

In vitro acute skin irritancy of chemicals using the validated EPISKIN model in a tiered strategy Results and performances with 184 cosmetic ingredients

José Cotovio¹, Marie-Hélène Grandidier¹, Damien Lelièvre¹, Roland Roguet¹,
Estelle Tinois-Tessonnaud² and Jacques Leclaire¹

¹L'Oréal Resaerch, Life Sciences, ²SkinEthic Laboratories

Corresponding author: José Cotovio

L'Oréal Research

1 Avenue Eugène Schueller 93 600 Aulnay Sous Bois, France

jcotovio@rd.loreal.com

Abstract

Following considerable interest in the development of in vitro model substitutes, reconstructed human skin or epidermis available as standardized kits, allow the measurement of parameters linked to the safety of topically applied products. Able to mimic in vivo situation, the EPISKIN™ 3-D model (SkinEthic Laboratories) was involved in the formal ECVAM sponsored acute skin irritation validation. EPISKIN™ 0.38 cm² was validated as a full replacement method by the ESAC (April 2007 statement).

The protocol was based on a short treatment time (15 min) followed by an extended 42 hour post-treatment incubation period. The prediction model mainly based on the validated viability measurement (MTT) combined with the useful adjunct IL-1 α and appropriate cut-off values allowed the drawing up of 2 chemical classes: Irritants (EU risk, R38) and non-irritants (EU risk, no classification). Applied to a set of 184 raw ingredients covering diverse physical-chemical categories, sensitivity, specificity and accuracy of the prediction model were 85%, 86% and 86% respectively. Results demonstrated the usefulness of the tiered strategy combining two complementary end points MTT + IL-1 α as a decision-making tool for skin irritancy hazard identification. The present work is part of a strategy assessing performances on a large number of chemical realistically relevant to industrial needs in order to concretely determine applicability domains while understanding the remaining gaps.

Keywords: skin irritation, EPISKIN, reconstructed epidermis, alternative method, validation

Abbreviations

ECVAM: European Centre for the Validation of Alternative Methods; ESAC: ECVAM scientific advisory committee; EU: European Union; GHS: Globally Harmonized System of classification; I: irritant; IL-1 α : interleukin 1 α ; MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NI: non irritant; PM: prediction model; PII: primary irritation index; SIFT, skin integrity function test; UN, United Nations; OECD: Organization for Economic Co-operation and Development; MSDS: material safety data sheet; SOP: standard operating procedure.

Introduction

Acute skin irritation is an important consideration when establishing procedures for safe handling, packing and transport of chemicals in the EU and

UN. The identification of new chemicals potential to induce skin irritation is of first importance. The assessment of acute skin irritation is usually performed on rabbit skin (OECD guideline 404, 2002. Draize, 1944) and involves determination of clinical signs as erythema, and oedema after single application on skin. Regarding potential misclassification (York, 1996) and ethical considerations, animal data could be questionable and are subjected to regulatory issues (EU, 2003). Consequently, alternatives to these animal test procedures have been developed in different fields (Botham, 1998; Liebsch, 2002; Zuang, 2005). Acute irritation is characterized by local and reversible non immunological inflammatory response of living skin following injury or single contact with an irritant. Great improvements were achieved in the bioengineering of 3-D skin models. Due to both their industrialization and easy availability,

and because systemic reactions play a minor role in skin inflammatory processes, these complex tissue systems able to express cells reactivity and to mimic barrier function were largely evaluated (Faller, 2002; Perkins, 1999; Roguet, 1998, 1999).

Due to the increasing need to identify non animal tests for predicting of human skin irritation, the European Centre for the Validation of Alternative Methods (ECVAM) focused on the evaluation of suitable *in vitro* assays (Botham, 1998). The final ECVAM sponsored formal validation study on acute skin irritation focused on two *in vitro* test systems (EPISKINTM, EpiDermTM) and one *ex vivo* test (SIFT) and involved a set of 58 chemicals. The main goal was to define tests and prediction models able to correctly predict *in vivo* classifications according to EU rules (R38 and No label) with good associated performances. The ECVAM advisory committee endorsed in April 2007 the validity of the EPISKIN test as a replacement for the rabbit skin irritation method (ESAC, 2007). Final performances combining the MTT test and the release of IL-1 α gave a sensitivity close to 91% with a specificity of 79% (Spielmann, 2007). Following this validation and previous published studies applying similar protocols to different chemicals population (Cotovio, 2005; Kandarova, 2005) the question of defining applicability domains of the validated test is crucial. Indeed these robust 3-D models enabled a topical application of various types of chemicals with wide physical-chemical properties (liquids, insoluble powders, gels, sticky compounds etc...). But industrial specificities and diversity pointed out more complex questions. It is of first importance to define the ability of tests, validated on small sets of chemicals, to keep similar performances when the question is addressed to new chemical families. In this study we evaluated the performance of the validate EPISKIN protocol applied to a new set of 184 "real life" cosmetics chemicals extracted from l'Oréal chemicals library with historical *in vivo* data. The aim was to start the understanding of the application fields covered by the validated test and to defined preliminary applicability domains, not only in relation to chemical categories but also linked to physical varieties. The outcome should enable to start drawing up the real actual applicability zone and above all the future adaptations needed, if necessary, in order to cover the largest field of chemicals.

