EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection
European Union Reference Laboratory for Alternative Methods to Animal Testing (ECVAM)

Human Cell Line Activation Test (h-CLAT) Validation Study Report

Summary

The evaluation of skin sensitisation potential represents an important component of the safety assessment of new and existing substances. Traditionally this has required animal tests such as the murine Local Lymph Node Assay (LLNA) and/or guinea pig tests (Buehler Test and Guinea Pig Maximisation Test).

There is a pressing need for alternative non-animal methods to reduce and ultimately replace animal tests for this endpoint as required by some European regulations aiming at the protection of human health and the environment (i.e. Cosmetics Regulation and REACH).

Several mechanistically-based non-animal test methods for the assessment of skin sensitisation are currently under development/evaluation. The validation of the human Cell Line Activation Test (h-CLAT) was part of a larger study coordinated by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL-ECVAM), including two other methods: the Direct Peptide Reactivity Assay (DPRA) and the Myeloid U937 Skin Sensitisation Test (MUSST).

The study reported here was designed to allow sound conclusions to be drawn on the transferability and reproducibility (within- and between-laboratories) of the h-CLAT in view of its potential for future use as part of non-animal integrated testing strategies (ITS) for skin sensitisation hazard assessment.

Having reviewed the information generated in the course of this study, the Validation Management Group (VMG) concluded that the findings satisfy the information requirement for the evaluation of validation modules 1-4 (test definition, within laboratory reproducibility, transferability, between laboratory reproducibility) and, in addition, contributes to module 5 (predictive capacity) and module 6 (applicability domain) of the ECVAM modular approach to validation.

The VMG concluded that the information generated in the study shows that the h-CLAT is a robust and reliable test method. Therefore, the VMG supports the consideration and use of the h-CLAT in non-animal Integrated Testing Strategies (ITS)/approaches to support regulatory decision making.

The VMG also considers the h-CLAT merits further evaluation as a component in ITS for replacing the *in vivo* assays for skin sensitisation hazard identification and for the role it might play in the determination of skin sensitisation potency.

Background

Skin sensitisation is the toxicological endpoint associated with substances that have the intrinsic ability to cause skin allergy, leading to the disease allergic contact dermatitis (ACD) in humans, and which represents the most common manifestation of immunotoxicity to chemicals. The identification of skin sensitisation potential represents an important component of the safety assessment of new and existing substances including cosmetic ingredients. Current regulatory predictive tests for skin sensitisation rely on the use of animals. These include guinea pig tests (Buehler Test and Guinea-pig Maximisation Test (OECD TG 406; TM B06, EU Regulation 440-2008), and the murine Local Lymph Node Assay (LLNA, OECD TG 429; TM B42, EU Regulation 440-2008) including non-radio-isotopic variants (OECD TG 422A and OECD TG 422B). Summary details of the animal tests currently in use are provided in Table 1. The LLNA is considered a reduction/refinement method with respect to the traditional guinea-pig tests and of greater value in generating skin sensitisation potency information which is required for full risk assessment to establish safe levels of human exposure. With the entry into force of the 7th Amendment to the Cosmetics Directive and the new European chemicals regulation REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) there is an urgent need for validated alternative non-animal methods for this endpoint.

Skin sensitisation is a delayed-type hypersensitivity reaction induced by low molecular weight reactive chemicals. It develops in two distinct phases; the induction phase which sensitises the immune system for an allergic response and the elicitation phase which occurs following a subsequent contact with the allergen and which leads to the clinical signs and symptoms of allergic contact dermatitis (ACD) in humans or contact hypersensitivity (CHS) in the rodent models. The key underlying biological mechanisms of the induction of skin sensitisation are relatively well understood as recently documented by the OECD in its report on: "The Adverse Outcome Pathway (AOP) for Skin Sensitisation Initiated by Covalent Binding to Proteins" (OECD 2012a; 2012b). These include: 1) the ability of the chemical to penetrate the skin and reach the site of haptenation (skin bioavailability), 2) the covalent binding of the chemical sensitiser to the skin protein (haptenation), 3) the release of pro-inflammatory signals by epidermal keratinocytes (skin inflammation) 4) the activation and maturation of Dendritic cells (DC) the skin immunocompetent cells, 5) the migration of DC from skin to the regional lymph nodes, 6) the presentation by DCs of the haptenated protein to T cells and the clonal expansion of memory T cells (lymphocytes capable of being stimulated and activated specifically by the haptenated protein). Progress has been made in recent years in the development of mechanistically-based alternative methods for hazard identification some of which might also be able to contribute to potency characterisation. However none of the mechanistically-based non-animal tests currently under development/evaluation is considered to have the potential to function as a stand-alone method to fully replace the currently used animal tests. Instead it is proposed that a combination of such mechanistic tests, addressing the key biological events of skin sensitisation, will be needed to achieve this goal. A comprehensive overview of the currently available methods targeting the key mechanisms described above was published in 2010 (Adler et al., 2010).

In the first quarter of 2009 three partial replacement methods for skin sensitisation testing were formally submitted to ECVAM for evaluation. These methods, namely the Direct Peptide

Reactivity Assay (DPRA), the human Cell Line Activation Test (h-CLAT) and the Myeloid U937 Skin Sensitisation Test (MUSST) were developed by the European Cosmetics Association (Cosmetics Europe, former Colipa) associated industries and optimised within Cosmetics Europe ring trials. These three test methods were proposed as potential candidates for regulatory use as part of integrated alternative testing approaches for replacing the current regulatory *in vivo* tests. How the information generated with these test methods could potentially be integrated to achieve this goal still has to be fully determined, and did not form part of the validation study reported here.

Following the evaluation of the information provided on these test methods and after review of the submitted protocols (i.e. SOPs) ECVAM concluded that they were sufficiently developed and standardised to be included in the ECVAM validation process.

Table 1: Overview of regulatory accepted animal tests

OECD Test Guidelines for Skin Sensitisation	Skin sensitisation phases covered	Animal species	Adjuvant	Exposure	Dose levels	N° of animals in control group	N° of animals in treatment group	Test duration (days)	Endpoint	Classification criteria
406: Guinea Pig Maximisation Test (GPMT)	Induction + elicitation	Guinea pig	Y (Freund's Complete Adjuvant-FCA)	*Induction*: intradermal injections (day 0) and topical application (day 5-7 and day 6-8) *Challenge:* topical application (day 20-22) by occluded patch *Re-challenge:* possible	*Induction*: 1 dose (highest concentration to cause mild-to moderate-skin irritation) *Challenge:* 1 dose (highest non-irritant dose)	5	10	23-25	Skin reactions (erythema/ oedema)	Positive reaction in at least 30% of the animals in the treatment group
406: Buehler Test	Induction + elicitation	Guinea pig	N	*Induction*: topical application (day 0, day 6-8 and day 13-15) *Challenge:* topical application (day 27-29) *Re-challenge:* possible	*Induction*: 1 dose (highest concentration to cause mild skin irritation) *Challenge:* 1 dose (highest non-irritant dose)	10	20	30-32	Skin reactions (erythema/ oedema)	Positive reaction in at least 15% of the animals in the treatment group
429: Local Lymph Node Assay	Induction	Mouse	N	Topical application	At least 3 doses (highest dose should not give systemic toxicity and/or excessive local irritation)	4	4	6	Cellular proliferation in auricular lymph nodes measured by radioactive labelling	Stimulation Index (SI) >3 at any dose.
442A: Local Lymph Node Assay: DA	Induction	Mouse	N Pre-treatment with 1% Sodium Lauryl Sulphate (SLS)	Topical application	At least 3 doses (highest dose should not give systemic toxicity and/or excessive local irritation)	4	4	8	Cellular proliferation in auricular lymph nodes quantified by determination of ATP content	Stimulation Index (SI) ≥1.8 at any dose.
442B: Local Lymph Node Assay: BrdU Elisa	Induction	Mouse		Topical application	At least 3 doses (highest dose should not give systemic toxicity and/or excessive local irritation)	4	4	6	Cellular proliferation in auricular lymph nodes quantified by determination of BrdU incorporation	Stimulation Index (SI) ≥1.6 at any dose.

Management of the Study

Reference documents:

- **Project Plan (Appendix 1)**
- **List of additional available documents filed for the study and available on request (Appendix 15)**

1. Study objectives

1.1 Primary objectives

In September 2009 a formal validation study of the h-CLAT was launched, with the **primary overall objective of evaluating its transferability and reliability (reproducibility within- and between-laboratories) with a view to its future use in integrated non-animal approaches for replacing the currently used regulatory animal tests**.

1.2 Secondary objectives

As secondary goals of the study, the information and experimental data generated during the evaluation of transferability and reliability were used to perform:

a) A preliminary evaluation of the ability of the h-CLAT to reliably discriminate skin sensitising (S) from non-sensitising (NS) chemicals as defined by the Globally Harmonised System (GHS) of Classification and Labelling of Chemicals for skin sensitisation (category 1; no category) and as implemented in the European Commission Regulation (EC) No 1272/2008 (EC, 2008) on classification, labelling and packaging (CLP) of substances and mixtures.

b) Where possible, a preliminary consideration of the ability of the h-CLAT to contribute to sub-categorisation of skin sensitising chemicals, e.g. into Sub-category 1A and Sub-category 1B as adopted in the 3rd revised version of the GHS (UN, 2009).

1.3 Other considerations

The current report, which was prepared by EURL ECVAM with the support of the Validation Management Group (VMG), presents the outcome of the validation study of the h-CLAT, during which the transferability and reliability were independently evaluated in four separate laboratories.

Evaluation of how the data generated with the h-CLAT might be accommodated within future testing strategies and data integration activities was outside the scope of the current study, though it is recognised that the availability of high quality non-animal data, such as those generated in the ECVAM study, is a prerequisite for such activities.

2. Project Plan

Prior to the start of the study, a Project Plan was approved and issued by the Validation Management Team (VMT). This document was reviewed as required at each VMT meeting; the final version is annexed to this report (see Appendix 1). The Project Plan documents the objectives, coordination and sponsorship of the study; the nature and roles of the study personnel at each testing site; the minimum quality assurance systems required in the case of non-GLP laboratories; the nature and deliverables of the different study phases; as well as instructions regarding the reception, handling and storage of the test chemicals.

Prior to the start of the training phase, the Project Plan was sent to all testing sites for their information, approval and implementation. They each returned a signed declaration attesting that they had read and understood the project plan, and that their testing facilities would work in compliance with the provisions set out in the Project Plan.

a. Structure of the study

This validation study was organised to generate information relevant to modules 1-4 (1: test definition, 2: within laboratory reproducibility, 3: transferability, 4: between laboratory reproducibility) of the ECVAM modular approach to validation (Hartung et al., 2004) in line with the study's stated objectives. In addition, the experimental data produced in the study also contribute to module 5 on predictive capacity and to module 6 on applicability domain, which were in part addressed by the information generated and submitted to ECVAM by the test method submitter. However, the number of chemicals used in this validation study, which was set to satisfy the primary goal of the study, is not sufficient on its own to draw robust or definitive conclusions on these last two modules.

The study was entirely coordinated by ECVAM with participation from NICEATM-ICCVAM and JaCVAM via the VMT with regard to study design, chemical selection, and test method SOPs. **Figure 1** illustrates how the validation study was organised with respect to the management, the test methods included, the participating laboratories, the selection, coding and supply of the test chemicals and the data collection and statistical analyses. Full details of the management, sponsorship, coordination, timings, responsibility and overall set-up of the study are provided in the Project Plan (Appendix 1). The organisation and conduct of the study was performed in compliance with the principles laid down in the OECD guidance document on test method validation (OECD, 2005).

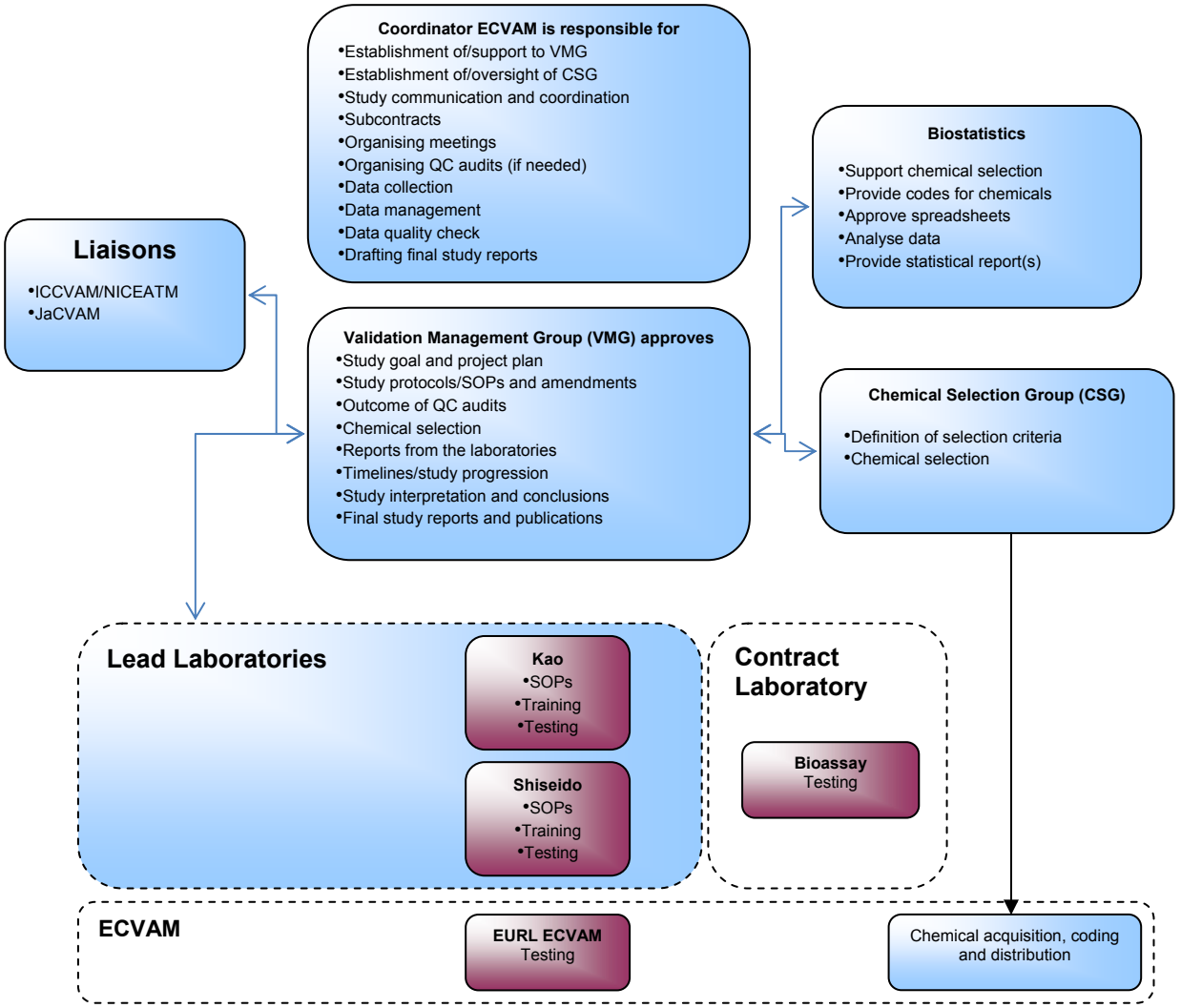

Figure 1: Schematic representation of the study structure and organisation.

b. Validation Management Group

An expert independent Validation Management Group (VMG) was established by ECVAM. Its role was to ensure that the study objectives and goals were clearly defined, to guide and facilitate the validation process, to take study management decisions as the study progressed, to evaluate the results and to draw conclusions regarding the outcome of the study with respect to the study goals. David Basketter was appointed as chair of the VMG because of his acknowledged expertise in the field.

In addition to the VMG, representatives from the lead laboratories were involved in a subset of the discussions, together with liaisons from other validation bodies, ICCVAM/NICEATM (USA), and JaCVAM (Japan). This extended group is referred to as the "Validation Management Team" (VMT).

The strategic decisions, including the selection and approval of the test chemicals to be used in the study, were taken by the VMG only. The liaisons were involved in all discussions. The lead laboratories representatives were not involved in discussions related to the selection of the chemicals.

Validation Management Team Composition:

Validation Management Group

Chair	David Basketter (DABMEB Consultancy Ltd)
Co-chair	Silvia Casati (ECVAM)
Study Coordinator	Alexandre Angers (ECVAM)
Chair of the Chemical Selection Group	Thomas Cole (ECVAM)
Statistician	André Kleensang (ECVAM, up to September 2010), Anna Compagnoni (ECVAM, up to January 2011), Els Adriaens (Adriaens Consulting, from June 2011)
Industry representative	Pierre Aeby (Cosmetics Europe (ex Colipa))
External expert	Sebastian Hoffmann (seh consulting + services)
External expert	Jon Richmond (dr.jonrichmond: Advice & Consultancy)

Lead laboratories Representatives

Shiseido (h-CLAT)	Takao Ashikaga
Kao Corporation (h-CLAT)	Hitoshi Sakaguchi

Liaisons

JaCVAM	Hajime Kojima; alternate Yasuo Ohno
NICEATM	William S. Stokes; alternate Eleni Salicru (up to May 2011); alternate Judy Strickland (since June 2011)
ICCVAM	Joanna M. Matheson; alternate Abigail Jacobs

c. Laboratories

Four laboratories listed below participated in the validation of the h-CLAT. Kao Corporation and Shiseido, in whose laboratories the test method was jointly developed and which have the most experience in performing it, acted as lead laboratories.

Laboratory 3 was involved in the study as a naïve laboratory which had experience with flow cytometry.

Laboratory 4 was selected through an open call for tender published by the Institute for Health and Consumer Protection of the JRC. The criteria for selection were primarily based on the technical merit of the tender including the information provided on the laboratories awareness of the nature of the work involved.

Laboratory 1 (Study Director: Hitoshi Sakaguchi)

Global R&D, Safety and Microbial
Kao Corporation
2606 Akabane
Ichikai-Machi, Haga-Gun
Tochigi 321-3497 Japan

Laboratory 2 (Study Director: Takao Ashikaga)

Shiseido Quality Assessment Center
2-12-1, Fukuura, Kanazawa-ku, Yokohama
236-8643, Japan

Laboratory 3 (Study Director: Ingrid Langezaal)

The European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) GLP Test Facility (as from May 2012, former IVMU and VAM).
Institute for Health and Consumer Protection
Joint Research Centre
European Commission
Ispra, Italy

Laboratory 4 (Study Director: Axel Hohenstein)

Bioassay GmbH
Im Neuenheimer Feld 515
69120 Heidelberg
Germany

d. Quality Systems of the Participating Laboratories

Laboratories 1 and 2 were not OECD Good Laboratory Practice (GLP) compliant. Laboratory 3 was in the formal process of requesting the OECD GLP compliance status during the conduct of the study and used this study as a pilot for the OECD GLP compliance monitoring. However, as Laboratory 3 had not yet been monitored by the OECD GLP monitoring authorities, it cannot be deemed to have been OECD GLP compliant at the time the study was performed. Only Laboratory 4 was fully OECD GLP compliant and subject to inspections by relevant agencies however the study was not conducted under full GLP compliance at this laboratory.

