# EpiSensA

# Report of the Peer Review Panel

on

a JaCVAM coordinated study program addressing the validation status of the EpiSensA for prospective identification of skin sensitizing chemicals

Report completed by the Peer Review Panel on *November 3, 2022*.

## Table of Contents

## Executive Summary

The EpiSensA has been proposed as an in vitro skin sensitization alternative method, providing information on key event 2 (KE2) in the adverse outcome pathway. The EpiSensA Peer Review Panel (PRP) believes the validation study led by test developers and Japanese Center for the Validation of Alternative Methods (JaCVAM) was successful. The assay is a reconstructed human epidermis (RhE) model that demonstrates advantages for testing of lipophilic compounds and pre- and pro-haptens. The validation study conducted for the EpiSensA demonstrated the assay's applicability, transferability, and inter- and intra-laboratory reproducibility. Both within- and between-laboratory reproducibility easily met the pre-specified 85% and 80% thresholds, respectively. The proficiency chemicals and individual assay performance standards have been provided along with a detailed Standard Operating Procedure (SOP). The EpiSensA provides similar predictivity as other validated methods for the same endpoint with increased capacity to correctly identify pre/pro-haptens and lipophilic compounds.

## Peer Review Panel Composition

G. Frank Gerberick, PhD (PRP Chair) – GF3 Consultancy, USA
Emily N. Reinke, PhD – Integrated Laboratory Systems, LLC an Inotiv Company, Contractor supporting the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), USA
Emanuela Corsini, PhD – Università degli Studi di Milano, Italy
Atsushi Ono, PhD – Okayama University, Japan
Tomoki Fukuyama, DVM, PhD – Azabu University, Japan

Support from Takao Ashikaga, PhD and Hajime Kojima, PhD – JaCVAM, Japan

## Background

Significant progress has been achieved in developing New Approach Methods (NAM) for assessing the skin sensitization potential of chemicals over the last two decades. The OECD developed the adverse outcome pathway for skin sensitization divided into mechanistic key events (OECD 2014). Three OECD guidelines have been published that cover these mechanistic events (covalent binding to protein, keratinocyte activation, and dendritic cell activation). Eight non-animal test methods are approved in OECD TG 442C, 442D, and 442E (OECD 2018a; 2018b; 2021b). However, each of the test methods has some limitations, especially with testing challenging test materials such as highly lipophilic compounds.

The current focus is to find ways to combine *in vitro, in chemico,* and *in silico* assessments with read-across predictions from similar chemicals to generate integrated approaches to testing and assessment (IATA) or Defined Approach (DA). The OECD has published a new guideline (No. 497) that describes three simple DAs for assessing skin sensitization (OECD 2021a). Combining NAM data will enhance the vigor of skin sensitization hazard identification calls on chemicals and further address the critical need for predicting a chemical's potency. Potency prediction is needed for the UN Global Harmonized System (GHS) subclassification of sensitizers into 1A (strong sensitizers) and 1B (other sensitizers). For the specific need of GHS subclassification into sub-category 1A, the kDPRA assay is accepted as a standalone assay (Natsch et al. 2020; OECD 2021b). However, assessing potency is also required for conducting Next Generation Risk Assessments on new chemical entities where only non-animal information is available (Api et al. 2020; Bernauer et al. 2021; Dent et al. 2018; Gilmour et al. 2020).

Another focus area in NAM development is to have reconstructed human epidermis (RhE) models available for testing materials that are incompatible with solution-based systems or may require metabolic activation of the test materials under evaluation. The RhE EpiSensA assay permits the ability to evaluate more lipophilic compounds. The model also has some metabolic capacity, so pre- and pro-haptens can be assessed. The assay addresses KE2 of the skin sensitization AOP (OECD 2014).