Materials and methods

In vitro Tissue culture

The EpiSkinSM model, manufactured by EPISKIN S.N.C. (Lyon, France) and supplied by SkinEthic Laboratories (www.skinethic.com) is a reconstructed organotypic culture of human adult keratinocytes reproducing a multilayered and differentiated human

epidermis. Briefly, human adult keratinocytes were seeded on a dermal substitute consisting of a collagen I matrix coated with a layer of collagen IV fixed to the bottom of a plastic chamber. This insert was specifically designed to keep the epidermis in a fixed position in the culture wells, allowing appropriate gaseous exchanges and ready access to the culture media during assays. Epithelial differentiation was obtained by an air-exposed step leading to a 3-dimensional epidermis construct (0.38 cm² surface) after 13 days of culture. Main layers are present (e.g. basal, spinous, granular layers and a stratum corneum (Tinois, 1991). The kit system was presented as 12 units pack with all necessary culture media, sterile additional plates and control certificate (SDS IC50, cell viability by MTT conversion test, histology scoring). Tissues were manufactured according ISO 9001 norms. Upon receipt, tissues were transferred from the nutrient agar transport medium to sterile 12-well culture plates (provided with the kit) containing 2 m/well of 37°C pre-warmed maintenance medium. During all the assay period, tissues were maintained in classical conditions and at 37°C, 5% CO₂ and 95% humidity (Cotovio, 2005).

Test chemicals:

All 184 chemicals tested came from l'Oréal chemicals library. Among this set, 61 chemicals were *in vivo* irritants (EU risk phrase: R38) and 123 chemicals were *in vivo* non irritants (EU risk phrase: no classification). *In vivo* data origin was multiple: l'Oréal historical databases, European chemical substance information system and ECETOC databases. The distribution *in vivo* data sources is the following: 71% of the data were documented by OECD TG 404 reports with individual rabbit data; 8% of the data were documented by OECD TG 404 abstracts plus material suppliers data sheet (MSDS); 8% of the data were documented by Draize test (Draize, 1944) assays with individual rabbit data; 5% of the data was documented by ECETOC databases; 5% of the data were documented by MSDS alone and 3 % of the data were documented by both ECETOC databases and European chemical substance information system.

Categories tested: 12 categories were evaluated. Silicones (7 chemicals); vegetal extracts (4 chemicals); surfactants (35 chemicals); Fatty (6 chemicals); solvents (6 chemicals); polymers (78 chemicals); actives (15 chemicals); miscellaneous (16 chemicals); colorants (7 chemicals); preservatives (5 chemicals); filters (4 chemicals); vitamins (1 chemical).

Reagents:

MTT: (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (CAS 298-93-1) was

purchased from Sigma, M-2128 (France). Phosphate buffer saline with calcium and magnesium (D-PBS) was from Gibco (France). Sodium dodecyl sulphate and Meyer hemalun were from BDH (England), isopropanol was from SDS, (13124 Peypin, France) and HCl 12 N from Prolabo (France). Acidic isopropanol was made of 0.04 N HCl in isopropanol (498.2 ml of isopropanol plus 1.8ml of 12N HCl). Store up to 3 weeks at 4°C protected from light.

ELISA kits for IL-1 α (ref: SLA50) were obtained from R&D Systems (Minneapolis, USA). Adenylate kinase kits (ToxiLight Bio Assay kits, ref: LT07-217) were from Cambrex Bio Science. Rockland Inc. (USA). IL-1 α Standard Reference 86/632 was purchased from the National Institute for Biological Standards and Control (NIBSC). 117,000 International Units (IU) per ampoule. standards@nibcs.ac.uk or enquiries@nibcs.ac.uk

Treatment protocol:

Evaluation of direct MTT reduction by chemicals

Prior to the experiments, each chemical was evaluated for its intrinsic ability to directly reduce MTT. 10 μ l (liquids) or 10 mg (solids) of test chemical were added to 2 ml of 0.3 mg/ml MTT solution (in assay medium), incubated for 3 hours at 37°C, 5% CO₂ protected from light. MTT interacting chemicals induced a blue staining of the solution. They were submitted to specific protocol (using dead epidermis) during the study in order to subtract non specific color due to the interaction.

Topical application of chemicals

10 μ l (for liquid to slightly viscous chemicals) were applied onto the surface of the epidermis. 10 mg of solid, powder or sticky material were applied and gently spread using a curved spatula (5 μ l of water were previously spread on the epidermis in order to facilitate the contact with powders). Solids and coarse powders were thinly crushed before application. Samples were tested in triplicate on 2 batches.

Controls

PBS was used as the negative control (10 μ l topically applied in triplicate) and SLS 5% in water was the positive control (10 μ l topically applied in triplicate).

