Organisation for Economic Co-operation and Development

ENV/JM/MONO(2019)15

Unclassified

English - Or. English

28 August 2019

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

# REPORTS OF THE PEER REVIEWS OF THE epiCS AND SKIN+ TEST METHODS IN VIEW OF THEIR INCLUSION IN TEST GUIDELINE 439 ON IN VITRO SKIN IRRITATION

**Series on Testing and Assessment**
**No. 299**

Document only available in PDF

JT03450439

SERIES ON TESTING AND ASSESSMENT
NO. 299

# REPORTS OF THE PEER REVIEWS OF THE epiCS AND SKIN+ TEST METHODS IN VIEW OF THEIR INCLUSION IN TEST GUIDELINE 439 ON IN VITRO SKIN IRRITATION

IOMC
INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS
A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate
ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT
Paris 2019

**About the OECD**

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 35 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in eleven different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents;** and **Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/chemicalsafety/).

*This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.*

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

**This publication is available electronically, at no charge.**

**For this and many other Environment, Health and Safety publications, consult the OECD's World Wide Web site (www.oecd.org/chemicalsafety/)**

**or contact:**

**OECD Environment Directorate,**
**Environment, Health and Safety Division**
**2 rue André-Pascal**
**75775 Paris Cedex 16**
**France**

**Fax: (33-1) 44 30 61 80**

**E-mail: ehscont@oecd.org**

# FOREWORD

This document contains two reports related to the Peer Reviews of the Epics and Skin+ Test Methods contained in Test Guideline 439:

- Annex 1: epiCS Skin Irritation Test Methods ( CellSystems);
- Annex 2: Skin+ Reconstructed human Epidermis assay ( Sterlab)

This document includes the peer-review reports of the validation of epiCS® and Skin+® test methods. The two methods were validated following the Performance Standards for similar methods or modified in vitro Reconstructed Human Epidermis (RHE) test methods for skin irritation (No. 219 in the Series on Testing and Assessment). In 2018, Germany and France jointly proposed the project to include these test methods in the Test Guideline 439 on *in vitro* skin irritation (TG 439). Two independent peer-reviews were organised in March and September 2018 for epiCS® and Skin+®, respectively. The supporting data for each test method and the draft updated TG 439 were circulated for review and comments in October. Comments received were addressed at the meeting of the dedicated Expert Group in November 2018. The elements supporting the validation of the test methods are currently available in different files available as individual annexes, upon request to the OECD Secretariat.

In April 2019, the Working Group of the National Coordinators of the Test Guidelines Programme approved the updated TG 439, now including epiCS® and Skin+®, and endorsed the peer-review reports presented in this document.

# Annex 1: epiCs Skin Irritation Test Methods (Cellsystems)

# epiCS® Skin Irritation Test method (CellSystems)

***Expedite Peer-Review of the Performance Standard (PS)- based validation study according to OECD GD 220 for similar or modified in vitro Reconstructed human Epidermis (RhE) test methods for skin irritation testing as described in TG 439***

**Date: 21 March 2018**

## Summary

The epiCS® RhE model produced by CellSystems® (Germany) is proposed as a similar assay to the validated reference models falling within the OECD TG 439, and has undergone a performance standard (PS)-based validation study for skin irritation testing according to OECD GD 220. The assay underwent a first independent peer-review by the EURL-ECVAM Scientific Advisory Committee (ESAC) which published its opinion and recommendations in June 2016. Based on the ESAC recommendations, CellSystems® gathered additional information (in particular on technical issues encountered during the PS-based validation study and on trans-continental tissue transportation). Following availability of these additional information, a further expedited independent peer-review took place. The peer-review was coordinated by SeCAM and seh consulting+services, in close collaboration with the National Coordinators representing France and Germany at the OECD Working Group of the National Coordinators of the Test Guidelines Programme (WNT). The similarity of the two me-too assays to the validated reference method, as well as their scientific validity were assessed by the international peer-review panel (PRP) composed of:

- Dave Allen (ILS, Contractor supporting NICEATM, USA)
- Hajime Kojima (JaCVAM, NIHS, Japan)
- Ilyoung Ahn (KoCVAM, NIFDS, Republic of Korea)
- Uwe Pfannenbecker (Beiersdorf, Germany)
- John Harbell (JHarbell Consulting LLC, USA)

The criteria for peer-review evaluation were prepared by SeCAM and by the National Coordinators from France and Germany and were revised by the PRP members. The PRP members external to the OECD WNT provided with a declaration of interest that was made available to the WNT. The peer-review took place from November 2017 to February 2018, during which one initial teleconference took place between the PRP and CellSystems (where additional information was requested to CellSystems), and two additional teleconferences took place in the absence of CellSystems.

**Based on its evaluation (see detailed criteria and evaluation below), the PRP is of the opinion that the information made available to them do support the scientific similarity of the epiCS® RhE assay to the validated reference method both in terms of the essential test method components and of assay performance regarding its reproducibility and predictive capacity as described within the GD 220.**

***Evaluation criterion 1*: Rationale for the test method, including a description of the advantages of the similar or modified test method in terms of i) mechanistic advantages, applicability, predictive capacity, technical advances, reduction in hazardous reagents, ii) IP rights, geographical availability and animal welfare, iii) costs, analysis time, sample amount, competitiveness, iv) others.**

An additional model can favour market competitiveness and ensure the availability of models e.g., in case of cease of production or regional unavailability of a model. The epiCS® RhE assay offers the same mechanistic, technical and animal welfare advantages as the currently adopted test methods falling within TG 439. Based on the information provided by the tissue developers, its costs are also within the range of costs from the currently adopted RhE models falling within TG 439.

***Evaluation criterion 2*: A detailed protocol for the similar or modified test method should be available.**

A detailed protocol of the similar test method is available and was considered to be adequate by the Peer Review Panel (PRP).

***Evaluation criterion 3:* Adherence to the essential test method components as described in paragraphs 6 to 23 of GD 220 should be demonstrated for the similar or modified test methods regarding i.e., the general conditions, the functional conditions and the procedural conditions.** *Note: For specific parameters or modified procedures, adequate values or procedures should be provided for the proposed similar or modified test method (e.g. Tables 1 and 2 of GD 220).*

The PRP considered the similarity of the epiCS® RhE assay in terms of essential test method components, to be adequate and sufficient and recommends the tissue developer to continue to assess, as a tissue batch production release criteria, the barrier function of the epiCS® RhE tissue batches based on at least two time-points (one above and one beneath the $ET_{50}$ value), as implemented in January 2018 and reported in the certificate of analysis provided to the PRP (Annexes 1 and 2).

***Evaluation criterion 4*: In addition, for modified test methods, the toxicological mechanisms and the relationship between the test method endpoint(s) with the biological effect as well as the toxicity of interest should be addressed, describing limitations of the test method.**

The criterion is not applicable as the epiCS® RhE model does not represent a modified test method, but a functionally and mechanistically similar assay.

***Evaluation criterion 5:* At least the 20 recommended reference chemicals within GD 220 should be tested with the similar or modified test method according to recommendations of paragraphs 24, 25, 26 and 31 of GD 220, to demonstrate reliability and accuracy.** *I.e., testing should be conducted in at least three independent laboratories if the similar or modified test method is proposed to be used in several laboratories. In each laboratory, each chemical should be tested in three independent runs performed with different tissue batches and at sufficiently spaced time points. Each run should consist of at least three concurrently tested tissue replicates. Furthermore,*

*additional chemicals representing other chemical classes and for which adequate in vivo reference data are available may be tested in addition to the minimum list of reference chemicals.*
*Note: the use of the reference chemicals for development/optimization of the proposed similar or modified test method should be avoided to the extent possible.*

Based on the information provided to the PRP, a total of 44 test chemicals were used for the development of the epiCS® RhE assay. The list provided to the PRP encompassed 19 reference chemicals (9 Cat. 2, 3 Cat. 3 and 7 No Cat. according to the UN GHS classification scheme), and 25 additional chemicals distributed as follows:

- 4 UN GHS Cat. 2 chemicals
- 10 UN GHS Cat. 3 chemicals (EU CLP No Cat.)
- 7 UN GHS No Cat. chemicals
- 4 chemicals with unclear identification and/or *in vivo* UN GHS classification

From these 25 chemicals, 22 corresponded to chemicals used in the validation, optimization and follow-up studies of the test methods adopted within TG 439[1,2,3] whereas 3 of the 4 chemicals having unclear identification and/or *in vivo* classification came from other sources. Out of the test chemicals having a clear identification and *in vivo* classification (19 reference chemicals and 21 additional chemicals), a total of 7 were solid test chemicals (4 of which are reference chemicals) representing circa 18% (7 out of 40). For comparison, solids represent 20% of the reference chemicals (4 out of 20) and 25% of the chemicals used in the skin irritation validation study (25 out of 60). Finally, although based on a different dataset, the predictive capacity obtained with the 40 chemicals having a clear identification and *in vivo* classification was of the same order of magnitude as the currently adopted RhE models (based on the dataset collected from literature).

Based on the above information and the testing of 21 chemicals having clear identification and *in vivo* classification in addition to the reference chemicals, the PRP considers the information provided as sufficient. It would have been ideal however, to have a number of chemicals, different from the chemicals used to develop the model, tested in a blind manner.

***Evaluation criterion 6*: The reliability obtained with the reference chemicals should be equal to or better than the defined minimum target values for the similar or modified test method as specified in paragraphs 28 and 29 of GD 220, and the analyses of reliability should be conducted according to the specifications described in paragraph 27 of GD 220.**

In its evaluation, ESAC considered that the epiCS did not meet the criteria for within-laboratory reproducibility in one of the participating laboratory situated across ocean (89% (16/18) versus 90% as required by the performance standards) whereas the within-laboratory reproducibility in the two other laboratories and the between-laboratory reproducibility were considered appropriate. This was due to the fact that GD 220 allows for a maximum of 5 test runs for each chemical, whereas in the laboratory in question, two chemicals had more than 5 test runs. As a consequence, ESAC did not account for the additional runs, and considered the two chemicals to have incomplete test

[1] Eskes C, Cole T, Hoffmann S, Worth A, Cockshott A, Gerner I, Zuang V (2007). The ECVAM international validation study on *in vitro* tests for acute skin irritation: selection of test chemicals. *ATLA* **35**, 603-619.

[2] Kandárová H, Hayden P, Klausner M, Kubilus J, Kearney P., Sheasgreen J. (2009). In Vitro Skin Irritation Testing: Improving the Sensitivity of the EpiDerm Skin Irritation Test Protocol. *ATLA* **37**: 671–689.

[3] Cotovio J., Grandidier M.-H., Portes P., Roguet R., Rubinstenn G. (2005). The *in vitro* acute skin irritation of chemicals : optimisation of the EPISKIN prediction model within the framework of the ECVAM validation process. *ATLA* **33**, 329-349.

sequences (Annex 3). However, the trans-continental laboratory was able to demonstrate that the additional runs were performed due to technical errors (e.g., contamination and non-exposure to the positive control [Annexes 4 and 5]).

The present PRP is of the opinion that technical issues (such as contamination) are not a reflection of the performance of the RhE model, so that it should not be taken into account when comparing the scientific similarity between assays. Furthermore, as suggested within the originally proposed performance standards developed by EURL-ECVAM from 2009, a margin of tolerance shall be accepted when comparing the performance of similar assays especially when scientifically justified.

When the test runs having technical issues are not considered for the comparison of similarity, the within-laboratory reproducibility of the trans-continental laboratory and between-laboratory reproducibility do fall within the requirements of the GD 220 (see table below as extracted from Annex 6). Furthermore, following the ESAC peer-review, the tissue developer has optimized the conditions for trans-continental transportation, leading to controlled temperature conditions, and simulation experiments indicated appropriate cell viability with 1% Triton X-100, used to assess barrier function (Annex 7).

