Risiken erkennen – Gesundheit schützen

Validation of the EPISKIN and EpiDerm assays and of the Skin Integrity Function Test for acute Skin Irritation Testing

JRC-ECVAM Contract N° 21323-2003-10F 1ED ISP

Summary Report of the ECVAM Skin Irritation Validation Study (SIVS)

Note: In addition to reprints of publications cited (1)-(7) in this summary report other core documents of the Validation Study are mentioned that are part of the Peer Review Documents. These are cited in the summary report [PR document xy].

1 EXECUTIVE SUMMARY

To replace the Draize skin irritation test (OECD TG 404), ECVAM has sponsored a formal validation study of three *in vitro* test systems, two employing reconstituted human epidermis models (EPISKIN, EpiDerm) and the skin integrity function test (SIFT) employing *ex vivo* mouse skin. The objectives were to conduct a validation study to assess the relevance (predictive ability) and reliability (reproducibility within and between laboratories) of these test systems with a set of 60 coded test chemicals for which high quality *in vivo* data were available. It was the goal of the study to assess if the *in vitro* tests would predict *in vivo* classification according to the two classes of the EU classification "R 38" and "non-irritant". In addition, the chemical selection was representative for the three categories of the GHS classification system. Test chemicals were selected by an independent Chemical Selection Sub Committee (CSSC). The validation study was conducted according to the principles and criteria documented in the draft OECD *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment* (No. 34). To ensure high quality of the commercially produced human skin models, the facilities of the producers of the human skin models EPISKIN and EpiDerm were evaluated by independent auditors at the beginning of the ECVAM Skin Irritation Validation Study (SIVS).

In phase 1 of the ECVAM SIVS, 20 chemicals (9 irritant, 11 non irritant) from the New Chemicals Database (NCD) hosted by the ECB, backed by high quality *in vivo* rabbit skin irritation data were tested under blind conditions in the lead laboratories (EPISKIN - L'Oréal, EpiDerm - ZEBET, SIFT - Syngenta). The methods applied (with Standard Operating Procedures, SOPs) were the refined, optimised protocols developed after the ECVAM prevalidation study. When cell viability (MTT reduction) was used as endpoint, the two skin models met the acceptance criteria set by the Management Team (MT) of the study: within laboratory *reproducibility*: identical predictions were obtained in each independent test run with the same chemical with both models. Also, the predictive performance was acceptable: *accuracy:* EpiDerm 75%, EPISKIN 80%; *sensitivity*: EpiDerm 56%, EPISKIN 67%; *specificity*: EpiDerm 91%, EPISKIN 91%. For both skin models, false predictions were only obtained around the *in vivo* classification border (dominant median score 2). In contrast, the SIFT test did not meet the acceptance criteria set by the MT.

The results of phase 1 of the SIVS indicated that when applying the MTT protocols false negative results were the major problem. Meanwhile the lead lab of the EPISKIN test, L'Oréal had developed a promising protocol, in which the release of the cytokine IL-1α was determined in test samples providing a negative MTT result (7) [PR document 19]. Since the SIFT test was not proceeding to phase 2 of the SIVS, the MT decided to add IL-1α determination to the protocols of the two human skin models and established a tiered testing strategy, in which MTT was determined in tier 1 and IL-1α in tear 2 in samples from chemicals providing MTT results below the threshold of 50% viability. Taking into account the results obtained with IL-1α in the EPISKIN model, the lead lab of the EpiDerm test, ZEBET, developed a protocol, in which IL-1α release was also determined in samples from chemicals, which produced a negative MTT result.

In phase 2 of the ECVAM SIVS, 58 test chemicals (18 from phase 1 and 40 chemicals selected by the CSSC, including both new chemicals from the NCD and existing chemicals) were tested under blind conditions with the two human skin models. Each chemical was tested on three parallel tissue replicates per test in three independent tests with the MTT test in each laboratory and test samples were frozen to allow for IL-1α determination at the end of the study.

The EpiDerm test was conducted in the following laboratories: ZEBET (lead lab) Germany, Institute for *In vitro* Sciences (IIVS) USA and BASF Germany. The EPISKIN test was conducted in the following laboratories L'Oréal (lead lab) France, Unilever UK and Sanofi-Synthélabo France. The prediction model (PM) applied in the formal validation study used

the following endpoints: MTT - threshold of 50% reduction of cell viability; IL-1α release - threshold of 60 pg/ml. The PM for IL-1α release was improved taking into account the results of the formal validation study.

Since the IL-1α release protocol of the EpiDerm test was introduced rather late in the study, there was not sufficient time to allow for optimising the protocol. When IL-1α-release was determined in the lead lab of the EpiDerm test, it did not contribute to improve the predictive capacity of the EpiDerm test. Therefore, IL-1α release was not analysed in test samples from the other two labs conducting the EpiDerm Test.

Thus testing of 58 chemicals in phase 2 of the ECVAM SIVS provided the following predictive capacity of the two *in vitro* models for the EU classification system:

EPISKIN (MTT)	SENSITIVITY	77.6%
	SPECIFICITY:	80.7%
EPISKIN (MTT + IL-1α)	**SENSITIVITY:**	**90.7%**
	SPECIFICITY:	**78.8%**
EpiDerm (MTT)	SENSITIVITY:	60.1%
	SPECIFICITY:	88.8%
EpiDerm (MTT+ IL-1α)	no improvement of the predictive capacity	

The predictive capacity of the two skin model assays for the GHS system was insufficient.