Treated epidermis were incubated for 15 min at room temperature and then rinsed with 25 ml of PBS. Rinsed tissues were incubated for a post-treatment incubation period of 42-hours at 37 °C, 5 % CO₂ in humidified atmosphere. At the end of the post treatment incubation period, the culture media of each tissue were collected individually and stored at -20°C for IL-1 α measurement (Fig. 1).

Viability measurement – MTT assay:

MTT assay (Mossman, 1983) is based on the reduction of the yellow tetrazolium salt

3-(4,5-dimethylthiazol-2-yl) 2,5-diphenyl tetrazolium bromide into purple insoluble formazan crystals by mitochondrial dehydrogenase activity of viable cells. Tissues were washed with D-PBS and placed in 12 well plates containing 2 ml/well of MTT working solution (0.3 mg/ml in assay medium) and incubated for 3 hours in a humidified chamber at 37°C, 5% CO₂ protected from light. At the end of the incubation period, the epidermis were punched and separated from their collagen matrix. To extract formazan crystals, matrix and tissues were both placed in 500 μ l of acidic isopropanol (0.04 N) overnight at room temperature or alternatively 3 days at 4°C (protected from light). At the end of the extraction period, 200 μ l of each sample were transferred to 96 flat bottom well plates for absorbance measurements. Optical densities (OD) were measured at 570 nm, with acidic isopropanol as the blank. For each chemical treated epidermis, viability was expressed as a percentage relative to negative controls (PBS-treated epidermis).

Direct MTT reducers as well as color interfering chemicals can produce false MTT reduction signal or unspecific OD if they remain bound to the tissue after washing. To measure this effect, the chemical should be applied on dead tissues using the standard procedure. Dead tissues were obtained by incubating epidermis 24 hours with 2 ml of water, dried and stored at -20°C until use. The protocol used to treat these samples should be the same as the one for viable tissues. Untreated dead tissues were used as controls. OD₅₇₀ values measured with dead-treated epidermis should be equal or less than 30% of those measured in living PBS-treated epidermis. To calculate the percentage of viability, OD₅₇₀ of treated living tissues were adjusted by subtracting the OD₅₇₀ values measured in treated dead tissues (EPISKIN® -MTT reduction SOP)

IL-1 α measurements:

IL-1 α release in the culture medium was determined by a classic quantitative sandwich enzyme immunoassay technique (Quantikine, R&D Systems, UK).

At the end of the 42 hours post-incubation period, shake the plates containing the treated epidermis (lids on) on a plate shaker for 15 minutes in order to homogenize the released mediators in the medium before sampling (Media from each tissue can be kept frozen (-20°C) until measurements). Briefly, 200 μ l of samples were added into a 96 well ready to use microtiter plate (pre coated with monoclonal antibodies specific for human IL-1 α). Unbound proteins were removed by washing with buffered solution. 200 μ l/well of horse radish peroxidase conjugated polyclonal antibody against IL-1 α was added for 1 hour. Afterwards, a solution of hydrogen peroxide and tetramethylbenzidine was added. A blue color developed and was further stabilized as a yellow

product by adding 2N sulphuric acid. Absorbance was measured at 450 nm. Cytokine concentrations were then calculated using the standard curve. Detection limits of the methods were 3.9 pg/ml. When necessary, dilutions were made for samples with the maintenance medium (if samples have been diluted, results must be multiplied by the dilution factor).

Preliminary calibrations of the ELISA kit with Reference Standard IL- α

Note: The ELISA kit used during the ECVAM validation (Spielmann, 2008) was supplied by R&D System. However, it is possible to use kits from other suppliers if kit is calibrated and the cytokine is expressed as International Units (IU). In all cases, according to specific laboratory conditions, it is necessary to calibrate the ELISA kit response against IL-1 α standard reference (provided by NIBSC).

It is necessary to check 1- that the dose-response curve is linear when using the range of IL-1 α standard concentrations recommended by the supplier, 2-that the measured converting factor (permitting the conversion of pg/ml to the corresponding value in IU/ml) determined in the laboratory conditions is closed to the supplier defined converting factor.

Preparation of the IL-1 α reference standard dilutions (maintenance medium was used as solvent for all steps): Reconstitute IL-1 α reference Standard with the maintenance medium in order to obtain a stock concentration of 58500 IU/ml. Prepare a mother solution of 2000 IU/ml and make the necessary dilutions in order to obtain the following recommended concentrations (IU/ml): 0; 0.39; 0.78; 1.6; 3.12; 6.25; 12.5; 25; 50; 100; 200.

Calculations of the IL-1 α Measurement Converting Factor (MCF)

IL-1 α reference standard dilutions are treated as samples with the ELISA kit. Determine the corresponding concentrations in pg/ml using the ELISA kit standard curve. Using the curve IL-1 α Reference Standard in pg/ml on the x-axis against IU/ml on y-axis, determine the best fit line. From the equation $y = ax$ determine the curve hillside "a" corresponding to the MCF. The MCF determined in the laboratory must be closed to the Converting Factor given by the ELISA kit supplier (i.e. 0.153 for R&D System SLA 50 kits). To be acceptable, the CV of the MCF mean must be < 5%.

