223/2009, the responsible person shall keep a product information file (PIF) for it.

The procedure for the safety evaluation of a cosmetic product is defined in detail by “The SCCS Notes of Guidance for the testing of cosmetic ingredients and their safety evaluation 10th revision” (SCCS, 2018). Table 1 shows the workflow that a risk assessor follows and explains how these steps have been translated into SpheraCosmolife.

The risk assessment procedure followed by the regulators, including the four main steps (hazard identification, exposure assessment, dose-response assessment, and risk characterization) is largely replicated by the software. The *in silico* structure of the tool represents a forward-looking and future-oriented idea and supports the assessor with remarkable time-saving as well as with a considerable amount of data and information that facilitate the decision-making procedure. The assessor must still evaluate the provided values and their uncertainty (also considering the remarks indicated by the software), so the software does not substitute the work of the assessor.

The software will continue to be improved by including more data and information in the future. New technical and scientific issues, for example new toxicological endpoints and evaluation approaches (e.g., read-across) will be implemented in future versions in order to better cover the evaluation procedure and offer a more complete and innovative assessment. The following steps describe the SpheraCosmolife software system structure.

2.1 Identification of the ingredients, products and exposure scenarios

The ingredients of a product are identified through the International Nomenclature of Cosmetic Ingredients (INCI), Simplified Molecular Input Line Entry System (SMILES) format, or the Chemical Abstract Service (CAS) number. The identification of the product types derives from Tables 2A, 2B and 3 of the 10th revision of the SCCS Notes of Guidance (NoG) (SCCS, 2018). The exposure scenarios are those defined in this document, for each product type.

2.2 The database

A key component of the software system is its database. It was populated with data retrieved from the COSMOS Cosmetic Inventory⁵ (Worth et al., 2012) and the information in the Annexes to Regulation (EC) 1223/2009 (EC, 2009), in particular:

- Annex II: List of substances prohibited in cosmetic products
- Annex III: List of substances which cosmetic products must not contain except subject to the restrictions laid down
- Annex IV: List of colorants allowed in cosmetic products
- Annex V: List of preservatives allowed in cosmetic products
- Annex VI: List of UV filters allowed in cosmetic products

Further sources of data refer to these repositories:

- CLP harmonized classification⁶
- Safer Chemical Ingredients List (SCIL)⁷

These data sources refer to official repositories from the EU and the US Environmental Protection Agency (EPA). Overall, the database of SpheraCosmolife currently contains data related to about 5000 substances, but it is regularly updated with new and upgraded information and data. Most of these substances are cosmetic ingredients, and the following categories are particularly well represented: skin conditioning, skin protecting, surfactants, emulsifying, and perfuming. From a chemical point of view, alcohols, amines, ketones, and substances with an aliphatic chain of at least 8 carbons are well represented.

The information in the Annexes includes the maximum concentration, the product type in which the ingredient is allowed, and the wording of conditions of use and warnings. With regard to inorganic and organometallic compounds, polymers, and data related to mixtures of chemicals, SpheraCosmolife recognizes them if they are present in the database; otherwise, the software is not able to process these kinds of substances. The original structure of salts is stored in the database, so the user can use a salt as an input. However, the final assessment is done using the neutralized structure, without the cation or anion. Semi-automated data curation and quality checking workflow using the KNIME platform (Gadaleta et al., 2018) and expert-based knowledge were used to retrieve neutralized structures in the database. However, for substances not in the database, the user should insert the structure of the neutral substance, without ions.

2.3 Risk prediction models

The software evaluates a number of properties as described below. For these properties, the source of the data is the VEGA platform⁸, and information on a specific property can be found in the description of the *in silico* models for the individual property, available from VEGA. New *in silico* models have also been developed (see below).

The SpheraCosmolife system uses several equations and *in silico* models. Some models estimate exposure, others predict hazard. By combining the results of these models, the SpheraCosmolife system can assess the risk associated with the ingredients of a product. If available in the system database, experimental values are used, otherwise the system provides predictions. The software can deal with multiple ingredients of a given cosmetic product, for different categories of products. Although the system treats multiple ingredients, interactions between the ingredients are not considered, since each cosmetic product is considered an individual combination of cosmetic substances, as described in the 9th revision of SCCS NoG (SCCS, 2015).

2.3.1 Exposure evaluation

The list of product types

The user is asked to indicate the product type, as defined in Tables 2A, 2B and 3 of the 10th revision of SCCS NoG (SCCS, 2018), to identify an exposure scenario.

⁵ <https://cosmosdb.eu/cosmosdb.v2/comptools/>

⁶ <https://echa.europa.eu/it/information-on-chemicals/cl-inventory-database>

⁷ <https://www.epa.gov/saferchoice/safer-ingredients>

⁸ www.vegahub.eu



External exposure

External exposure is obtained as indicated by the SCCS NoG (SCCS, 2018). SpheraCosmolife implements Equation 1:

$$\text{External Dermal Exposure (mg/kg bw/day)} = E_{\text{product}} * C/100 \quad (\text{Eq. 1})$$

where E_{product} is the calculated relative daily exposure (mg/kg bw/day), which is tabulated in the tables of the SCCS NoG (SCCS, 2018). C is the concentration of the ingredient (%).

Systemic exposure dose (SED)

SED is obtained as indicated by the SCCS NoG (SCCS, 2018), from Equation 2:

$$\text{SED(mg/kg bw/day)} = E_{\text{product}} * \frac{C}{100} * \frac{DA}{100} \quad (\text{Eq. 2})$$

This equation takes into account the amount of the finished cosmetic product applied per day (E_{product}), the concentration (C) of the substance under study in the product category expressed as a percentage, and the dermal absorption (DA) expressed as a percentage.

A tiered approach is followed to calculate the SED. The SED value is provided by SpheraCosmolife considering three scenarios, where absorption is taken (i) as 100% (for oral and inhalation exposure), or (ii) 50% (for dermal exposure, a default value defined by the SCCS NoG (SCCS, 2018)), or (iii) 10, 40 or 80% (for dermal exposure, in a more accurate way, calculated according to the Kroes approach (Kroes et al., 2007), which requires information on skin permeation, see below).

Skin permeation

Two models are used that implement the models described by Potts and Guy (1992) and ten Berge (ten Berge, 2009; Vecchia and Bunge, 2002) according to Equations 3 and 4, respectively. They provide the constant of permeation, K_p .

$$\log K_p = 0.71 \log(K_{ow}) - 0.0061 MW - 2.7 \text{ (cm/h)} \quad (\text{Eq. 3})$$

$$\log K_p = \left[\frac{1}{\frac{1}{K_{lip}} + \frac{1}{K_{aq}}} \right] \quad (\text{Eq. 4})$$

where:

$$- K_{pol} = \frac{b4}{MW^{b5}}$$

$$- K_{aq} = \frac{b6}{MW^{b7}}$$

$$- K_{lip} = 10^{[b1 + b2 \cdot \log(K_{ow}) + b3 \cdot MW]}$$

- K_{lip} = permeation coefficient lipid medium

- K_{pol} = permeation coefficient corneocytes [proteins]

- K_{aq} = permeation coefficient epidermis [aqueous]

- K_{ow} = octanol/water partition coefficient

- MW = molecular weight

- $b_1, b_2, b_3, b_4, b_5, b_6, b_7$ = regression coefficients:

$$b_1 = -2.694; b_2 = 0.9809; b_3 = -7.868 \cdot 10^{-3}; b_4 = 0.05523;$$

$$b_5 = 1.383; b_6 = 1.121 \cdot 10^3; b_7 = 1.957$$

Once K_p is obtained, the worst-case scenario, i.e., the most conservative value, is employed to proceed with the workflow, and the software calculates the maximum flux of the substance, J_{max} , according to Equation 5:

$$J_{\text{max}} \text{ (mg/cm}^2\text{/h)} = K_p * C_{\text{water,sat}} \quad (\text{Eq. 5})$$

where $C_{\text{water,sat}}$ is the saturated water solubility, in mg/cm³, obtained using the model implemented in the VEGA platform.