Against this background, the Japanese company Kao Corporation undertook to address certain limitations of current methods to develop a RhE model consisting of normal human-derived epidermal keratinocytes. The EpiSensA evaluates the gene expression of four relevant markers, including (i) the encoding activating transcription factor 3 (*ATF3*), (ii) the glutamate-cysteine ligase, modifier subunit (*GCLM*), (iii) the DnaJ (Hsp40) homolog, subfamily B, member 4 (*DNAJB4*), and (iv) interleukin-8 (*IL-8*). The expression of these genes reflects the keratinocyte response to the early phase of skin sensitization, including (i) the induction of cyto-protective gene pathways (e.g., the antioxidant response element (ARE)-dependent pathways) and (ii) the inflammatory response (e.g., the induction of inflammatory cytokines) (OECD 2014)(VMT Report, May 2022). The EpiSensA method, like any new method under evaluation, must be judged independently, in terms of its relevance and reliability via the process of validation

The PRP held its kick-off meeting on June 1, 2022 and continued to interact through November 1, 2022. A face-to-face meeting was held in Tokyo, Japan, on September 8-9, 2022. The PRP focused on reviewing the detailed report on the EpiSensA prepared by the Validation Management Team (VMT). Following the commentary on this work by the PRP, the VMT clarified some minor details in the protocol and made some adjustments to their validation report. The PRP engaged in assessing the assay's validation status, leading to the generation of this EpiSensA Validation Report.

## Evaluation Criterion

**Evaluation Criterion 1: A rationale for the test method should be available, including a description of the human health effect, a clear statement of scientific need, and regulatory application.**

The EpiSensA is an *in vitro* test method designed to be part of a non-animal test battery or an integrated testing strategy for assessing the skin sensitization potential of chemicals (Mizumachi et al. 2021; Mizumachi et al. 2018; Saito et al. 2013; Saito et al. 2017). The validation report does provide a rationale for the test method as well as the published references by the test developer. The report states that while there are already validated alternative methodologies for Key Events 1-3 of the skin sensitization adverse outcome pathway, each has defined limitations, especially concerning low solubility of lipophilic compounds and limited metabolic capabilities to detect pre- and pro-haptens. The EpiSensA assay highlights that the use of reconstructed human epidermis (RhE) allows the ability to assess more lipophilic compounds. As the model also has some metabolic capacity, pre- and pro-haptens can also be evaluated. The assay is designed to address KE2 in the AOP, and the current validation effort was designed to assess technical transferability and reliability for GHS hazard classification in the labelling of skin sensitization and as part of an ongoing effort for it to be used as part of the already existing Defined Approaches for skin sensitization. There are immediate applications and improvements in using the 3D epidermal EpiSensA model and it will be of great interest to regulators worldwide.

**Evaluation Criterion 2: The toxicological mechanisms and the relationship between the test method endpoint(s) with the biological effect and the toxicity of interest should be addressed, describing the limitations of the test method.**

The EpiSensA prediction model evaluates the gene expression of four relevant markers, including (i) the encoding activating transcription factor 3 (*ATF3*), (ii) the glutamate-cysteine ligase, modifier subunit (*GCLM*), (iii) the DnaJ (Hsp40) homolog, subfamily B, member 4 (*DNAJB4*), and (iv) interleukin-8 (*IL-8*). The expression of these genes reflects the keratinocyte response to the early phase of skin sensitization, including (i) the induction of cyto-protective gene pathways (e.g., the antioxidant response element (ARE)-dependent pathways) and (ii) the inflammatory response (e.g., the induction of inflammatory cytokines) (OECD 2014)(VMT Report, May 2022). Thus, the EpiSensA test method addresses a biological mechanism: inflammatory and cytoprotection-related gene expression of keratinocytes, which is known to be associated with underlying mechanisms of skin sensitization defined in the AOP (Sullivan et al. 2017).

Analyses showed that the expression of the four genes (*GCLM*, *DNAJB4*, *IL-8*, *ATF3*) provides reliable and predictive biomarkers for the discrimination of sensitizers from non-sensitizers (Saito et al. 2013; Saito et al. 2017). In an early study, 16 chemicals (8 sensitizers, 4 non-sensitizers, and 4 pre-/pro-haptens) were used to assess the predictive performance of the final four marker gene candidates. Based on testing 72 test chemicals containing 54 sensitizers and

18 non-sensitizers, the cut-off values of the respective marker genes were maximized. In the end, the cut-off values were set (ATF3: 15-fold, GCLM: 2-fold, DNAJB4: 2-fold, and IL-8: 4-fold). The cut-offs were found to be appropriate when analyzing the data from 136 chemicals.