The MTof the SIVS concluded, therefore, that in this study

- **the sensitivity and specificity of the EPISKIN skin iritation test (MTT + IL-1α-release) were acceptable, and that the method can therefore be recommended as a replacement for the Draize skin irritation test (EU Annex V B.4; OECD TG 404),**
- **only the specificity of the EpiDerm assay (MTT) was acceptable and, therefore, the assay cannot be recommended as a replacement for the Draize skin irritation test but could be considered for use within a testing strategy, and**
- **further work may be needed to establish the suitability of the methods for classifying chemicals under the new GHS**.

2 HISTORY, CHRONOLOGY and STUDY MANAGEMENT

In 1998, the ECVAM Skin Irritation Task Force published a report on the actual status of *in vitro* skin irritation testing and proposed 10 "challenge chemicals" for which promising, concordant *in vivo* data from the rabbit test, *in vivo* data from 4hr human patch test, and *in vitro* data from the human skin model EpiDerm were available. Proponents of new *in vitro* test systems were encouraged to submit data obtained with new *in vitro* skin irritation test protocols for these chemicals (1) [PR document 14] for assessment whether these tests could be considered in an ECVAM prevalidation study. At the same time, the suitability of various endpoints for prediction of human skin irritation was evaluated in an EU 4th framework collaborative project in several human reconstructed skin models, revealing cell viability reduction (MTT reduction) and IL-1α release the most promising endpoints. Because MTT reduction and IL-1α release showed a high inter-correlation, and IL-1α release was more variable, MTT-reduction was proposed to be the best endpoint for human skin models (2) [PR document 15].

Of the test systems for which data were submitted to the ECVAM TF, five tests were promising for participation in the ECVAM prevalidation study [perfused pig-ear, Prediskin, SIFT, EPISKIN, EpiDerm]. However, during the prevalidation study, two tests failed already in phase 2 due to insufficient reproducibility, whereas the other tests [SIFT, EPISKIN and EpiDerm] showed a sufficient intra- and inter-laboratory reproducibility, but failed in their ability to correctly predict the skin irritation potential of 20 chemicals that were tested in phase 3 of the ECVAM prevalidation study (3) [PR document 16]. The ECVAM Management Team of the study therefore proposed refinement and optimisation of these three tests before approaching further to formal validation.

In 2001, the ECVAM Skin Irritation Task Force and the laboratories responsible for the refinement of the tests met again, discussed ways forward to approach formal validation. In addition, since a post hoc analysis of prevalidation data for MTT reduction for EPISKIN and EpiDerm revealed similar sensitivity, it was recommended to develop a common test protocol for both skin models before start of a formal validation study (4) [PR document 17] .

In November 2002, the ECVAM Skin Irritation Task Force (TF) discussed the refinements of the SIFT and the skin model tests, and came to the conclusion that a formal validation study could be recommended. However, because all refinements were made using the 20 chemicals from the prevalidation study, the TF recommended to perform the SIVS in two phases, where in phase 1 the refinements are confirmed by the leading labs Syngenta (SIFT), L'ORÉAL, and ZEBET, by testing new chemistry in a controlled way under blind conditions. If the outcome of phase 1 was still promising, the SIVS should enter phase 2, a blind trial with three laboratories per test.

During 2003, the EPISKIN test was further refined by L'OREAL by extending the post incubation period of the tissues (after 15 min chemical exposure) to 42 hours which allowed significant effects to increase, and recovery from weak effects.

In May 2003, an ECVAM stakeholder Workshop recommended to conduct a formal validation study and to concentrate on the predictions of the EU classification system (R38 vs. no label), because the tests were developed and optimised for these predictions. L'ORÉAL and ZEBET collaborated then in developing a common test protocol to be used in the ECVAM SIVS, and evaluated it first with the 20 "challenge" chemicals of the ECVAM prevalidation study. Upon request by the ECVAM SIVS Management Team, in parallel to performing phase 1 of the SIVS, in 2004, the data base was further increased by testing all non-corrosive chemicals recommended in the ECETOC reference data base (ECETOC report No. 66). The data obtained in both skin models with the optimised common protocol were very promising, and published back-to back in 2005 (5)(6) [PR document 18] and [PR document 19].

After announcement of the ECVAM SIVS by the European Commission in June 2003 and a tender of the BfR in July 2003, the BfR was contracted in November 2003. The study started formally with the 1st Meeting of the SIVS Management Team (MT) on 17-18 November 2003, and discussion and approval of a draft project plan provided by the BfR [PR document 2]. During the ECVAM SIVS (11/2003 - 6/2006), six face-to face MT Meetings were held as well as six MT Teleconferences, and about 1.200 Emails were exchanged to manage the study. A summary of the study chronology including all important management decisions is given at the end of this summary report (**Table A**).

Table A and the **Project Plan** [PR document 2] are crucial to understand the timelines of the SIVS and changes made during the study. Main points to mention here are the following decisions which were taken:

- to audit the skin model production facilities to address requirements of OECD GLP Guidance Document No.5,
- not to proceed with the SIFT after completion of phase 1 because of insufficient correct predictions
- to include a second endpoint (IL-1α release) in the skin model tests in a tiered manner: all laboratories collect and freeze media of each treated tissue, and if this endpoint provides improvement of the predictivity, all laboratories shall determine under blind conditions, IL-1α on the media samples kept frozen.

3 EVALUATION OF THE ECVAM SIVS ACCORDING TO THE MODULAR APPROACH TO VALIDATION

According to international agreements (*c.f.* OECD Guidance Document No. 34) the validity of a new, revised, or updated method can be assessed by evaluation of seven modules, proposed in the ECVAM modular approach to validation (1) [PR document 14]. The main advantage of this approach is that it allows both evaluations of well-structured experimental formal validation studies as well as retrospective evaluations of existing literature by applying formalised criteria.