Once J_{max} is obtained, the software applies the Kroes approach (Kroes et al., 2007) according to Equation 6, which provides the percentages of absorption (%A):

$$\%A = \begin{cases} 10\% & \text{if } J_{\text{max}} \leq 0.1 \text{ } \mu\text{g/cm}^2\text{/h} \\ 40\% & \text{if } 0.1 < J_{\text{max}} \leq 10 \text{ } \mu\text{g/cm}^2\text{/h} \\ 80\% & \text{if } J_{\text{max}} > 10 \text{ } \mu\text{g/cm}^2\text{/h} \end{cases} \quad (\text{Eq. 6})$$

The SpheraCosmolife software provides the percentages of absorption for these three cases, which represent low, medium and high dermal absorption, as defined by Kroes (Kroes et al., 2007; Shen et al., 2014).

2.3.2 Hazard evaluation

The hazard assessment of chemicals is performed with several *in silico* models and profilers available on the VEGA platform. A brief description of each model is presented below.

Mutagenicity, bacterial reverse mutation test – Ames test

The consensus model available in VEGA is used for Ames mutagenicity prediction. It combines the results of four models, taking into account the reliability of each prediction for the target substance (Manganelli et al., 2018). It integrates models using both expert- and statistical-based tools, following the recommendation of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) (ICH, 2017).

Chromosomal aberration

The integrated model available in VEGA is used for chromosomal aberration. It was built with the CORAL software using SMILES-based attributes. The classification model is based on a dataset of 477 organic compounds (223 active and 254 inactive in chromosomal aberration tests). The data were collected from the Genotoxicity OASIS Database and from the Toxicity Japan MHLW, which include experimental data for chromosomal aberrations determined by *in vitro* testing using Chinese hamster lung (CHL) and ovary (CHO) cells, with and without S9 metabolic activation (Toropov et al., 2019).

In vitro micronucleus genotoxicity test

The model developed with the SARpy software, based on structural alerts, is used for the *in vitro* micronucleus genotoxicity

ty test. The model provides a qualitative prediction of genotoxicity as induction of micronucleus in mammalian cells *in vitro* (MNvit). The model was built on a dataset containing 380 organic chemicals with genotoxicant and non-genotoxicant MNvit experimental data (153 inactive and 227 active chemicals). The experimental data were collected, according to the OECD 487 guideline, from eChemPortal inventory, peer-reviewed literature, SCCS and European Food Safety Authority (EFSA) opinions, European Centre for Validation of Alternative Methods (ECVAM) guidelines and review. The fragment-based model uses 138 structural alerts, including 82 active and 56 inactive fragments (Baderna et al., 2020).

Skin sensitization

Two models are used, which are available in VEGA. The first is the CAESAR model (Chaudhry et al., 2010). The second is a new model, which is described here: It is a decision tree (DT) based on a data set of 332 chemicals with data from the local lymph node assay (LLNA) (226 sensitizers and 106 non-sensitizers). Data were collected from the CAESAR model database (209 substances) and from Asturiol et al. (2016) (269 substances). The dataset was split into a training (80%) and a test (20%) set. To partition the chemicals into training set and test set, assuring high diversity and keeping the ratio sensitizer/non-sensitizer, the following procedure was used. The chemicals were initially separated into sensitizer and non-sensitizer. The subsequent steps were carried out separately for sensitizer and non-sensitizer. Each group was clustered based on chemical similarity defined by the chemicals' fingerprints (RDKit atomic pairs) into as many clusters as the number of chemicals divided by 10. Subsequently, 80% of clustered chemicals were assigned randomly to the training set, using the assigned cluster as stratification variable. The remaining 20% of chemicals were assigned to the test set. The chemicals were structurally diverse, and the distribution between sensitizers and non-sensitizers was preserved. Structural diversity refers to diversity in terms of chemical classes. Details are described by Asturiol et al. (2016). The RDKit atom pairs fingerprints⁹ were used to cluster similar chemicals with a kNN algorithm. The chemicals were randomly selected from each cluster in the same proportions of sensitizers/non-sensitizers as in the dataset. The DT model was built using the Recursive PARTitioning (rpart) module included in R software¹⁰ (Therneau and Atkinson, 2015) to develop CART (Classification and Regression Trees) models. The model is based on 2D descriptors calculated using Dragon (Dragon v. 7.0.8, Kode srl¹¹) and a stepwise variable selection using linear discriminant analysis (LDA). A bootstrap technique (based on balanced resampling) was used for validation to select the best variables, with an in-house code implemented in R (Manganelli et al., 2019).

Molecular descriptors used in the DT are:

- nDB, number of double bonds
- IC1, information content index

- SIC2, structural information content index (neighborhood symmetry of 2-order)
- GATS8s, Geary autocorrelation of lag 8 weighted by I-state
- nRCHO, number of aldehydes (aliphatic)
- CATS2D_01_NL, CATS2D negative-lipophilic at lag 01
- F04[C-O], frequency of C - O at topological distance 4
- MLOGP, Moriguchi octanol-water partition coefficient (logP).

No observed adverse effect level (NOAEL)

The NOAEL model available in VEGA, which was built with the CORAL software using SMILES attributes, was used. Repeated dose 90-day oral toxicity study data in rodents was considered to build the model. Studies with a treatment duration of 28 days were also considered; this data was divided by a factor of 3 to approximate the 90-day value, as specified in the SCCS NoG (SCCS, 2018). The dataset consists of 140 organic compounds with experimental values collected from the US EPA's Integrated Risk Information System (IRIS) database, the Hazard Evaluation Support System (HESS) and Munro databases. The regression model is based on optimal descriptors calculated by the Monte Carlo method with SMILES attributes and the graph of atomic orbitals (Toropov et al., 2015).

2.3.3 Risk characterization

As the last step, SpheraCosmolife software carries out the risk characterization of chemicals or formulations following the process defined within the SCCS NoG (SCCS, 2018). The margin of safety (MoS) and the threshold of toxicological concern (TTC) are calculated.