Importantly, limitations of the assay have been highlighted and include: (i) if the chemical is not soluble or does not form a stable dispersion at 0.0122% in any vehicle, the chemical is not applicable for testing with EpiSensA; (ii) chemicals that affect the expression of *GAPDH* might not be suitable for testing; (iii) Chemicals that directly interfere with both LDH and MTT; and (iv) chemicals that affect RNA isolation. It is also noted that increases in the gene expression of *ATF3* and *IL-8* are also observed, along with a decrease in cell viability when some surfactants (such as sodium lauryl sulfate and benzalkonium chloride) are tested on EpiSensA. In addition, the EpiSensA assay has an inherent limitation in detecting some weak sensitizers, the same as the other threshold-based prediction models.

**Evaluation Criterion 3: A detailed test method protocol should be available. Transferability to other labs should be demonstrated.**

A detailed test method is available, written in such a way that a naïve facility should be able to easily pick up the test method for their use. The validation report highlighted that the method was transferable from the lead laboratory to three supporting facilities with some minor adjustments to the protocol, which were captured in subsequent versions. Participating laboratories underwent training to learn the method and returned to their home facilities to test 4 chemicals for the potential to cause skin sensitization. All test chemicals were appropriately classified by the testing facilities, with one test facility conducting testing with two independent technicians, both ultimately successful. However, one technician did require more tests to meet the criteria. The technology transfer portion of this validation study also allowed for additional refinement of the SOP to alter the positive control test chemicals due to a variety of factors, adjust the viability acceptance criteria, and add endogenous control gene (GAPDH) Ct criteria.

Specific changes to the SOP included changing the positive control from 6.25% w/v bisphenol A diglycidyl ether (BADGE) to 0.78 w/v% clotrimazole and 0.10% w/v 4-nitrobenzylbromide (4NBB) because of the poor stability and the inaccurate concentration of the commercially available BADGE. The acceptance criteria for the mean viability of the vehicle control have changed from 80% to 90%. The new acceptance criterion (requiring a ≥95% cell viability in each vehicle control well) was set based on the statistical analysis of the cell viability distribution of the vehicle control obtained by the lead laboratory.

If there is a scientific rationale with supporting experimental data, other housekeeping genes or RNA extraction isolation techniques could be used. It is recommended that such changes to the SOP will require an assessment of defined proficiency chemicals and performance controls.

**Evaluation Criterion 4: The within and between laboratory reproducibility of the test method should be demonstrated.**

Both within and between laboratory reproducibility have been demonstrated in the validation project for EpiSensA, as measured using concordance of classification. Phase I was conducted with 15 chemicals tested in three independent experiments, and the concordance ranged between 86.7 to 93.3% for the participating laboratories. Combining these results with the 12 Phase II chemicals, tested once, resulted in a between-laboratory reproducibility of 88.9%. These are acceptable levels of concordance for both within and between laboratory reproducibility and exceed the targeted 85% for within-laboratory and 80% for between-laboratory reproducibility.

Following some adjustments in the acceptance criteria, the results support the EpiSensA to be reproducible within and between laboratories, satisfying the performance criteria for the EpiSensA.

**Evaluation Criterion 5: Demonstration of the test method's performance should be based on testing a diversity of chemicals, preferably coded reference chemicals.**

JaCVAM oversaw test chemical distribution and coding. Coded chemicals were distributed to participating laboratories and pertinent safety information was provided to a chemical master at each facility. At the conclusion of the validation studies, the code was not broken by any test facility. The validation management team selected chemicals by utilizing published in vitro skin sensitization studies and previously conducted validation studies for in vitro skin sensitization. Chemicals needed to meet a range of criteria for inclusion, such as background information on mode or site of action, quality of the reference data, coverage of relevant chemical and product classes, a range of potencies, including non-sensitizers, information on pre-and pro-hapten potential, physical and chemical properties, commercial availability, and high purity. Therefore, chemical coding, distribution, and handling were conducted correctly and supervised by JaCVAM.