Therefore, the following chapters are structured according to the seven modules that have to be addressed in a validation study. In cases where the modules are well addressed in detail in one of the PR documents [02 - 13] only cross-references to these documents are made; otherwise (e.g. study chronology and chemical selection procedure) the text of this summary report contains the relevant information.

3.1 MODULE 1: DEFINITION OF TEST METHODS

3.1.1 Human Skin Model Skin Irritation Tests (EPISKIN SIT and EpiDerm SIT)

The two human reconstituted skin model skin irritation tests (SIT) evaluated in the SIVS are well-defined test methods that have undergone pre-validation and succeeding refinements. Before start of the SIVS, the lead labs L'ORÉAL and ZEBET have collaborated in developing a common test protocol for the endpoint MTT(6)(7) [PR documents 19, 20]. As a consequence, the SOPs used for the two tests were identical to the extent possible (experimental design, application and rinsing procedures, applied amount per area of tissue, post-incubation period before determination of MTT reduction), and differed only in skin model specific treatment details like the conditioning of the tissues after transport and separation of the EPISKIN tissues from the thick collagen layer before performing the MTT test. Details can be obtained from the EPISKIN SIT SOP [PR document 03] and the EpiDerm SIT SOP [PR document 04].

For the secondary endpoint IL-1α-release, the SOP developed by L'ORÉAL was used also with the EpiDerm SIT.

For both tests, identical acceptance criteria were defined based on the outcome of the concurrently tested positive controls (5% SLS) and negative controls (water or PBS) and on a maximum standard deviation (SD) obtained from the three replicate tissues treated identically. Assays revealing an inter-tissue SD >18% were rejected as non-qualified and repeated. However, if a fourth run did again reveal a NQ result, no further repetition was performed.

3.1.2 Skin Integrity Function Test (SIFT)

The SIFT is a well-defined test employing a positive control (10% SLS) and negative controls as acceptance criteria, which had undergone refinements after the ECVAM prevalidation study. The method used by the lead laboratory Syngenta CTL in phase 1 is described in the study protocol [PR document 05]. Because the SIFT did not progress to phase 2, literature references describing refinements are not given in this summary report; they can be found in the Syngenta study protocol.

3.1.3 Test Chemicals and Selection Criteria

In the ECVAM prevalidation study and in the following test optimisation phases, existing chemicals proposed by ECETOC had been extensively used (ECETOC Report No. 66). The ECVAM SIVS therefore needed to make use of new sources of test materials. As crucial criteria for the selection of chemicals were the availability and the high quality of in vivo data, the first source of chemicals used was the New Chemicals Database (NCD) from the European Chemicals Bureau. The NCD comprises new commercial chemicals registered after 1981 and for which skin irritation testing has been performed according to regulatory standards, including GLP and official test methods. Therefore, ECVAM and the ECB derived information relevant for the chemical selection from the NCD which served as a basis for the work of the Chemicals Selection Sub-Committee (CSSC). At the time of the selection, the NCD contained approximately 5600 notifications representing around 3500 substances. A primary extraction was carried out by applying the following exclusion criteria:

- Chemicals notified before 1995
- Repeated notifications
- Market volume < 0.1 t per year (no skin irritation data)
- Gases & vapours
- Corrosives

The following information was extracted from the NCD for the selected chemicals:

- Substance ID (EC number, dossier number)
- Physical State (solid/liquid) and Mixture (Y/N)
- Purity (%) Typical, Lower limit, Upper limit
- MW (incl. components in mixtures)
- MP, BP, Vapour Pressure
- Water Solubility, Octanol-Water partition coefficient
- Skin Irritation scores (erythema and oedema)
- Classification & Labelling (R38, Xi, Xn, C)
- Desired Effect and Use Categories
- Producer/Notifier names (incl. country of origin)

To further reduce the number of potential candidates, the following secondary exclusion criteria were applied by the CSSC:

- Typical purity < 94%, lower purity < 90% or purity unknown

- Mixture with > 3 components (> 4 for isomeric mixtures)
- Mixtures with unknown component proportions
- Complex mixtures with unidentified components
- MW > 1000, MW ranges or MW unknown

After applying these criteria, 845 chemicals remained, amongst which 731 were non-irritant chemicals according to the GHS classification scheme (GHS NI) so that a further reduction of these was necessary. A pragmatic reduction of the GHS NI chemicals was carried out by selecting (1) those substances supplied by the same companies as those supplying the GHS mild-irritant and irritant chemicals, and (2) those substances that were notified to be used as cosmetic ingredients. A total of 218 GHS NI chemicals were short listed resulting in a total of 332 selected chemicals. These chemicals were then screened further by applying the following exclusion criteria:

- Dangerous chemicals (e.g. explosives, carcinogens)
- Chemicals presenting testing difficulties:
  - hydrolysing chemicals
  - polymerising chemicals
- Chemicals presenting data interpretation difficulties:
  - classified from non-standard test or by read-across
  - classified on the basis of persistent effects
  - classification inconsistent with Draize scores (e.g. classification made by subjective judgement)
- Chemicals only available in preparations
- Chemicals no longer in production

Furthermore, in agreement with a proposal of US representatives from ICCVAM and NICEATM, who acted as observers on the SIVS MT, care was taken that the GHS class of "mild irritants" was represented in equal portions in the groups of chemicals with "no label" and "R38" according to the EU classification system. The final selection of the 20 test chemicals (9 irritant and 11 non-irritant) used in SIVS phase 1 was as shown in **Table 1.**