Based on the above considerations, the present PRP considers that the reliability of the epiCS® RhE assay obtained with the reference chemicals (summarized below) to be sufficient. Furthermore, it suggests that in case future trans-continental transportation is conducted, that the barrier function of the tissues is assessed by measuring viability to 1% Triton X-100 for at least two time-point (i.e., one above and one beneath the $ET_{50}$) in order to confirm the barrier properties of the tissues transported across continents.

| | **epiCS® RhE** | **PS target values** |
|---|---|---|
| Within-laboratory reproducibility | 95% (18/19), 100% (20/20), 90% (18/20) | ≥ 90% |
| Between-laboratory reproducibility | 84% (16/19) | ≥ 80% |
| % of complete run sequences / lab. | 95%, 100%, 100% | ≥ 85% |
| % of complete run sequences over the 3 labs | 98% | ≥ 90% |

***<u>Evaluation criterion 7</u>: The predictive capacity obtained with the reference chemicals should be equal to or better than the defined minimum target values for the similar or modified test method as specified in paragraph 30 of GD 220, and the analyses of predictive capacity should be conducted according to the specifications described in paragraph 27 of GD 220.***

The PRP considered the predictive capacity of the epiCS® RhE assay obtained with the reference chemicals (summarized below) to be adequate and sufficient.

| | **epiCS® RhE (average 3 labs*)** | **PS target values** |
|---|---|---|
| Sensitivity (n=10) | **87%** | ≥ 80% |
| Specificity (n=10) | **80%** | ≥ 70% |
| Accuracy (n=20) | **83%** | ≥ 75% |

* as calculated within the EURL-ECVAM template (Annex 6)

***Evaluation criterion 8*: The applicability domain of the new or modified test method should be defined.**

A total of 21 test chemicals with clear identification and *in vivo* classification have been tested in addition to the 20 reference chemicals, providing information on the applicability domain of the epiCS® RhE similar assay, so that the provided information was considered to be sufficient by the PRP.

***Evaluation criterion 9:* All data from the PS-based validation study supporting the validity of the similar or modified test method should be obtained in accordance with the principles of Good Laboratory Practice (GLP).**

According to the information provided to the PRP, the study was conducted according to GLP principles.

***Evaluation criterion 10:* Completeness of all data and documents supporting the assessment of the validity of the similar or modified test method.**

The information provided by the test method developer was considered to be sufficient for the assessment of the validity of the epiCS® RhE assay.

***Evaluation criterion 11:* PS-based validation study management and conduct.**

The PRP considered the information provided on the study management conduct to be adequate and sufficient.

***Evaluation criterion 12: Other considerations: A) quality control procedures for lot release; B) audit of tissue production.***

A) The PRP considered the quality control procedures for lot release sufficient provided that the tissue developer continues to assess the barrier function of the epiCS® RhE tissue batches, based on at least two time-points (i.e., one above and one below the $ET_{50}$ value), similarly to the certificate of analysis provided to the PRP from the 8 and 22 January 2018 (Annexes 1 and 2).

B) Based on the information provided to the PRP, the production of the epiCS® RhE tissues has been independently audited within the framework of an ISO 9001:2008 certification (Annex 8).

***Evaluation criterion 13:* All data should adequately support the peer review assessment that the proposed test method is structurally and mechanistically similar to the validated reference method, and demonstrate sufficient reliability and relevance for the proposed specific testing purpose i.e., that the proposed similar or modified test method is scientifically valid.**

The data assessed by the peer-review panel supports the scientific similarity of the epiCS® RhE assay to the validated reference method both in terms of the essential test method components and the performances of the assay in terms of reproducibility and predictive capacity as described within the GD 220.