The EpiSensA performed well with a diversity of chemicals, including those with different reactivity domains and physiochemical properties. As stated previously, the assay provides advantages over other alternative methods in that it applies to a broader range of lipophilic chemicals and more pre/pro-haptens. The performance of the EpiSensA using 69 lipophilic compounds (LogKow >3.5) was better than 3 other *in vitro* assays (DPRA, KertinoSens, h-CLAT) with sensitivity, specificity, and accuracy values of 83%, 65%, and 78%, respectively. The predictivity of 37 pre/pro-haptens was even better, with a sensitivity of 95%.

**Evaluation Criterion 6: Predictive capacity should be demonstrated using representative chemicals.**

Predictive capacity was demonstrated with representative chemicals across a range of chemical classes, including lipophilic and pre-/pro-haptens, and was comparable to other chemical sets utilized in other validation studies. Regarding the actual results for the assay, 136 chemicals were tested by the lead laboratory using EpiSensA. As stated in the validation report, the sensitivity, specificity, and accuracy were 88.1%, 65.7%, and 82.4%, respectively, compared to the LLNA. These were comparable to other existing methods. Additionally, with the ability to test pre/pro-haptens, the EpiSensA had a 97.3% sensitivity for predicting these chemicals. The cumulative sensitivity, specificity, and accuracy for the three participating labs were 92.6%, 63%, and 82.7%, respectively, compared to the LLNA. These are similar to the larger dataset assessed by the lead lab. Against human data, the cumulative sensitivity, specificity, and accuracy were 92.2%, 61.9%, and 83.3%.

EpiSensA performed adequately when used to predict potency (GHS 1A/1B). While this was not a primary endpoint in the validation study, the assay demonstrated a sensitivity of 92.6%, a specificity of 63.0%, and an accuracy of 82.7% compared to LLNA results for the three participating laboratories. This is comparable to the predictive performance found following the lead laboratory's assessment of 136 chemicals (88.1%, 65.7%, and 82.4% for sensitivity, specificity, and accuracy). However, when the accuracy is broken down a bit further, there is some concern that the model may be missing 1A chemicals, with 29% of chemicals defined as 1B and 6% as NC from the set of 1A chemicals. Approximately 15% of the 1B chemicals were under classified as NC in the assay as well. While overall accuracy compared with human data, i.e., GHS category, is not high (64.4%), this is the same accuracy for LLNA as compared with human data. However, there is some concern for the chemicals underpredicted when compared with 1A/1B/NC class by LLNA. Overall underprediction values were reported at 16.1%. Compared to the kDPRA, the EpiSensA results seem comparable, with 31% of Cat. 1A chemicals underpredicted compared to 36% for the kDPRA (Natsch et al. 2020). A comparison of the predictive capacity of the LLNA GHS classification of 1A chemicals between EpiSensA and kDPRA for the same set of chemicals (72) demonstrated similar accuracy values of 88.9% and 91.7%, respectively. A comparison of the predictive capacity of the strong human sensitizers between the EpiSensA and kDPRA for the same set of chemicals (57) also showed similar accuracy values of 82.5% and 77.2%, respectively.

EpiSensA has enough predictive capacity in numerical terms compared with other methods such as LLNA, human tests, *in vitro* and *in vivo* studies. But there is a slight concern about the predictability of 1A and 1B classification. The aim of this validation is not the evaluation of the potency predictability of this method; therefore, discussion of applying EpiSensA to *in vitro* potency prediction methods of skin sensitizers should be considered as a next step.