		NCD
R38	GHS I	4
	GHS Mild	5
Non-R38	GHS Mild	2
	GHS NI	9
Total		20

Table 1: Distribution of phase 1 test chemicals with respect to EU and GHS classification

Due to the low prevalence (7%) of chemicals labelled "R38" in the NCD, the CSSC had to make use of other sources of chemicals for phase 2, which resulted in a longer selection procedure than the experimental conduction of phase 1.The first source used were the chemicals registered in the NCD before 1995, these represented 54 chemicals out of which only 5 met the selection criteria applied by the CSSC. As a consequence, additional sources of existing chemicals were used. These sources were:

(1) 3400 candidate chemicals from the TSCA (Toxic Substance Control Act) database maintained by the US Environmental Protection Agency (EPA);

(2) 124 candidate chemicals obtained from an ICCVAM public call for the submission of dermal irritancy chemical and protocol information/test data which was published at the U.S. Federal Register; and as ultimate choice

(3) additional chemicals from the ECETOC database, preferentially those which had not been used in the earlier prevalidation and protocol optimisation efforts.

With regard to the selection and exclusion of chemicals, the same criteria as those for phase 1 were applied by the CSSC. In total, 60 materials were selected for phase 2 comprising 18 test materials from phase 1 (20 were initially foreseen but 2 could not be reused due to short shelf-lives) and 42 chemicals selected from the NCD and from the other sources mentioned above. The 60 chemicals were distributed to the laboratories in 2 deliveries of 30 chemicals each in September 2004 and in February 2005. Confidentiality of chemical identity prevented using the results obtained from 2 chemicals. The final selection of phase 2 is shown in **Table 2**.

Source		R38 (Skin irritants)		Non irritants		Totals
		GHS Irritants	GHS Mild Irritants	GHS Mild Irritants	GHS Non Irritants	
New Chemicals Database (NCD)	Post 95	6	5	3	14	28
	Pre 95	1	4	n.s.	n.s.	5
ECETOC		5	2	2	10	19
TSCA		1	1	0	4	6
Totals		25		33		58

Table 2: Distribution of phase 2 test chemicals with respect to EU and GHS classification

Throughout the selection procedure, care was also taken that chemicals were representing different gradual degrees of irritancy to the best possible extent in order to mimic the prevalence observed in the NCD.

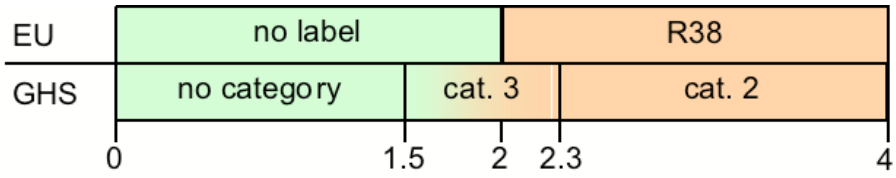

Figure 1: Relation between dominant median score and EU or GHS classifications

Based on the dominant median *in vivo* rabbit score (29) **[DOCUMENT 20]**, that perfectly represents regulatory classification of "no label" (scores < 2.0) and "R 38" (scores ≥ 2.0) (**Figure 1**) the distribution of *in vivo* scores of the 20 chemicals used in phase 1 of the SIVS (**Figure** 2) and the 58 chemicals used in phase 2 (**Figure 3**) showed almost identical distributions, thus not providing an explanation why in phase 1 EPISKIN and EpiDerm performed identical, and in phase 2 showed differences.

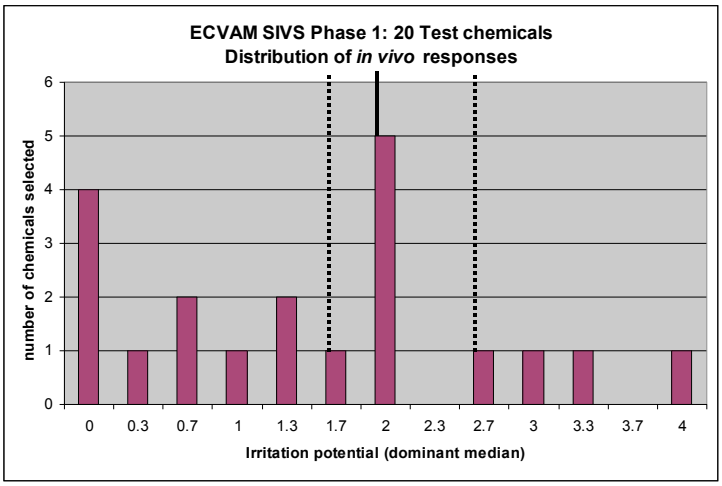

Figure 2: Distribution of *in vivo* responses of 20 chemicals used in phase 1

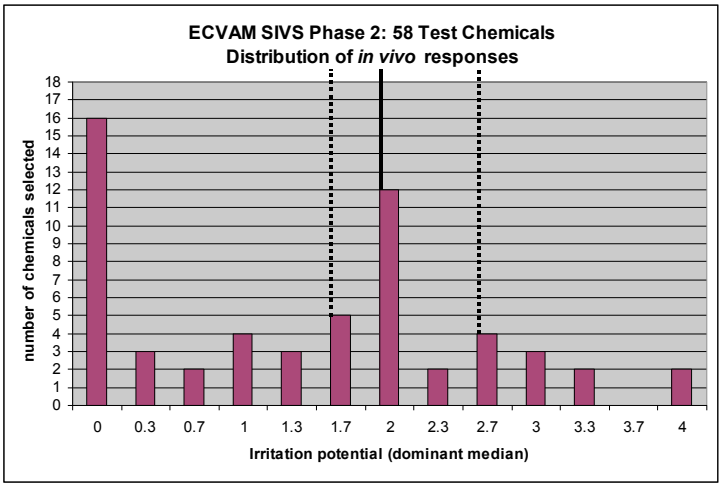

Figure 3: Distribution of *in vivo* responses of 58 chemicals used in phase 2

3.1.4 Definition of Prediction Models

For EPISKIN SIT and EpiDerm SIT, the SOPs [PR document 03] [PR document 04] defined the following common prediction model (PM):

An irritation potential of test materials according to EU classification (R38 or no label) is predicted, if the mean relative tissue viability of three individual tissues exposed to the test substance is reduced below 50% of the mean viability of the negative controls.