Calculation of the margin of safety

The MoS is the ratio of the POD and SED. It is commonly used in human health risk assessment and, in particular, in cosmetics risk assessment, and it is a key value within the PIF, which is essential to be able to put the product on the European market. The software uses the NOAEL as a POD. The MoS is calculated according to Equation 7:

$$\text{MoS} = \frac{\text{POD (point of departure)}}{\text{SED (systemic exposure dose)}} \quad (\text{Eq. 7})$$

Within SpheraCosmolife, the POD can be an experimental, if available, or a predicted NOAEL. To consider a substance safe, the MoS must be higher than 100, as indicated in the SCCS NoG (SCCS, 2018).

Threshold of toxicological concern (TTC)

The software includes an assessment that considers the TTC approach, which refers to the establishment of a level of exposure below which there is no appreciable risk to human health (Kroes et al., 2004). It uses the Cramer decision tree, implemented in VE-

⁹ Landrum, G. (2016). RDKit: Open-source cheminformatics. <https://www.rdkit.org/>

¹⁰ R Core Team (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>

¹¹ Kode srl. (2017). Dragon (Software for Molecular Descriptor Calculation) Version 7.0.8. <https://chm.kode-solutions.net>



GA. The tool for TTC uses the algorithm implemented in Toxtree v. 3.1.0 (Patlewicz et al., 2008), with the original Cramer classes (Cramer et al., 1978; Munro et al., 1996; Patlewicz et al., 2008). The TTC assessment refers to the classes listed in Table S1¹².

3 Results

We developed SpheraCosmolife, a unified software system intended to support risk assessors of cosmetic products and to assist companies to ensure the safety of their products, avoiding ingredients that may be of concern at a given concentration. For this purpose, some new models were developed and implemented on the VEGA website⁸. The overall scheme follows the workflow of the assessor. The evaluation is easy to perform: First, the user is asked to provide information regarding each ingredient by inserting it using the SMILES, the INCI, or the CAS format, and adding its concentration and the product type. Then, the software checks whether the ingredient is listed in the Annexes of Regulation (EC) 1223/2009 (EC, 2009). Next, the software searches for values in its database that are useful for the chemical's risk characterization, as described in the Methods section. If there are no experimental values, the software predicts the properties of interest. The focus is on systemic toxicity, and the MoS is calculated. The user is assisted by the provision of a series of outcomes related to the product of interest. We describe the components of the new software and its outcomes below.

3.1 Performance of new skin sensitization model

The new decision tree for the prediction of skin sensitization potential as expressed by the LLNA had good results (Tab. 2): The results on the training set were quite balanced, while on the test set the model resulted in more false negatives than false positives. In contrast, the CAESAR model is over-conservative, so it is appropriate to look at the results of both models. To facilitate

Tab. 2: Performance of the decision tree model for skin sensitization

	Training set	Test set
No. of compounds	264	68
True positives (TP)	145	30
True negatives (TN)	66	18
False positives (FP)	18	4
False negatives (FN)	35	16
Accuracy	0.80	0.71
Specificity	0.79	0.82
Sensitivity	0.81	0.65
Matthews correlation coefficient (MCC)	0.57	0.44

the use of the results, a system has been developed to integrate the two models, following the approach for the integration of the results of the models for mutagenicity described above.

The results of the *in silico* model for skin sensitization were analyzed considering the true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). We then calculated the statistical parameters sensitivity, specificity, accuracy, and Matthews correlation coefficient (MCC) as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})}$$

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{\text{Total}}$$

$$\text{MCC} = \frac{(\text{TP} * \text{TN}) - (\text{FP} * \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}$$

The reported statistics refer to the performance on the training set. However, cross-validation (CV) was performed during model building using the recursive partitioning (rpart) module included in R software¹⁰ (Therneau and Atkinson, 2015).

3.2 The results of the other *in silico* models

SpheraCosmolife is a novel software, which includes a new integrated software system for skin sensitization and a new model for this endpoint, as described previously. Other *in silico* models have been recently developed (e.g., *in vitro* micronucleus genotoxicity test model), and they are individually available in VEGA. They are organized into a unified workflow here. The detailed information on each VEGA model is available within the VEGA platform⁸. Tables 3 and 4 summarize the performance of each model, in classification and regression, respectively.

3.3 User inputs for SpheraCosmolife

In the first form, the user must insert the product type, the ingredients of the product, and their concentrations. Figure 1 shows where to insert the information. The product type needs to be selected from a list of possible product types (e.g., body lotion, face cream, hand cream, shampoo, shower gel, deodorant, etc.), according to the product types in Tables 2A, 2B and 3 of SCCS NoG (SCCS, 2018), to define an exposure scenario. The ingredients can be inserted using the SMILES, the INCI, or the CAS format. If INCI or CAS are provided as the only input, the molecule can be processed only if a match is found in the database, otherwise the SMILES of the ingredient is needed. The user can insert a single ingredient or a set of ingredients corresponding to a specific formulation. The concentration of each ingredient is required. Then, the software searches for the ingredients in the internal database.

3.4 The SpheraCosmolife software summary results

The software provides a summary table of the results for all the ingredients in the product in html format. The results depend on the product type and on the concentrations of the ingredients en-

¹² doi:10.14573/altex.2010221s

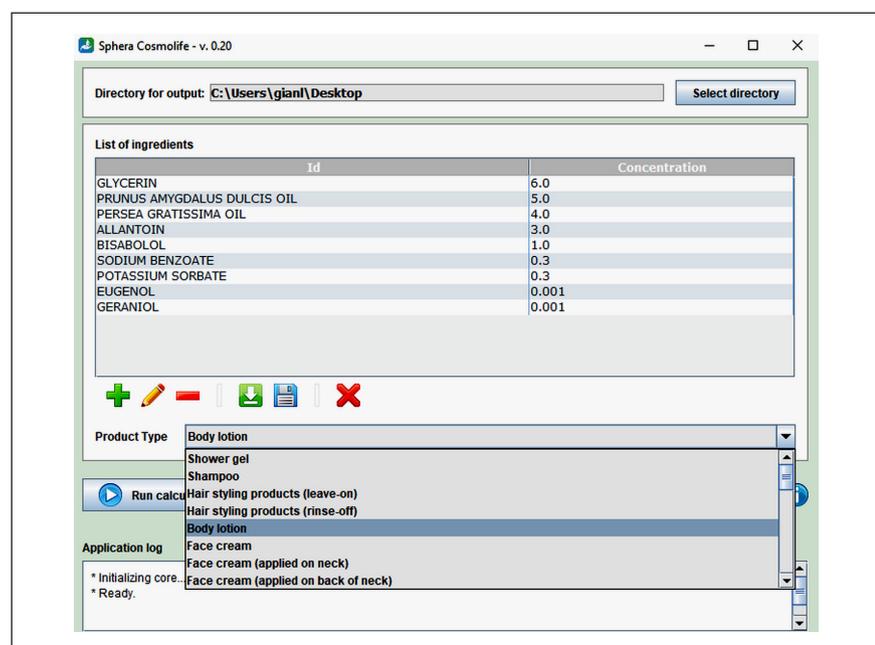
Tab. 3: Summary of the statistical parameters of the classification models implemented in SpheraCosmolife

Model		n	Accuracy	Specificity	Sensitivity
Skin sensitization model CAESAR	Training set	167	0.91	0.74	0.95
	Test set	42	0.93	0.75	0.97
Mutagenicity, bacterial reverse mutation test – Ames test*	Training set	NA	NA	NA	NA
	Validation set 1 (Cassano et al., 2014)	532	0.80	0.61	0.84
	Validation set 2 (Benfenati et al., 2018)	~2000	0.70	0.59	0.72
	Validation set 3 (Carneseccchi et al., 2020)	673	0.72	0.72	0.72
Chromosomal aberration	Training set	407	0.75	0.80	0.70
	Calibration set	35	0.94	0.95	0.94
	Validation set	35	0.94	1	0.87
<i>In vitro</i> micronucleus genotoxicity test	Training set	293	0.88	0.73	0.97
	Test set	87	0.83	0.63	0.94

*This is an integrated model that combines the predictions of four models (Cassano et al., 2014). The statistics of each individual model are reported in VEGA. This integrated model has been used to predict substances not present in the training set within three studies (see references), and the values are reported here.