Determination of IL-1 α concentration in treated samples

Only culture media of test-chemicals presenting viabilities > 50% were selected for cytokine assays. IL-1 α mean release concentrations from triplicate tissues were first expressed as pg/ml. For each independent experiment, in order to minimize

intra and inter-variability, data were normalized by correcting values with controls (mean [IL-1 α (pg/ml)] = mean [IL-1 α treated tissues] - mean [IL-1 α negative control tissues]). Corrected IL-1 α concentrations of samples were then expressed as IU/ml according to the NIBSC IL-1 α Reference Standard by using the MCF value: IL-1 α (IU/ml) sample = MCF x mean [IL-1 α (pg/ml)].

Acceptance criteria:

OD₅₇₀ of the negative control (PBS-treated) reflects the viability of the tissues used in the test conditions after shipping. Based on historical data, OD₅₇₀ of negative controls (NC) using 200 μ l of extract/well for readings, should be > 0.6 (using 500 μ l extracts). In the same manner, we can consider that an SDS (5%) Positive Control (PC) is acceptable if the mean viability expressed as % of the NC is \leq 40% with an SD \leq 18.

Defined Prediction Model (PM):

Irritant potential of test chemicals according to EU classification was predicted by the mean viability of tissues exposed to the test chemical. A chemical is part of the R38 Irritant class (I) if the mean viability is \leq 50 % relative to NC or if the mean viability is > 50% with a release of IL-1 α \geq 7.65 IU/ml. If the mean viability is > 50 % and the release of IL-1 α < 7.65 IU/ml, the chemical is classified as Non Irritant (NI). [with the R&D systems IL-1 α ELISA kit, 7.65 IU/ml corresponded to 50pg/ml (control value subtracted)] (Fig. 2).

Results

The set of 184 chemicals tested broadly covered 12 categories of ingredients used in the cosmetic industry. This categorization defined subclasses ranging from 1 individual to 78 individuals reflecting the heterogeneousness of the portfolio. Two categories were over represented namely polymers and surfactants with 78 and 35 chemicals respectively. Some categories showed a balanced population in terms of irritancy potential (e.g. actives, miscellaneous, solvents), while others were highly unbalanced or lacked one category (e.g. colorants, preservatives, fatty, vegetal extracts, filters, polymers, surfactants, vitamins, silicones) (Table 1). 16.3% of the chemicals belonged to powders and solids families while 79.9% of the chemicals were liquids. Only 3.4% belonged to more specific physical categories (gels and pastes). These materials obtained from l'Oréal chemicals library, possess *in vivo* data extracted from different databases. The majority of the chemicals (71%) were classified by using data coming from historical databases and extracted from individual detailed reports. Remaining classifications were based on other sources as follows: OECD TG 404 abstracts coupled with MSDS (8%), Draize

Table 1: Set of chemicals tested and in vivo classifications

Chemical Family	physical state* Number chem.	EU class R38 Number chem.	EU class NC Number chem.	total number Number chem.
Actives	9xL, 6xP	6	9	15
Colorants	5xL, 2xP	1	6	7
Preservatives	1xL, 3xP, 1xPa	5	0	5
Fatty	6xL	1	5	6
Vegetal extract	4xL	1	3	4
Filters	2xL, 2xP	0	4	4
Polymers	68xL, 7xP, 1xPa, 2xG	11	67	78
Miscellaneous	12xL, 4xP	8	8	16
Solvents	6xL	3	3	6
Surfactants	27xL, 6xP, 1xG, 1xW	25	10	35
Vitamines	1xL	0	1	1
Silicones	6xL, 1xG	0	7	7
Total	147xL, 30xP, 2xPa, 1xW, 4xG	61	123	184

*: L=liquid; P=powder; S=solid; Pa=paste; W=wax, G=gel

Table 2: Overall performances applied to the full tested set of chemicals (184).

Total set 184 chemicals	Combined viability and IL-1 α
SENSITIVITY(+)	85 %
SPECIFICITY(-)	86 %
CONCORDANCE	86 %

Table 3: Performances of the protocol applied to the sub set of liquid chemicals (147).

Subset of liquids (147)	Combined viability and IL-1 α
SENSITIVITY(+)	88.7 %
SPECIFICITY(-)	86.4 %
CONCORDANCE	87.0 %

test-primary dermal irritation index reports (8%), ECETOC databases (5%), MSDS (5%) and MSDS coupled with information from the European chemical substances information system (3%) (Fig. 3). All chemicals were first tested for their ability to directly interact with the MTT reagent. 11/ 184 directly reacted with the MTT and followed the specific protocol using dead epidermis in order to define the non specific OD and correct the final results. None of the chemicals of the all set were declared as non compatible with the test (unspecific OD was always < 30% of the PBS-treated controls). Interleukin was measured as the second parameter when MTT% was > 50%. Preliminary assays demonstrated the usefulness of the normalization of IL-1 α into international units (IU) by using calibration curves with IL-1 α standard (see materials and methods section). This approach was established to enable the use of different ELISA kits while having a common unit to define the cut-off

Table 4: Performances of the protocol applied to the sub set of powders (30).