***In vitro* result**	***In vivo* prediction**
mean tissue viability ≤ 50	Irritant (I), R38,
mean tissue viability > 50	non-irritant (NI)

For IL-1α-release, (tier 2) L'ORÉAL pre-defined before entering into phase 2, a five fold increase compared to the NC as cut-off for a chemical to be irritant (applied only to chemicals that are predicted non-irritant with MTT reduction). However, this cut-off value turned out to be laboratory-specific, so that only the absolute IL-1α-release of ≥ 60 pg/mL medium (established post-hoc by ROC analysis by ECVAM) revealed a promising result with acceptable inter-laboratory reproducibility.

For the SIFT, the PM was pre-defined in the study protocol [PR document 05]: Two endpoints (TEWL) and electrical resistance (ER) are determined in each test, and used in an either/or condition to predict skin irritancy:

> Test chemicals are classified as potentially irritant to skin if either the post application mean TEWL value is >10g/m²/h or the mean ER value is <4kΩ.

3.1.5 SOP Definitions of Applicability Domain

Restrictions of chemistry that cannot be tested in any of the three tests had been identified before the SIVS started.

For the EPISKIN SIT and EpiDerm SIT, for chemicals that directly reduce MTT, correction techniques were developed, see SOPs [PR document 03] [PR document 04]. Also, testing volatile chemicals is possible when plates are covered and these chemicals are tested on separate plates.

However, post-hoc, it turned out that chemicals that react with the plastic material of the transwells may be predicted false negative. Here, the polystyrene used with EpiDerm provided a larger problem than the polypropylene used with EPISKIN transwells (false negative classification of bromohexane with EpiDerm).

For the SIFT, no restrictions were pre-defined before the SIVS. However, an analysis of the disappointing outcome of the SIFT in phase 1 revealed that in particular irritant solids provided false negative results. The limitations of the SIFT were therefore further investigated by Syngenta during phase 2 - these data will be presented elsewhere.

3.1.6 Explanation of Mechanistic Basis

Due to the fact that reconstituted skin models are lacking of vascularisation, the most important endpoints defined for *in vivo* irritation (erythema and oedema) cannot be used in the *in vitro* tests. However, analysis of literature data (2) [PR document 15] and evaluation of several *in vitro* endpoints (3) [PR document 16] revealed cell viability (MTT reduction)

as most promising *in vitro* substitute endpoint, followed by IL-1α-release which showed a higher variability, but was slightly more sensitive than MTT-reduction.

It is well-known that of the two endpoints determined in the SIFT (TEWL and ER), the TEWL determined *in vivo* correlates with skin irritation potential. The basis of using the two endpoints in combination is desribed in the SIFT protocol [PR document 05] in paragraph 5.2.

3.2 MODULE 2: WITHIN-LABORATORY VARIABILITY

The within - laboratory variability was determined twice, once in phase 1 in the lead labs of the SIFT, EPISKIN SIT and EpiDerm SIT (see ANNEX I of biostatistical report [PR document 06] for 20 chemicals, and then in phase 2 in all skin model labs for 60 test chemicals (see below).

In phase 1, the within-lab reproducibility with regard to consistency of classifications obtained in three independent test runs was acceptable for all three tests. However, the SIFT did not progress to phase 2 because of lacking predictivity.

Within-laboratory variability for EpiDerm and EPISKIN (MTT) is shown on the following pages of PR document 06:

ZEBET	page 15 ff	L'ORÉAL	page 58 ff
IIVS	page 19 ff	Unilever	page 61 ff
BASF	page 23 ff	Sanofi	page 65 ff

The within-lab reproducibility was acceptable in all cases, however there was a significant difference in the number of NQ tests in the lead lab of the EpiDerm test and in the two participating labs.

3.3 MODULE 3: TRANSFERABILITY

Both, the EPISKIN SIT and the EpiDerm SIT were successfully transferred to laboratories that had never been using the test protocols before: The lead labs L'ORÉAL and ZEBET performed face-to-face meetings in Paris and Berlin, in which chemicals were tested and classified consistently across labs. The training and method transfer is reported as part of the phase 1 reports of L'ORÉAL and ZEBET [PR documents 10 and 11].

3.4 MODULE 4: BETWEEN LABORATORY VARIABILITY

For EpiDerm and the endpoint MTT, the between laboratory variability in terms of classifications obtained is shown in the biostatistical report [PR document 06] on page 34 (Table 19). **For the non-irritating chemicals, the inter-laboratory concordance was 78.8% and for the R38 chemicals the inter-laboratory concordance was 74.1%.**

For EPISKIN and the endpoint MTT, the between laboratory variability in terms of classifications obtained is shown in the biostatistical report [PR document 06] on page 76 (Table 44). **For the non-irritating chemicals, the inter-laboratory concordance was 90.9% and for the R38 chemicals the inter-laboratory concordance was 80.0%.**

The overall reproducibility (positive and negative predictions) was 74.1% for EpiDerm and 86.2% for EPISKIN.