Tab. 4: Summary of the statistical parameters of the regression models implemented in SpheraCosmolife

Model		n	R ²	RMSE
NOAEL	Training set	97	0.53	0.61
	Test set	16	0.73	0.49
	Validation test	27	0.60	0.43
Skin permeation logKp model (Potts and Guy, 1992)	Training set	271	0.51	0.79
Skin permeation logKp model (ten Berge, 2009)	Training set	271	0.58	0.72


Fig. 1: An example of the input for a cosmetic product that is to be assessed, with hypothetical ingredients



Processed product

Product type: **Body lotion**

Ingredients:

	Ingredient Id	CAS	INCI	Conc. %	Annex	Mutagenicity	Skin Sensitization	Dermal abs.	MoS	TTC
Details	GLYCERIN	56-81-5	GLYCERIN	8.0	-	NON-Mutagen (EXPERIMENTAL value)	NON-Sensitizer (EXPERIMENTAL value)	80%	32.73	0.03 mg/kg bw/day
Details	PRUNUS AMYGDALUS DULCIS OIL		PRUNUS AMYGDALUS DULCIS OIL	5.0	-	-	-	-	-	
Details	PERSEA GRATISSIMA OIL		PERSEA GRATISSIMA OIL	4.0	-	-	-	-	-	
Details	ALLANTOIN	97-59-6	ALLANTOIN	3.0	-	NON-Mutagen (moderate reliability)	Sensitizer (low reliability)	40%	268.64	0.0015 mg/kg bw/day
Details	BISABOLOL	23089-26-1	BISABOLOL	1.0	-	NON-Mutagen (good reliability)	Sensitizer (moderate reliability)	40%	133.5	0.0015 mg/kg bw/day
Details	SODIUM BENZOATE	532-32-1	SODIUM BENZOATE	0.3	V	NON-Mutagen (EXPERIMENTAL value)	NON-Sensitizer (EXPERIMENTAL value)	80%	93.55	0.03 mg/kg bw/day
Details	POTASSIUM SORBATE	24634-61-5	POTASSIUM SORBATE	0.3	V	NON-Mutagen (EXPERIMENTAL value)	Sensitizer (good reliability)	40%	3382.03	0.03 mg/kg bw/day
Details	EUGENOL	97-53-0	EUGENOL	0.001	III	NON-Mutagen (EXPERIMENTAL value)	Sensitizer (EXPERIMENTAL value)	80%	233360.39	0.03 mg/kg bw/day
Details	GERANIOL	106-24-1	GERANIOL	0.001	III	NON-Mutagen (EXPERIMENTAL value)	Sensitizer (EXPERIMENTAL value)	40%	158441.56	0.03 mg/kg bw/day

Fig. 2: An example of the output table, with the summary of the hazard and exposure features of hypothetical ingredients of a cosmetic product

tered by the user. Figure 2 gives an example of the output table, with the summary of the hazard and exposure features of the ingredients indicated in the input file. It shows whether an ingredient is present in an Annex of the cosmetics regulation (EC, 2009), whether it is mutagenic (Ames test), whether it is a sensitizer, the dermal absorption according to the Kroes approach, the MoS, and the TTC. No assessment is provided for substances for which no SMILES is retrieved, e.g., for inorganic compounds and mixtures (unless present in the database). All the assessments refer to the neutralized structure of the compounds as they are neutralized during the process.

3.5 The SpheraCosmolife software detailed results

Clicking on the “Details” box for each ingredient in the first column of the summary table reveals the detailed results for each ingredient as an html page. First, the information on the substance is presented, i.e., the SMILES, the CAS and INCI, the information on the product type selected, and the concentration of the ingredient, as a percentage (%) and mg/g. Then, SpheraCosmolife reports the values from the SCCS NoG (SCCS, 2018) related to the product type selected by the user, i.e., the relative daily exposure, the surface area involved, the type of exposure, and the exposure time. These values are specific for each product type and serve to calculate the exposure, as described in the next section.

The presence of a substance in one of the Annexes does not preclude the evaluation. However, the substances in Annex II cannot be used in cosmetic products. This, for example, is the

case for formaldehyde, and the software provides a warning by showing a red cell if an Annex II substance is included as an ingredient. In other cases, the Annexes give a threshold above which the substances cannot be used. For instance, phenoxyethanol, which is in Annex V, may be used at a maximum concentration of 1%. The user must enter a concentration that is lower than 1% for the product to be safe and in compliance with the law.

Moreover, in this section the software checks whether the ingredient is classified according to the Classification, Labelling and Packaging (CLP) regulation (EC, 2008) and whether it is contained in the Safer Chemical Ingredients List (SCIL). This is a list of chemicals, arranged by functional-use class, that the Safer Choice Program has determined are safer than traditional chemical ingredients¹³.

Figure S1¹² gives an example of the described information.

3.5.1 Information on exposure

SpheraCosmolife provides information regarding the exposure to each ingredient, based on the inputs of the user (product type and concentration). Figure S2¹² shows an example of the exposure information provided by the software.

The exposure is calculated for different scenarios. External exposure is obtained using the parameters defined by the SCCS NoG for the product type, as described in the Methods section. The SED is obtained for three scenarios: (1) absorption of 100%, oral or inhalation exposure; (2) absorption of 50%, a default value for dermal exposure, as indicated by SCCS NoG (SCCS,

¹³ <https://www.epa.gov/saferchoice/safer-ingredients>

2018); or (3) the more realistic scenario for dermal absorption based on the models for skin permeation. SpheraCosmolife provides the output of two models for skin permeation, and then chooses the worst case, using the most conservative of these two values. As explained above, for these models and all the others, whenever the database contains an experimental value for the ingredient of interest, the software shows and uses it. Water solubility and the J_{\max} are also shown.