**List of Annexes**

**Annex 1:** Certificate of Analyses of the epiCS® RhE from 8 January 2018

**Annex 2:** Certificate of Analyses of the epiCS® RhE from 22 January 2018

**Annex 3:** ESAC opinion on the epiCS® RhE performance-based validation study for in vitro skin irritation testing from 24 June 2016 (with relevant text highlighted in yellow)

**Annex 4:** Letter from US-based laboratory explaining technical issues encountered during testing from 1st September 2016

**Annex 5:** Pictures of the invalid runs (December 2017)

**Annex 6:** epiCS® RhE results according to the EURL-ECVAM reporting template (from October 2015)

**Annex 7:** CellSystems' responses to the PRP questions regarding transportation issues & quality of the model from November 2017

**Annex 8:** ISO 9001:2008 certificate including 'production of research products'

# Annex 2: Skin+ Reconstructed human Epidermis assay (Sterlab)

# Skin+® Reconstructed human Epidermis assay (Sterlab)

***Expedite Peer-Review of the Performance Standard (PS)- based validation study according to OECD GD 220 for similar or modified in vitro Reconstructed human Epidermis (RhE) test methods for skin irritation testing as described in TG 439***

**Date: 13 September 2018**


## Summary

The Skin+® Reconstructed human Epidermis (RhE) model produced by Sterlab (France) is proposed as a similar assay to the validated reference models falling within the OECD TG 439, and has undergone a performance standard (PS)-based validation study for skin irritation testing according to OECD GD 220. The PS-based validation study was submitted in June 2015 to EURL-ECVAM, which provided an assessment report of the test submission in December 2015 (Annex 1). In this report a number of questions were raised including the need to have a statistical report of the PS-based study, to explain the discrepancies observed in data entry and data analyses and the need to clarify which chemicals were used to define the final version of the SOP. Based on these recommendations, Sterlab prepared a reply (Annex 2), and contracted out an independent biostatistician that went over the PS-based validation study results, to check and correct for any data discrepancy, and to provide an independent biostatistical report (Annex 3).

Following availability of this additional information, the present expedite independent peer-review took place. The peer-review was coordinated by SeCAM and seh consulting+services, in close collaboration with the National Coordinators representing France and Germany at the OECD Working Group of the National Coordinators of the Test Guidelines Programme (WNT). The similarity of the me-too assay to the validated reference method, as well as its scientific validity was assessed by the international Peer Review Panel (PRP) composed of:

- Dave Allen (ILS, Contractor supporting NICEATM, USA)
- Hajime Kojima (JaCVAM, NIHS, Japan)
- Ilyoung Ahn (KoCVAM, NIFDS, Republic of Korea)
- Uwe Pfannenbecker (Beiersdorf, Germany)
- John Harbell (JHarbell Consulting LLC, USA)

The criteria for peer-review evaluation were prepared by SeCAM and by the National Coordinators from France and Germany and were revised by the PRP members. The PRP members external to the OECD WNT provided with a declaration of interest that was made available to the WNT. The peer-review took place from November 2017 to July 2018, during which two teleconferences took place between the PRP and Sterlab (where additional information was requested to Sterlab), and three additional teleconferences took place in the absence of Sterlab.

**Based on its evaluation (see detailed criteria and evaluation below), the PRP is of the opinion that the information made available for peer-review support the scientific similarity of the Skin+® RhE assay, when conducted within Europe, to the validated reference method both in terms of the essential test method components and of assay performance regarding its reproducibility and predictive capacity as described within the GD 220.** In case trans-continental transportation is foreseen, it is recommended that the barrier function of the tissues is

assessed after the trans-continental transportation in order to confirm the barrier properties of the tissues transported across continents[1].