3.5 MODULE 5: PREDICTIVE CAPACITY

Only the predictive capacity of the EPISKIN SIT, when MTT and IL-1α-release were applied in a strategic combination, met the criteria for a stand-alone replacement of the *in vivo* test, see biostatistical report [PR document 06], Table 68 (page 130-132) and Table 70, page 133.

3.6 MODULE 6: APPLICABILITY DOMAIN

No clear applicability restrictions could be defined so far: In the CSSC report on possible reasons for misclassifications [PR document 08] no reasons with regard to the type of chemistry could be identified. Further information on the range of properties covered by the selected chemicals are given in the detailed publication of the CSSC.

3.7 MODULE 7: PERFORMANCE STANDARDS

A document drafted by L'ORÉAL, ECVAM and the contractor (BfR-ZEBET) proposes performance standards [PR document 09].

N.B. A final document on Performance Standards for applying human skin models to in vitro skin irritation testing was approved on 25 May 2007 and can be downloaded from the ECVAM website: http://ecvam.jrc.ec.europa.eu/ under “Download study documents”.

4 REFERENCES

1. Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Janusch Roi, A., Prieto, P., Sabbioni, E., Scott, L., Worth, A. and Zuang, V. (2004). A Modular Approach to the ECVAM Principles on Test Validity. ATLA 32, 467-472.
2. Botham, P.A., Lesley, K.E., Fentem, J.H., Roguet, R and J.J.M. van de Sandt (1998) Alternative Methods for Skin Irritation Testing: the Current Status. ECVAM Skin Irritation Task Force Report 1
   ATLA 26, 195-211
3. Faller, C., Bracher, M., Dami, N. and R. Roguet (2002) Predictive ability of reconstructed human epidermis equivalents for assessment of Skin Irritation of cosmetics.
   Toxicology *In vitro* 16, 557-572
4. Fentem, J.H., Briggs, D., Chesne, C., Elliott, G.R., Harbell, J.W., Heylings, J.R., Portes, P., van de Sandt, J.J.M. and P.A. Botham (2001) A prevalidation study on *in vitro* tests for acute Skin Irritation: results and evaluation by the Management Team.
   Toxicology *In vitro* 15, 57-93
5. Zuang, V., Balls, M., Botham, P.A., Coquette, A., Corsini, E., Curren, R.D., Elliott, G. R., Fentem, J.H., Heylings, J.R., Liebsch, M., Medina, J., Roguet, R., van de Sandt, J.J.M., Wiemann, C. and A.P Worth (2002) Follow-up to the ECVAM prevalidation study on the *in vitro* tests for acute Skin Irritation. ECVAM Skin Irritation Task Force Report 2.
   ATLA 30, 109-129
6. Kandarova, H., Liebsch, M., Gerner, I., Schmidt, E., Genschow, E., Traue, D. and H. Spielmann (2005) The EpiDerm Test Protocol for the Upcoming ECVAM Validation Study on the Skin Irritation Tests - An assessment of the performance of the optimised Test.
   ATLA 33, 351-367
7. Cotovio, J., Grandidier, M.-H., Portes, P., Roguet, R. and G. Rubinsteen (2005) The *in vitro* acute Skin Irritation of chemicals: Optimisation of the EPISKIN Prediction Model within the Framework of the ECVAM Validation Process.
   ATLA 33, 329-249
8. Hoffmann, S., Cole, T. and T. Hartung (2005) Skin irritation: prevalence, variability, and regulatory classification of existing *in vivo* data from industrial chemicals.
   Regulatory Toxicology and Pharmacology 41, 159-166