Since all these models are implemented in VEGA, the output of the predictions considers the reliability of the prediction, measured using the applicability domain index (ADI) calculated by VEGA. ADI is a value that evaluates the reliability of the prediction by considering several parameters, such as the predictions done on similar compounds, the agreement between the predicted value for the target compound and the experimental values of the most similar compounds, the presence of unusual fragments, how similar the related substances are, etc. Full details for each model are given on the VEGA website⁸, including the user guidance¹⁴. Thus, for each prediction, SpheraCosmolife reports the level of reliability as low, medium or good. This indication should be used as a warning, and the assessor should evaluate whether the value can be used or not based on the level of reliability; in case the assessor needs more information to take the decision, they should use the original VEGA model of interest, which provides full details, while the SpheraCosmolife report only provides a summary.

3.5.2 Information on hazard values and TTC

The software reports values (experimental, if available, or predicted) related to mutagenicity (Ames integrated model), genotoxicity (*in vitro* micronucleus and chromosomal aberration), skin sensitization (CAESAR, DT model and the integrated model), and NOAEL. Figure S3¹² gives an example of the information on hazard assessment. All the values provided are highlighted with a color to assist the user. The cell with the prediction is red if the compound is predicted as mutagenic/genotoxic/sensitizer, or green if it is safe, i.e., if the prediction has good reliability or if an experimental value is available. The color is orange and yellow, respectively, if the reliability of the predictions is only moderate or low.

The software carries out a risk characterization, considering the systemic toxicity and the MoS.

The software also includes an assessment considering the TTC approach. In this case, the external and internal exposure values are compared with the TTC threshold for the specific ingredient in order to incorporate dermal bioavailability into the use of TTC for cosmetics. A decision tree is applied (Williams et al., 2016). The output shows a red or green cell, depending on whether external and internal exposure values are above or below the TTC threshold, using the Cramer decision tree implemented in VEGA. Figure 3 shows an example of the output for the TTC value for the three different SED scenarios. The TTC value is of interest for substances that are impurities. In general, the evaluation of the other endpoints, if reliable, should be considered more relevant.

7. TTC

Cramer class (ToxTree/Vega model prediction)	Low (Class I)
TTC according to Cramer classification	0.03 mg/kg bw/day

TTC vs External / Internal exposure	
TTC	0.03 mg/kg bw/day
External exp.	0.0012 mg/kg bw/d
Internal exp. (SED 100%)	0.0012 mg/kg bw/d
Internal exp. (SED 50%)	0.0006 mg/kg bw/d
Internal exp. (SED 80%)	0.001 mg/kg bw/d

Fig. 3: The output of the SpheraCosmolife software for the TTC for a given ingredient (Eugenol in a body lotion scenario)

3.5.3 Examples: comparison of the procedures and application of the software to a real case

As shown in the previous sections, SpheraCosmolife represents an innovative and robust system to guide the assessor during their evaluation analysis. The tool can be used to evaluate either a single ingredient or multiple ingredients typically used in a cosmetic product. It is important to note that the tool can process a great number of substances in a few seconds, saving time and money.

A first example is the evaluation of a hypothetical cosmetic product (body lotion) with a list of ingredients as described in Figure 2. Table 5 shows how the formulation is evaluated with or without the tool. Figure 4 reproduces the list of ingredients shown in Figure 2.

Another use of the tool is for the evaluation of impurities. Article 17 of Regulation (EC) 1223/2009 clearly explains that the non-intended presence of a small quantity of a prohibited substance (impurity) that is technically unavoidable in good manufacturing practice, shall be permitted, providing that such presence is in conformity with safety requirements. As an indicative real-case example, we consider the impurities detected during analytical checks of a volatile organic compound (VOC) in a nail product. Figure 5 shows the analytical results for a nail polish. The VOC analysis detected methanol (CAS 67-56-1), styrene (CAS 100-42-5), toluene (CAS 108-88-3) and dichloromethane (CAS 75-09-2).

Looking at Figure 5, the regulatory context is immediately clear. Methanol and toluene are substances included in Annex III (list of restricted ingredients). Therefore, even if they are not intentionally added, concern related to their safety is limited, since all the substances included in this list have already been evaluated by SCCS. Toluene can be used in nail polish up to 25% (Fig. 6), while methanol is allowed as an impurity of ethanol or isopropyl alcohol up to 5% (Fig. 7).

¹⁴ <https://www.vegahub.eu/download/vega-interpretation/>



Tab. 5: Differences between manual and automated evaluation

Steps	Manual procedure	SpheraCosmolife	Comparison consideration
Input information	Definition of: - product type - ingredients - concentration	The user inserts this information. The product type is selected from a picklist. The software automatically retrieves the ingredients from the database. If not present in the database, input structure is based on SMILES strings.	Same conceptual scheme. Same equations and rules to define exposure scenarios. Process is facilitated and much faster. Database is checked automatically, not manually.
Output assessment	The assessment is based on <i>in vitro</i> tests and data retrieved from the literature; <i>in silico</i> inputs are usually not used, for lack of experience. Assessment is time consuming. Reliability and robustness of the data are hard to assess. Regulatory information needs to be manually retrieved. This may result in an incomplete safety report, for lack of data.	The output is obtained within seconds. Valid and consolidated <i>in silico</i> data are available for all endpoints. Reliability of the results is presented using a specific algorithm. The user can visualize data for exposure, hazard, MoS and TTC in the same table. Complete and exhaustive regulatory data (presence in the Annexes, maximum threshold, etc.) are automatically retrieved from the internal database.	The safety assessment process is facilitated. Additional information is available from <i>in silico</i> models, with its related uncertainty.
Regulatory information	The assessor must browse through the regulation to retrieve information.	All detailed information on the cosmetic regulation is retrieved from the internal database.	A complete overview of the regulatory conditions of the different ingredients is offered, and information from US EPA is also provided.
Hazard and exposure assessment	Data retrieval is time-consuming. Data retrieved from the literature needs to be curated and correctly interpreted.	The process is much faster. Experimental data have been curated and the predictions report reliability to facilitate interpretation of the results.	Models for hazard and exposure are integrated within the same platform. Risk assessment is done immediately and includes information commonly not addressed (internal exposure, Kroes approach, etc.), which improves the safety assessment process. Some experience in the interpretation of the VEGA results is needed.

List of ingredients	
Id	Concentration
GLYCERIN	6.0
PRUNUS AMYGDALUS DULCIS OIL	5.0
PERSEA GRATISSIMA OIL	4.0
ALLANTOIN	3.0
BISABOLOL	1.0
SODIUM BENZOATE	0.3
POTASSIUM SORBATE	0.3
EUGENOL	0.001
GERANIOL	0.001

Fig. 4: List of a hypothetical cosmetic product

Styrene and dichloromethane are prohibited as cosmetic ingredients (included in Annex II), and are considered substances of concern. However, it is possible to check the specific exposure of the users and define their risk in use. In this case, concentrations are low enough (0.002% for styrene and 0.0003% for dichloromethane) to consider them as “traces”, and thus an evaluation using the TTC approach is possible.

Figure 5 shows that, according to the Cramer decision tree, styrene is in class I (low level of concern; TTC = 0.03 mg/kg bw/day), while dichloromethane is in class III (high level of concern; TTC = 0.0015 mg/kg bw/day). Comparing these TTC thresholds with the exposure calculated by SpheraCosmolife (see Fig. 8, 9), an indication of the risk associated with the presence of an unexpected substance is shown, helping the safety assessor during the evaluation process.