**Evaluation criterion 1: Rationale for the test method, including a description of the advantages of the similar or modified test method in terms of i) mechanistic advantages, applicability, predictive capacity, technical advances, reduction in hazardous reagents, ii) IP rights, geographical availability and animal welfare, iii) costs, analysis time, sample amount, competitiveness, iv) others.**

An additional model can favour market competitiveness and ensure the availability of models e.g., in case of cease of production or regional unavailability of a model. The Skin+® RhE assay offers the same mechanistic, technical and animal welfare advantages as the currently adopted test methods falling within TG 439. The Peer Review Panel (PRP) could not make an economical comparison as it had no access to the costs of the Skin+® RhE assay. Furthermore, regarding the geographical availability of the tissue, the PRP had no access to information regarding the functional conditions of the tissues after trans-continental shipment[1]. It is therefore recommended that in case trans-continental transportation is foreseen, that the barrier function of the tissues is assessed by measuring viability to 1% Triton X-100 for at least two time-points (i.e., one above and one beneath the $ET_{50}$) after the trans-continental transportation (as described under evaluation criterion 12.A) in order to confirm the barrier properties of the tissues transported across continents[1].

**Evaluation criterion 2: A detailed protocol for the similar or modified test method should be available.**

A detailed protocol of the similar test method is available and was considered to be adequate by the PRP (Annex 4).

**Evaluation criterion 3: Adherence to the essential test method components as described in paragraphs 6 to 23 of GD 220 should be demonstrated for the similar or modified test methods regarding i.e., the general conditions, the functional conditions and the procedural conditions.** *Note: For specific parameters or modified procedures, adequate values or procedures should be provided for the proposed similar or modified test method (e.g. Tables 1 and 2 of GD 220).*

The PRP considered the similarity of the Skin+® RhE assay in terms of essential test method components, to be adequate and sufficient.

**Evaluation criterion 4: In addition, for modified test methods, the toxicological mechanisms and the relationship between the test method endpoint(s) with the biological effect as well as the toxicity of interest should be addressed, describing limitations of the test method.**

The criterion is not applicable as the Skin+® RhE assay does not represent a modified test method, but a similar assay.

[1] Post-evaluation note: the PRP was informed that a study on transatlantic shipment of the tissues is being conducted.

**Evaluation criterion 5: At least the 20 recommended reference chemicals within GD 220 should be tested with the similar or modified test method according to recommendations of paragraphs 24, 25, 26 and 31 of GD 220, to demonstrate reliability and accuracy.** *I.e., testing should be conducted in at least three independent laboratories if the similar or modified test method is proposed to be used in several laboratories. In each laboratory, each chemical should be tested in three independent runs performed with different tissue batches and at sufficiently spaced time points. Each run should consist of at least three concurrently tested tissue replicates. Furthermore, additional chemicals representing other chemical classes and for which adequate in vivo reference data are available may be tested in addition to the minimum list of reference chemicals. Note: the use of the reference chemicals for development/optimization of the proposed similar or modified test method should be avoided to the extent possible.*

The test method developer states in its answers to PRP questions from 21 November 2017, that the 20 reference chemicals listed in the OECD GD 220 have been used to develop the RhE test method, and that no additional chemicals to these 20 reference chemicals have been tested with the model. GD 220 states however that "*the exclusive use of the reference chemicals for development/optimization of the proposed similar or modified test method should be avoided to the extent possible*". Furthermore, TG 439 (both versions from 2013 and 2010) states that "*The Reference Chemicals represent the minimum number of chemicals that should be used to evaluate the accuracy and reliability of a proposed similar or modified test method, but should not be used for the development of new test methods*."

Based on the above considerations, the PRP recommended in February 2018 that additional chemicals are tested with the Skin+® RhE assay such as the test chemicals used during the validation, optimization and follow-up studies of the test methods adopted within TG 439[2,3,4] in at least one laboratory and preferentially in a blind manner, in order to gain further information on the predictive capacity and applicability domain of the assay.

Following this recommendation, Sterlab proceeded with the testing of additional 17 test chemicals that were selected by the PRP in collaboration with an independent biostatistician (Els Adriaens) and an independent test chemical distribution facility (VITO, Belgium). These test chemicals encompassed 6 UN GHS Cat. 2 chemicals, 3 UN GHS Cat. 3 (EU CLP No Cat.) chemicals and 8 UN GHS No Cat. chemicals. The 17 additional chemicals were tested under blind conditions by Sterlab and results were summarized by the independent biostatistician (Annex 5).