**Evaluation Criterion 7: All data from the validation study supporting the validity of a test method should be obtained in accordance with the principles of Good Laboratory Practice (GLP).**

Studies were carried out in non-GLP laboratories in accordance with the principles of GLP, with quality controls in place. The lead lab created data reporting templates and record sheets and distributed them to the participating labs. The record sheets included important information such as start/end time of operation, lot number, and manufacturer of reagents; amount of reagents used; plate design and measurement conditions; and operator/technician signature and performance dates. All the records (data sheets and record sheets) from the participating laboratories were carefully checked by QC experts at the lead lab who are proficient in the principles of GLP. JaCVAM also reviewed the records, which are available on the JaCVAM website for review.

**Evaluation Criterion 8: The applicability domain of the test method should be defined.**

The chemical applicability domain is similar in makeup to the chemical set utilized in the Defined Approach for Skin Sensitization (TG497) for both reactivity domains and physicochemical properties.

The EpiSensA dataset contains many pre/pro-haptens and lipophilic chemicals, which can be considered more challenging for skin sensitization NAMs. The EpiSensA showed good predictive performance, with only a few limitations described regarding applicability domains. For example, if the chemical is not soluble or does not form a stable dispersion at 0.0122% in any specified vehicle, the chemical is not applicable for testing with EpiSensA. However, another vehicle outside of the protocol specified vehicles could be employed if sufficient scientific rationale can be provided. In addition, testing chemicals at high concentrations may lead to high osmotic stress conditions and lead to a false positive response.

The misclassification of lauryl gallate, a known pre/pro-hapten, suggests that chemical exposure time may not be optimal for some chemicals requiring time for autooxidation or metabolism. Therefore, if an unknown chemical is predicted *in silico* to be metabolized to a potential skin sensitizer and negative EpiSensA results are obtained, it may be necessary to increase incubation time for these challenging materials. The Lead Lab shared that a positive response was observed with lauryl gallate when 6 hr treated cultures were allowed to incubate for a total of 24 hours. However, it is important to point out again that EpiSensA correctly identified 35 out of 37 (95%) pre/pro-haptens.

Chemicals that affect the expression of GAPDH or interfere with viability assays, LDH, and MTT will be problematic for the assay. Limitation of the chemicals that may influence the RNA extraction and qPCR performance might be a future concern to spreading this method

widely. Like most available alternative assays, the EpiSensA has an inherent limitation in detecting very weak sensitizers. On the other hand, analysis of the DASS human set showed that most false predicted chemicals might have weak sensitization potential.

**Evaluation Criterion 9: Proficiency chemicals should be set up in the proposed protocol.**

The Proficiency chemicals have been shared by the test developers and were adapted from chemicals used in the JaCVAM coordinated validation study to demonstrate the successful implementation of the protocol. The chemicals contained in the list are commercially available and provide a range of responses. Many chemicals are similar to the proficiency chemicals used for other skin sensitization NAMs.

The chemicals are intended for method optimization in a lab performing the assay using the EpiSensA LabCyte RhE model described in the most current version of the SOP, Version: 2.6. Successful implementation of the assay in a new laboratory is demonstrated by correctly identifying 8 or more of the 10 chemicals. This success criterion is identical to that set for OECD accepted skin sensitization NAMs.

**Evaluation Criterion 10: Performance standards should be set up with the proposed protocol.**

Performance standards that refer to the acceptability of an individual test run or set of runs, also referred to as acceptance criteria, were available for the test method. Clotrimazole and 4-nitrobenzyl bromide are used as positive controls. Clotrimazole dissolved to a final concentration of 0.78 w/v% in AOO, and 4-nitrobenzyl bromide (4NBB) dissolved to a final concentration of 0.10 w/v%. The positive controls are unique to this assay compared to other NAMs currently being used. There is no negative control chemical tested in each run. The vehicle control is used as the negative control.

The SOP (Appendix 7) does provide requirements for qualified testing and data acceptance that are inclusive of cell viability limits, concordance in controls, minimums for positive control gene expression, and requirements for endogenous control gene variability. The mean cell viability of both positive controls should be equal to or greater than 80%. For clotrimazole, the mean values of fold induction for ATF3 and IL-8 should exceed the cut-off values. The mean values of fold induction for GCLM and DNAJB4 should exceed the cut-off values for 4-Nitrobenzyl bromide.