Processed product

Product type: **Nail polish**

Ingredients:

	Ingredient Id	CAS	INCI	Conc. %	Annex	Mutagenicity	Skin Sensitization	Dermal abs.	MoS	TTC
Details	methanol	67-56-1	METHYL ALCOHOL	0.2	III	NON-Mutagen (EXPERIMENTAL value)	NON-Sensitizer (good reliability)	80%	62500	0.03 mg/kg bw/day
Details	styrene	100-42-5	styrene	0.002	II	NON-Mutagen (EXPERIMENTAL value)	Sensitizer (moderate reliability)	80%	69500	0.03 mg/kg bw/day
Details	toluene	108-88-3	toluene	6.0E-4	III	NON-Mutagen (EXPERIMENTAL value)	Sensitizer (moderate reliability)	80%	290416.67	0.03 mg/kg bw/day
Details	DICHLOROMETHANE	75-09-2	DICHLOROMETHANE	3.0E-4	II	Mutagen (EXPERIMENTAL value)	NON-Sensitizer (low reliability)	80%	4681666.67	0.0015 mg/kg bw/day

Fig. 5: Analytical results (VOC) for a nail polish

Molecules found in the annex lists: 1

Found molecule no. 1	
CAS	108-88-3
INCI	TOLUENE
Found in annex	ANNEX III: LIST OF SUBSTANCES WHICH COSMETIC PRODUCTS MUST NOT CONTAIN EXCEPT SUBJECT TO THE RESTRICTIONS LAID DOWN
Annex details	ANNEX III: Product type, body part: Nail products ANNEX III: MAX Concentration in ready for use preparation: 25% ANNEX III: Wording of conditions of use and warnings: Keep out of reach of children To be used by adults only ANNEX III: NOTES: Reprotoxic 2 >1029/06 - Opinion on Toluene (its use as a solvent in nail cosmetics) >1170/08 - Opinion on Toluene (its use as a solvent in nail cosmetics)
Found Safer Chemicals classification	No
Found in CLP list	Yes
CLP details	CLP HARMONIZED CLASSIFICATION (Link): https://echa.europa.eu/it/information-on-chemicals/cl-inventory-database/-/discli/details/30426 HAZARD CLASS AND CATEGORY CODE (CLASSIFICATION): Flam. Liq. 2 Repr. 2 Asp. Tox. 1 STOT SE 3 STOT RE 2 * Skin Irrit. 2 HAZARD STATEMENT CODE (CLASSIFICATION): H225 H361d *** H304 H336 H373 ** H315 PICTOGRAM, SIGNAL WORD CODE (LABELLING): GHS02 GHS08 GHS07 Dgr HAZARD STATEMENT CODE (LABELLING): H225 H361d *** H304 H373 ** H315 H336

Fig. 6: Regulatory information provided by SpheraCosmolife for toluene

A last example of the use of the tool is the evaluation of a botanic derivative ingredient. Botanical extracts are commonly used in the cosmetic industry because they are relatively easy to obtain, they are generally considered safe (as they are of natural origin), and they are used commercially for marketing descriptions of cosmetic products. SpheraCosmolife allows a rapid evaluation of the botanic ingredient, as shown in the example described in Figure 10. The result is a safety profile of the botanic derivative ingredient that can be taken into consideration by the safety assessor to define the safety of the cosmetic product in which the botanic ingredient is used.

4 Discussion

4.1 The novelty of the tool

In silico models offer a unique opportunity for the assessment of cosmetic products, as they can process a very large number of ingredients quickly, without the need for new, additional tests, saving money and time (Raitano et al., 2019; Gellatly and Sewell, 2019; Taylor and Rego Alvarez, 2020). The assessment of cosmetic ingredients can benefit from the hundreds of *in silico* models available that predict the properties of interest for many endpoints. Some of these models are freely available, others are



Molecules found in the annex lists: 1	
Found molecule no. 1	
CAS	67-56-1
INCI	METHYL ALCOHOL
Found in annex	ANNEX III: LIST OF SUBSTANCES WHICH COSMETIC PRODUCTS MUST NOT CONTAIN EXCEPT SUBJECT TO THE RESTRICTIONS LAID DOWN
Annex details	ANNEX III: Product type, body part: Denaturant for ethanol and isopropyl alcohol ANNEX III: MAX Concentration in ready for use preparation: 5 % (as a % ethanol and isopropyl alcohol)
Found Safer Chemicals classification	No
Found in CLP list	Yes
CLP details	CLP HARMONIZED CLASSIFICATION (Link): https://echa.europa.eu/it/information-on-chemicals/cl-inventory-database/-/discli/details/37212 HAZARD CLASS AND CATEGORY CODE (CLASSIFICATION): Flam. Liq. 2 Acute Tox. 3 * Acute Tox. 3 * Acute Tox. 3 * STOT SE 1 HAZARD STATEMENT CODE (CLASSIFICATION): H225 H331 H311 H301 H370 ** PICTOGRAM, SIGNAL WORD CODE (LABELLING): GHS02 GHS06 GHS08 Dgr HAZARD STATEMENT CODE (LABELLING): H225 H331 H311 H301 H370 **

Fig. 7: Regulatory information provided by SpheraCosmolife for methanol

7. TTC	
Cramer class (ToxTree/Vega model prediction)	Low (Class I)
TTC according to Cramer classification	0.03 mg/kg bw/day
TTC vs External / Internal exposure	
TTC	0.03 mg/kg bw/day
External exp.	0.0001 mg/kg bw/d
Internal exp. (SED 100%)	0.0001 mg/kg bw/d
Internal exp. (SED 50%)	0.00005 mg/kg bw/d
Internal exp. (SED 80%)	0.00008 mg/kg bw/d

Fig. 8: TTC evaluation for styrene

7. TTC	
Cramer class (ToxTree/Vega model prediction)	High (Class III)
TTC according to Cramer classification	0.0015 mg/kg bw/day
TTC vs External / Internal exposure	
TTC	0.0015 mg/kg bw/day
External exp.	0.000015 mg/kg bw/d
Internal exp. (SED 100%)	0.000015 mg/kg bw/d
Internal exp. (SED 50%)	0.000007 mg/kg bw/d
Internal exp. (SED 80%)	0.000012 mg/kg bw/d

Fig. 9: TTC evaluation for dichloromethane

Processed product							
Product type: Face cream							
Ingredients:							
	Ingredient Id	CAS	INCI	Annex	Mutagenicity	Skin Sensitization	Dermal abs.
Details	CI 75100	27876-94-4	CI 75100	IV	Mutagen (moderate reliability)	Sensitizer (low reliability)	40%
Details	Ingredient n.2			-	NON-Mutagen (EXPERIMENTAL value)	Sensitizer (low reliability)	10%
Details	Ingredient n.3			-	Mutagen (moderate reliability)	NON-Sensitizer (low reliability)	10%
Details	Ingredient n.4	116-26-7	2,3-DIHYDRO-2,2,6-TRIMETHYLBENZALDEHYDE	-	NON-Mutagen (good reliability)	Sensitizer (EXPERIMENTAL value)	40%
safranal							

Fig. 10: Saffron (*Crocus sativus* stigma extract) components



commercially available, or require a fee. For instance, models to predict the Ames test are the most numerous, and they have been reviewed elsewhere (Cassano et al., 2014; Honma et al., 2019). Despite the number of models available, their application can be complicated, as the user must run separate models on different platforms, which may require different input formats, specific instructions, or provide outputs that may be difficult to integrate and compare. Furthermore, independent models for hazard and exposure assessment need to be run for risk assessment. All these difficulties represent a barrier to the use of *in silico* models for cosmetics.