For the laboratories participating in the validation study which were not formally GLP compliant, the VMG defined and requested the application of a minimum set of quality assurance requirements considered essential for the acceptance of information and data produced in the validation process. These requirements formed part of the Project Plan which was sent to, and accepted by, all participating laboratories prior to the initiation of the study.

These minimum requirements were:

• Qualified personnel, and appropriate facilities, equipment and materials shall be provided.

• Records of the qualifications, training and experience, and a job description for each professional and technical individual, shall be maintained.

• For each study, an individual with appropriate qualifications, training and experience shall be appointed to be responsible for its overall conduct and for any report issued.

• Instruments used for the generation of experimental data shall be inspected regularly, cleaned, maintained and calibrated according to established SOPs, if available, or to manufacturers' instructions. Records of these processes shall be kept, and made available for inspection on request.

• Reagents shall be labelled, as appropriate, to indicate their source, identity, concentration and stability. The labelling shall include the preparation and expiry dates, and specific storage conditions.

• All data generated during a study shall be recorded directly, promptly and legibly by the individual(s) responsible. These entries shall be attributable and dated.

• All changes to data shall be identified with the date and the identity of the individual responsible, and a reason for the change shall be documented and explained at the time.

3. Study Experimental Design and sample size

Reference document: Experimental design (Appendix 2)

The number and nature of the chemicals involved in the training and transfer phases were at the discretion of the lead laboratories.

For the blind testing phase, the ECVAM biostatisticians calculated the minimum number of chemicals that would be required to properly evaluate the performance of the tests in light of the primary study goal, i.e. to perform a statistically sound evaluation of the Within Laboratory Reproducibility (WLR) and of the Between Laboratory Reproducibility (BLR).

To ensure coverage of the range of sensitisation potencies, and since it was expected that weak and moderate sensitisers would be more informative for the evaluation of the reproducibility of the test methods under consideration; the ratio of sensitisers to non-sensitisers was set to 2:1.

Preliminary parameters for the statistical analysis and evaluation of the WLR and the BLR were defined at the beginning of the process taking into consideration the expected proportion of concordant classifications derived from the data reported in the test submissions to ECVAM (see Appendix 2 for full details). From these parameters, the number of chemicals required was calculated to be **at least 21** chemicals for the evaluation of the BLR and **at least 13** chemicals for the evaluation of the WLR.

Informed by the calculations performed by the biostatistician, the VMG felt it prudent, as the chemical set was used to evaluate three different *in vitro* test methods, to increase the number

of chemicals with respect to the calculated minimum number. This would allow for potential reduction of the sample size, e.g. the exclusion of chemicals that might fall outside the applicability of some of the test methods (e.g. metals in case of the DPRA).

On the basis of the above considerations, the following design was approved by the VMG:

- **BLR -** For evaluation of the BLR, **24 chemicals** tested once in every laboratory (16 sensitisers and 8 non-sensitisers).
- **WLR -** For evaluation of the WLR, a subset of **15 chemicals** (10 sensitisers and 5 non-sensitisers) from those used for the evaluation of the BLR, tested two additional times in each laboratory, the same subset being tested at every site.

The VMG considered this experimental design appropriate and adequate for providing the information needed to perform a sufficiently robust assessment of the WLR and BLR based on the concordance of classification, sensitizer (S) versus non-sensitisers (NS).

Study Phases:

The study was structured and conducted in two sequential phases, with the performance of the test system being reviewed by the VMG between each phase (and sub-phase):

Phase A: training of the participating laboratories (phase A1), test method transfer to the trained laboratories and verification of the Test Method Protocols (phase A2).

Phase B: assessment of the protocol performance by testing chemicals, under blind conditions, in all the laboratories.

Since 15 of the 24 chemicals selected would be tested three times in each laboratory (to evaluate the WLR, see above), and 9 would be tested only once, it was decided to separate the Phase B into two phases: first, Phase B Stage I with the 9 chemicals being tested once at each site, and Phase B Stage II for the remaining 15 chemicals tested three times at each site. The laboratories were required to prepare and submit an interim report at the end of the Phase B Stage I, with the results being evaluated by the VMG before allowing the laboratories to proceed to the next phase. This provided an additional review and control point, following the initiation of the testing phase, in order to verify that no serious issues were arising before the bulk of the testing was performed.

4. Selection of Test Chemicals

Reference documents:

- **Chemical Selection report (Appendix 3)**
- **Selection of the Phase B Stage I chemicals (Appendix 4)**

The test chemicals for validation study were selected by an independent Chemicals Selection Group (CSG) appointed by ECVAM and chaired by Dr. Thomas Cole (ECVAM). In addition to the chair, the CSG was composed of Dr. Luca Tosti (ECVAM); Dr. David Basketter (chair

of the VMG) and Dr. Bill Stokes (NICEATM/ICCVAM). The general strategy for the chemical selection was presented and approved by the VMG at its 1st meeting. The final list of chemicals was presented and endorsed by the VMG at its 4th meeting, before the initiation of the blind testing phase.

To facilitate the chemical selection process and in view of the intended use of the experimental data generated in this study to support future activities on data integrations to achieve ultimately full replacement, the VMG decided to use a common set of test chemicals for the three test methods under evaluation.

Two recognised databases provided a convenient master source of authoritative data for selection of eligible substances:

1) The ICCVAM database of 103 chemicals, subsequently supplemented with unpublished additions, provided by NICEATM.

2) The LLNA database of 341 chemicals, extracted by the CSG from published compilations (Gerberick et al., 2005; Kern et al., 2010).

A primary eligibility criterion for the chemical selection was the availability of robust *in vivo* data to allow a proper comparative evaluation of the *in vitro* results. In particular, availability of both LLNA and GPMT *in vivo* data, with concordance of their corresponding skin sensitisation classification as an assurance of quality, formed the basis for short-listing candidate chemicals. Availability of accepted human data was adopted as a secondary criterion, in cases where there would otherwise be an insufficient number of sub-sets of eligible chemicals as determined by the primary criterion.

During their respective development and optimisation, the three *in vitro* methods had been used to evaluate certain of the chemicals listed in these databases, as described in the original submissions of the methods to ECVAM. Acknowledging this, the chemical selection for this study was designed to include:

- A small quota of "tested" substances (i.e. substances reported as being previously tested by the method in the original submission to ECVAM)

- A majority of "untested" substances (i.e. substances not being reported as previously tested in the original submission to ECVAM)

The preferred ratio of tested to untested chemicals was set in advance at around 1:2 by the VMG. The final selection was composed of 9 tested chemicals and 15 untested chemicals for the h-CLAT.

To ensure parity between the three *in vitro* methods (DPRA, h-CLAT and MUSST) evaluated in the validation study, the only "tested" chemicals that were considered were those already tested by all three methods and that have been correctly predicted by each method with respect to the *in vivo* classifications. The only exception to this criterion was the inclusion in the final list of 2-Mercaptobenzothiazole, a chemical previously tested in the DPRA and in the h-CLAT but not in the MUSST.

Inclusion in the final list of a proportion of chemicals already successfully tested provided an opportunity to confirm the reproducibility of the test method with these chemicals when tested under blind conditions and by other laboratories.

Applying the primary criterion (available and concordant LLNA and GPMT data), the source databases yielded 11 eligible chemicals reported as previously tested in all three methods, all of which were correctly classified by the three methods. Applying the primary criterion to the 215 untested chemicals found in the original lists identified another 24 eligible substances. Therefore, in collaboration with NICEATM, 8 additional untested chemicals were identified from an unpublished updated version of the ICCVAM database, increasing the total number of candidates to 43, to provide an adequate and practical shortlist.

This list of 43 substances was reduced to 24 for different pragmatic reasons (see Appendix 3) and by expert judgment by the CSG. The final selected chemicals covered the range of sensitisation potency (*i.e.*: extreme, strong, moderate, weak). Inclusion of a small subset of substances known to be misclassified or classified inconsistently by the *in vivo* tests made provision for evaluating whether these limitations were shared by the *in vitro* tests. Furthermore, the chemical selection aimed for a balance of physical states(solid versus liquid) and avoided structural analogues, unless contrasting skin sensitisation potential was evident (e.g. 1-Thioglycerol (S) and Glycerol (NS)).

The shortlist of 43 chemicals and its refinement to the final selection of 24 chemicals (together with detailed comments about the choices made) can be found in Appendix 3.

The final selection included 9 LLNA performance standards (PS) reference chemicals (OECD, 2010). In particular, nickel chloride and xylene (both with ambiguous or inconsistent *in vivo* classification from LLNA and GPMT, but known human response) were considered eligible under the secondary selection criterion. Nickel chloride (human positive, GPMT positive, LLNA negative) is accepted as a PS true positive reference chemical (i.e., LLNA false negative). Xylene (human negative, LLNA positive) is accepted as a PS true negative reference chemical (i.e., LLNA false positive). In addition, Kathon CG, a commercial aqueous mixture including 1.2% CMI (5-chloro-2-methyl-4-isothiazolin-3-one) was selected, making exception to a general preference for pure substances with discrete chemical composition. CMI is a LLNA PS reference chemical of extreme sensitisation potency, and the Kathon CG commercial preparation is a recognised source.

Table 2: List of the 24 chemicals selected for the coded testing phase.

	Chemical Name	CAS#	State	LLNA	LLNA potency category	GP	EC3	GHS potency category	DPRA R&D result	hCLAT R&D result	MUSST R&D result
SENSITISERS	Beryllium sulfate	7787-56-6	Solid	+	extreme	+	0.001	1A			
	Kathon CG (1.2% CMI)	26172-55-4	Liquid	+	extreme	+	0.009	1A	+	+	+
	Benzoquinone	106-51-4	Solid	+	extreme	+	0.0099	1A	+	+	+
	4-Phenylenediamine	106-50-3	Solid	+	strong	+	0.11	1A	+	+	+
	Chlorpromazine HCl	69-09-0	Solid	+	strong	+	0.14	1A			
	Chloramine T	127-65-1	Solid	+	strong	+	0.4	1A			
	Formaldehyde	50-00-0	Liquid	+	strong	+	0.61	1A	+	+	+
	2-Mercaptobenzothiazole	149-30-4	Solid	+	moderate	+	1.7	1A	+	+	
	Benzylsalicylate	118-58-1	Liquid	+	moderate	+	2.9	1B			
	1-Thioglycerol	96-27-5	Liquid	+	moderate	+	3.6	1B			
	Dihydroeugenol	2785-87-7	Liquid	+	moderate	+	6.8	1B			
	Nickel chloride	7718-54-9	Solid	-	no category[1]	+		1B			
	Benzylcinnamate	103-41-3	Solid	+	weak	+	18.4	1B			
	Imidazolidinylurea	39236-46-9	Solid	+	weak	+	24	1B	+	+	+
	R(+)-Limonene	5989-27-5	Liquid	+	weak	+	69	1B			
	Methylmethacrylate	80-62-6	Liquid	+	weak	+	90	1B			
NON-SENSITISERS	Glycerol	56-81-5	Liquid	-	no category	-		NC[3]	-	-	-
	2,4-Dichloronitrobenzene	611-06-3	Solid	-	no category	-		NC			
	Benzyl alcohol	100-51-6	Liquid	-	no category	-		NC			
	Methylsalicylate	119-36-8	Liquid	-	no category	-		NC	-	-	-
	Isopropanol	67-63-0	Liquid	-	no category	-		NC	-	-	-
	Dimethylisophthalate	1459-93-4	Solid	-	no category	-		NC			
	4-Aminobenzoic acid	150-13-0	Solid	-	no category	-		NC			
	Xylene	1330-20-7	Liquid	+	weak[2]		95.8	NC			

[1] False negative in the LLNA

[2] False positive in the LLNA

[3] NC: Not Classified

The subset of 9 of the 24 chemicals which were to be tested only once at each site (as 15 chemicals were shown to be sufficient for the evaluation of the WLR, see above) was selected by a stratified random sampling procedure, to ensure a consistent distribution of potencies in the subsets of 15 and 9 chemicals (see Figure 2 and Appendix 4 for details). For the purpose of the stratification, the chemical Nickel chloride, false-negative in the LLNA, was assigned to the "moderate" category by weight of evidence. Similarly, Xylene, a false-positive in the LLNA, was assigned the "no category" class (see Appendix 4).

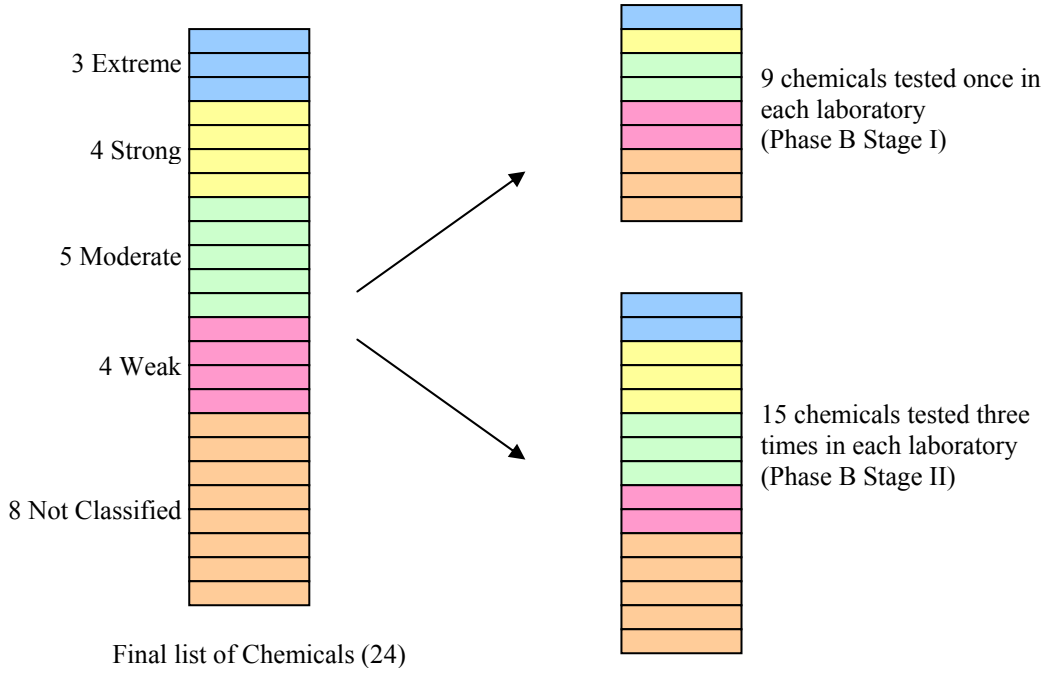

Figure 2: Stratified random sampling of the 24 chemicals to identify the 9 chemicals that were tested once and the 15 chemicals that were tested three times each.

Table 3: List of the phase B1 chemicals tested only once in each laboratory.

	Chemical Name	CAS#	State	LLNA	LLNA potency category	GP	EC3	GHS potency category	DPRA R&D result	hCLAT R&D result	MUSST R&D result
SENSITISERS	Benzoquinone	106-51-4	Solid	+	extreme	+	0.0099	1A	+	+	+
	4-Phenylenediamine	106-50-3	Solid	+	strong	+	0.11	1A	+	+	+
	1-Thioglycerol	96-27-5	Liquid	+	moderate	+	3.6	1B			
	Dihydroeugenol	2785-87-7	Liquid	+	moderate	+	6.8	1B			
	Imidazolidinylurea	39236-46-9	Solid	+	weak	+	24	1B	+	+	+
	Methylmethacrylate	80-62-6	Liquid	+	weak	+	90	1B			
NON-SENS.	Glycerol	56-81-5	Liquid	-	no category	-		NC	-	-	-
	2,4-Dichloronitrobenzene	611-06-3	Solid	-	no category	-		NC			
	Benzyl alcohol	100-51-6	Liquid	-	no category	-		NC			

Table 4: List of the phase B2 chemicals tested in three independent experiments by each laboratory.

	Chemical Name	CAS#	State	LLNA	LLNA potency category	GP	EC3	GHS potency category	DPRA R&D result	hCLAT R&D result	MUSST R&D result
SENSITISERS	Beryllium sulfate	7787-56-6	Solid	+	extreme	+	0.001	1A			
	Kathon CG (1.2% CMI)	26172-55-4	Liquid	+	extreme	+	0.009	1A	+	+	+
	Chlorpromazine HCl	69-09-0	Solid	+	strong	+	0.14	1A			
	Chloramine T	127-65-1	Solid	+	strong	+	0.4	1A			
	Formaldehyde	50-00-0	Liquid	+	strong	+	0.61	1A	+	+	+
	2-Mercaptobenzothiazole	149-30-4	Solid	+	moderate	+	1.7	1A	+	+	
	Benzylsalicylate	118-58-1	Liquid	+	moderate	+	2.9	1B			
	Nickel chloride	7718-54-9	Solid	-	no category	+		1B			
	Benzylcinnamate	103-41-3	Solid	+	weak	+	18.4	1B			
	R(+)-Limonene	5989-27-5	Liquid	+	weak	+	69	1B			
NON-SENS.	Methylsalicylate	119-36-8	Liquid	-	no category	-		NC	-	-	-
	Isopropanol	67-63-0	Liquid	-	no category	-		NC	-	-	-
	Dimethylisophthalate	1459-93-4	Solid	-	no category	-		NC			
	4-Aminobenzoic acid	150-13-0	Solid	-	no category	-		NC			
	Xylene	1330-20-7	Liquid	+	weak		95.8	NC			

It is important to note that in view of the primary objective (i.e. assessment of the reliability of the test method) of the blind testing phase, no consideration was given to the suggested applicability domain of the h-CLAT (or the other two *in vitro* methods evaluated in the validation study) in the chemical selection. For example the list contains two well characterised pre-haptens (4-Phenylendiamine and R(+)-Limonene) and a well characterised pro-hapten (Dihydroeugenol).

Wherever possible, the selection of chemicals was predicated on a consistency of evidence from human experience, guinea pig tests and the LLNA. Thus for the majority of chemicals chosen as skin sensitisers, clear human evidence of allergic contact dermatitis was supported by positive guinea pig maximisation test and/or Buehler test results coupled with a positive LLNA. Two clear exceptions to this were xylene, a false positive in the LLNA and nickel chloride, a false negative in the same assay. For xylene, the human evidence of skin sensitisation, while not entirely absent, is of a similar scale to petrolatum and does not meet the criteria for classification as a skin sensitiser. For nickel chloride, the human evidence of skin sensitisation is abundant.

Two other points should be noted. First, for a few substances (e.g. benzoquinone and dichloronitrobenzene) human data is very limited, but does not contradict the results from animal tests. Second, for some of the chemicals which have a "Not Classified" categorisation, there is evidence that they are human sensitisers (benzyl alcohol, isopropanol, xylene, methyl salicylate, 4-aminobenzoic acid) and this may be kept in mind when interpreting the test results with respect to the test's predictive capacity. However, important also to remember that this evidence is of limited nature (e.g. isolated case reports), that they are considered to be

substances with insufficient human evidence to warrant classification according to REACH and United Nations GHS guidelines.

5. Chemicals purchase, coding and distribution

Reference document:

- **Chemical coding, sampling and shipping procedures (Appendix 5)**

The chemicals used for the transfer phase (study phase A Stage II) were purchased by the trained laboratories on the basis of the instructions reported in the training and transfer plan issued by Kao and Shiseido and approved by the VMG.