Based on the above information and the blind testing of the additional 17 chemicals having clear identification and *in vivo* classification in addition to the reference chemicals, the PRP considered the information provided as sufficient.

---

[2] Eskes C, Cole T, Hoffmann S, Worth A, Cockshott A, Gerner I, Zuang V (2007). The ECVAM international validation study on *in vitro* tests for acute skin irritation: selection of test chemicals. *ATLA* **35**, 603-619.

[3] Kandárová H, Hayden P, Klausner M, Kubilus J, Kearney P., Sheasgreen J. (2009). In Vitro Skin Irritation Testing: Improving the Sensitivity of the EpiDerm Skin Irritation Test Protocol. *ATLA* **37**: 671–689.

[4] Cotovio J., Grandidier M.-H., Portes P., Roguet R., Rubinstenn G. (2005). The *in vitro* acute skin irritation of chemicals : optimisation of the EPISKIN prediction model within the framework of the ECVAM validation process. *ATLA* **33**, 329-349.

**<u>Evaluation criterion 6</u>: The reliability obtained with the reference chemicals should be equal to or better than the defined minimum target values for the similar or modified test method as specified in paragraphs 28 and 29 of GD 220, and the analyses of reliability should be conducted according to the specifications described in paragraph 27 of GD 220.**

The RPR considered the reliability of the Skin+® RhE assay obtained with the reference chemicals (summarized below) to be adequate and sufficient.

| | **Skin+® RhE assay** | **PS target values** |
|---|---|---|
| Within-laboratory reproducibility | 95% (19/20), 95% (19/20), 90% (18/20) | ≥ 90% |
| Between-laboratory reproducibility | 100% (20/20) | ≥ 80% |
| % of complete run sequences / lab. | 100%, 100%, 100% | ≥ 85% |
| % of complete run sequences over the 3 labs | 100% | ≥ 90% |

Furthermore, when combining the results of reference chemicals with the additional 17 coded test chemicals, a within-laboratory concordance of predictions of 91.7% (33/36) was found within the Sterlab laboratory.

**<u>Evaluation criterion 7</u>: The predictive capacity obtained with the reference chemicals should be equal to or better than the defined minimum target values for the similar or modified test method as specified in paragraph 30 of GD 220, and the analyses of predictive capacity should be conducted according to the specifications described in paragraph 27 of GD 220.**

The PRP considered the predictive capacity of the Skin+® RhE assay obtained with the reference chemicals (summarized below) to be adequate and sufficient.

| | **Skin+® RhE assay (average 3 labs)** | **PS target values** |
|---|---|---|
| Sensitivity (n=10) | **100%** | ≥ 80% |
| Specificity (n=10) | **70%** | ≥ 70% |
| Accuracy (n=20) | **85%** | ≥ 75% |

* as calculated within the EURL-ECVAM template

Furthermore, when combining the results of the reference chemicals with the additional 17 chemicals tested in a blind manner, a sensitivity of 93.8% (15/16), a specificity of 76.2% (16/21) and an accuracy of 83.8% (31/37) were found.

**<u>Evaluation criterion 8</u>: The applicability domain of the new or modified test method should be defined.**

The PRP is of the opinion that, based on the data presented with the reference chemicals and on the results obtained with the additional 17 test chemicals tested in a blind manner (see criterion 5, 6 and 7), there is no reason to consider the applicability domain of the Skin+® RhE to be different than the applicability domain of the test methods adopted within TG 439.

**Evaluation criterion 9: All data from the PS-based validation study supporting the validity of the similar or modified test method should be obtained in accordance with the principles of Good Laboratory Practice (GLP).**

Based on the information provided to the PRP, the study was conducted according to GLP principles.

**Evaluation criterion 10: Completeness of all data and documents supporting the assessment of the validity of the similar or modified test method.**

The information provided by the test method developer was considered to be sufficient for the assessment of the validity of the Skin+® RhE assay.