Within the EC project VERMEER, we aimed to integrate models for hazard and exposure. The SpheraCosmolife software system presented here is one of the tools we are developing to increase the use of *in silico* tools for risk assessment, at the same time improving the robustness and the reliability of the results.

A major innovative aspect introduced by SpheraCosmolife is that the system does not provide the output for one single endpoint, as in the usual situation; instead, a battery of models that are specific for the application are wrapped into a single system. The user is not required to learn multiple programs, because the input is the same for all the models, which run automatically in the background.

A second innovative aspect is that SpheraCosmolife integrates models for hazard and exposure within the same platform. Historically, models for exposure have been developed separately, typically within commercial platforms, while platforms for hazard models do not contain tools for exposure. SpheraCosmolife offers a novel approach because it integrates both.

A third innovative aspect is that this system is dedicated to the specific sector of cosmetics. Thus, it contains legislative thresholds and specific references to the European regulation, proceeding in the direction of a practical application. Usually, the existing platforms are generic and intended to be used for multiple purposes.

A fourth innovative aspect of this system is that it can be used to assess products, not only ingredients, and thus it simultaneously addresses all ingredients contained in the cosmetic product. This also helps the user who is interested in the practical case of a product in which multiple substances are used.

The safety evaluation is performed easily: The user provides the structure of the chemicals of interest, the type of product in which the chemicals will be used as ingredients, and the concentrations in the final formulations. Then, SpheraCosmolife automatically executes several analyses and predictions, providing an overall evaluation of the formulation in a structured output report.

SpheraCosmolife is opening up new avenues to the use of *in silico* models, moving towards real case application, facing the practical problems of a sector urgently demanding solutions to assess products, and introducing novel topics to take into account the specific needs of a focused sector. The effort to apply models to practical cases will provide solutions, lowering the barriers to the use of *in silico* models, which tend to be viewed more as theoretical tools. Too often the developers oblige the user to learn programs that only partially assist them. Instead, a dialogue must be established between the users and the developers so that the *in*

in silico models can be designed to fulfill the users' requirements. This has been accomplished within the VERMEER project. Development of the overall architecture should start at the very initial phases. Here, we codified the different endpoints used by an assessor of cosmetic products and analyzed how the assessment is done without the *in silico* model. Then, we tried to replicate these steps using different software modules, of which some already existed and others were developed for the required purpose. This process is close to the development of an expert system software, and we inserted a number of tools that are statistical-based, not only expert-based rules.

4.2 The limitations of the tool

The novel system should be considered as support for the assessor, and not as a replacement of the human evaluation. The major novelty is in the organized series of *in silico* models. In case experimental values are available in the database, the evaluation is straightforward. Otherwise, the missing values are predicted. However, particularly in this case, the user should consider the reliability of the values, indicated by the system as the uncertainty, in reaching a final assessment. This demonstrates the objectivity of the assessment, which systematically shows the different levels of uncertainty of the experimental and calculated evidence. The assessor should use expert knowledge, particularly when the uncertainty is high. Since SpheraCosmolife is supported by the VEGA platform, the ADI tool to evaluate the reliability of the prediction can be used, and further information can be obtained using the individual VEGA models, which can be downloaded from the VEGA website. These additionally show similar compounds, which can be used for read-across, and provide more details on the ADI.

The ADI is a quantitative value based on three fundamental components at the basis of all *in silico* models: the chemical information, the toxicological/property information, and the algorithm. Each of these components is used within the ADI. VEGA evaluates similar chemicals and the uncommon features from a chemical point of view. The ADI also investigates the toxicological profiles and features of the related substances (thus not only based on the chemical similarity) using, for instance, toxicological alerts. Finally, the uncertainty of the algorithm is addressed too. Combining all these analyses regarding the applicability domain, VEGA provides an assessment of the reliability of the predictions, which is reported by SpheraCosmolife.

The user should be aware that the uncertainty is higher for some of the endpoints, and Tables 2 and 3 provide a first indication. An example of an endpoint with higher uncertainty is the predictive tool for NOAEL, which is a difficult endpoint because it is affected by natural variability and by the choice of the doses within the experimental test. The current version of the model is based on a limited training set. We are working to reduce uncertainty, and new models for NOAEL will be added in the future. Skin sensitization is another endpoint with uncertainty. Both models for this endpoint implemented in the system are quite conservative, and therefore result in some false positives. Also in this case, we are developing new models, which will be implemented in a new version. As stated above, in case of un-



certainty the user can run the individual models within VEGA and look at the similar compounds, taking advantage of the read-across approach. In the future, read-across will be implemented into SpheraCosmolife, and further endpoints will be added.

5 Conclusions

We introduced for the first time a single software system, SpheraCosmolife, to facilitate and harmonize the safety assessment of cosmetic products. The tool is specific for cosmetic products and refers to the respective thresholds and requirements. The software system provides the MoS, based on the systemic exposure dose, and includes a number of models for both exposure and hazard prediction. Different scenarios are provided based on the use or not of the results of the skin permeation models. The software aims to be user-friendly, requiring a limited number of inputs from the user, and uses its internal database and models to provide an evaluation even when experimental values are lacking. The system was developed based on user requirements rather than aiming to offer a set of prebuilt tools which could be useful. The level of uncertainty of the results is given, and the assessor is provided with supporting information for the final evaluation. SpheraCosmolife has been designed with flexible capabilities for future extensions in mind, and more features will be added in future versions. New functionalities will be added to accommodate user requests. The frequency of the updates depends on the progress of the work necessary to create new models and evaluate the consistency of new toxicological endpoints, but a version 2.0 is planned for the end of 2021.