ECVAM was responsible for purchasing, coding and distributing the chemicals for the blind testing phase (Phase B) to the laboratories participating in the study. The selected chemicals were sourced from Sigma Aldrich. Aliquots of the chemicals were prepared, coded, labelled and properly stored in the chemical repository of EURL ECVAM before distribution to the test laboratories.

a. Solvent Compatibility assessment

To avoid possible problems with the solubility of the test chemicals during the blind testing phase, all chemicals underwent an assessment of solubility at ECVAM, following the "solvent selection" procedures as described in the test methods' SOP. In all cases, at least one suitable solvent was identified. The solvent(s) identified for each chemical was not communicated to the testing facilities, since the SOP required the testing laboratory to be responsible for the identification of the suitable solvent(s) (see below).

b. Coding/Decoding

A randomly generated code was assigned to each aliquot, unique for each method, laboratory and experiment. For the assessment of the WLR, three separate vials of each chemical from the list of 15 were sent to the laboratories, each assigned a different two-letter code. A number added to the code (1, 2 or 3) distinguished the three sets of 15 chemicals to be used for the individual experiments, and the laboratories were instructed never to mix chemicals labeled with different experiment numbers in the same run. This ensured that the three evaluations of the corresponding chemicals were performed in independent runs in order to best provide data suitable for a proper evaluation of the WLR.

The codes for all chemical aliquots were recorded in a database (Excel spreadsheet format) prepared and maintained by the Chemical Selection Group. The identity of the chemicals to which the codes were assigned was not disclosed to the laboratories, and was kept confidential from the VMG and the biostatisticians until the end of the study. Copies of the tables, showing the codes assigned to the same chemicals but which do not identify the chemicals by name, were prepared to be given to the biostatistician for analysis of the reproducibility, and to the ECVAM study coordinator(s) to assist the VMG consideration of the experimental data. A detailed description of the chemical coding and distribution procedures applied is provided in Appendix 5.

The Chemical Selection Group provided the VMG with the final decoding list for the chemicals only once all the experimental data had been generated by the laboratories, quality checked by ECVAM and analysed by the biostatistician for the assessment of the WLR and BLR. The decoding list was then used by the VMG and the biostatistician to analyse and assess the information generated in this study on the predictive capacity of the h-CLAT.

c. Emergency procedure implemented at the laboratories during the blind testing phase.

An emergency procedure was established to allow the laboratories to obtain the necessary chemical safety information in the event of an accident. Individual sealed envelopes, each containing a Material Safety Data Sheet (MSDS) relating to one specific chemical and labelled with the corresponding code, were sent with the test chemicals to a named recipient not associated with the testing (typically the Safety Officer) at each laboratory, with instructions to return the unopened MSDSs to ECVAM upon completion of the testing phase.

During the testing of the h-CLAT, no such incident was reported and none of the envelopes had to be unsealed, and all envelopes were returned to ECVAM, sealed, upon completion of the validation study.

The laboratories were instructed to treat all coded test chemicals as potential sensitisers and to dispose of laboratory waste as toxic waste.

6. Data management

Reference documents:

- **h-CLAT reporting templates (Appendix 6 and 7)**
- **h-CLAT QC template (Appendix 8)**

Prior to the start of the study standard reporting templates were prepared and distributed to the participating laboratories. These templates were developed by the lead laboratories, and comprised three different Excel documents:

- A template to report the cytotoxicity assays, and calculate the CV75 values (Appendix 6)
- A template to report the evaluation runs and calculate the RFI values for all tested concentrations (Appendix 7)

In addition the laboratories were provided a template they could use to compile the results of the three evaluation runs required to apply the prediction model (summary template). Following a request from the laboratories, another template was prepared by the lead laboratories to assist in the calculations of the EC150 and EC200 values

The templates each contained formulae verified by the ECVAM biostatistician in a documented exercise before they were distributed. The laboratories were asked to use the templates during the Transfer phase to compile and communicate their results to the lead laboratories. In this way they could familiarise themselves with them and this ensured that the

instructions included in the templates were clear and understood before the start of the blind testing.

All the results from the Phase B were submitted by the laboratories directly and exclusively to ECVAM by e-mail. Since the templates were not sent protected/locked, upon receipt at ECVAM each completed template was formally quality controlled according to a checklist (see Appendix 8). The quality check focused on the acceptance criteria for the run and for each of the chemicals' data to ensure the results were valid, and confirm the correct selection of the CV75 concentration. Once completed, the checklists were scanned as a PDF file. The templates and checklists were then added to the official results folder of the study.

For the statistical analyses, a summary template was designed by the statistician, and the results were transferred to this template by ECVAM. Preparation of this summary template contained internal checks that ensured that no transcription errors were made in the transfer of the results. As an additional check, the final conclusions/outcomes for each chemical were then compared to the conclusions/outcomes in the reports sent by the laboratories.

7. Statistical Analysis of Experimental Data

Reference document: Experimental Design (Appendix 2)

a. Analysis performed on the experimental data

A detailed statistical analysis plan was produced and agreed by the VMG before the start of the testing phase (see Appendix 2 for full details). It was stipulated that only data from the valid experiments would be considered for statistical analyses. Failed runs and experiments were also reported in order to assess their occurrence and frequency.

The statistical analysis on the test method's reproducibility focused on the concordance of classification, sensitizer (S) versus non-sensitisers (NS). Reproducibility was evaluated with respect to both WLR and BLR.

Additionally the raw data were analysed using descriptive statistical analysis however the information generated by this analysis was not used by the VMG, on its own, to draw conclusions on the WLR and BLR of the h-CLAT.

b. Criteria for Assessing the Study Outcome

The VMG considered it appropriate to define in advance indicative assessment criteria to be used to enable a transparent judgment and decision on the performance of the test methods in consideration of the study primary goal.

When defining such indicative criteria the following factors were considered important:

1. the background and specific objectives of the validation study;
2. the standards of performance that can realistically be expected from an *in vitro* test and standards of performance which have been considered acceptable in previous validation studies;

3. the proposed used of the *in vitro* tests (i.e. as a partial replacement method to become part of a testing strategy or a toolbox of tests to be used in combination); and
4. the statistical power of the design of the validation study.

In consideration of the above, and in particular the anticipated use of such test methods within integrated testing strategies or a toolbox of tests to be used in combination, the target performance for this study was set at 80% for the Between Laboratory Reproducibility and 85% for the Within Laboratory Reproducibility.

The target performance informed the sample size calculation and is consistent with the reproducibility values, in terms of reproducibility in the concordance of predictions, elaborated from the data submitted by the test developers, in which the h-CLAT showed a between laboratory reproducibility of 87% (see Appendix 2 for further details).

Test definition (Module 1)

Reference Documents:

- **h-CLAT SOPs (Appendices 9 and 10)**

1. Intended purpose of the test method

The h-CLAT is a potential partial replacement *in vitro* test method designed to be part of a non-animal test battery or integrated testing strategy for assessing the skin sensitisation potential of chemicals. As such it could contribute to the reduction of the number of animals used for skin sensitisation testing and to the replacement of current regulatory *in vivo* tests for skin sensitisation hazard classification and labelling (OECD TG 406, OECD TG 429, OECD TG 442A, OECD TG 442B). The information the h-CLAT generates can already be used in a weight of evidence approach to support regulatory decision making (e.g. to better characterise equivocal responses in *in vivo* studies). For example, REACH permits the use of methods under validation for this purpose.

2. Evidence demonstrating the need of the test method

This test method is of importance in regard of the European Union ban on *in vivo* testing of cosmetic and toiletry ingredients and products implemented by the Seventh Amendment to the Cosmetics Directive and for the European Regulation on Registration, Evaluation, and Authorisation of Chemicals (REACH) that requires evaluation of a large number of chemicals. The successful validation of the h-CLAT and other *in vitro* assays (e.g. assessing peptide reactivity) would support their use in the assessment of a chemical's skin sensitisation potential and reduce or eliminate the need for animal testing for skin sensitisation

.

3. Status of development of the test method

The h-CLAT was developed in a collaborative effort by Kao & Shiseido.

Extensive work had been undertaken since 2003 to optimise the h-CLAT protocol. This included four inter-laboratory studies, during which critical aspects of the h-CLAT protocol were investigated and optimal procedural conditions (including criteria for the selection of a suitable lot of cells and Fetal Bovine Serum (FBS) were established (Sakaguchi *et al*., 2005; Ashikaga *et al.,* 2008; Kosaka *et al*., 2008; Sono *et al.,* 2008; Mizuno *et al*., 2008). In the context of this optimisation work the h-CLAT has previously been successfully transferred to 5 Japanese laboratories. Subsequently, the technology was transferred to an additional 3 laboratories for the purpose of a ring trial organised by Cosmetics Europe (formerly Colipa) (Sakaguchi et al., 2010). This study aimed at providing insights in the likely within-laboratory reproducibility, transferability, and between laboratory reproducibility of the test method by evaluating a set of 15 substances with varying degrees of sensitisation potency, including non-sensitisers.

In their test submission to ECVAM (received in April 2009) Kao and Shiseido reported test data for 100 chemicals (72 sensitisers: 8 extreme, 16 strong, 25 moderate, 23 weak and 28

non-sensitisers according to LLNA results) with an accuracy of 84% (sensitivity 88% and specificity 75%). When compared to human data the h-CLAT was reported to have an accuracy of 80% (sensitivity 83% and specificity 69%).

4. Scientific basis – biological and/or mechanistic relevance of the h-CLAT

Immature epidermal dendritic cells, the Langerhans cells (LCs), and dermal dendritic cells (DCs) are important mediators in the skin sensitisation process since they are capable of presenting the hapten-protein conjugate to responsive T lymphocytes in the lymph nodes draining the site of exposure (Kimber and Cumberbatch, 1992). The maturation process of LCs and DCs from antigen processing cells to antigen presenting cells is considered a Key Event in the acquisition of skin sensitisation as also documented in the Skin Sensitisation Adverse Outcome Pathway recently developed by the OECD (OECD, 2012a, 2012b).

This maturation process involves the modulation of the expression of cell surface phenotypic markers, those most commonly reported being CD54, CD80, CD86 and major histocompatibility complex [MHC] class II (Galvao dos Santos et al., 2009). This knowledge has been exploited in the development of *in vitro* tests based on the use of DC-like immortalised cell-lines to screen the skin sensitization potential of chemicals.

As described in detail below, the h-CLAT measures by flow-cytometry the modulation of CD86 and CD54 protein markers on the surface of THP-1 cells following controlled exposure to the chemical. THP-1 is a human monocytic leukemia immortalised cell line used as a surrogate of DCs.

5. Protocol of the test method

The detailed updated h-CLAT test method protocol used during the Phase BII blind testing phase is described in Appendix 9.

The h-CLAT is an *in vitro* cell based assay in which the expression of CD86 and CD54 surface markers in THP-1 cells are quantified by flow cytometric analysis following 24 hour exposure to 8 concentrations of the test substance. A propidium iodide (PI) viability assay is performed prior to the main experiment for the calculation of the CV75 value (estimated concentration of test substance yielding 75% cell viability) which is then used for the selection of the final concentrations.

a) Selection of the solvent

The h-CLAT SOP prescribes a procedure to select the appropriate solvent.

In this study, since the three vials containing each of the chemicals that were tested three times in each laboratory (10-24) were coded independently, the laboratories had to repeat the solvent selection procedures three times for these chemicals. The choice of solvent was, for the

majority of cases, consistent within and between the laboratories (see Table 15), suggesting that the procedure in the SOP is adequately described.

When they occurred, differences in solvent selection did not correlate with differences in predictions obtained.

b) Determination of the CV75

To conclude on the sensitising potential of a test substance based on the induction of cell surface markers, the h-CLAT requires the evaluation of a specific range of chemicals concentrations, which is based on the CV75 value, i.e. the concentration of chemical that allows 75% of THP-1 cell survival (25% cell toxicity).

The h-CLAT SOP contains a section that explains how this value should be derived, using flow cytometry and Propidium Iodine (PI) staining.

c) Evaluation runs

One complete experiment requires the generation of three independent evaluation runs. Each evaluation run of the h-CLAT measures the induction of two independent cell-surface markers: CD54 and CD86.

Eight concentrations of chemicals are tested, based on the calculated CV75, for a final range that corresponds to 1.2xCV75 (highest dose) to 0.335xCV75 (smallest dose). The content of each well of treated cells is split in three aliquots for staining with three different antibodies:

- FITC-labelled anti-CD86, to measure the expression (Mean fluorescence Intensity, MFI) of the CD86 cell marker
- FITC-labelled anti-CD54, to measure the expression (MFI) of the CD54 cell marker
- FITC-labelled mouse IgGI ("isotype"), used to correct the fluorescence CD86 and CD54

For each of the chemical concentration, two values are then calculated: the CD86 and the CD54 RFI (relative fluorescence intensities compared to an untreated control well), using the following formula:

$$\text{RFI} = \frac{\text{MFI of chemical-treated cells} - \text{MFI of chemical treated isotype control cells}}{\text{MFI of vehicle control cells} - \text{MFI of vehicle isotype control cells}} \times 100$$

The whole evaluation run (each with the 8 test chemical concentrations and the staining for the two markers) is repeated THREE TIMES, in independent runs, and a final prediction is derived based on the obtained RFI values using the following rule:

If the RFI of CD86 is ≥ 150 at any dose in at least 2 of 3 independent evaluation runs

AND/OR

If the RFI of CD54 is ≥ 200 at any dose in at least 2 of 3 independent evaluation runs,

THEN

The chemical is considered to be as a sensitiser.

Otherwise, the chemical is considered to be a non-sensitiser.

The following two tables show examples how the prediction model is applied:

	conc.	CD86			CD54			Cell viability(IgG)		
sample code	(µg/mL)	1	2	3	1	2	3	1	2	3
D3SS	139.5	96	102	87	183	177	149	92	95	94
	167.4	104	103	82	160	209	166	92	94	96
	200.9	95	125	85	188	320	148	87	91	92
	241.1	99	129	104	158	200	189	85	89	85
	289.4	115	107	85	151	220	118	83	86	81
	347.2	108	85	81	150	271	156	76	85	81
	416.7	92	102	89	110	283	127	72	81	73
	500.0	122	126	102	129	249	127	72	81	65

	conc.	CD86			CD54			Cell viability(IgG)		
sample code	(µg/mL)	1	2	3	1	2	3	1	2	3
D3QX	136.4	100	95	90	113	112	136	96	96	97
	163.6	110	102	96	115	114	132	96	96	97
	196.4	113	109	106	130	114	142	96	95	96
	235.6	100	93	116	128	121	144	95	96	95
	282.8	104	98	123	138	149	177	93	93	85
	339.3	116	99	166	150	163	212	87	94	60
	407.2	136	144	145	177	308	233	85	67	81
	488.6	106	100		167	267		69	71	18

For chemical D3SS, the three runs were negative for CD86, and only the second run showed induction of CD54 above the threshold (200) indicated by the red numbers. As the conditions to be considered a sensitiser are not met, the h-CLAT classifies this chemical as non-sensitiser. In contrast, for chemical D3QX, the chemical is labelled as a sensitiser since two of the three runs showed induction of CD54 above the threshold.

For chemicals considered to be sensitisers, this threshold-based prediction model allows the calculation of two Effective Concentrations (EC) values, EC150 for CD86 and EC200 for CD54, i.e. the concentration at which the chemical induces a RFI of 150 or 200. Note that since the prediction model requires only at least one of the two markers to be positive, it is not always possible to derive both EC150 and EC200 values for positive chemicals.

The SOP submitted to ECVAM did not contain descriptions or formulae to calculate the relevant EC values. A proposal was subsequently made by the lead laboratories and, with the consent of the VMG, included in the SOP for the blind testing phase covering most of the possible concentration-response scenarios. Although the h-CLAT SOP was not developed or adapted with the aim of a precise derivation of EC values, the EC values obtained during the studies are reported in the within- and between-laboratory reproducibility sections, below.

d) Modifications made to original protocol prior to the initiation of the validation study

Following the ECVAM assessment of the original h-CLAT protocol, revisions to the SOP were made in collaboration with Kao and Shiseido prior to the initiation of the validation study to clarify sections of the SOP which were felt not clear enough and could thus impact on the interpretation and performance of the protocol. This work resulted in the release of SOP Version 2 which Kao and Shiseido then used to train the naïve laboratories (Study Phase AI). SOP Version 3 aimed to clarify procedural aspects following the training stage and subsequent discussions with the laboratories and was released for use during the transfer of the method (Study Phase AII). SOP Version 4, then 5, were prepared and issued based on the experience of the transfer phase at IVMU and Bioassay. Version 5 was released for use during the blind testing phases. However, following the experience of Phase B1, one clarification, resulting in Version 6, was added to Version 5 to address special cases that occurred during the testing phase but were not correctly covered by Version 5 of the SOP.

The SOP includes a requirement for the adoption of the evaluation runs that produced a negative response: the viability at the highest concentration should be less than 90%. This is meant to prevent the possibility that, in the event of an incorrect CV75 determination, inappropriate concentrations are used in the evaluation runs. When cell viability criteria are not met, the SOP requires the run to be considered invalid and the CV75 re-evaluated.

However, some chemicals are not cytotoxic even at the maximum soluble or allowed concentrations (1000 μg/ml in DMSO, and 5000 μg/ml in saline). The instructions to re-evaluate the CV75 for these chemicals in the event they produced negative results was questioned (the issue was reported by one of the study directors). Communications with the lead laboratories clarified that, historically, the results in these cases were accepted, and this was clarified in the SOP (version 6) and applied for this study.

All modifications to the SOP were made in close collaboration with the lead laboratories to make sure that such modifications did not affect the outputs of the h-CLAT procedure itself, in ways which could invalidate or compromise the historical data generated with the h-CLAT prior to the submission to ECVAM.

Version 7 was prepared and released by the VMG at the end of the validation study to include the provision to allow performing more than one run with the same chemical on the same day and the conditions for this. This version of the SOP is the one suggested for future use of the h-CLAT.

A summary of the protocol revisions is outlined in Table 5.

Table 5: Description of the main modifications made to each version of the SOP.