Subset of solids (30)	Combined viability and IL-1 α
SENSITIVITY(+)	84.5 %
SPECIFICITY(-)	94.1 %
CONCORDANCE	90.0 %

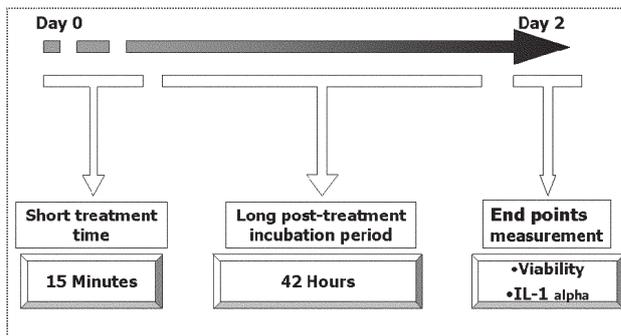


Fig. 1: Key points of the protocol used in this study (ECVAM validate protocol).

of the PM. Most of the kits commercially available can be used (if calibrated) as they have different sensitivities but linear responses (data not shown). All the results generated in this work were based on R&D System kits. Among the 184 chemicals tested, 147 were assayed for IL-1 α . Distribution of the chemicals based on their MTT and IL-1 α were showed in Figs. 4 and 5. In order to improve the detection of I chemicals missed by the MTT test (first sort, see Fig. 2) it appeared worthwhile to evaluate a "cross combination" with IL-1 α (second sort, see Fig. 2). Among the Non irritant population (123 chemicals), the first sort, based on the 50% viability cut-off selected 117 chemicals well classified as NI and generated 6 false positives. After the second sort based on the 7.65 IU/ml cut-off of IL-1 α release, 11 additional false positives were generated (Fig. 4).

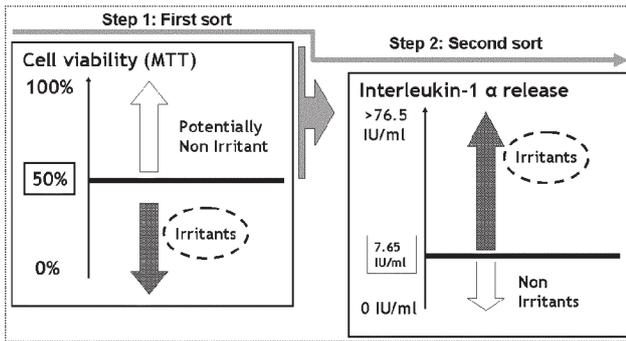


Fig. 2: Tiered strategy using two complementary end points: the viability level and the IL-1 α release in the culture medium using specific predefined cut-off. The MTT (step 1) was used as the first sort to pre select two populations: a population with chemicals classified as irritants and a population with chemicals potentially not Irritant that needed further sorting (step 2) by using the level of released IL-1 α . The second parameter enabled to identify some of the false negatives present in the "potentially non irritant population".

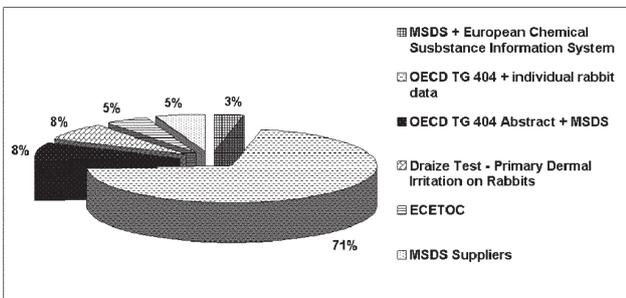


Fig. 3: Distribution of the in vivo data sources qualifying the set of chemicals tested in the study.

Among the Irritant population (61 chemicals), the first sort based on the 50% viability cut-off selected 31 chemicals well classified as I and generated 30 false negatives. After the second sort based on the 7.65 IU/ml cut-off of IL-1 α release, 21 irritants were recovered amongst the false negatives generated by the MTT (Fig. 5). The overall predictive capacities taking into account the tiered strategy combining MTT and IL-1 α and the defined PM for the tested population is given in table 2. Performances are characterized by balanced capacities for predicting the I chemicals (85% sensitivity) and NI chemicals (86% specificity) as well as a good 86% accuracy. Dividing the population tested into powders/solids and liquids enabled to calculate performances related to the physical state of the products. Results for liquids (147 chemicals) are given in table 3 and show very similar performances as compared to the overall population (88.7% sensitivity, 86.4% specificity). Comparable results were obtained with the set of 30 powders/solids (84.5% sensitivity, 94.1% specificity) as showed in table 4. Since different chemicals categories were defined, specific performances were calculated and are given in figure 5. Overall accuracies were calculated for each category whatever