References

- Asturiol, D., Casati, S. and Worth, A. (2016). Consensus of classification trees for skin sensitisation hazard prediction. *Toxicol In Vitro* 36, 197-209. doi:10.1016/j.tiv.2016.07.014
- Baderna, D., Gadaleta, D., Lostaglio, E. et al. (2020). New in silico models to predict in vitro micronucleus induction as marker of genotoxicity. *J Hazard Mater* 385, 121638. doi:10.1016/j.jhazmat.2019.121638
- Benfenati, E., Golbamaki, A., Raitano, G. et al. (2018). A large comparison of integrated SAR/QSAR models of the Ames test for mutagenicity. *SAR QSAR Environ Res* 29, 591-611. doi:10.1080/1062936X.2018.1497702
- Berggren, E., White, A., Ouedraogo, G. et al. (2017). Ab initio chemical safety assessment: A workflow based on exposure considerations and non-animal methods. *Comput Toxicol* 4, 31-44. doi:10.1016/j.comtox.2017.10.001
- Carneseccchi, E., Raitano, G., Gamba, A. et al. (2020). Evaluation of non-commercial models for genotoxicity and carcinogenicity in the assessment of EFSA's databases. *SAR QSAR Environ Res* 31, 33-48. doi:10.1080/1062936X.2019.1690045
- Cassano, A., Raitano, G., Mombelli, E. et al. (2014). Evaluation of QSAR models for the prediction of Ames genotoxicity: A retrospective exercise on the chemical substances registered under the EU REACH regulation. *J Environ Sci Health C* *Environ Carcinog Ecotoxicol Rev* 32, 273-298. doi:10.1080/10590501.2014.938955
- Chaudhry, Q., Piclin, N., Cotterill, J. et al. (2010). Global QSAR models of skin sensitizers for regulatory purposes. *Chem Cent J* 4, Suppl 1, S5. doi:10.1186/1752-153X-4-S1-S5
- 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. doi:10.1016/s0015-6264(76)80522-6
- EC – European Commission (2008). Regulation (EC) No. 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. *OJ L* 353, 31/12/2008, p. 1-1355. <http://data.europa.eu/eli/reg/2008/1272/oj>
- EC (2009). Regulation (EC) No. 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products. *OJ L* 342, 59-209. <http://data.europa.eu/eli/reg/2009/1223/2020-05-01>
- Gadaleta, D., Lombardo, A., Toma, C. et al. (2018). A new semi-automated workflow for chemical data retrieval and quality checking for modeling applications. *J Cheminformatics* 10, 60. doi:10.1186/s13321-018-0315-6
- Gellatly, N. and Sewell, F. (2019). Regulatory acceptance of in silico approaches for the safety assessment of cosmetic-related substances. *Comput Toxicol* 11, 82-89. doi:10.1016/j.comtox.2019.03.003
- Honma, M., Kitazawa, A., Cayley, A. et al. (2019). Improvement of quantitative structure-activity relationship (QSAR) tools for predicting Ames mutagenicity: Outcomes of the Ames/QSAR international challenge project. *Mutagenesis* 34, 3-16. doi:10.1093/mutage/gy031
- ICH (2017). ICH guideline M7 (R1) on assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk. EMA/CHMP/ICH/83812/2013. <https://bit.ly/3DoCgxw>
- Kroes, R., Renwick, A. G., Cheeseman, M. et al. (2004). StructuRe re-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food Chem Toxicol* 42, 65-83. doi:10.1016/j.fct.2003.08.006
- Kroes, R., Renwick, A. G., Feron, V. et al. (2007). Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food Chem Toxicol* 45, 2533-2562. doi:10.1016/j.fct.2007.06.021
- Manganelli, S., Schilter, B., Benfenati, E. et al. (2018). Integrated strategy for mutagenicity prediction applied to food contact chemicals. *ALTEX* 35, 169-178. doi:10.14573/altex.1707171
- Manganelli, S., Roncaglioni, A., Mansouri, K. et al. (2019). Development, validation and integration of in silico models to identify androgen active chemicals. *Chemosphere* 220, 204-215. doi:10.1016/j.chemosphere.2018.12.131
- Munro, I. C., Ford, R. A., Kennepohl, E. et al. (1996). Correlation of structural class with no-observed-effect levels: A proposal for establishing a threshold of concern. *Food Chem Toxicol* 34, 829-867. doi:10.1016/s0278-6915(96)00049-x
- Patlewicz, G., Jeliakova, N., Safford, R. J. et al. (2008). An

- evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR QSAR Environ Res* 19, 495-524. doi:10.1080/10629360802083871
- Potts, R. O. and Guy, R. H. (1992). Predicting skin permeability. *Pharm Res* 9, 663-669. doi:10.1023/a:1015810312465
- Raitano, G., Roncaglioni, A., Manganaro, A. et al. (2019). Integrating in silico models for the prediction of mutagenicity (Ames test) of botanical ingredients of cosmetics. *Comput Toxicol* 12, 100108. doi:10.1016/j.comtox.2019.100108
- Rogiers, V., Benfenati, E., Bernauer, U. et al. (2020). The way forward for assessing the human health safety of cosmetics in the EU-Workshop proceedings. *Toxicology* 436, 152421. doi:10.1016/j.tox.2020.152421
- SCCS – Scientific Committee on Consumer Safety (2015). The SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation. 9th revision. Revised version of 25 April 2016, SCCS/1564/15. https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_190.pdf
- SCCS (2018). The SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation. 10th revision. 24-25 October 2018, SCCS/1602/18. https://ec.europa.eu/health/sites/health/files/scientific_committees/consumer_safety/docs/sccs_o_224.pdf
- Shen, J., Kromidas, L., Schultz, T. et al. (2014). An in silico skin absorption model for fragrance materials. *Food Chem Toxicol* 74, 164-176. doi:10.1016/j.fct.2014.09.015
- Taylor, K. and Rego Alvarez, L. (2020). Regulatory drivers in the last 20 years towards the use of in silico techniques as replacements to animal testing for cosmetic-related substances. *Comput Toxicol* 13, 100112. doi:10.1016/j.comtox.2019.100112
- ten Berge, W. (2009). A simple dermal absorption model: Derivation and application. *Chemosphere* 75, 1440-1445. doi:10.1016/j.chemosphere.2009.02.043
- Therneau, T. M. and Atkinson, E. J. (2015). An Introduction to Recursive Partitioning Using the RPART Routines. Mayo Foundation, Rochester. <https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>
- Toropov, A. A., Toropova, A. P., Pizzo, F. et al. (2015). CORAL: Model for no observed adverse effect level (NOAEL). *Mol Divers* 19, 563-575. doi:10.1007/s11030-015-9587-1
- Toropov, A. A., Toropova, A. P., Raitano, G. et al. (2019). CORAL: Building up QSAR models for the chromosome aberration test. *Saudi J Biol Sci* 26, 1101-1106. doi:10.1016/j.sjbs.2018.05.013
- Vecchia, B. E. and Bunge A. L. (2002). Skin absorption databases and predictive equations. R. H. Guy and J. Hadgraft (eds.), *Transdermal Drug Delivery*. Chapter 3. CRC Press. <https://bit.ly/2V0yWHy>
- Williams, F. M., Rothe, H., Barrett, G. et al. (2016). Assessing the safety of cosmetic chemicals: Consideration of a flux decision tree to predict dermally delivered systemic dose for comparison with oral TTC (threshold of toxicological concern). *Regul Toxicol Pharmacol* 76, 174-186. doi:10.1016/j.yrtph.2016.01.005
- Worth, A., Cronin, M., Enoch, S. et al. (2012). Applicability of the Threshold of Toxicological Concern (TTC) approach to cosmetics – Preliminary analysis. European Union. doi:10.2788/5059

Conflicts of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

This work was developed under the framework of the LIFE+ Projects VERMEER [LIFE16/ENV/IT/000167] and CONCERT REACH [LIFE17/GIE/IT/000461].