	Version 1	**Version 2**	**Version 3**	**Version 4**	**Version 5**	**Version 6**	**Version 7**
I. Establishing the method	Added list of reagents Added cells thawing, freezing and maintenance instructions Added cell doubling time evaluation procedures	Clarified concentrations of Penicillin and Streptomycin	Added BSA, volumetric flasks, and glass vials suppliers Clarified centrifugation speeds	Added that more than one concentration of chemicals can be used in the reactivity checks, with instructions	Added that if in the first reactivity check SLS is positive, the reactivity check can be repeated before dismissing the batch of cells.	Changed the negative chemical in the reactivity check from SLS to Lactic acid although testing SLS was still requested	Removed requirement to test SLS in the reactivity check.
2. Dose finding assay (PI assay)	Clarified reactivity check procedure Added sample plate setup for CV75 determination assay Added cell viability calculation formula Added example flow cytometry	Added minimuml chemical concentrations for dose-finding assay Added maximum number of the chemicals that can be tested at the same in the dose-finding assay	Clarified chemicals solubilisation procedure Clarified that the cells and reagents should be kept in the cold and in the dark during the staining procedures				

	plots Set the number of CV75 determination assays to be performed						
3. Experiment (measuring CD86/CD54 expression)	Clarification of the CD86/CD54 staining procedure, including sample plate setup and master mixes preparations Clarification of the flow cytometry acquisition procedure, including Excitation and Emission wavelengths Description of the cytometer plot and gate settings, including examples. Instructions on	Clarified starting dose in CD86/CD54 staining when CV75 could not be determined	Clarified chemicals solubilisation procedure Clarified that the cells and reagents should be kept in the cold and in the dark during the staining procedures Clarified that the two independent PI assays should be performed on different days Clarified that the negative controls should include both a medium control sample and a vehicle	Clarification that the CD86/CD54 induction measurement should be done using the same batch of THP-1 cells as used for the CV75 determination Added that more than one dose of positive control can be used for the experiment. Added setting instructions for the FACSAria (Becton Dickinson) flow cytometer Added instructions of			Clarification on the possibility to perform more than one run per chemical on the same day and the conditions for this.

	what to do in case of unqualified tests		control	proper maintenance of the flow cytometers			
4. Data analysis and interpretation	Clarification that the cell viability at the highest does should be more than 90% only in the case of negative results. Inclusion of a maximum allowed number of runs allowed to test a given chemical.			Added requirement for qualified testing: In the DMSO vehicle control, RFI values compared to the medium control of both CD86 and CD54 should not exceed the positive criteria ("CD86 > 150 and CD54 > 200") Added section on "abnormal values"	Added requirement for qualified testing: For both medium and DMSO controls, the MFI ratio of both CD86 and CD54 to isotype control should be > 105%.	Clarified that when 5000 µg/mL in saline, 1000 µg/mL in DMSO, or highest soluble dose is used as the maximal test concentration instead of CV75-based dose, the data for test chemical are accepted regardless of the cell viability	
5. Data reporting	Instructions on how to calculate EC150 and EC200 Instructions for data reporting.						

e) Acceptance criteria

The h-CLAT SOP, as submitted to ECVAM by the lead laboratories, contained a set of acceptance criteria for the evaluation runs to determine whether the results obtained are valid.

Following the VMG evaluation of the SOP prior to the start of the study, it was not felt necessary to propose additional criteria and no modification was made to this section of the SOP.

The testing laboratories raised no issues with respect to meeting these criteria during the study as demonstrated by the relatively low rate of invalid runs which was consistent between the laboratories (see Tables 19 and 20).

f) Requirement for a third evaluation run

The h-CLAT SOP requires three independent evaluations runs to be performed to apply the prediction model. Since the classification is driven by the majority of the conclusions of the three runs (i.e. at least two out three) the VMG discussed prior to the start of the study whether the third run was required when the first two runs give consistent conclusions (i.e. two negatives or two positives runs) since in that case the result of the third run can have no effect on the final classification. Although it was appreciated that eliminating the requirement of the third run in such cases would decrease the amount of work involved, the VMG did not consider necessary to modify this element of the h-CLAT SOP, thus maintaining the requirement for a third run. This allowed the generation a full dataset for the chemicals evaluated in the study.

Additional analysis requested and performed in parallel by Japanese biostatisticians, (see Appendix 13) suggests that the third run is important in case of positive predictions to refine the derivation of the EC values. This may be important for the purpose of using the h-CLAT for informing potency predictions.

However, the VMG agreed that, for future use of the h-CLAT in cases where it is used for hazard identification purposes, the SOP could be adapted to eliminate the need for the third evaluation run in case the first two runs are consistent without compromising test sensitivity and specificity.

6. Known technical limitations and drawbacks of the test method

The following limitations were reported by the lead laboratories in their submission to ECVAM:

- The test method requires expensive equipment (flow cytometer).
- When cell viability is less than 50%, flow cytometric analysis cannot be conducted correctly because of artefacts arising from diffuse labelling cytoplasmic structures due to cell membrane destruction.
- To avoid osmotic stress, the maximum concentration of test chemical cannot exceed 5000 μg/mL.
- Since the h-CLAT uses a fluorescently labelled antibody, self-fluorescence of test chemicals might interfere with the flow cytometry acquisition (e.g. Abietic acid).

- In the h-CLAT test method the test substance needs to be dissolved in a solvent compatible with cell culture conditions (i.e. saline or DMSO). In the submission to ECVAM chemicals reported not to be correctly classified because of solubility problems include: Hexil cinnamic aldehyde, Abietic acid and Phthalic anhydride.

In relation to this last point it is important to note that for the purpose of the current study none of the chemicals selected had to be discarded because of solvent incompatibility consistent with the preliminary solubility assessment performed by ECVAM (see Section 5a above). This was also consistent with the fact that no major problem was subsequently encountered by the laboratories during the solubility assessment for the blind testing phase.

7. Limitations in applicability

Some chemicals, designated "pro-haptens", require a metabolic transformation step to act as haptens capable of activating the immune system. If the *in vitro* test system (i.e. the THP-1 cells in the h-CLAT assay) does not provide adequate metabolic activity, the detection of such pro-haptens may be impaired. The THP-1 cells used as the test system in the h-CLAT assay have indeed been reported by the protocol submitter and others (for example Hennen *et al.,* 2011; Chipinda *et al*., 2011) to possess only a limited or no metabolic activity. Moreover, some chemicals (designated "pre-haptens") need a non enzymatic oxidation step to act as haptens. Depending on the chemistry involved and the oxidation conditions pertaining during the performance of the assay, such molecules might not be transformed into active haptens and will thus not be detected as potential sensitizers by the h-CLAT assay. For example, Benzoyl peroxide, Isoeugenol and Abietic acid are reported to be false negatives in the h-CLAT.

Assays using a threshold based prediction model (h-CLAT uses EC150 for CD86 or EC200 for CD54) have an inherent limitation in the detection of some weak sensitisers. The cell activation induced by very weak sensitizers may remain just below the thresholds set for a positive result and the molecules will not be correctly classified (e.g. 1-Bromohexane, Cyclamen aldehyde, Butyl glycidyl ether).

Finally, a general limitation of assays using suspension cell culture techniques concerns the solubility of the test substance in an aqueous environment compatible with cell survival. Therefore, when no toxic effect is observed up to the highest test concentration allowed in the assay, a careful examination of the solubility data should be performed to ensure that the test substance has indeed been solubilised. Negative test results obtained with poorly or insoluble substances should not be considered sufficient for classification decisions.

Conclusion of the Validation Management Team on Module 1

The h-CLAT test method addresses a biological mechanism, the modulation of the expression of cell surface phenotypic markers associated with the maturation of DCs, which is considered to be a key step in the skin sensitisation AOP.

Information provided by the h-CLAT on a chemical's potential to activate dendritic-like cells in vitro can complement information generated by other non animal approaches (*in silico, in chemico, in vitro*) designed to address chemical/biological mechanisms preceding and following the activation of DC in the sequence of events leading to the induction of skin sensitisation. This supports the use of the h-CLAT as mechanistically relevant element of a testing strategy for skin sensitsation.

The need for the method in the context of current EU regulatory requirements is evident. The proposed use of the methods as mechanistically-based partial replacement method to be used in conjunction with other data for hazard identification purposes is plausible.

The h-CLAT protocol (as revised prior to the initiation of the blind testing phase) proved to be generally robust for the purposes of this study, only minor clarifications were made to the SOP during the course of the study in relation to specific elements of the procedure and the data interpretation to minimise the sources of variability.

The amendments introduced were largely to resolve ambiguities and minor omissions in the original SOP in order to improve clarity and consistency of data generation and interpretation.

The VMG believes that the procedural clarifications to the h-CLAT original SOP, the supporting documents (including the original submission to ECVAM and associated scientific publications) and the current study findings adequately demonstrate the status of development and optimisation, the mechanistic basis, the intended use and, the regulatory relevance of the h-CLAT.

In conclusion, the VMG believes that Module 1, Test Method Definition, is now satisfied.

Transferability (Module 3)

Reference documents:

- **h-CLAT Training and Transfer Plan (11)**
- **List of additional available documents filed for the study and available on request (Appendix 15)**

1. General aspects

The h-CLAT can be performed in laboratories equipped with and experienced in cell culture and flow cytometry techniques. All apparatus/instruments and reagents needed for the performance of the method are readily available commercially.

Kao and Shiseido, being the lead laboratories and having jointly developed the method, were responsible for both the training of the personnel at the other testing facilities participating in the study and for overseeing and providing advice during the test method transfer in order to make sure that the procedure for performing the h-CLAT as described in the SOP was clearly understood and properly implemented.

Specialists from both Kao and Shiseido provided and evaluated training to the study personnel of Bioassay and EURL ECVAM at the Food and Drug Safety Center of the Hatano Research Institute, Kanagawa, Japan (Phase A Stage I). The newly trained personnel were then responsible for the transfer of the test method to their own laboratories under the supervision of the lead laboratories (Phase A Stage II).

The schedule for the training of these laboratories as well as the details of the transfer experiments were drafted by the lead laboratories on the basis of their experience with the test method. To demonstrate successful method transfer, the laboratories had to perform the test method procedure by testing in-house a number of chemicals and meet the transfer acceptance criteria as defined in the Transfer Plan (Appendix 11). The chemicals used for these qualification runs, as well as the criteria for a successful achievement of the transfer of the method, were selected by the lead laboratories, and approved by the VMG prior to the initiation of the training.

The chemicals used for this phase A were not supplied by ECVAM, but were purchased by the trained laboratories on the basis of the instructions set out in the training and transfer plan (Appendix 11). All chemicals were tested uncoded, and the results sent directly to the lead laboratories for evaluation.

2. Training

Bioassay and EURL ECVAM personnel received theoretical and procedural training at the Hatano Research Institute. The training sessions took place between March 1st and March 5th, 2010. The training included a theoretical component with discussions on different theoretical aspects (SOP; critical aspects of the procedure; data analysis and reporting), and practical sessions where the trainees observed and performed the different elements of the h-CLAT

SOP (PI assays, cell passaging, plating and treatment, flow cytometry staining and acquisition).

For the practical training sessions, two test chemicals were chosen by the lead laboratories for testing by the trainees: DNCB and SLS

Dose finding assay (cytotoxicity test):
To determine the appropriate test concentrations of the two sample chemicals (DNCB as positive chemical and SLS as negative chemical), cytotoxicity testing was conducted by each lab following a demonstration. The trainees from each laboratory conducted this experiment and the mean values were calculated to derive the test application concentrations. The CV75 of DNCB was 6.9 (6.6 and 7.3) mg/mL. The CV75 of SLS was 39.9 (38.1 and 41.8) mg/mL. The value for EURL ECVAM and Bioassay were very similar, and consistent with lead laboratories' historical data.

CD86 and CD54 measurements:
The trainees from each laboratory conducted CD86/CD54 measurement of DNCB and of SLS-treated THP-1 cells. In terms of DNCB, the requirement for qualified testing was satisfied in all experiments. The results of both naïve laboratories were, correctly, positive. For SLS, the expression levels of both CD86 and CD54 were below the cut-off points.

Trainings sessions for Bioassay and EURL ECVAM were judged to be successful by the trainers in a Statement on training outcome in the Training Report.

3. Transfer of the test method to the naïve laboratories

As part of the Training and Transfer plan, Kao and Shiseido requested that trainees then acquire the following list of chemicals to evaluate the establishment of the assay in their laboratories:

Table 6: List of transfer chemicals for the h-CLAT.

Chemical name	abbreviation	Cas#	Commercial source	Cat#	LLNA EC3	Potency category
2,4-Dinitrochlorobenzene	DNCB	97-00-7	Aldrich	237329	0.04	Extreme
Nickel sulfate	$NiSO_4$	10101-97-0	Sigma-Aldrich	N4882	4.8	Moderate
Phenylacetoaldehyde	PAA	122-78-1	Aldrich	107395	3	Moderate
Sodium lauryl sulfate	SLS	151-21-3	Sigma-Aldrich	L6026	14	Weak (false positive)
Lactic acid	LA	50-21-5	Sigma	L6661	NC	Non-sensitizer

The transfer of the h-CLAT was separated into phases covering the different steps of the SOP. After each phase, the laboratories were asked to report the results to the lead laboratories for analyses and permission to proceed.

The schedule was planned as follows:

Week	Experiments
1-2	Doubling time, cell viability, reactivity
3-4	Cytotoxicity test for the 5 chemicals
5	CD86 and CD54 expression (MFI raw data) and cell viability of control (untreated) cells
6-7	Valid data (3 runs) for DNCB and SLS
8-9	Valid data (3 runs) for the other three chemicals

The success of the transfer was judged based on the results obtained for the five chemicals, with the following criteria:

All three sensitizers (DNCB, Nickel Sulfate and phenylacetaldehyde) must produce a positive response in <u>both</u> CD86 and CD54 (i.e. for both CD86 and CD54, at least 2 out of 3 runs should be positive).

Both non-sensitizers (SLS, Lactic acid) must produce a negative response in both CD86 and CD54 (i.e. for both CD86 and CD54, at least 2 out of 3 runs should be negative).

Note that, in the case of the positive chemicals, the transfer criteria is different from, and more stringent than, the h-CLAT prediction model, where only one of the two markers has to be positive for the chemical to be declared positive (sensitiser).

The sections below summarize the results of the transfer phases at Bioassay and EURL ECVAM. Details on the results can be found on the corresponding transfer reports, available upon request (see Appendix 15).

1. <u>Transfer to Bioassay.</u>

Having allowed the laboratory the necessary time to acquire all the reagents and buy/expand/bank the THP-1 cells, the transfer experiments (described in the table above) started on the 10th of May, 2010. No major issues arose during the first steps of the transfer phase. DNCB, SLS, NiSO4 and PAA were correctly evaluated in consecutive runs.

However**,** problems arose in the testing of Lactic Acid, as the second and third runs showed unexpected CD54 activation. Eight additional runs were required to solve the issue and obtain the data required in the transfer plan. Comparison of different lactic acid batches did not suggest that this was an issue with a particular batch. A site visit of lead laboratories representatives was organised to investigate the problem at Bioassay. The visit did not identify major problems in the performance of the assay.

The main procedural issue arising from the site visit related to the setup of the flow cytometry at Bioassay, which used a micro plate autosampler. Since with this system the time lap between samples acquisition is longer than without the autosampler, the problem encountered might have been due to prolonged sample storage at room temperature leading to variations in the RPMI control results that might have caused the unexpected results with Lactic Acid.

To remedy this problem, Bioassay subsequently tested a single substance per plate, with its own control wells. While one plate was being analysed, the other one(s) were kept on ice in the dark. This setup reduced the throughput of the testing at Bioassay, but allowed the completion of the transfer phase, and was maintained during the whole testing phase.

Additional runs performed following the visit confirmed the consistent negative results with Lactic Acid, leading to approval of the transfer at Bioassay by the lead laboratories. Based on this recommendation of the lead laboratories, the VMG approved the transfer of the h-CLAT at Bioassay on the 16th of December 2010.

2. Transfer to EURL ECVAM

Having allowed time to acquire all the reagents and buy/expand/bank the THP-1 cells, the transfer experiments (described in the table above) started on 26th of April 2010. No major issues arose during the first steps of the transfer phase and the testing of the sensitising chemicals. However, difficulties in obtaining negative results for the non-sensitising substances (in particular SLS) led to additional experiments and required a site visit from lead laboratories representatives in October 2010. The visit failed to identify the cause of the problem.

It was suggested that the problem could be due to the higher sensitivity of the FACS Aria flow cytometer used by EURL ECVAM. An important difference between the FACS Aria and other flow cytometers used in the validation study is that the FACS Aria uses a fully digital processing of the signals, while the flow cytometers used by the other participating laboratories are analog. Despite this difficulty in obtaining consistent negative results with SLS, the lead laboratories felt that, as the issue was limited to this chemical, the h-CLAT had been successfully transferred at EURL ECVAM. Based on this recommendation, the VMG approved the transfer on the 24th of January, 2011.

The issue of the digital versus analog cytometer was subsequently investigated by additional studies from the lead laboratories (see additional considerations).

3. Second Transfer to EURL ECVAM

At the time they were given permission to proceed on the basis of the results above, EURL ECVAM was in the process of acquiring a new flow cytometer, a FACS CANTO II.

EURL ECVAM requested permission to delay the start of the testing phase in order to use this machine instead of the FACS Aria, as they felt it would greatly increase the throughput of the testing and still allow EURL ECVAM to complete the testing in the requested timeframe. This request was discussed by the VMG during the 5th VMG meeting, and agreed to, with the following conditions:

- All the testing for the blind testing phases (BI and BII) should be performed using the same flow cytometer
- To use the new cytometer for phase B, a full set of successful transfer experiments should first be performed with the new FACS CANTO II machine.

- Initiation, and completion, of Phase B testing should not be unreasonably delayed.

EURL ECVAM successfully transferred the h-CLAT to the new flow cytometer between the 14th of February and the 11th of March 2011, and was given the official permission to start the testing on April 11th, 2011. No problems surfaced during this second set of transfer experiments.

4. Additional considerations

A) Although the exact value of the CV75 concentrations determined by the laboratories were not part of the criteria established for successful transfer, it was observed that the values were very reproducible, as shown in Table 7.

Table 7: Summary of the CV75 values obtained for the 5 transfer chemicals during the Transfer phase of the study.

Chemical	Bioassay	EURL ECVAM (FACS Aria)	EURL ECVAM (FACS CANTO II)
DNCB	5.8	9.7 / 13.5	8.7
SLS	46.7 / 60.6	60.2 / 58.1	54.7
NiSO4	160.2	198.0 / 216.5	156.7
Lactic Acid	2707	2822.8 / 2990.2	3010.6
PAA	33.1	52.6	46.6

B) In its transfer report, EURL ECVAM recommended the VMG to consult a flow cytometry expert to consider in more detail the possible differences in sensitivity between analogue and digital instruments, as this was felt to be a factor potentially causing the frequent positive results obtained by EURL ECVAM with the non-sensitising chemicals (all the other laboratories in the study, as well as the laboratories that developed the h-CLAT, use analog flow cytometers).

Following discussions during the 6th VMT meeting, Kao and Shiseido agreed to investigate this issue and to discuss with the Japanese branch of Becton Dickinson. In addition, experiments were conducted in the lead laboratories to compare the results from an analog flow cytometer (BD FACS Calibur) and a digital flow cytometer (BD FACS Canto II), testing DNCB and SLS.

The results of these discussions and study was presented during the 9th VMT meeting, where the lead laboratory representatives concluded that no significant differences were observed when comparing the results generated by both cytometers (see

Figure 3).

Interestingly the results show that for SLS higher concentrations result in a positive signal, emphasizing the importance of correctly determining the CV75 for this chemical.