5 Table A: Study Chronology and Management

2001	17-18 May	**ECVAM TF Meeting**: decision about way forward after finalisation of ECVAM SI Prevalidation Study (*cf. Report: Zuang et al., ATLA 2002*)
2002	all year	**test optimisation** EPISKIN and SIFT (L'OREAL, Syngenta)
	20-21 Nov	**ECVAM TF Meeting**: recommendation of a 2-phased Validation Study
2003	07-08 May	**ECVAM Stakeholder Workshop**, positive decision for start of SIVS
	Jun	**EC/JRC call for tenders** on conduct and management of SIVS
	04 Jul	**tender** of main contractor (BfR) including 9 partner institutes
	Jul-Oct	L'OREAL and ZEBET: development of optimised, common test protocol
	05 Aug	**1st SIVS MT Meeting**: report on common skin model protocol; decision to test further chemicals at L'OREAL and ZEBET; agreement on study goal: prediction of EU classification system. However, chemical selection shall equally represent the 3 categories of the GHS classification system
	10 Nov	**contract** signed between JRC and BfR
	17-18 Nov	**2nd SIVS MT Meeting**: discussion of and agreement on project plan provided by BfR and subcontracts with Syngenta and L'OREAL
2004	21 Jan	**1st MT Teleconference**: final agreement on BfR project plan. Report on status of CSSC chemical selection.
	13 Feb	**2nd MT Teleconference:** for phase 1: approval of 20 test chemicals, SOP's (SIFT EpiDerm, EPISKIN,) and data spreadsheets. Decision of GLP/GMP audits by BfR at the skin model production facilities.
	Mar-Apr	two distributions of 10 test chemicals each to L'OREAL, Syngenta and ZEBET
	02-07 Apr	**GLP/GMP Audits** (MatTek Corp, Ashland, USA; EPISKIN SNC, Lyon, F)
	Apr-Jun	**Phase 1 testing** (20 chemicals, 3x) at L'OREAL, Syngenta and ZEBET
	21-22 Jun	**3rd MT Meeting:** discussion of results of phase 1 and audits of skin model production facilities. Conclusion: overall performance of EpiDerm and EPISKIN (not SIFT) promising to progress to phase 2. SIFT needs investigation of test limitations. Chemical selection: re-use of 19-20 chemicals of phase 1, 40 new chemicals.
	28-30 Jun	**training** of laboratories for phase 2 (EPISKIN and EpiDerm)
	26 Aug	**3rd MT Teleconference:** approval of: training reports, blinded biostatistical analysis of phase 1 incl. IL-1α. Approval of chemical selection.
	14 Sep	**4th MT Teleconference**: agreement on IL-1α endpoint inclusion (tiered: all labs keep frozen media, lead labs test, if promising: all labs test IL-1α.
	29 Sep	**1st distribution of 30 chemicals** for phase 2 to 6 laboratories
	Oct	**start of phase 2 testing** L'ORÉAL, Sanofi, Unilever, ZEBET, IIVS, BASF
2005	16 Feb	**2nd distribution of 30 chemicals** for phase 2 to 6 laboratories
	13 May	**training** of ZEBET by L'ORÉAL: IL-1α-determination
	Jun	**submission of phase 2 data**: lead labs: MTT + IL-1α. Other labs: MTT only
	12-13 Jul	**4th MT Meeting**: discussion of preliminary phase 2 analysis: EpiDerm: MTT not sensitive enough (special study at ZEBET with extended exposure) and IL-1α not promising. EPISKIN: MTT more balanced prediction and IL-1α promising. Thus, testing of frozen samples needed by Sanofi and Unilever.
	Aug - Oct	**IL-1α training and testing** by EPISKIN labs 2 & 3 (Sanofi & Unilever)
	08 Nov	**Submission of IL-1α data** by EPISKIN labs 2 & 3 (Sanofi & Unilever)
2006	17 Feb	**5th MT Meeting**: discussion of 1st drafts of ECVAM Biostatistical Report and CSSC Report on misclassifications. ANOVA not adequate. Data retrieval at ECVAM needs independent audit by BfR, data used need approval by testing labs.
	Feb	**data QC** by labs, both for MTT and IL-1α
	15 Mar	**5th MT Teleconference** on open actions needed for study finalisation (e.g. communication with EpiDerm labs 2 and 3 about reasons for NQ assays
	16 May	**6th MT Meeting**: Conclusion of the MT based on 3rd version of Biostatistical Report: both tests sufficiently reproducible; because of high specificity & low sensitivity, EpiDerm usable in tiered strategies; because of balanced predictivity, EPISKIN (in particular when MTT + IL-1α is used) validated as stand alone replacement test.
	30 Jun	**6th MT Teleconference**: agreement on study communication and actions needed for submission of documents that allow ECVAM Peer Review

MT = Management Team; TF = Task Force; SIFT = Skin Integrity Function Test

6 Table B Test Chemicals, *In vivo* Classifications and *in vitro* Classifications (58 chemicals of phase 2 incl. 18 chemicals tested also in phase 1)