**Evaluation Criterion 11: Advantages in terms of time, cost, and animal welfare should be described.**

The animal welfare advantages are straightforward and similar to other NAMs that have been developed for evaluating the skin sensitization potential of chemicals. The benefits in terms of cost and time are not addressed here since a comparison would need to be made by contract laboratories or individual laboratories that include the cost of personnel time for conducting the assay, not just materials. The expectation would be that the cost and time would be reasonable and comparable to other NAMs, and the cost being easily justified based on the advantages the EpiSensA provides. It is important to note that the test developers have begun to assess cost comparisons of EpiSensA compared to both *in vivo and in vitro* test methods.

**Evaluation Criterion 12: Completeness of all data and documents supporting the assessment of the validity of the test method.**

This validation study aimed to assess the transferability, the within-laboratory reproducibility, and the between-laboratory reproducibility of EpiSensA. The study was coordinated by the JaCVAM in consultation with the VMT regarding the study design, the test chemical selection, and the SOP. The study was undertaken using coded test chemicals, enabling the VMT to assess the scientific validity and appropriateness of EpiSensA in anticipation of the development and issue of an OECD test method guideline for its use.

Raw data sheets, GLP notations sheets, and other supporting documents for the validation of EpiSensA were made available as part of the validation report. They did appear to be complete and in support of the validity of the test method.

**Evaluation Criterion 13: Validation Study should be managed and conducted adequately.**

The primary goals of this validation study were to assess the test definition, transferability, and within- and between-laboratory reproducibility while additionally using the data produced to evaluate predictive capacity and the applicability domain compared to the LLNA. The management and conduct of the validation study were adequate.

## Recommendations and Further Considerations

EpiSensA Considerations for Longer Incubations - The EpiSensA was demonstrated by test developers to show remarkable predictivity of skin sensitization with respect to pre-/pro-haptens, with 35/37 chemicals accurately predicted. The two missed chemicals were benzo[a]pyrene and lauryl gallate. The latter tested negative in the validation study. Further assessment by the test developers regarding the incorrect prediction of lauryl gallate showed that lower concentrations, as tested by participating laboratories, as well as longer incubation times, resulted in positive predictions for the chemical. Several factors appear to affect the

ability of EpiSensA to predict lauryl gallate, inclusive of the lipophilicity of the compound (high Log $K_{OW}$) and the rapidity with which the chemical may undergo metabolism or auto-oxidation. With lauryl gallate, the test developers showed that a 6-hour exposure, followed by an additional 18-hour post-wash incubation period, resulted in a strong response and a positive skin sensitization prediction in the EpiSensA. With these factors in mind, it is recommended that the test developers address the potential for additional follow-on testing options to assess chemicals that are predicted to be pre-/pro-haptens and require longer incubation times to activate and/or also have physical-chemical properties that could indicate longer incubation times are necessary, such as a high log $K_{OW}$, which may affect absorption into the metabolically active layers of the model.

Development of 'Me Too' Methods Based on EpiSensA - Prior to the routine use of the EpiSensA test method, laboratories should demonstrate technical proficiency, using the ten Proficiency Chemicals suggested by the test developer. These chemicals can be used by laboratories to demonstrate technical competence with a standardized test method or to demonstrate the reproducible performance of the test method over time. Proficiency chemicals represent a subset of the reference chemicals included in Performance Standards as defined by the OECD, which include essential test method components, a list of reference substances and accuracy and reliability performance values, which are distinct from the performance standards referred to in Evaluation Criterion 10 above. Test developers should determine an expanded set of reference chemicals to allow the method to be transferred to another model of reconstituted epidermis (colloquially referred to as "me-too" test methods). This is important for a greater distribution of the method worldwide, where accessibility of the different epidermal models is varied. Performance standards must be available to facilitate the validation of new or modified EpiSensA test methods similar to the original EpiSensA, which will allow for timely amendment of the Test Guideline for their inclusion, which is crucial for the Mutual Acceptance of Data.