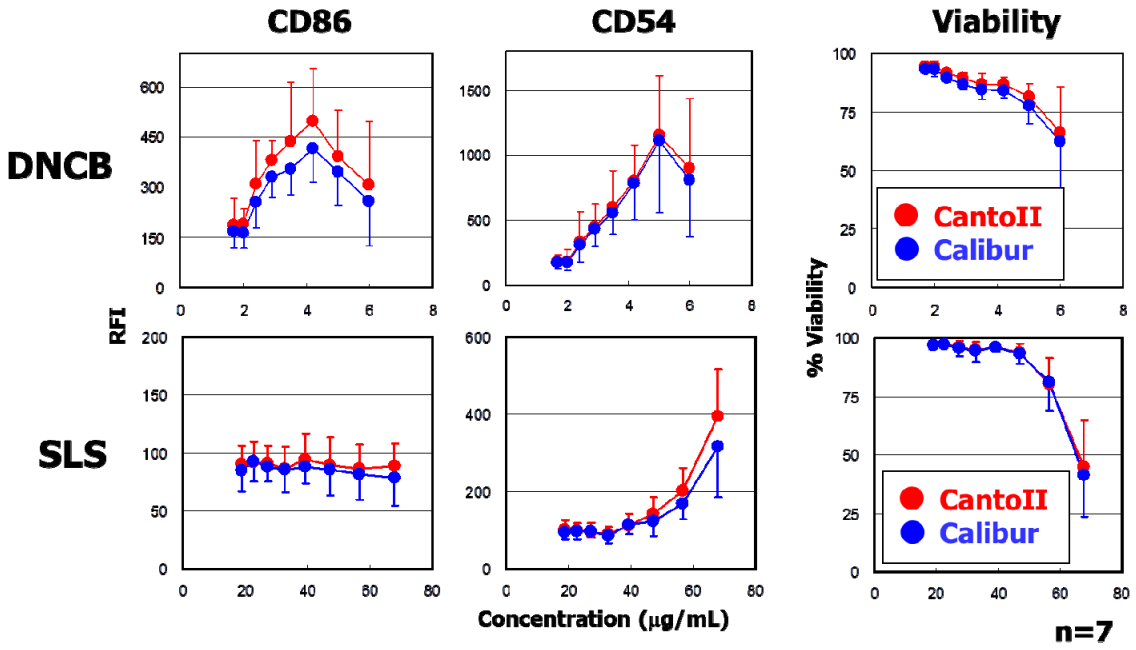

Figure 3: Results of the study comparing the FACS CantoII (digital) and FACS Calibur (analog) cytometers.

DNCB and SLS are also used in the h-CLAT for the qualification of freshly thawed cultures of THP-1 cells ("reactivity checks") prior to using them for experiments. In view of the dose response curves observed for these chemicals, i.e. the fact that SLS can become positive and DNCB negative at high doses, the VMG, at the suggestion of the lead laboratories, authorized a modification of the reactivity checks procedure to allow each laboratory to determine their optimal chemical concentration to be used, based on an evaluation of the CV75. In addition, the lead laboratories proposed the use of Lactic Acid instead of SLS in the reactivity checks, as Lactic Acid showed, historically, a lower rate of false positive results with the h-CLAT.

Conclusion of the Validation Management Team on Module 3

The VMG agreed that both laboratories succeeded in transferring the h-CLAT protocol to their facilities and, in the case of EURL ECVAM, using two different flow cytometers.

The settings of the flow cytometer posed no particular challenges, and were set after the initial reactivity checks experiments. The different batches of THP-1 cells purchased by the laboratories caused no concerns for the success of the transfer, consistent with published information (Kosaka *et al.,* 2008).

Both laboratories generated the required experiments in 3 successive runs for the three sensitisers. The difficulties at both sites related to the negative chemicals, SLS and Lactic Acid, which were solved following site visits by representatives of the lead laboratories. This emphasizes the importance of transfer experiments in new laboratories that would include these chemicals to ensure correct implementation of the protocol.

In general the VMG recommends the use of a careful step-wise approach similar to the one used in the study for future implementations of the test method in new laboratories.

Within-laboratory reproducibility (Module 2)

Reference documents:

- **Statistical report (Appendix 12)**
- **List of additional documents filed for the study and available on request (Appendix 15)**

The within laboratory reproducibility was assessed using the data generated with a subset of 15 chemicals tested in three independent experiments in each laboratory (Study Phase B2). As described in the h-CLAT SOP, each of the independent experiments is composed of three valid evaluation runs to derive a final conclusion.

To minimise bias in the generation of the results for the assessment of the WLR the laboratories were provided with 3 different vials for each chemical (with different codes), one for each experiment.

For this reason the laboratories had to perform a separate solvent determination (see section on BLR) and CV75 determination prior to the performance of the evaluation runs; that is, it was done on three separate occasions for each chemical.

The WLR of the h-CLAT was assessed on the concordance of classification, sensitiser (S) versus non-sensitiser (NS) between the three independent experiments. Additionally, descriptive analyses were performed to evaluate the reproducibility of the three respective CV75 values and on the calculated EC values, where appropriate. The descriptive analysis is presented for information only and was not used by the VMG, on its own, to draw conclusions on the WLR of the h-CLAT

The data are presented laboratory by laboratory and the three independent assessments of each chemical required for the evaluation of the WLR are referred to as experiment 1, experiment 2, and experiment 3.

1. Kao

a. Reproducibility (concordance of predictions)

In relation to the primary objective, the reproducibility in terms of classification S versus NS, the same prediction in the 3 independent experiments was obtained for 13 of the 15 chemicals resulting in a WLR of 86.7%. The two chemicals for which the classification was not concordant across experiments were Benzyl salicylate (chem.16) and Methyl salicylate (chem. 19).

Chemical		**Prediction (run1, run2, run3)**			**Agreement** Between experiment
		Exp 1	Exp 2	Exp 3	
10	Kathon CG (CMI/MI)	PPP	PPP	PPP	Yes

11	Beryllium sulfate	NNN	NNN	NNN	Yes
12	Formaldehyde	PPP	PPP	PPP	Yes
13	Chloramine T	PPP	PPP	PPP	Yes
14	Chlorpromazine HCl	PPP	PPP	PPP	Yes
15	2-Mercaptobenzothiazole	PPP	PPP	PPP	Yes
16	**Benzyl salicylate**	**NNN**	**PPN**	**NPN**	**No**
17	Benzyl cinnamate	NNN	NNN	NNN	Yes
18	R(+)- Limonene	PPP	PPP	PPP	Yes
19	**Methyl salicylate**	**PPN**	**NNN**	**NPP**	**No**
20	Isopropanol	NNN	NNN	NNN	Yes
21	Dimethyl isophthalate	NPN	NNN	NPN	Yes
22	4-Aminobenzoic acid	NNN	PNN	NNN	Yes
23	Nickel chloride	PPP	PPP	PPP	Yes
24	Xylene	NNN	NNN	NNN	Yes

Table 8: concordance in the classifications sensitisers (S) verus non-sensitiser (NS) obtained in three independent experiments at Kao with the 15 chemicals (chem. 10-24) used for the evalution of the WLR.

b. Reproducibility of CV75 values

Figure 4 shows the distribution of the CV75 determinations performed as part of the three independent experiments for each of the 15 chemicals. CV75 values for chemicals 11 (Beryllium sulphate), 17 (Benzyl cinnamate), 20 (Isopropanol), 21 (Dimethyl isophthalate), 22 (4-Aminobenzoic acid) and 24 (Xylene) could not be derived because of their low cytotoxicity and for this reason the highest concentration used for these chemicals was the maximum one allowed by the SOP (i.e. 5000 µg/mL in medium/saline or 1000 µg/mL in DMSO), or the maximum soluble concentration (chemical 21, Dimethyl isophthalate, whose maximum concentration tested was 500 µg/mL in DMSO). For more information refer to Appendix 12 (Statistical report).

Notwithstanding that different cell batches and passage numbers have been used, the calculated CV75 are consistent across experiments.

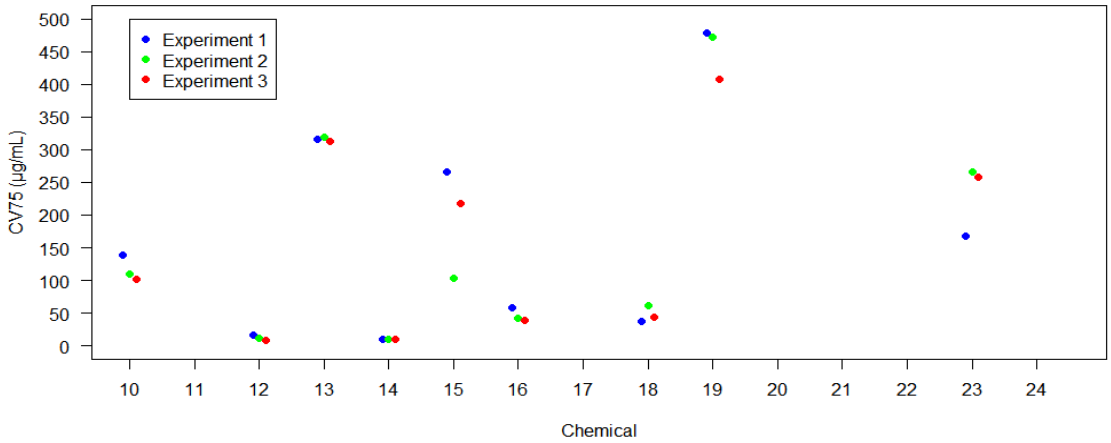

Figure 4: CV75 determinations expressed in µg/mL performed within three individual experiments for the 15 chemicals (chem. 10-24). Each dot corresponds to the mean of 2 or more independent CV75 determinations.

c. Reproducibility of EC values

The concentration estimated to produce an RFI of 150 or 200 for CD86 (EC150) and CD54 (EC200), respectively are presented in Figure 5. For more information, refer to Appendix 12 (Statistical report).

For CD86, 10/15 (66.7%) chemicals were consistently negative (RFI<150 for all concentrations) or positive (RFI ≥150 for at least one concentration) in the three independent experiments. For CD54, 12/15 (80%) chemicals were consistently negative (RFI<200 for all concentrations) or positive (RFI ≥200 for at least one concentration) in the three independent experiments.

Where it was possible to calculate them, the EC150 and EC200 values were very reproducible.

Raw RFI data and the corresponding viability for all concentrations for each chemical are presented in Annex II of the statistical report (see Appendix 12).

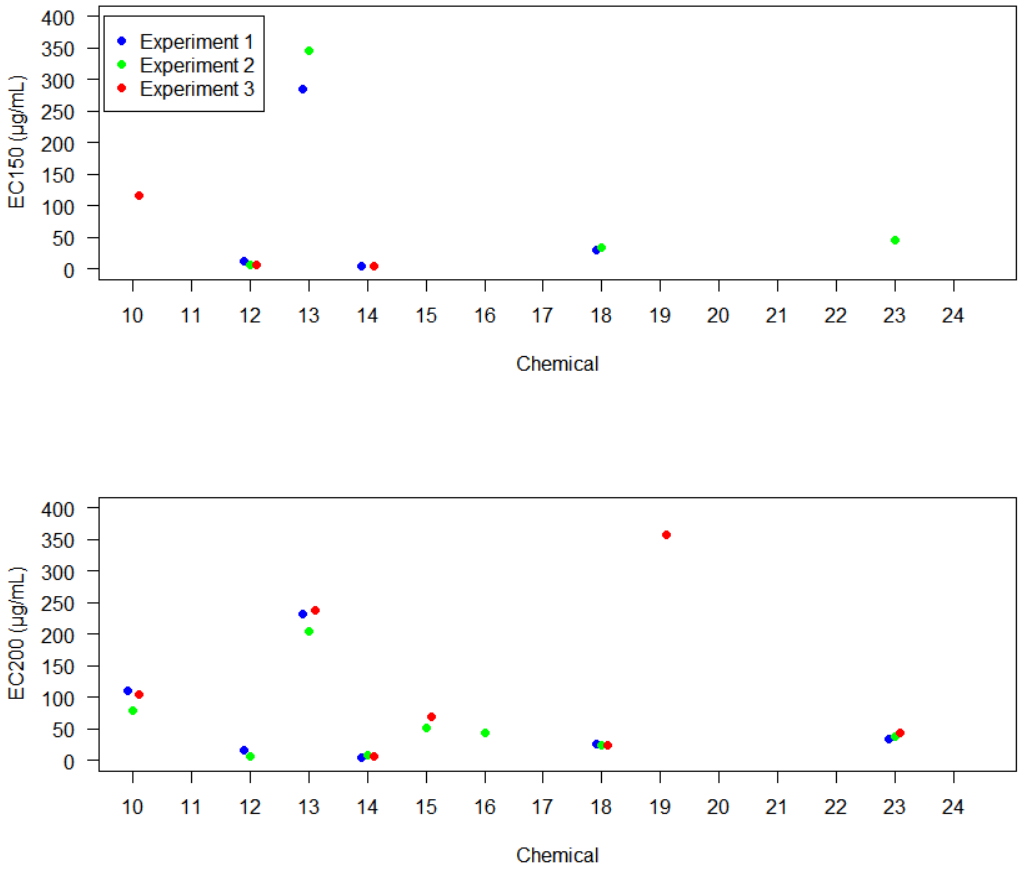

Figure 5: EC150 (CD86) and EC200 (CD54) concentrations for the three independent experiments performed at Kao

2. Shiseido

a. Reproducibility (concordance in predictions)

In relation to the primary aim, the reproducibility in terms of classification S versus NS, the same prediction in the 3 independent experiments was obtained for 12 of the 15 chemicals resulting in a WLR of 80%. The three chemicals for which the classification was not concordant across experiments were Kathon CG (chem. 10), Beryllium Sulphate (chem. 11) and Formaldehyde (chem. 12).

Chemical		Prediction (run1, run2, run3)			Agreement Between Experiments
		Exp 1	Exp 2	Exp 3	
10	**Kathon CG (CMI/MI)**	**PPN[A]**	**PPP**	**PPP**	**No**
11	**Beryllium sulfate**	**NPP**	**PPN**	**NPN**	**No**
12	**Formaldehyde**	**NNP**	**PPP**	**PPP**	**No**
13	Chloramine T	PPP	PPP	PPN	Yes

14	Chlorpromazine HCl	PPP	PPP	PPP	Yes
15	2-Mercaptobenzothiazole	PPP	PPP	PPP	Yes
16	Benzyl salicylate	NNN	PNN	NNP	Yes
17	Benzyl cinnamate	NPP	PPN	PPP	Yes
18	R(+)- Limonene	PPP	PPP	PPP	Yes
19	Methyl salicylate	NNP	PNN	NNN	Yes
20	Isopropanol	NNN	PNN	NNN	Yes
21	Dimethyl isophthalate	NNN	NNN	NNN	Yes
22	4-Aminobenzoic acid	NNN	NNN	NNN	Yes
23	Nickel chloride	PPP	PPP	PPP	Yes
24	Xylene	NNP	NNN	NNN	Yes

Table 9: concordance in the classifications sensitisers (S) verus non-sensitiser (NS) obtained in three independent experiments at Shiseido with the 15 chemicals (chem. 10-24) used for the evalution of the WLR. [A] This chemical resulted positive in one experiment for the CD54 an in another for the CD86, therefore the final prediction corresponds with a non-sensitizer.

b. Reproducibility of CV75 values

Figure 6 shows the distribution of the CV75 determinations performed as part of the three independent experiments for each of the 15 chemicals. CV75 values for chemical 11 (Beryllium sulphate), could be determined only in experiments 2 and 3 since in experiment 1 the same chemical did not show sufficient cytotoxicity. Also for chemicals 20 (Isopropanol), 22 (4- Aminobenzoic acid) and 24 (Xylene) the CV75 values could not be derived because of their low cytotoxicity and for this reason the highest concentration used for these chemicals was the maximum one allowed by the SOP (i.e. 5000 µg/mL in medium/ saline or 1000 µg/mL in DMSO). For more information, refer to Appendix 12 (Statistical report).

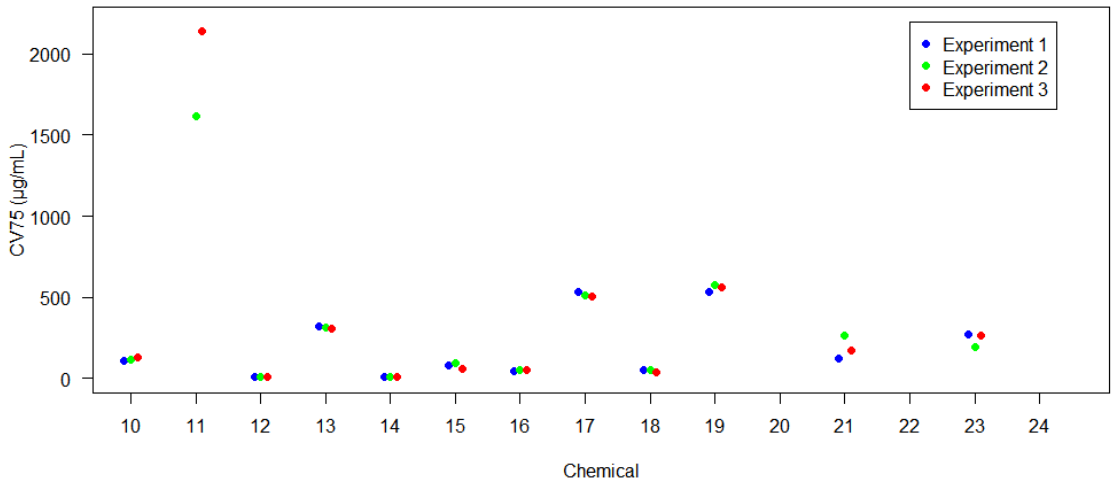

Figure 6: CV75 determinations expressed in µg/mL performed within three individual experiments for the 15 chemicals (chem. 10-24). Each dot corresponds to the mean of 2 or more independent CV75 determinations.

c. Reproducibility of EC values

The concentration estimated to produce an RFI of 150 or 200 for CD86 (EC150) and CD54 (EC200), respectively are presented in Figure 7. For more information, refer to Appendix 12 (Statistical report).

For CD86, 9/15 (60%) chemicals were consistently negative (RFI<150 for all concentrations) or positive (RFI ≥150 for at least one concentration) in the three independent experiments. For CD54, 11/15 (73%) chemicals were consistently negative (RFI<200 for all concentrations) or positive (RFI ≥200 for at least one concentration) in the three independent experiments.

Where it was possible to calculate them, the EC150 and EC200 values were very reproducible.

Raw RFI data and the corresponding viability for all concentrations for each chemical are presented in Annex II of the statistical report (see Appendix 12).