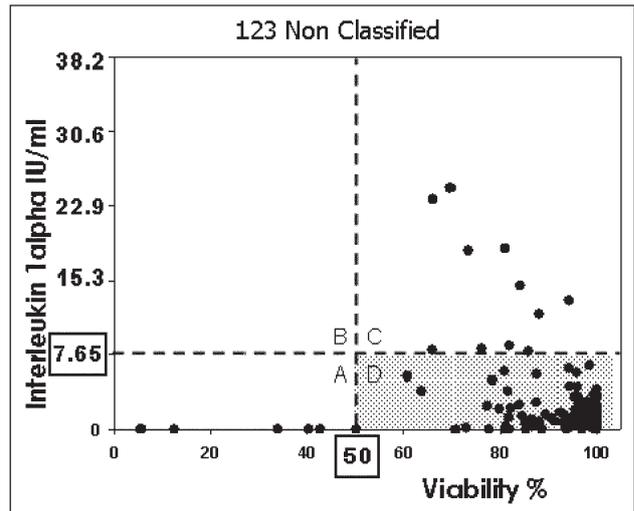


Fig. 4: Distribution of the non classified chemicals (subset of 123 chemicals) using the tiered strategy combining Viability (MTT) and IL-1 α levels. Dot lines represent the two defined cut-off: 50% viability and 7.65 IU of IL-1 α release in the culture medium. Grey area D, represents the well classified chemicals (106 products). Areas A and C represent the false positive classifications (17 products). Chemicals showing viabilities under 50 % are not dosed for IL-1 α release as defined in the prediction model. Results represent the mean of 2 batches, 3 tissues per batch.

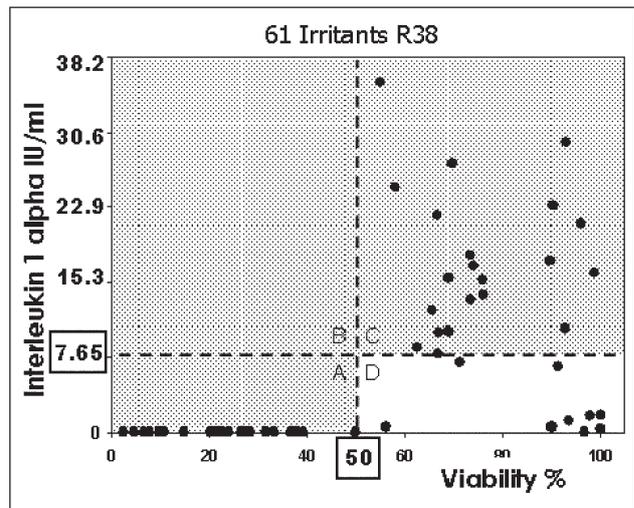


Fig. 5: Distribution of the irritant (R38) chemicals (subset of 61 chemicals) using the tiered strategy combining Viability (MTT) and IL-1 α levels. Dot lines represent the two defined cut-off: 50% viability and 7.65 IU of IL-1 α release in the culture medium. Grey areas A and C represent the well classified chemicals (52 products). Area D represents the false negative classifications (9 products). Chemicals showing viabilities under 50 % are not dosed for IL-1 α release as defined in the prediction model. Results represent the mean of 2 batches, 3 tissues per batch.

the population size. Due to the predominance of small populations these results have to be taken as a tendency. However, results indicate that the highest and the lowest accuracies were obtained mainly for categories with small number of individuals (<

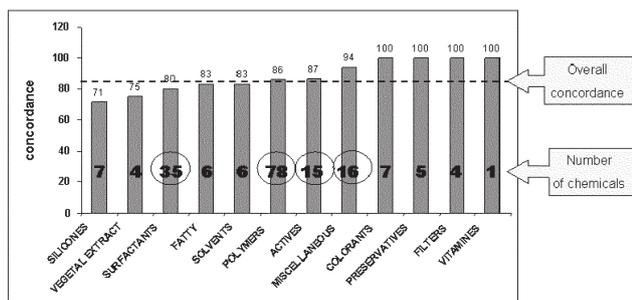


Fig. 6: Comparison of the performances per chemicals categories.

10 chemicals). Categories containing at least 15 or more chemicals (surfactants, polymers, actives, miscellaneous) showed accuracies very closed to that of the all set tested.

Discussion

Because a cosmetic manufacturer placing a cosmetic product on the EU market must be able to demonstrate that the product is safe for the consumer, sufficient safety data should be generated on the ingredients. Due to the ban of animal experiments taking place in March 2009 (7th amendment of the cosmetic directive, EU 2003), it is obvious that *in vitro* tests as well as other systems based on QSARs, analyze of physicochemical properties, read across ... integrated in new strategies are of first importance. For their ability to mimic the human skin (André, 2005; Nohynek, 2005; Poumay, 2007; Schreiber, 2005; Dreher, 2002) *in vitro* skin models are important actors of these new alternative strategies (Zuang, 2005). In April 2007 the ESAC endorsed the scientific validity if the EPISKIN test as a replacement for the rabbit skin irritation method (ESAC, 2007). As this method is the first use of an *in vitro* toxicity testing for acute skin irritation, it is urgent for the industry to know the outcome of the validated test when applied to the chemical diversity of the existing industries. The 12 chemical categories evaluated in this work were characterized by very different population sizes per category (from 1 to 78 individuals). Amongst the 184 chemicals tested, structure diversity covered large fields of application and included 39% of small molecular weight chemicals (MW<1000) and 61% of high molecular weight chemicals (MW>1000) essentially composed of polymers and silicones. This heterogeneity illustrates some of the possible gaps existing between an industrial reality somewhat different when compared to very well defined datasets used in validation processes. Indeed the molecules involved in the ECVAM formal validation set covered wide families and structures but high molecular weight chemicals were not represented (Spielmann, 2007). This example of gap could be considered as part of the remaining application domains to explore. The first results on polymers presented in our work