Performance standards are typically based on a scientifically valid and accepted test method and can be used to evaluate the reliability and relevance of other analogous test methods that are based on similar scientific principles and measure or predict the same biological or toxic effect. Based on previous experience (moving from EpiDerm to LabCyte EPI-MODEL 24), the test developer is encouraged to provide the performance standards, together with the criteria for a limited assessment of reliability, sensitivity, specificity, and relevance without needing to undergo a formal validation. A similar approach used by the KeratinoSens, DPRA, and LLNA 'me too' methods should be adopted, where 20 Reference Substances have been identified, along with other factors to determine the performance of the proposed new method.

Borderline Responses – The test developer and participating laboratories demonstrated that the EpiSensA test method has enough sensitivity, specificity, and accuracy to predict typical strong skin sensitizers compared to the LLNA and other *in vitro* methods (e.g., EpiSensA and kDPRA for the same set of chemicals of GHS1A demonstrated similar accuracy values of 88.9% and 91.7%, respectively). On the other hand, there is a slight concern about the ability of the test to discriminate between 1A and 1B, and 1B and NC. Approximately 15% of the 1B chemicals were underclassified as NC with the EpiSensA test method. While the accuracy

of classification of GHS 1A/1B/NC class by EpiSensA is comparable with the kDPRA, there are concerns for the chemicals underpredicted when compared with 1A/1B/NC class by LLNA. The test developers did provide some additional analysis of setting borderline thresholds of the cut-off value ±10%, which resulted in 3 false negative chemicals (3%, 3/101) in the original dataset being classified as borderline under these criteria. However, the developers do state that the EpiSensA has fewer borderline results than the TG497 methods KeratinoSens and h-CLAT (5.9%; 8/135 and 6.7%, 9/134 respectively), demonstrating that the method already has adequate retesting criteria to prevent false negative judgement. Regardless, while discerning borderline responses is not the aim of the validation, discussion of how borderline responses should be classified/handled with EpiSensA should be undertaken by the test developers if inclusion into a defined approach such as TG497 is to be consideration.

## Conclusion

The peer review panel (PRP) for the EpiSensA validation study has assessed the method for appropriate rationale, mechanistic applicability, scientific need, test method description, transferability, within- and between-laboratory reproducibility, predictive capacity based on a relevant set of reference chemicals, the defined domain of applicability, and overall conduct of the study (GLP, completeness of protocol, etc). The PRP has found that the EpiSensA test method is robust, meets specified criteria for validity, and additionally fills in testing gaps with the ability to correctly assess pre-/pro-haptens and lipophilic chemicals. The PRP has made some minor recommendations related to development of me-too test method performance standards, further assessment of the feasibility of using the test method for potency predictions, how to address borderline responses, and potential follow-on confirmatory testing when predicted pre-/pro-haptens or highly lipophilic compounds test as negative (e.g., longer incubation time). The recommendations are meant to help strengthen the assay but are not limiting factors in the use and application of the assay as it is currently designed.

Overall, the assay performed well against pre-specified assessment criteria, easily exceeding the 85% and 80% thresholds for within- and between-laboratory reproducibility (>86.7% and 88.9% respectively). Transferability between labs was also successful, with a well-defined SOP for conduct of the assay utilizing the LabCyte EPI-MODEL 24 skin model. The assay's sensitivity, specificity, and accuracy for a hazard classification were comparable to other existing skin sensitization assays, inclusive of other OECD guideline-based in vitro test methods. The assay shows promise to be applied for potency prediction, but further assessment is necessary. Regarding the applicability domain of the assay, its ability to better predict lipophilic compounds and pre-/pro-haptens is a distinct advantage over other existing in vitro assays.

Overall, the PRP agrees that this model has performed well in the validation study and can be applied to predict skin sensitization hazard, and potentially potency, with the intent to eventually apply it into a defined approach for skin sensitization.