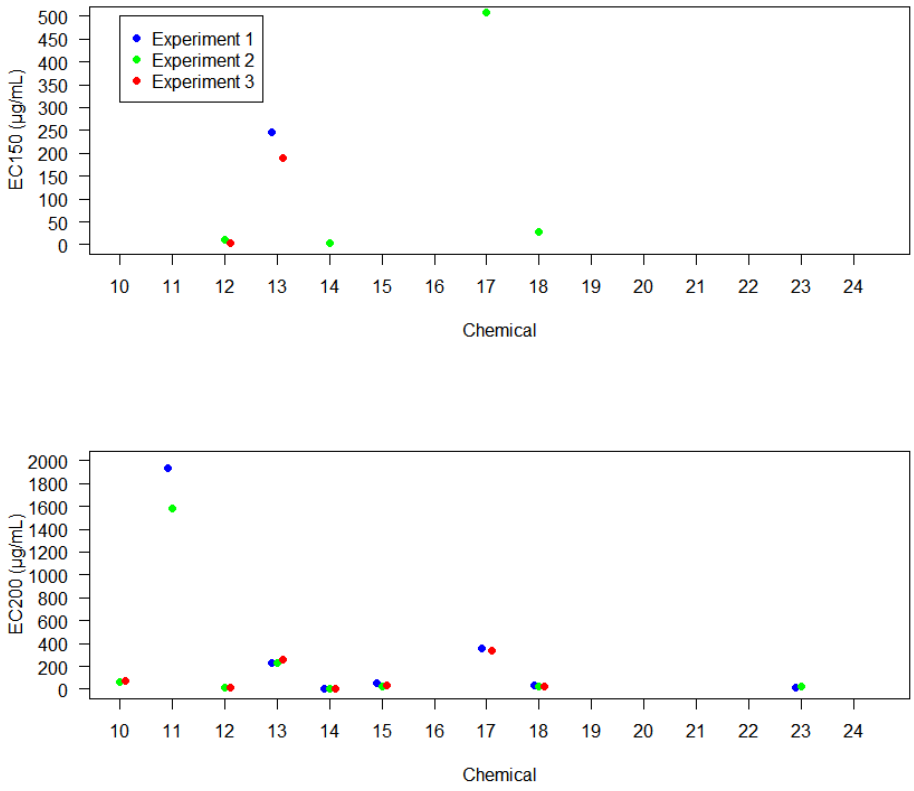

Figure 7: EC150 (CD86) and EC200 (CD54) concentration for the three independent experiments performed at Shiseido.

3. Bioassay

a. Reproducibility (concordance in predictions)

In relation to the primary aim, the reproducibility in terms of classification S versus NS, the same prediction in the 3 independent experiments was obtained for 11 of the 15 chemicals resulting in a WLR of 73.3%. The four chemicals for which the classification was not concordant across experiments were Formaldehyde (chem. 12), Benzyl salicylate (chem. 16), Methyl salicylate (chem. 19) and Xylene (chem. 24).

Chemical		Prediction (run1, run2, run3)			Agreement Between Experiments
		Exp 1	Exp 2	Exp 3	
10	Kathon CG (CMI/MI)	PPP	PPP	PPN	Yes
11	Beryllium sulfate	NNP	NNN	NNN	Yes
12	**Formaldehyde**	**NPN**	**PPP**	**PPP**	**No**
13	Chloramine T	PPP	PPP	PPP	Yes
14	Chlorpromazine HCl	PPP	PPP	NPP	Yes
15	2-Mercaptobenzothiazole	PPP	PPP	PPP	Yes
16	**Benzyl salicylate**	**NPN**	**NNN**	**PPN**	**No**
17	Benzyl cinnamate	NNN	NNN	NNN	Yes
18	R(+)- Limonene	PPP	PPP	PPP	Yes
19	**Methyl salicylate**	**PPP**	**NNN**	**NPN**	**No**
20	Isopropanol	NNP	NNN	NNN	Yes
21	Dimethyl isophthalate	PNN	NNN	NPN	Yes
22	4-Aminobenzoic acid	NNN	NNN	NNN	Yes
23	Nickel chloride	PPP	PPP	PPP	Yes
24	**Xylene**	**PPN**	**NNP**	**PNN**	**No**

Table 10: concordance in the classifications sensitisers (S) verus non-sensitiser (NS) obtained in three independent experiments at Bioassay with the 15 chemicals (chem. 10-24) used for the evalution of the WLR.

b. Reproducibility of CV75 values

Figure 8 shows the distribution of the CV75 determinations performed as part of the three independent experiments for each of the 15 chemicals. CV75 values for chemicals 11 (Beryllium sulphate), 17 (Benzyl cinnamate), 20 (Isopropanol), 22 (4-Aminobenzoic acid) and 24 (Xylene) could not be derived because of their low cytotoxicity and for this reason the highest concentration used for these chemicals was the maximum one allowed by the SOP (i.e. 5000 µg/mL in medium/saline or 1000 µg/mL in DMSO), or the maximum soluble concentration (chemical 11, Beryllium sulphate, whose maximum concentration tested was 500 µg/mL in DMSO). CV75 values chemical 21 (Dimethyl isophthalate) could be

determined only in experiments 2 and 3 since in experiment 1 the chemical did not show sufficient cytotoxicity. For more information, refer to Appendix 12 (Statistical report).

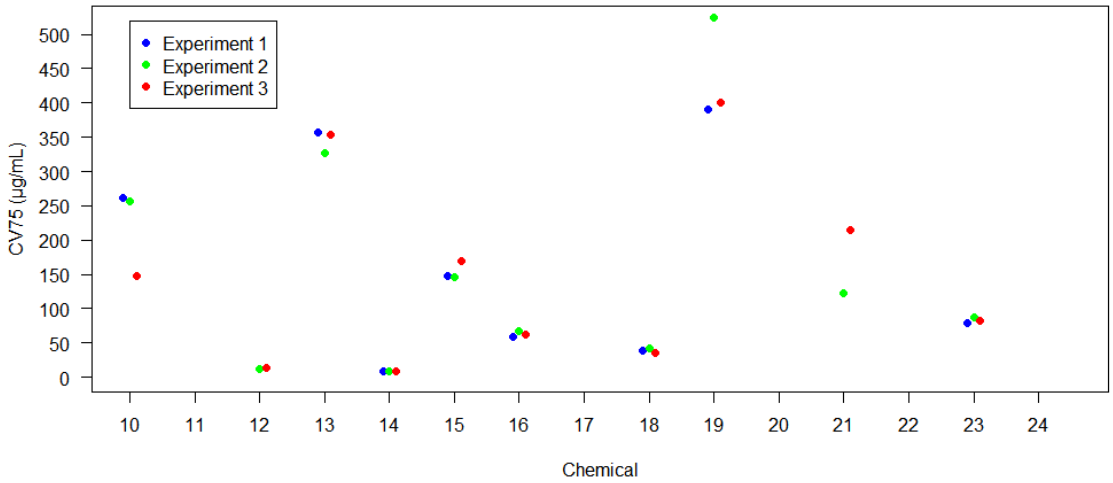

Figure 8: CV75 determinations expressed in µg/mL performed within three individual experiments for the 15 chemicals (chem 10-24). Each dot corresponds to the mean of 2 or more independent CV75 determinations.

c. Reproducibility of EC values

The concentration estimated to produce an RFI of 150 or 200 for CD86 (EC150) and CD54 (EC200), respectively are presented in Figure 9. For more information, refer to Appendix 12 (Statistical report).

For CD86, 11/15 (73.3%) chemicals were always negative (RFI<150 for all concentrations) or positive (RFI ≥150 for at least one concentration) in the three independent experiments. For CD54, 9/15 (60%) chemicals were always negative (RFI<200 for all concentrations) or positive (RFI ≥200 for at least one concentration) in the three independent experiments.

Raw RFI data and the corresponding viability for all concentrations for each chemical are presented in Annex II of the statistical report (see Appendix 12).

Where it was possible to calculate them, the EC150 and EC200 values were very reproducible.

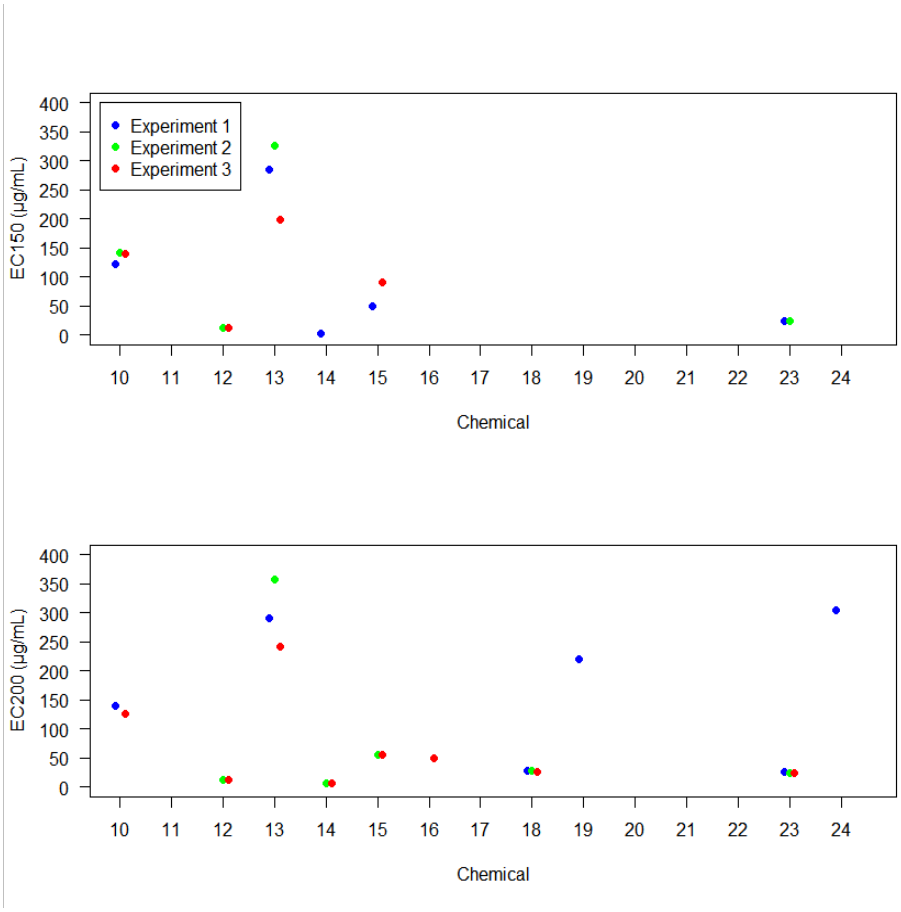

Figure 9: EC150 (CD86) and EC200 (CD54) concentration for the three independent experiments performed at Bioassay.

4. EURL ECVAM

a. Reproducibility (concordance in predictions)

In relation to the primary aim, the reproducibility in terms of classification S versus NS, the same prediction in the 3 independent experiments was obtained for 12 of the 15 chemicals resulting in a WLR of 80%. The three chemicals for which the classification was not concordant across experiments were Chlorpromazine hydrochloride (chem. 14), Benzyl cinnamate (chem. 17), and Dimethyl isophthalate (chem. 24).

Chemical		Prediction (run1, run2, run3)			Agreement Between Experiments
		Exp 1	Exp 2	Exp 3	
10	Kathon CG (CMI/MI)	PPP	PPN	PPP	Yes
11	Beryllium sulfate	PNP	PPP	PPN	Yes
12	Formaldehyde	PPP	PPP	PPP	Yes
13	Chloramine T	PPP	PPP	PPP	Yes
14	**Chlorpromazine HCl**	**PPP**	**PPP**	**NPN**	**No**

15	2-Mercaptobenzothiazole	PPP	PPP	PPP	Yes
16	Benzyl salicylate	PPP	PPP	PPP	Yes
17	**Benzyl cinnamate**	**PPP**	**NNP**	**NNN**	**No**
18	R(+)- Limonene	PPP	PPP	PPP	Yes
19	Methyl salicylate	PPP	PPP	PNP	Yes
20	Isopropanol	NNN	NNN	NNN	Yes
21	**Dimethyl isophthalate**	**NNP**	**PPN**	**NPP[A]**	**No**
22	4-Aminobenzoic acid	NNN	NNN	NNP	Yes
23	Nickel chloride	PPP	PPP	PPP	Yes
CHEM 24	Xylene	PPP	PPP	PPP	Yes

Table 11: concordance in the classifications sensitisers (S) verus non-sensitiser (NS) obtained in three independent experiments at EURL ECVAM with the 15 chemicals (chem. 10-24) used for the evalution of the WLR. [A]This chemical resulted positive in one experiment for the CD54 an in another for the CD86, therefore the final prediction corresponds with a non-sensitizer**.**

b. Reproducibility of CV75 values

Figure 10 shows the distribution of the CV75 determinations performed within each of the three independent experiments for each of the 15 chemicals. CV75 values for chemicals 18 (R(+)- Limonene), 20 (Isopropanol), 22 (4-Aminobenzoic acid) and 24 (Xylene) could not be derived because of their low cytotoxicity and for this reason the highest concentration used for these chemicals was the maximum one allowed by the SOP (i.e. 5000 μg/mL in medium/ saline or 1000 μg/mL in DMSO), or the maximum soluble concentration (chemical 18, R(+)- Limonene, whose maximum concentration tested was 120 μg/mL in DMSO). CV75 values for chemical 11 (Beryllium sulphate), chemical 17 (Benzyl cinnamate) and chemical 21 (Dimethyl isophthalate) could be determined only in experiments 2, 1-2, and 1-2 respectively since in the other experiments the chemicals did not show sufficient cytotoxicity. For more information, refer to Appendix 12 (Statistical report).

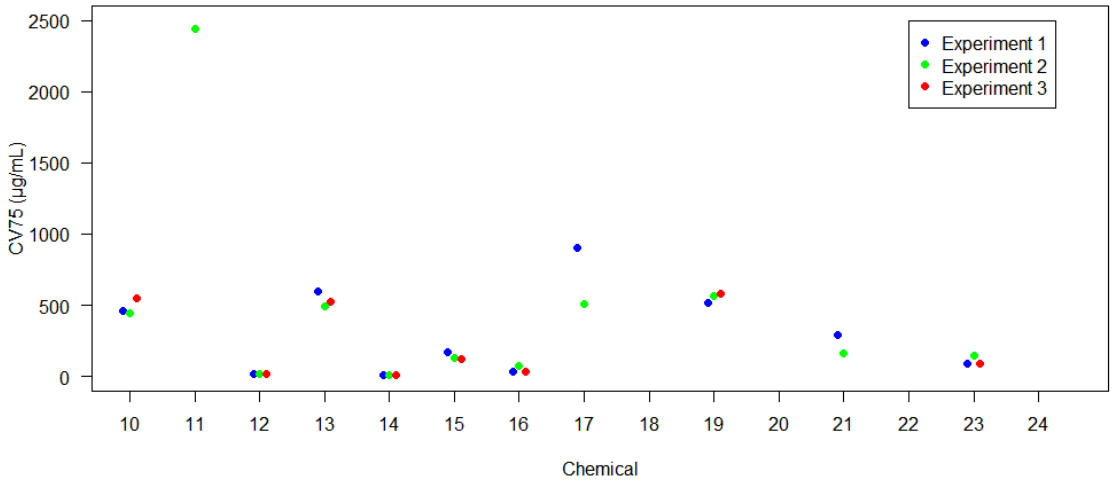

Figure 10: CV75 determinations expressed in µg/mL performed within three individual experiments for the 15 chemicals (chem. 10-24). Each dot corresponds to the mean of 2 or more independent CV75 determinations.

c. Reproducibility of EC values

The concentration estimated to produce an RFI of 150 or 200 for CD86 (EC150) and CD54 (EC200), respectively are presented in Figure 11. For more information, refer to Appendix 12 (Statistical report)

For CD86, 12/15 (80%) chemicals were always negative (RFI <150 for all concentrations) or positive (RFI ≥150 for at least one concentration) in the three independent experiments. For CD54, 14/15 (93.3%) chemicals were always negative (RFI <200 for all concentrations) or positive (RFI ≥200 for at least one concentration) in the three independent experiments.

Raw RFI data and the corresponding viability for all concentrations for each chemical are presented in Annex II of the statistical report (see Appendix 12).

Where it was possible to calculate them, the values of the EC150 and EC200 values were very reproducible.

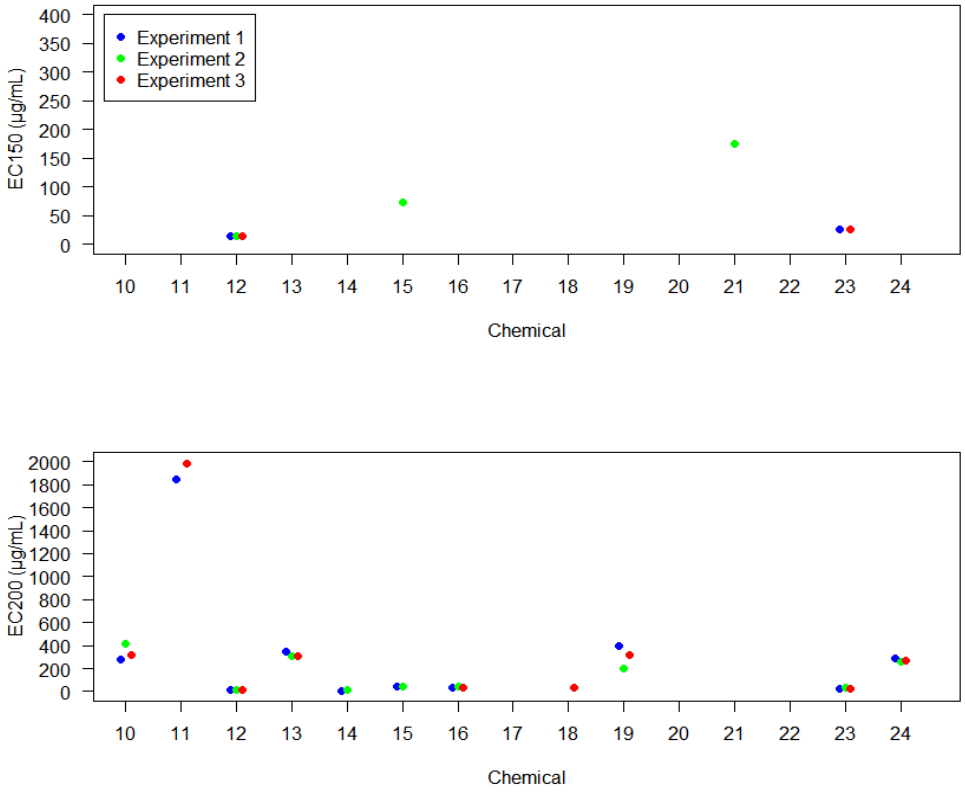

Figure 11: EC150 (CD86) and EC200 (CD54) concentration for the three independent experiments performed at EURL ECVAM.

Summary

The following table shows the chemicals that were not consistently classified within the laboratories. Of the eight chemicals which were not consistently classified in one laboratory three of them (Formaldehyde, Benzyl salicylate and Methyl salicylate), gave discordant results also in a second laboratory.

Table 12: chemicals with inconsistent classifications within laboratories.

WLR			
Kao	**Shiseido**	**Bioassay**	**EURL ECVAM**
	Kathon CG		
	Beryllium sulphate		
	Formaldehyde	Formaldehyde	
			Chlorpromazine HCl
Benzyl salicylate		Benzyl salicylate	
			Benzyl cinnamate
Methyl salicylate		Methyl salicylate	
		Xylene	
			Dimethyl isophtalate

Conclusion of the Validation Management Team on Module 2

The main focus of the within laboratory reproducibility (WLR) for the subset of 15 chemicals in each laboratory was on the concordance of the predictions sensitiser (S) versus non-sensitiser (NS) as determined by the results of three independent experiments. The WLR for Kao, Shiseido, EURL ECVAM and Bioassay for the S/NS predictions were 86.7%, 80%, 80% and 73.3% respectively.

Only one of the participating laboratories met the target performance proposed at the outset of the study (85%).