No	Chemical identification	CAS No	L/S*	EU class	GHS class	EpiDerm Phase 1 MTT	EpiDerm Phase 2 MTT	EPISKIN Phase 1 MTT	EPISKIN Phase 2 MTT	EPISKIN com-bined
1	2-chloromethyl-3,5-dimethyl-4-methoxypyridine hydrochloride	86604-75-3	S	R38	I	1	1	1	1	1
2	1-bromo-4-chlorobutane	6940-78-9	L	NC	NI		1		1	1
3	1-bromohexane	111-25-1	L	R38	I		0		1	1
4	1-decanol	112-30-1	L	R38	I		1		1	1
5	3-chloro-4-fluoronitrobenzene	350-30-1	S	NC	NI		0		1	1
6	3-diethylaminopropionitrile	02/04/51	L	NC	NI		1		0	0
7	3-mercaptohexanol	51755-83-0	L	NC	NI		0		1	1
8	4-methylthio-benzaldehyde	3446-89-7	L	NC	NI		0		1	1
9	2,6-dimethyl-4-nitrobenzeneamine	16947-63-0	S	NC	NI	0	0	0	0	0
10	allyl heptanoate	142-19-8	L	NC	mild		0		0	0
11	allyl phenoxyacetate	7493-74-5	L	NC	NI		0		0	0
12	2-ethylhexyl 4-aminobenzoate	26218-04-2	S	NC	NI	0	0	0	0	0
13	1-[4-(2-dimethylaminoethoxy)phenyl]-2-phenylbutan-1-one	68047-07-4	S	R38	mild	0	1	1	1	1
15	a-terpineol	98-55-5	L	R38	I		0		1	1
16	capryl-isostearate	209802-43-7	L	NC	NI	0	0	0	0	0
17	2-methyl-3-[(1,7,7-trimethylbicyclo[2.2.1]hept-2-yl)oxy]-1-propanol, bornyl isomer	128119-70-0	L	NC	mild	1	1	1	1	1
18	butyl methacrylate	97-88-1	L	R38	I		0		1	1
19	2,5-dimethyl-4-oxo-4,5-dihydrofuran-3-yl acetate	4166-20-5	L	NC	NI		0		0	0
20	cyclamen aldehyde	103-95-7	L	R38	I		1		1	1
21	A mixture of: 5-exo-decylbicyclo[2.2.1]hept-2-ene; 5-endo-decylbicyclo[2.2.1]hept-2-ene	22094-85-5	L	NC	mild		0		0	0
22	diethyl phthalate	84-66-2	L	NC	NI		0		0	0
23	di-n-propyl disulphide	629-19-6	L	R38	I		0		0	1
24	di-propylene glycol	25265-71-8	L	NC	NI		0		0	0
25	dipropylene glycol monobutyl ether	29911-28-2	L	NC	NI		0		0	0
26	3,4-dimethyl-1H-pyrazole	2820-37-3	S	NC	NI		1		1	1
27	2-isopropyl-2-isobutyl-1,3-dimethoxypropane	129228-21-3	L	R38	I		0		0	1
28	ethyl cis-4-[4-[[2-(2,4-dichlorophenyl)-2-(1H-imidazol-1-ylmethyl)-1,3-dioxolan-4-yl]methoxy]phenyl]piperazine-1-carboxylate	67914-69-6	S	NC	NI	0	0	0	0	0
29	Mixture of: 2-methyl-4-(2',2',3'-trimethyl-3'-cyclopenten-1'-yl)-4-penten-1-ol 56% (1'R,2R) & 40%(1'R,2S) isomer	014864-90-6	L	R38	mild	1	1	1	1	1
30	Mixture of:diethyl cis-1,4-cyclohexanedicarboxylate, diethyl trans-1,4-cyclohexanedicarboxylate	0072903-27-6	L	NC	NI	0	0	0	0	0
31	A mixture of isomers: ethyl exo-tricyclo[5.2.1.0(2,6)]decane-endo-2-carboxylate; ethyl endo-tricyclo[5.2.1.0(2,6)]decane-exo-2-carboxylate	80657-64-3	L	R38	mild		1		1	1
32	2S-(2-furyl)-5R-hydroxy-4R-(1R,2-dihydroxy)ethyl-6S-hydroxymethyl-1,3-dioxane	7089-59-0	S	NC	NI	0	0	0	0	0
33	heptyl butyrate	5870-93-9	L	NC	mild		0		0	0
34	hexyl salicylate	6259-76-3	L	R38	mild		0		0	0
35	cyclohexadecanone	2550-52-9	S	NC	NI	0	0	0	0	0
36	isopropanol	67-63-0	L	NC	NI		0		0	0
37	[2-(cyclopentyloxy)ethyl]benzene(cyclopentyl 2-phenylethyl ether)	not allocated.	L	R38	I	1	1	1	1	1
39	methyl stearate	112-61-8	S	NC	NI		0		0	0
40	1-methyl-3-phenyl-1-piperazine	5271-27-2	S	R38	I	1	1	1	1	1
41	naphthalene acetic acid	86-87-3	S	NC	NI		0		0	0
42	disodium 2,2'-(1,4-phenylene)bis-(1H-benzimidazole-4,6-disul fonic acid or monosulfonic acid, monosulfonate or disulfonate	180898-37-7	S	NC	NI	0	0	0	0	0
43	A mixture of isomers: 1-(1,1-dimethylpropyl)-4-ethoxy-cis-cyclohexane; 1-(1,1-dimethylpropyl)-4-ethoxy-trans-cyclohexane	181258-87-7 (cis), 181258-89-9 (trans)	L	R38	mild		0		1	1
44	phenylethylalcohol	60-12-8	L	NC	NI		0		0	0
45	(+/-) trans-3,3-dimethyl-5-(2,2,3-trimethyl-cyclopent-3-en-1-yl)-pent-4-en-2-ol	107898-54-4	L	R38	I		1		1	1
46	4-methyl-8-methylenetricyclo[3.3.1.1(3,7)]decan-2-ol	122760-84-3	S	R38	Mild		1		1	1
47	4-methyl-8-methylenetricyclo[3.3.1.1(3,7)]dec-2-yl acetate	122760-85-4	L	R38	Mild		1		1	1
48	2-(formylamino)-3-thiophenecarboxylic acid	43028-69-9	S	NC	NI		0		0	0
49	isostearic acid monoisopropanolamide	152848-22-1	L	R38	mild	0	0	0	0	1
50	2-phenylhexanenitrile	3508-98-3	L	NC	mild		0		0	0
51	Mixture of isomers:1-(2-isopropylphenyl)-1-phenylethane (CAS# 191044-60-7) 1-(3-isopropylphenyl)-1-phenylethane (CAS# 191044-59-4) 1-(4-isopropylphenyl)-1-phenylethane (CAS# 2320-06-1)	52783-21-8	L	R38	mild	0	0	0	0	1
52	propyl (2S)-2-(1,1-dimethylpropoxy)-propanoate	0319002-92-1	L	NC	NI	0	0	0	0	0
53	silane A-1430	2530-87-2	L	NC	NI		0		0	0
54	Mixture of isomers: 1-(spiro[4.5]dec-7-en-7-yl)pent-4-en-1-one (CAS# 224031-70-3) 1-(spiro[4.5]dec-6-en-7-yl)pent-4-en-1-one (CAS# 224031-71-4)	224031-70-3	L	NC	NI		0		0	1
55	terpinyl acetate	80-26-2	L	R38	mild		0		0	1
56	benzenethiol, 5-(1,1-dimethylethyl)-2-methyl *(NB: CAS name from company)*	7340-90-1	L	R38	I		1		1	1
57	triethylene glycol	112-27-6	L	NC	NI		0		0	0
58	tri-isobutyl phosphate	126-71-6	L	R38	mild		1		1	1
59	(E,E)-3,7,11-trimethyldodeca-1,4,6,10-tetraen-3-ol	125474-34-2	L	R38	I	1	1	1	1	1
60	bis[(1-methylimidazol)-(2-ethyl-hexanoate)], zinc complex	not allocated	L	R38	mild		1		0	0