**Evaluation criterion 11: PS-based validation study management and conduct.**

The PS-based validation study of the Skin+® RhE assay attempted to follow the currently accepted standards for validation study. The PRP considered the information provided on the study management and conduct to be sufficient. However, it is noted that the coding of test chemicals during the PS-based validation study could have been better randomized. In addition, the EURL-ECVAM evaluation from December 2015 stated that "*a discrepancy of viability values*" was found for two of the three participating laboratories. EURL-ECVAM indicated however that this discrepancy had "*no influence on the predictions yielded*" and that "*the differences are mostly in the decimal range*...", "...*however too large to be explained as rounding error*...". Sterlab answered that such a discrepancy may be due to the use of different softwares (Microsoft Excel and Libre Office). The present PRP was informed that following the EURL-ECVAM evaluation, all data templates were reviewed, and the viability values were corrected when necessary by Sterlab's technician. Furthermore, an independent statistician analyzed the data and prepared a statistical report (Annex 3). During its evaluation, the PRP noticed that the data reported in the independent statistical report did correspond to the data reported in the EURL-ECVAM template provided to the PRP by the test developer.

**Evaluation criterion 12: Other considerations: A) quality control procedures for lot release; B) audit of tissue production.**

A) The PRP considered the quality control procedures for lot release sufficient (Annexes 6a, 6b and 6c) provided that the tissue developer continues to:

- assess the barrier function of the Skin+® RhE tissue batches based on a sufficient number of time points that cover the $ET_{50}$ acceptability range ($4.0 \leq ET_{50}$ (hr) $\leq 9.0$) such as for example 0, 2, 4 and 9 hours;

- calculates the $ET_{50}$ based on one time-point just above and one time-point just below the $ET_{50}$ and not based on a linear correlation, since the correlation between time of exposure and cell viability may not always follow a linear correlation.

B) Based on the information provided to the PRP, the production of the Skin+® RhE tissues has been independently audited within the framework of an ISO 9001:2008 certification (Annexes 7 and 8).

**Evaluation criterion 13: All data should adequately support the peer review assessment that the proposed test method is structurally and mechanistically similar to the validated reference method, and demonstrate sufficient reliability and relevance for the proposed specific testing purpose i.e., that the proposed similar or modified test method is scientifically valid.**

The data assessed by the peer-review panel supported the scientific similarity of the Skin+® RhE assay, when the method is conducted within Europe, to the validated reference method both in terms of the essential test method components and of assay performance regarding its reproducibility and predictive capacity as described within the GD 220. In case trans-continental transportation is foreseen, it is recommended that that the barrier function of the tissues is assessed after the trans-continental transportation in order to confirm the barrier properties of the tissues transported across continents[5].

## List of Annexes

**Annex 1:** EURL-ECVAM assessment report on the Skin® RhE PS-validation study from 17 Dec. 2015

**Annex 2:** Reply from Sterlab to EURL-ECVAM dated from May 2016

**Annex 3:** Independent statistical report on the Skin® RhE PS-validation study from 1st July 2016

**Annex 4:** Standard Operating Procedures (SOP) of the Sterlab's Skin® RhE test method for skin irritation testing

**Annex 5:** Independent statistical report on the Skin® RhE testing of additional blind chemicals from May 2018

**Annex 6a:** Certificate of Analyses of the Skin® RhE model from 10 April 2018 and 19 March 2018

**Annex 6b:** ET50 calculations for the batch from 10 April 2018

**Annex 6c:** Calculations for the batch from 19 March 2018

**Annex 7:** ISO 9001:2008 certificate including 'production of human or animal tissues, rebuilt *in vitro*, in aim of tests' from 6 January 2012

**Annex 8:** ISO 9001:2008 certificate including 'production of human or animal tissues, rebuilt *in vitro*, in aim of tests' from 7 January 2015

[5] Post-evaluation note: the PRP was informed that a study on transatlantic shipment of the tissues is being conducted.