However, the VMG agreed, that the obtained WLR was sufficient considering the potential use of the method in a toolbox or integrated testing strategies, taking into consideration the fact that:

- The BLR met the target performance set by the VMG prior to the initiation of the blind testing phase (see section below)
- There is a consistency in the chemicals that showed discrepancies in the WLR and the BLR (see table 14).
- It was noted that some of the chemicals responsible for the discrepancies are suggested to fall outside the applicability domain of the h-CLAT (in relation to the Log Kow, see section on predictive capacity).

Between laboratory reproducibility (Module 4)

Reference documents:

- **Statistical report (Appendix 12)**
- **List of additional documents filed for the study and available on request (Appendix 15)**

1. Reproducibility (concordance in predictions)

The between laboratory reproducibility (BLR) was assessed on the basis of the results generated with 24 chemicals (9 chemicals tested once and 15 chemicals tested independently 3 times in each laboratory). The main focus of the evaluation of the between-laboratory reproducibility was on the concordance of the predictions sensitisers (S) versus non-sensitisers (NS). As discussed in the WLR section, descriptive statistical analyses were also performed on the EC values, but were not considered by the VMG, on their own, to determine to what extent the primary objective of the study had been satisfied.

For the evaluation of the BLR, the final prediction for the chemicals that were tested 3 times (chemicals 10 to 24) in each laboratory was based on the median classification obtained in that laboratory. For example, for Chlorpromazine hydrochloride at EURL ECVAM the predictions obtained were "sensitiser", "sensitiser" and "non-sensitiser", so the median classification was "sensitiser".

Nineteen of the 24 chemicals were consistently classified (S or NS) by the 4 laboratories resulting in a BLR of 79.2% (Table 13). The five chemicals which were not consistently classified by the laboratories were: Beryllium sulphate (chem. 11), Benzyl salicylate (chem. 16), Benzyl cinnamate (chem. 17), Methyl salicylate (chem. 19) and Xylene (chem. 24).

Since BLR is usually reported less stringently as the consistency of prediction between 3 laboratories, which is one fewer than the number of testing laboratories in this study due to the participation of 2 lead laboratories, the VMG suggested, in the experimental design document drafted prior to the initiation of the testing phase, that two BLR values would be derived: one comparing the consistency of the two naïve laboratories with the first lead laboratory (Kao) and the second comparing the consistency of the two naïve laboratories with the other lead laboratory (Shiseido). The values obtained in these comparisons (83.3% and 79.2%) are comparable to the reproducibility value calculated when all four laboratories are considered together.

Table 13: Concordance in S versus NS predictions between the laboratories.

		Kao	Shiseido	Bioassay	EURL ECVAM	Agree	Kao versus BA & EURL ECVAM	Shi versus BA & EURL ECVAM
1	Benzoquinone	S	S	S	S	Yes	Yes	Yes
2	PPD	S	S	S	S	Yes	Yes	Yes
3	Dihydroeugenol	S	S	S	S	Yes	Yes	Yes
4	Thioglycerol	S	S	S	S	Yes	Yes	Yes
5	Imidazolidinyl urea	S	S	S	S	Yes	Yes	Yes
6	Methyl methacrylate	NS	NS	NS	NS	Yes	Yes	Yes
7	Glycerol	NS	NS	NS	NS	Yes	Yes	Yes
8	DCNB	S	S	S	S	Yes	Yes	Yes
9	Benzyl alcohol	S	S	S	S	Yes	Yes	Yes
10	Kathon CG (CMI/MI)	S	S	S	S	Yes	Yes	Yes
11	**Beryllium sulfate**	**NS**	**S**	**NS**	**S**	**No**	**No**	**No**
12	Formaldehyde	S	S	S	S	Yes	Yes	Yes
13	Chloramine T	S	S	S	S	Yes	Yes	Yes
14	Chlorpromazine HCl	S	S	S	S	Yes	Yes	Yes
15	2-Mercaptobenzothiazole	S	S	S	S	Yes	Yes	Yes
16	**Benzyl salicylate**	**NS**	**NS**	**NS**	**S**	**No**	**No**	**No**
17	**Benzyl cinnamate**	**NS**	**S**	**NS**	**NS**	**No**	Yes	**No**
18	R(+)- Limonene	S	S	S	S	Yes	Yes	Yes
19	**Methyl salicylate**	**S**	**NS**	**NS**	**S**	**No**	**No**	**No**
20	Isopropanol	NS	NS	NS	NS	Yes	Yes	Yes
21	Dimethyl isophthalate	NS	NS	NS	NS	Yes	Yes	Yes
22	4-Aminobenzoic acid	NS	NS	NS	NS	Yes	Yes	Yes
23	Nickel chloride	S	S	S	S	Yes	Yes	Yes
24	**Xylene**	**NS**	**NS**	**NS**	**S**	**No**	**No**	**No**
						79.2%	83.3%	79.2%

Interestingly all of the five chemicals which were not consistently classified between the laboratories gave also discordant results, within at least one of the laboratories, in the three independent experiments performed for the evaluation of the WLR (see Table 12). As observed earlier, this tends to suggest that the observed performance may be enhanced by careful determination and consideration of the applicability domain of the h-CLAT.

Table 14: chemicals with inconsistent classifications within and between laboratories.

WLR				BLR
Kao	**Shiseido**	**Bioassay**	**EURL ECVAM**	
	Kathon CG			
	Beryllium sulphate			Berillium sulphate
	Formaldehyde	Formaldehyde		
			Chlorpromazine HCl	
Benzyl salicylate		Benzyl salicylate		Benzyl salicylate
			Benzyl cinnamate	Benzyl cinnamate
Methyl salicylate		Methyl salicylate		Methyl salicylate
		Xylene		Xylene
			Dimethyl isophtalate	

2. Other considerations

a) Solvent selection

It was decided by the VMG prior to the initiation of the study, and after the preliminary ECVAM confirmation of solubility had been performed, that the participating laboratories would not be instructed by the VMG on which solvent to use to solubilise the coded chemicals. The VMG considered the solvent selection procedures to be an integral part of the test method SOP, to be evaluated for reproducibility together with the rest of the procedure.

The SOP provided a tiered solvent selection strategy. Since the chemicals that were tested three times in each laboratory (9-24) were coded independently, the laboratories had to repeat the solvent selection procedures for each test (that is, three separate, independent determinations for each chemical).

The choice of solvent was, for the majority, consistent within and between the laboratories, suggesting that the procedure in the SOP is adequately described.

Discrepancies in the prediction obtained did not correlate with different solvents chosen.

Table 15: Solvent selection for each of the chemicals by the four laboratories. Cells with grey background correspond with different vehicles used for dissolving the same chemical within and/or between laboratories.

Chemical		Kao	Shiseido	Bioassay	VAM
1	Benzoquinone	DMSO	DMSO	Saline	DMSO
2	PPD	Saline	Saline	DMSO	DMSO
3	Dihydroeugenol	DMSO	DMSO	DMSO	DMSO
4	Thioglycerol	Saline	Saline	Saline	Saline
5	Imidazolidinyl urea	Saline	Saline	Saline	Saline
6	Methyl methacrylate	DMSO	DMSO	DMSO	DMSO
7	Glycerol	Saline	Saline	Saline	Saline
8	DCNB	DMSO	DMSO	DMSO	DMSO

9	Benzyl alcohol	DMSO	DMSO	DMSO	DMSO
10	Kathon CG (CMI/MI)	DMSO (1) Saline (2,3)	Saline	Saline	Saline
11	Beryllium sulfate	DMSO (2,3) Saline (1)	Saline	DMSO (2,3) Saline (1)	DMSO (3) Saline (1,2)
12	Formaldehyde	DMSO (3) Saline (1,2)	Saline	Saline	Saline
13	Chloramine T	Saline	Saline	Saline	Saline
14	Chlorpromazine HCl	DMSO (3) Saline (1,2)	Saline	Saline	Saline
15	2-Mercaptobenzothiazole	DMSO	DMSO	DMSO	DMSO
16	Benzyl salicylate	DMSO	DMSO	DMSO	DMSO
17	Benzyl cinnamate	DMSO	DMSO	DMSO	DMSO
18	R(+)- Limonene	DMSO	DMSO	DMSO	DMSO
19	Methyl salicylate	DMSO	DMSO	DMSO	DMSO
20	Isopropanol	Saline	Saline	Saline	Saline
21	Dimethyl isophthalate	DMSO	DMSO (1,3) Saline (2)	DMSO	DMSO
22	4-Aminobenzoic acid	DMSO	DMSO	DMSO	DMSO
23	Nickel chloride	Saline	DMSO (1) Saline (2,3)	Saline	Saline
24	Xylene	DMSO	DMSO	DMSO	DMSO

b) Determination of the CV75

To correctly conclude on the skin sensitisation potential of a test substance based on the induction of cell surface markers, the h-CLAT requires the exposure of the cells to a specific range of chemicals concentrations, the selection of which is based on the CV75 values for this chemical, i.e. the concentration of chemical that allows 75% of THP-1 cell survival (25% cell toxicity).

The SOP of the h-CLAT contains a section that explains how this value should be derived, using flow cytometry and Propidium Iodine (PI). At least two experiments are needed to determine a CV75 value to be used in the evaluation runs.

Although this value is not strictly part of the h-CLAT prediction model, it can have an effect on the final conclusion, as the induction of the cell markers by a chemical is highly dependent on its concentration. Efforts were made by the VMG prior to the initiation of the study to clarify and optimize this section in the SOP, and the CV75 determined by each of the laboratories for the blind tested chemicals were compiled. As for the solvent selection, since the replicates used to evaluate the WLR were coded independently, the laboratories had to apply the CV75 determination procedure for each of three independent experiments. Figure 12 show the CV75 obtained by each laboratory during the study. For more details, refer to Appendix 12 (Statistical report).

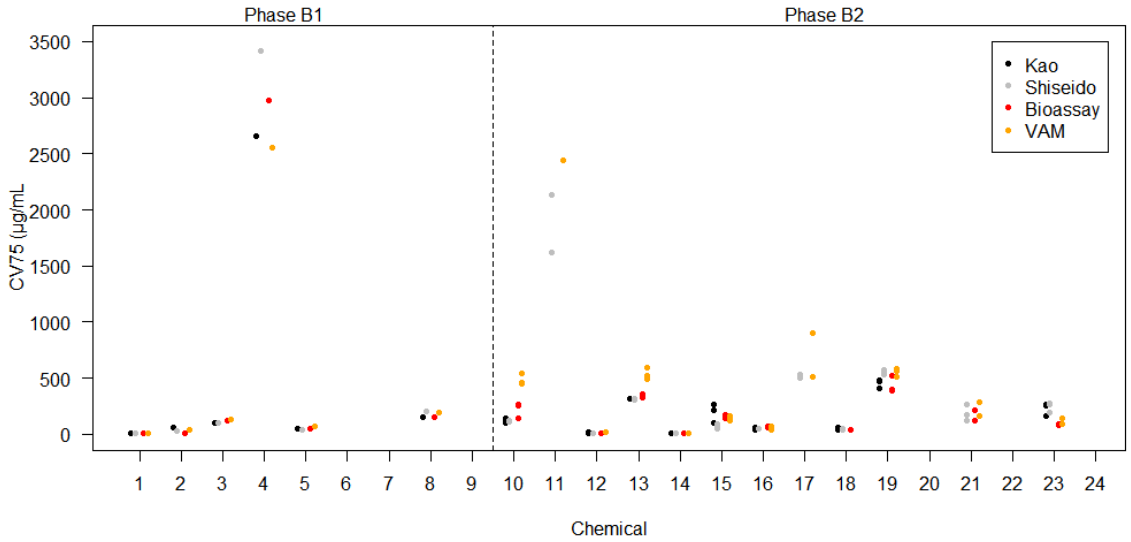

Figure 12: CV75 determinations expressed in µg/mL performed within the individual experiments by each laboratory for the 24 chemicals. Each dot corresponds to the mean of 2 or more independent CV75 determinations.

The determination of the CV75 was not considered by the VMG for the formal evaluation of the BLR. However, considering that the laboratories were using a different lot of the THP-1 cells, the calculated CV75 for each chemical proved to be highly consistent within and across laboratories suggesting that the procedure described in the SOP for defining the dose range to be used in the evaluation runs is robust and can be consistently reproduced when performed in different laboratories.

c) Reproducibility of the EC150 and EC200 values

A summary of the EC150 concentrations (CD86) and EC200 concentrations (CD54) is given in Figure 13. In case a chemical resulted in an RFI ≥ 150 for at least one concentration, the corresponding EC150 concentrations obtained in the four laboratories were generally of the same magnitude. For some chemicals the threshold value of RFI ≥ 150 was not obtained in all laboratories; this was the case for chemical 3, 10, 13, 15, and 18. The EC200 concentrations were also of the same magnitude in the four laboratories except for some chemicals were the threshold value of RFI ≥ 200 was not reached in all laboratories; this was the case for chemical 8, 11, 17, 19, and 24.

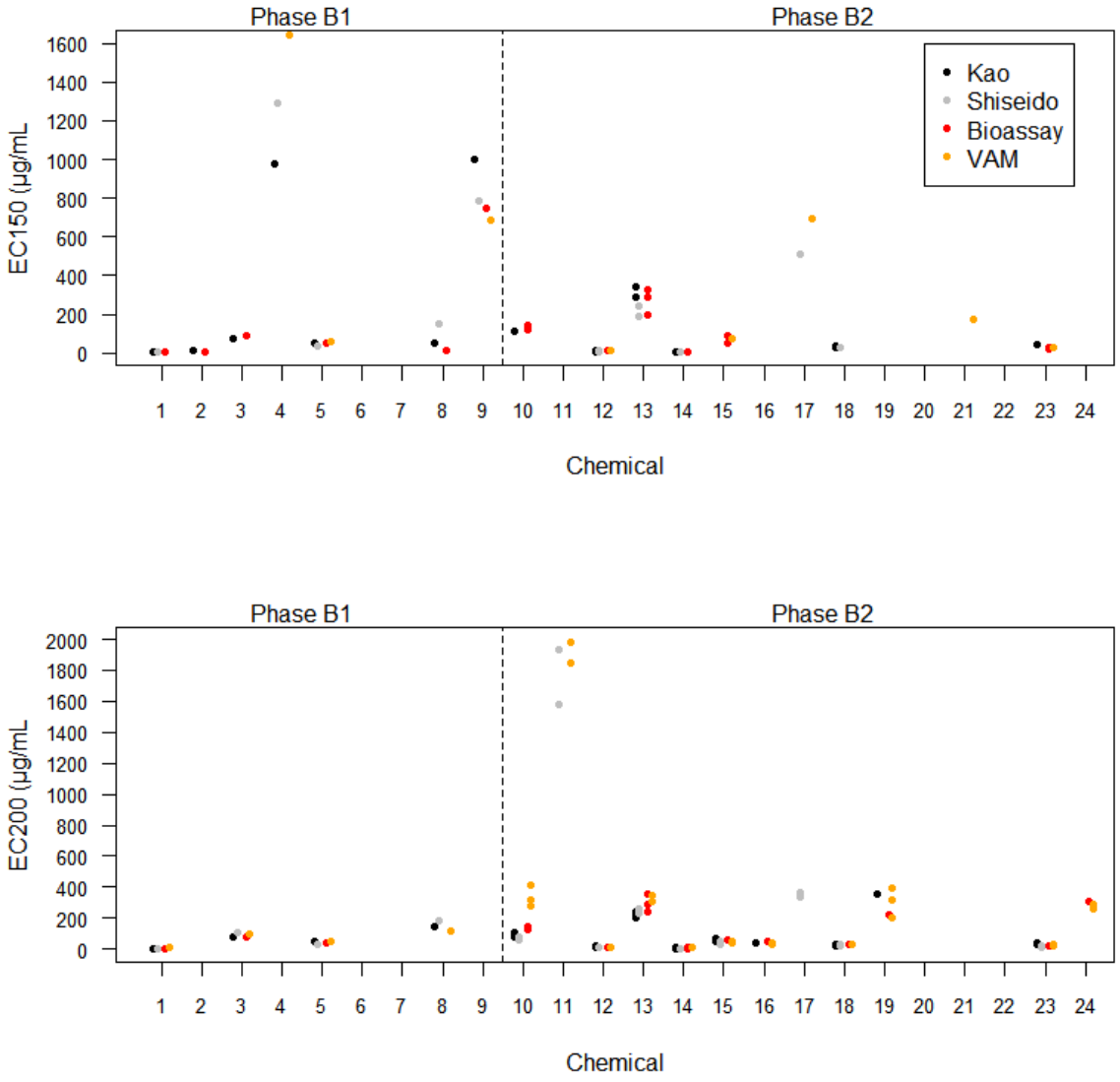

Figure 13: EC150 (CD86) and EC200 (CD54) concentration for the independent experiments of the different laboratories. Phase B1 (chemical 1 to 9), every chemical was tested in one experiment. Phase B2 (chemical 10 to 24), every chemical was tested in three independent experiments.

Conclusion of the Validation Management Team on Module 4

The main focus of the between laboratory reproducibility (BLR) for the 24 chemicals was on the concordance of the predictions sensitisers (S) versus non-sensitisers (NS) between the four laboratories.

Since two of the laboratories were lead laboratories, the VMG decided prior to the initiation of the testing phase to evaluate the BLR by comparing the two naïve laboratories with each of the two lead laboratories separately (see Appendix 2, Experimental Design). The BLR for the S/NS prediction of the two naïve laboratories and Kao was 83.3%, while the BLR for the two naïve laboratories and Shiseido was 79.2%, for an average of 81%.

The BLR comparing all four laboratories was 79.2%.

The VMG agreed that the BLR results were adequate for the proposed use of the test as a component of integrated testing strategies for skin sensitisation testing.

The values were also consistent with the target performance proposed at the onset of the study (80%).

Predictive Capacity (Module 5)

Reference documents:

- **Statistical report (Appendix 12)**
- **Analysis of h-CLAT historical data to categorise chemicals in three classes (Appendix 13)**
- **Analysis of h-CLAT validation study data to categorise chemicals in three classes (Appendix 14)**
- **List of additional documents filed for the study and available on request (Appendix 15)**

The assessment of the predictive capacity forms only a secondary goal of the present validation study, not least since the sample size was determined for the purpose of satisfying the primary study goal and does not permit robust conclusions to be drawn on the predictive capacity of the h-CLAT.

The analysis of the predictive capacity was performed using the results from all the 24 chemicals. An overview of the predicted classification and the reference classification is presented in Table 16.

The predictive capacity was evaluated for each laboratory (Table 17). As for the evaluation of the BLR, for chemicals that were tested three times in each laboratory, the median classification for each laboratory was chosen for this analysis.

This resulted in an accuracy for S/NS classifications of 76% for all laboratories (Table 18), with individual results ranging from 70.8% (Kao) to 83.3% (Shiseido). The two naïve laboratories each had an accuracy of 75%.

The nine chemicals previously tested in the h-CLAT were all consistently and accurately identified with the only exception of Methyl Salicylate (misclassified at Kao and EURL ECVAM) suggesting that the behaviour of the h-CLAT in this study was consistent with historical and published information. Accuracy for the 15 chemicals not previously tested was 65% (39 out of 60, 15 chemicals in 4 laboratories). The results for these latter chemicals should contribute to a future and more comprehensive assessment of predictive accuracy, since no meaningful or robust conclusions on this parameter can be drawn from the limited sample size of this study alone.