are encouraging as their prediction accuracy (86%) is equal to the entire set accuracy (86%). On the other hand, silicones exhibited the lowest performance (71% accuracy) as compared to the other categories (however acceptable) but the small population (7 individuals) needs to be extended to reinforce the analysis. This result points out the necessity to work with higher number of chemicals per category in order to ensure the robustness of the analyze. Indeed, very high or low performance levels coming from small populations must hardly be considered as tendencies (it is the case of 8 categories in our study, see results). Applicability domains not only cover chemical families and structures (possibly linked to mechanisms of action) but also all dimensions or aspects of a chemical likely to interfere with the biological system itself whatever the route. In this way, we compared the performances of the "liquid" population (supposed to have a close contact with the epidermis) with the performance of the "solid/powder" population (supposed to have discontinuous less efficient contact with the epidermis surface). Interestingly performances were very good and very similar for both categories (87% accuracy for liquids; 90% accuracy for powders). In spite of the fact that liquids population was 5 times that of powders, both sets could be considered as representative. We can hypothesize that specific humidification steps defined for powders or similar materials, during the elaboration of the standard operating procedure (<http://ecvam.jrc.it/>) was a major breakthrough helping to overcome some difficulties linked to powders/solids. Overall performances of the EPISKIN protocol applied to this "realistic" 184 chemicals population (85% sensitivity, 86% specificity) were close to those determined during the formal ECVAM validation with 58 chemicals (91% sensitivity, 79% specificity) (Spielmann, 2007) when using the tiered strategy approach combining the MTT assay (validated) and the IL-1 α release (recommended as an useful adjunct) (ESAC, 2007). The results of this study reinforced the necessity to combining MTT and IL-1 α (21 chemicals were correctly classified as irritants thanks to the second sort with IL-1 α) and confirm previous work based on the same protocol applied to a population of 48 chemicals (Cotovio, 2005). Furthermore, this work aims to draw a first follow-up of the successful validation starting to open runs for defining limits and applicability domains when shifting from defined validation sets of molecules to "wider" or more complex "real life" molecules.

Conclusion

The validated EPISKIN test for the prediction of acute skin irritation applied to an extended population of 184 industrial cosmetic chemicals confirmed:

- Good performances of the PM using the tiered

strategy combining MTT and IL-1 α (improvement of the detection of irritants).

- The relevance of the validated test when confronted to "real life" raw materials of the cosmetic industry.
- The necessity to define applicability domains step by step on the basis of experimental shared results.

This work is part of an underway strategy based on the assessment of large number of chemical ingredients. Besides the need of methods validated for regulatory purposes, industry also do need to construct strategies of alternative safety specific procedures linked to risk assessment.

Acknowledgments

Authors are grateful to: Frederic Amaral for interleukin calibrations and Corinne Potdevin for the excellent technical assistance; Sophie Loisel-Joubert, Aline Van Der Lee, Marie-Hélène Savary for in vivo databases analyze and review.