## Acknowledgments

The PRP is grateful to the members of the VMT for their hard work and excellent responsiveness to our questions. Thanks, are also due to JaCVAM for their support in setting up and hosting our meeting in Japan, as well as for setting up each of our telephone conferences. This work was supported by the Japanese Society for Alternatives to Animal Experiments (JSAAE). Finally, a special thanks to Tamaki Yoshikawa for her outstanding administrative support in helping us with our visas during these challenging times.

## References

Api AM, Basketter D, Bridges J, Cadby P, Ellis G, Gilmour N, Greim H, Griem P, Kern P, Khaiat A et al. 2020. Updating exposure assessment for skin sensitization quantitative risk assessment for fragrance materials. Regul Toxicol Pharmacol. 118:104805.

Bernauer U, Bodin L, Chaudhry Q, Coenraads PJ, Dusinska M, Ezendam J, Gaffet E, Galli CL, Granum B, Panteri E et al. 2021. The sccs notes of guidance for the testing of cosmetic ingredients and their safety evaluation, 11th revision, 30–31 march 2021, sccs/1628/21. Regulatory Toxicology and Pharmacology. 127:105052.

Dent M, Amaral RT, Da Silva PA, Ansell J, Boisleve F, Hatao M, Hirose A, Kasai Y, Kern P, Kreiling R et al. 2018. Principles underpinning the use of new methodologies in the risk assessment of cosmetic ingredients. Computational Toxicology. 7:20-26.

Gilmour N, Kern PS, Alepee N, Boisleve F, Bury D, Clouet E, Hirota M, Hoffmann S, Kuhnl J, Lalko JF et al. 2020. Development of a next generation risk assessment framework for the evaluation of skin sensitisation of cosmetic ingredients. Regul Toxicol Pharmacol. 116:104721.

Mizumachi H, LeBaron MJ, Settivari RS, Miyazawa M, Marty MS, Sakaguchi H. 2021. Characterization of dermal sensitization potential for industrial or agricultural chemicals with episensa. J Appl Toxicol. 41(6):915-927.

Mizumachi H, Sakuma M, Ikezumi M, Saito K, Takeyoshi M, Imai N, Okutomi H, Umetsu A, Motohashi H, Watanabe M et al. 2018. Transferability and within- and between-laboratory reproducibilities of episensa for predicting skin sensitization potential in vitro: A ring study in three laboratories. J Appl Toxicol. 38(9):1233-1243.

Natsch A, Haupt T, Wareing B, Landsiedel R, Kolle SN. 2020. Predictivity of the kinetic direct peptide reactivity assay (kdpra) for sensitizer potency assessment and ghs subclassification. ALTEX. 37(4):652-664.

OECD. 2014. The adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins.

OECD. 2018a. Test no. 442d: In vitro skin sensitisation.

OECD. 2018b. Test no. 442e: In vitro skin sensitisation.

OECD Guideline no. 497: Defined approaches on skin sensitisation. 2021a. [accessed]. https://www.oecd-ilibrary.org/content/publication/b92879a4-en.

OECD. 2021b. Test no. 442c: In chemico skin sensitisation.

Saito K, Nukada Y, Takenouchi O, Miyazawa M, Sakaguchi H, Nishiyama N. 2013. Development of a new in vitro skin sensitization assay (epidermal sensitization assay; episensa) using reconstructed human epidermis. Toxicol In Vitro. 27(8):2213-2224.

Saito K, Takenouchi O, Nukada Y, Miyazawa M, Sakaguchi H. 2017. An in vitro skin sensitization assay termed episensa for broad sets of chemicals including lipophilic chemicals and pre/pro-haptens. Toxicol In Vitro. 40:11-25.

Sullivan KM, Enoch SJ, Ezendam J, Sewald K, Roggen EL, Cochrane S. 2017. An adverse outcome pathway for sensitization of the respiratory tract by low-molecular-weight chemicals: Building evidence to support the utility ofin vitroandin silicomethods in a regulatory context. Applied In Vitro Toxicology. 3(3):213-226.