Table 16: Agreement between the predicted class and the reference class.

	Chemical	R&D Kao - Shiseido	Kao	Shiseido	Bioassay	EURL ECVAM	Reference
			Prediction based on 1 experiment				
1	Benzoquinone	S	S	S	S	S	+ (1A)
2	PPD	S	S	S	S	S	+ (1A)
3	Dihydroeugenol		S	S	S	S	+ (1B)
4	Thioglycerol		S	S	S	S	+ (1B)
5	Imidazolidinyl urea	S	S	S	S	S	+ (1B)
6	Methyl methacrylate		NS	NS	NS	NS	+ (1B)
7	Glycerol	NS	NS	NS	NS	NS	- (NC)
8	DCNB		S	S	S	S	- (NC)
9	Benzyl alcohol		S	S	S	S	- (NC)
			Prediction based on median of 3 experiments				
10	Kathon CG (CMI/MI)	S	S	S (2/3)	S	S	+ (1A)
11	Beryllium sulfate		NS	S (2/3)	NS	S	+ (1A)
12	Formaldehyde	S	S	S (2/3)	S (2/3)	S	+ (1A)
13	Chloramine T		S	S	S	S	+ (1A)
14	Chlorpromazine HCl		S	S	S	S (2/3)	+ (1A)
15	2-Mercaptobenzothiazole	S	S	S	S	S	+ (1A)
16	Benzyl salicylate		NS (2/3)	NS	NS (2/3)	S	+ (1B)
17	Benzyl cinnamate		NS	S	NS	NS (2/3)	+ (1B)
18	R(+)- Limonene		S	S	S	S	+ (1B)
19	Methyl salicylate	NS	S (2/3)	NS	NS (2/3)	S	- (NC)
20	Isopropanol	NS	NS	NS	NS	NS	- (NC)
21	Dimethyl isophthalate		NS	NS	NS	NS (2/3)	- (NC)
22	4-Aminobenzoic acid		NS	NS	NS	NS	- (NC)
23	Nickel chloride		S	S	S	S	+ (NA)
24	Xylene		NS	NS	NS (2/3)	S	- (NA)

Chemical 10 to 24: (2/3) means that two times the same prediction was obtained, in case nothing is mentioned the same prediction was obtained for the three independent experiments.

Table 17: Predictive capacity of the h-CLAT for each laboratory.

Reference result	Kao		Shiseido		Bioassay		EURL ECVAM	
	+	-	+	-	+	-	+	-
+ (n=16)	12	4	14	2	12	4	14	2
- (n=8)	3	5	2	6	2	6	4	4
Total	15	9	16	8	14	10	18	6
Sensitivity	75.0		87.5		75.0		87.5	
Specificity	62.5		75.0		75.0		50.0	

Accuracy	70.8	83.3	75.0	75.0

Table 18: Overall predictive capacity of the h-CLAT (cumulative over the four laboratories).

Reference result	**Cumulative**	
	+	-
+ (n=64)	52	12
- (n=32)	11	21
Total	63	33
Sensitivity	81.3	
Specificity	65.6	
Accuracy	76.0	

Potency subcategorisation

Within the GHS classification scheme, there is the possibility to refine the evaluation of skin sensitisers on the basis of their potency. Skin sensitisers can be assigned to subcategory 1A "strong sensitisers" or to subcategory 1B "other skin sensitisers" using a weight of evidence approach on the basis of reliable and good quality data from human cases or epidemiological studies and/or observations from appropriate studies in recognised and accepted animal tests. In the case of the LLNA a threshold defined for the EC3 value is used to subcategorise skin sensitisers into the two subcategories.

In order to address this secondary study goal the VMG made a request to Dr Omori and Dr Yoshimura, statisticians based in Japan, to analyse the lead laboratories' h-CLAT historical data and propose a prediction model for potency subcategorization. The proposal that was made uses the MIT, i.e. the lowest value between the calculated EC150 and EC200 of each chemical to assign it to a GHS subcategory (see Appendix 13). Specifically if the MIT is equal or less than 13µg/mL then the chemical would be classified as category 1A; otherwise it would be classified as 1B. Applying the MIT cut-off of 13µg/mL to the historical data the accuracy in correctly assigning a chemical to one of the three categories is reported to be 72%. This proposal is similar to the one that has been previously published (Nukada et al., 2012), however the MIT cut-off value used is different.

When applied to the results of the validation, a MIT cut-off value of 13µg/mL gave a concordant rate of 56.5% in assigning a chemical to one of the three categories (NS corresponding to the GHS NC category = Not Classified and subcategories 1A and 1B). A MIT of 16 µg/mL applied to the study results produced a slightly more accurate classification (58%) (see Appendix 14).

Conclusion of the Validation Management Team on Module 5

Secondary goal a) A preliminary evaluation of the ability of the h-CLAT to reliably discriminate skin sensitising (S) from non-sensitising (NS) chemicals

In our study, the sensitivity of the h-CLAT was 81.3% and the specificity was 65.6%, resulting in an accuracy of 76%. The VMG concludes that, in the current study, the predictive accuracy is lower than previously published information on the predictive capacity of the h-CLAT (Accuracy of 84%, with sensitivity of 88% and specificity of 75% (Ashikaga *et al.,* 2010). However, the VMG notes that, in the current study, all the GHS 1A chemicals (with the exception of Beryllium sulphate) were accurately and consistently identified as sensitisers.

It is important to note that the chemicals chosen for the study reported here included a subset of chemicals previously reported as tested in the h-CLAT, and that those chemicals gave consistent classifications compared to the published information.

The assessment of predictive capacity forms only a secondary goal of the present validation study, not least since the sample size was defined for the assessment of the within and between laboratory reproducibility and does not permit a robust conclusion to be drawn for predictive capacity.

In addition, it has been anticipated that the *in vitro* assays designed to address a specific event of the skin sensitisation pathway are expected to be used as part of a toolbox or an integrated testing strategy (ITS). It is envisaged that predictive capacity will need to be assessed on the basis of the information generated by a future ITS. The information presented in this report must be understood and considered in this context.

Secondary goal b) Preliminary consideration of the ability of the h-CLAT to contribute to sub-categorisation of skin sensitising chemicals, e.g. into Sub-category 1A and Sub-category 1B as adopted in the 3rd revised version of the GHS (UN, 2009).

Based on the accuracy of GHS subcategorisation of the proposal made by Dr Omori and Dr Yoshimura, the VMG agreed that the initial results were encouraging, and that further evaluations will be necessary to determine how information generated with the h-CLAT can successfully contribute to potency subcategorisation, as described above, in the context of an integrated testing strategy (ITS).

Additional observations

Log Kow and applicability domain of the h-CLAT

Consideration is being given to better define the applicability domain of the h-CLAT. A proposal was recently made by the lead laboratories to consider relatively low water soluble compounds, i.e. chemicals with a Log Kow higher than 3.5, to be prone to false negative results in the h-CLAT and thus to fall outside the applicability of the method (Takenouchi *et al.*, accepted for publication). In our study, the chemicals with a log Kow higher than 3.5 are tabulated below.

	Chemical	**Log Kow**[1]
14	Chlorpromazine hydrochloride	5.4
16	Benzyl salicylate	4.31
17	Benzyl cinnamate	4.06
18	R(+)-Limonene	4.57

[1]Predicted data generated using the US Environmental Protection Agency's EPISuite™, http://www.epa.gov/opptintr/exposure/pubs/episuite.htm

Of the four chemicals with a Log Kow higher than 3.5, Chlorpromazine hydrochloride and R(+)-Limonene (chem.14 and 18) have been consistently and accurately classified as sensitisers by all the four laboratories whereas Benzyl salicylate and Benzyl cinnamate (chem. 16 and 17) were in the majority of cases false negatives with discordant prediction in one of the four laboratories (EURL ECVAM and Shiseido respectively).

Run acceptance criteria

The h-CLAT SOP describes a set of acceptance criteria for the evaluation runs to determine whether the results are valid. For a given run to be considered valid the following conditions should be met otherwise the run should be discarded and the chemical re-tested.

Viability:

- Medium controls: viability should be > 90%.
- DMSO controls: viability should be > 90%.
- DNCB controls: viability should be > 50%

CD54 and CD86 RFI values:

- DMSO should be negative for both markers
- DNCB should be positive for both markers

The following table (Table 19) shows the occurrence of invalid runs per laboratory for the blind testing phases (phase BI and BII).

Table 19: Overview of valid and invalid runs by laboratory.

Laboratory	Number of runs			Proportion invalid
	Total	Valid	Invalid	
Kao	69	65	4	5.8%
Shiseido	102	99	3	2.9%
Bioassay	174	162	12	6.9%
EURL ECVAM	69	64	5	7.2%

The following table (Table 20) shows a breakdown of the reasons for the invalidation of the runs by laboratory and by acceptance criteria.

Table 20: Invalid runs - reason for rejection.

Laboratory	Viability			DMSO		DNCB	
	Medium	DMSO	DNCB	CD86	CD54	CD86	CD54
Kao			1	1			3
Shiseido	2					1	
Bioassay				4		3	8
EURL ECVAM				3		2	1
Total runs	**2**		**1**	**8**		**6**	**12**

The proportion of invalid runs was generally and consistently low between laboratories (<7.5%) and for the majority of the runs that were rejected the acceptance criteria were not met because the positive control (DNCB) gave RFI values below the threshold.

Control results

A complete analysis of the control results obtained for all the runs generated during the study can be found in Annex I of the statistical report (Appendix 12).

VMG overall conclusions and recommendations

Overall Conclusions

The primary aim of this validation study was to assess the transferability, within laboratory and between laboratory reproducibility of the h-CLAT with a number of relevant coded chemicals that were judged by the VMG to be suitable and sufficiently challenging to permit robust conclusions to be drawn.

The VMG considers that the information generated in the study completes and satisfies the information requirements for modules 1-4 (test definition, within laboratory reproducibility, transferability, between laboratory reproducibility) of the ECVAM modular approach to validation (Hartung et al., 2004). In addition the information generated contributes to, but does not on its own satisfy, module 5 (predictive capacity) and module 6 (applicability domain) for which a substantial body of information is already available as evidenced in the material submitted to ECVAM and published in the scientific literature.

The main conclusions of the VMG in relation to each module are set out in the table below:

Module		Summary & Conclusions
1	Test definition	Both the existing body of evidence (original submission to ECVAM including scientific publications) and the current study findings adequately demonstrate the intended purpose, the need for, the status of development, and the mechanistic basis of the h-CLAT test method. An improved, well-detailed and robust SOP is available.
2	Within laboratory reproducibility	The overall within laboratory reproducibility was considered to be acceptable for the proposed use of the h-CLAT (i.e. as part of an integrated testing strategy).
3	Transferability	The test method was shown to be transferable between laboratories. Training and demonstration of competence in the conduct of the assay is however considered important. In particular, the chemicals used during this study's transfer phase should be considered.
4	Between laboratory reproducibility	The between laboratory reproducibility was acceptable.
5	Predictive capacity	Complete evaluation of the predictive capacity was not one of the goals of this study.

Overall, the VMG concludes that the information generated in this validation study shows that the h-CLAT is a robust and reliable test method that can contribute to the determination of the sensitisation potential of substances. Consequently:

- Information generated by the h-CLAT can already be used in a weight-of-evidence approach to support regulatory decision making, e.g. to characterise equivocal responses in *in vivo* studies (e.g. conflicting results from multiple studies).

- For the purposes of some regulations (for example REACH in the EU) a positive h-CLAT result should be considered sufficient to classify a test material as a skin sensitiser.

- The h-CLAT is suitable for further evaluation as a component of a toolbox or an ITS for full replacement of the *in vivo* assays for skin sensitisation hazard identification.

- The consistent determination of EC values observed for the two markers suggest that they may play a role in the determination of skin sensitisation potency, including GHS sub-categorisation.

Recommendations

- The predictive accuracy of the h-CLAT should be evaluated in terms of its contribution to an integrated testing strategy for full replacement of current in vivo hazard identification assays.

- Considering the outcome of the study and in particular the consistency of the results obtained with the chemicals tested previously, and the fact that the submitted SOP was amended during the study solely to provide clarifications on the procedure that was described, it is suggested that existing/historical results are taken into account for future formal evaluations on the predictive capacity of this test methods.

- GHS sub-categorisation of sensitisers should form part of a wider assessment, as on the basis of these results it is envisaged that h-CLAT EC150 and EC200 values might provide useful information contributing to this purpose.

- For hazard classification purposes, the SOP can be adapted to reduce resource costs by eliminating the need for a third evaluation run in case the first two runs are consistent.

- Further consideration of the applicability domain of the h-CLAT and its place in integrated testing strategies are merited.

List of Appendices

Appendix 1	Project Plan version 5
Appendix 2	Experimental Design version 1
Appendix 3	Chemical Selection report
Appendix 4	Report on the selection of the 9 chemicals for Phase BI
Appendix 5	Chemical coding, sampling and distribution procedures
Appendix 6	Template for CV75 determination
Appendix 7	Template for evaluation runs
Appendix 8	h-CLAT results Quality Check template
Appendix 9	h-CLAT SOP version 6
Appendix 10	h-CLAT SOP version 7
Appendix 11	h-CLAT Training and Transfer Plan version 2
Appendix 12	Statistical report
Appendix 13	Analysis of h-CLAT historical data to categorise chemicals in three classes
Appendix 14	Analysis of h-CLAT validation study data to categorise chemicals in three classes
Appendix 15	List of additional documents filed for the study and available on request

List of References

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A, Zaldivar JM. (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* 85(5):367-485.

Ashikaga, T., Sakaguchi, H., Okamoto, K., Mizuno, M., Sato, J., Yamada, T., Yoshida, M., Ota, N., Hasegawa, S., Kodama, T., Okamoto, Y., Kuwahara, H., Kosaka, N., Sono, S., Ohno, Y. (2008). Assessment of the human Cell Line Activation Test (h-CLAT) for skin sensitization; results of the first Japanese inter-laboratory study. AATEX, 13: 27–35.

Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N, Itagaki H. (2010). A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). Altern Lab Anim, 38: 275-84.

Chipinda I, Ruwona TB, Templeton SP, Siegel PD. (2011). Use of the human monocytic leukemia THP-1 cell line and co-incubation with microsomes to identify and differentiate hapten and prohapten sensitizers. Toxicology. 27:135-43.

EC (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. *Official Journal No L 353, p. 0001-1355.*

Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, Patlewicz GY, Basketter DA. (2005). Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis. 16: 157-202.

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, pp. 467–472.

Hennen, J., Aeby, P., Goebel, C., Schettgen, T., Oberli, A., Kalmes, M., Blömeke, B., (2011. Cross talk between keratinocytes and dendritic cells: impact on the prediction ofsensitization. Toxicol Sci. 123: 501-10.

Kern PS, Gerberick GF, Ryan CA, Kimber I, Aptula A, Basketter DA. (2010). Local lymph node data for the evaluation of skin sensitization alternatives: a second compilation. Dermatitis. 21: 8-32.

Kimber I, Cumberbatch M. (1992). Dendritic cells and cutaneous immune responses to chemical allergens. Toxicol Appl Pharmacol. 117: 137-46. Review.

Kosaka, N., Okamoto, Y., Mizuno, M., Yamada, T., Yoshida, M., Kodama, T., Sakiko, S., Ashikaga, T., Sato, J., Ota, N., Hasegawa, S., Okamoto, Y., Kuwahara, H., Sakaguchi, H., Ohno, Y. (2008). A study of the criteria for selection of THP-1 cells in the human Cell Line Activation Test (h-CLAT). Results of 2nd Japanese inter-laboratory study. AATEX, 13: 55-62.

Mizuno, M., Yoshida, M., Kodama, T., Kosaka, N., Okamoto, K., Sono, S., Yamada, T., Hasegawa, S., Ashikaga, T., Kuwahara, H., Sakaguchi, H., Sato, J., Ota, N., Okamoto,Y., Ohno, Y. (2008). Effects of pre-culture conditions on the human Cell Line Activation Test (h-CLAT) results: Results of the 4th Japanese inter-laboratory study. AATEX 13: 70-82.

Nukada, Y., Ashikaga, T., Miyazawa, M., Hirota, M., Sakaguchi, H., Sasa, H., Nishiyama, N. (2012). Prediction of skin sensitization potency of chemicals by human Cell Line Activation Test (h-CLAT) and an attempt at classifying skin sensitization potency. Toxicology In Vitro. 26: 1150-60.

OECD (2005). Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment. OECD series on testing and assessment No. 34. ENV/JM/MONO(2005)14.Accessible at: http://www.oecd.org/

OECD (2010). *OECD Guideline for the Testing of Chemicals No. 429*: *Skin Sensitisation: Local Lymph Node Assay*. Paris, France: Organisation for Economic Cooperation and Development. Accessible at: http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788

OECD (2012a). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Series on Testing and Assessment No. 168. Accessible at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En

OECD (2012b). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 2: Use of the AOP to Develop Chemical Categories and Integrated Assessment and Testing Approaches. Series on Testing and Assessment No. 168. Accessible at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART2&docLanguage=En

Sakaguchi H, Ashikaga T, Miyazawa M, Yoshida Y, Ito Y, Yoneyama K, Hirota M, Itagaki H, Toyoda H, Suzuki H. (2006). Development of an in vitro skin sensitization test using human cell lines; human Cell Line Activation Test (h-CLAT). II. An inter laboratory study of the h-CLAT. Toxicol In Vitro. 20:774-84.

Sakaguchi H, Ryan C, Ovigne JM, Schroeder KR, Ashikaga T. (2010). Predicting skin sensitization potential and inter-laboratory reproducibility of a human Cell Line Activation Test (h-CLAT) in the European Cosmetics Association (COLIPA) ring trials. Toxicol In Vitro. 24:1810-20.

dos Santos GG, Reinders J, Ouwehand K, Rustemeyer T, Scheper RJ, Gibbs S. (2009). Progress on the development of human in vitro dendritic cell based assays for assessment of the sensitizing potential of a compound. Toxicol Appl Pharmacol. 236: 372-82.

Takenouchi, O., Miyazawa, M., Saito, K., Ashikaga, T., Sakaguchi, H. Predictive performance of the human Cell Line Activation Test (h-CLAT) for lipophilic chemicals with high octanol-water partition coefficients. Journal of Toxicological Sciences. *Accepted for publication.*

UN (2009). *Globally Harmonised System of Classification and Labelling of Chemicals (GHS). Part 3: Health Hazards*. New York, NY, USA, and Geneva, Switzerland. United Nations Economic Commission for Europe. Accessible at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev03/03files_e.html

Yamada, T., Kosaka, N., Okamoto, Y., Mizuno, M., Sato, J., Yoshida, M., Ota, N., Kodama, T., Okamoto, Y., Kuwahara, H., Sakaguchi, H., Hasegawa, S., Ashikaga, T., Ohno, Y. (2008). A study on

serum difference on test results in the human Cell Line Activation Test (h-CLAT): Results of the 3rd Japanese inter-laboratory study. AATEX 13: 63:69.