References

- André, V., Grenier, S., Pivard, F. and Perrier, E. (2005) 3D cellular models: powerful access to cutaneous physiology and innovative developments of cosmetic active compounds. *Pathologie Biologie*, 53, 618-626.
- Botham, P.A., Earl, L.K., Fentem, J.H., 30, R. and Van de Sandt, J.J.M. (1998) Alternative methods for skin irritation: the current status. ECVAM skin irritation task force report, 1. *ATLA* 26, 195-211.
- Cotovio, J., Grandidier, M.H., Portes, P., Roguet, R., Rubinstenn, G. (2005) The in vitro acute skin irritation of chemicals: optimisation of the EpiSkinSM prediction model within the Framework of the ECVAM Validation Process. *ATLA*, 33, 329-349.
- Draize, J.H., Woodward, G. and Calvery, H.O. (1944) Methods for the study of irritation and toxicity of substances applied directly on the skin and mucous membranes. *Journal of pharmacology and Experimental Therapeutics*, 82, 377-390.
- Dreher F., Fouchard F., Patouillet C., Andrian M., Simonnet J.T. and Benech-Kieffer F. (2002) Comparison of cutaneous bioavailability of cosmetic preparations containing caffeine or alpha-tocopherol applied on human skin models or human skin ex vivo at finite doses. *Skin Pharmacol Appl Skin Physiol* 15 Suppl. 1, 40-58.
- EPISKIN®-MTT reduction SOP. Scientifically validated as replacement to the Draize skin irritation test. Website: <http://ecvam.jrc.it/> (accessed on 29.12.2007)
- EPISKIN®- IL-1 α release SOP. As adjunct to the EPISKIN – MTT assay to increase the sensitivity of the assay. Website: <http://ecvam.jrc.it/> (accessed on 29.12.2007)
- ESAC (2007) Statement on the validity of in vitro tests for skin irritation. Website: <http://ecvam.jrc.it/> (accessed on 29.12.2007)
- EU (2003) Directive 2003/15/EC of the European Parliament and the council of 27 February 2003 amending council Directive 76/768/EEC on the approximation of the laws of the member states relating to cosmetic products. *Official Journal of the European Union* L66, 26-35.
- Faller, C. and Bracher, M. (2002) Reconstructed skin kits: reproducibility of cutaneous irritancy testing. *Skin Pharmacology and Applied Skin Physiology*, 15, 74-91.
- Gerhard J. Nohynek, Daniel Duche, Alexia Garrigues, Pierre-Alain Meunier, Herve Toutain and Jacques Leclaire (2005) Under the skin: Biotransformation of para-aminophenol and para-phenylenediamine in reconstructed human epidermis and human hepatocytes. *Toxicology Letters*, Vol. 158, 3,196-212.
- Liebsch, M. and Spielman, H. (2002) Currently available in vitro methods used in the regulatory toxicology. *Toxicology Letters*, 127:1-3, 127-134.
- OECD (2002) OECD Guidelines for the Testing of Chemicals No. 404: Acute Dermal Irritation/Corrosion, 13 pp. Paris, France: Organization for Economic Cooperation and Development.
- Perkins, M.A., Osborne, R., Rana, F.R., Ghassemi, A. and Robinson, M.K. (1999) Comparison of in vitro and in vivo human skin responses to consumer products and ingredients with a range of Irritancy Potential. *Toxicological Sciences*, 48, 218-229.
- Poumay, Y. and Coquette A. (2007) Modeling the human epidermis in vitro: tools for basic and applied research. *Arch. Dermatol. Res.* 298, 361-369.
- Roguet, R. (1999) Use of skin cells culture for in vitro assessment of corrosion and cutaneous irritancy. *Cell Biology and Toxicology*, 15, 63-75.
- Roguet, R., Cohen, C., Robles, C., Courtellement, P., Tolle, M., Guillot, J.P. and Pouradier Duteil, X. (1998) An inter laboratory study of the reproducibility and relevance of EPISKIN, a reconstructed human epidermis, in the assessment of cosmetics irritancy. *Toxicology In Vitro*, 12, 295-304.
- Schreiber S., Mahmoud A., Vuia A., Rübhelke M.K., Schmidt E., Schaller M., Kandárová H., Haberland A., Schäfer U.F., Bock U., Korting H.C., Liebsch M., Schäfer-Korting M. (2005) Reconstructed epidermis versus human and animal skin in skin absorption studies. *Toxicology in Vitro* 19:6, 813-822.
- Spielmann, H., Hoffmann, S., Liebsch, M., Botham, P., Fentem, J.H., Eskes, C., Roguet, R., Cotovio, J., Cole, T., Worth, A., Heylings, J., Jones, P., Robles, C., Kandarova, H., Gamer, A., Remmele, M., Curren, R., Raable, H., Cockshott, A., Gerner, I. and Zuang, V. (2007) The ECVAM International Validation on in vitro tests for acute skin irritation: report on the validity of the EMISKIN and EpiDerm assays and on the Skin integrity function test. *ATLA*, 35, 559-601.
- Tinois, E., Gaetani, Q., Gavraud, B., Dupond, D., Rougier, A. and Pouradier-Dupteil, X., (1994) The EPISKIN model: Successful reconstruction of human epidermis in vitro. Rougier, A., Goldberg, M. and Maibach, H.I. Eds. *In Vitro Skin Toxicology*. Liebert M.A. Inc. Publishers NY, 133-140.
- Kandárová H., Liebsch M., Gerner I., Schmidt E., Genschow E., Traue D., Spielmann H. (2005) The EpiDerm test protocol for the upcoming ECVAM validation study on in vitro skin irritation tests – an assessment of the performance of the optimized test. *ATLA*, 33, 1-17.
- York, M., Griffiths, H.A., Whittle, E. and Basketter, D.A. (1996) Evaluation of human patch test for the identification and classification of skin irritation potential. *Contact Dermatitis*, 34, 204-212.
- Zuang C., Alonso M.-A., Botham P.A., Eskes C., Fentem J., Liebsch M., van de Sandt J.J.M. (2005) Subchapter 3.2. Skin Irritation. In *Alternative (non-animal) Methods for Cosmetics Testing: Current Status and Future Prospects* (Eskes C., Zuang V. eds). *ATLA* 33-S1, 35-46.
- Zuang, V., Alonso, MA, Botham, P.A., Eskes, C., Fentem, J., Liebsch, M. and Van de Sandt, J.M. (2005) Skin Irritation and Corrosion, *ATLA*, 33, Suppl. 1, 35-46.