厚生労働行政推進調査事業費補助金（化学物質リスク研究事業）

OECD プログラムにおいて TG と DA を開発するための AOP に関する研究

平成30年度 分担研究報告書

AOP、TG、DAの開発、AOP国内マニュアルの作成

研究分担者　小島　肇

国立医薬品食品衛生研究所　安全性生物試験研究センター　安全性予測評価部　室長

研究要旨

動物実験 3Rs の国際的な浸透に加えて、実験動物とヒトとの種差等の克服のために、既存の毒性試験法の見直しが進んでいる。経済協力開発機構（OECD: Organisation for Economic Co-operation and Development）においても、反復投与毒性、生殖発生毒性、感作性、発がん性などの有害性発現経路（AOP: Adverse Outcome Pathway）を開発し、動物実験代替法（以下、代替法）を念頭においた試験法ガイドライン（TG: Test Guideline）の公定化や *in silico* 法の確立に AOP 情報を活用する戦略がある。一方で、毒性情報を網羅した IATA (Integrated Approaches to Testing and Assessment）を開発し、それに基づく DA (Defined Approach）により化学物質の安全性評価を推進する戦略がある。DA とは、単独の代替法ではなく、種々の試験データを組み合わせて化学物質の全身毒性を把握しようとする試みであり、OECD では DA の行政的利用が検討されている。このような国際的な潮流に乗り、日本が得意とする分野で主導権を握って、AOP や TG を公定化し、さらには IATA や DA の開発及び普及に協力することが本研究班の目的である。

昨年度からの継続した活動の中、本年度に OECD の TG や AOP が採択されたものはなかった。ただし、来年度に 3 試験法の TG、1 件の AOP を成立できる目途がたった。

研究分担者氏名・所属研究機関名及び所属研究機関における職名

相場節也	東北大学医学系研究科・医学部・皮膚科学分野教授
足利太可雄	国立医薬品食品衛生研究所 安全性生物試験研究センター 安全性予測評価部 主任研究官
大石　巧	日本免疫毒性学会試験法委員会 AOP 検討小委員会（株式会社ボゾリサーチセンター)
尾上誠良	静岡県立大学　薬学部・薬剤学分野　教授
笠原利彦	富士フイルム株式会社　安全性評価センター技術マネージャー

加藤雅一　株式会社ジャパン・ティッシュ・エンジニアリング（J-TEC）主任研究員
木村　裕　東北大学医学系研究科・医学部・皮膚科学分野准教授
久田　茂　日本免疫毒性学会試験法委員会（あすか製薬株式会社）

A. 研究目的

本研究班では、OECD の AOP 開発プロジェクトの中で、化学物質の毒性情報等を集積しながら、免疫毒性、生殖発生毒性、発がん性及び光安全性等に関する日本発の AOP 開発を進める。既存の AOP 情報をもとに開発された皮膚感作性試験代替法 ADRA （Amino acid Derivative Reactivity Assay）、免疫毒性試験 MITA（Multi-Immuno Toxicity Assay)、発生毒性試験スクリーニング Hand1-Luc EST（Embryonic Stem cell Test)、光安全性試験スクリーニング ROS（Reactive Oxygen Species）アッセイ、LabCyte EPI-MODEL24 を用いる腐食性試験代替法については、試験法毎に独立した国内外の専門家による第三者評価（peer review）を受けた後、TG を開発する。一方で、皮膚感作性 DA の開発に関与することを通じて、IATA や DA の国内での普及に務める。

B. 研究方法

B.1. AOP、TG、DA の開発、AOP 国内マニュアルの作成

OECDのAOP開発プロジェクト EAGMST（Extended Advisory Group on Molecular Screening and Toxicogenomics）及び、TGの開発プロジェクトWNT（Working Group of the National Coordinators of the Test Guidelines Programme）の進捗に合わせ、班員を支援した。

B.1.1. AOP開発

AOP に関しては、日本免疫毒性学会会員をメンバーとする同学会試験法委員会 AOP 検討小委員会に免疫毒性 AOP の開発を委託している。

文献調査の結果に基づいて、カルシニューリン阻害を Molecular initiating event（MIE）とし、T 細胞依存性抗体産生抑制（TDAR）を Adverse outcome（AO）とする AOP154 案 “Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response” を作成した。

B.1.2. TG 開発

日本から提案している試験法である皮膚感作性試験代替法 ADRA、光安全性試験 ROS アッセイ（尾上分担研究者との協同研究）、LabCyte EPI-MODEL24 を用いる腐食性試験代替法の TG 採択のために、電話会議や専門家会議にて交渉した。

分担研究者の相場が開発し、他の研究班でバリデーションを終了させた IL-2 を指標とした免疫毒性試験の TG を目指し、海外の専門家を招聘したバリデーション報告書の peer review 会議を企画した。

また、*in vitro* 免疫毒性試験に関する Detailed Review Paper(DRP)の SPSF を作成し、OECD に提案した。

B.1.3. DA の開発協力

足利分担研究者とともに、OECD における皮膚感作性試験の DA の開発に協力した。

C. 研究結果

C.1. AOP、TG、DA の開発、AOP 国内マニュアルの作成

C.1.1. AOP開発

AOP154 案 “Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response”に関しては、EAGMST の内部 peer reviewer のコメントに対応して、主に key event relationship (KER)における定量的な理解（quantitative understanding）を中心に修正・追記した。6 月 28 日の EAGMST 会議において更新した AOP154 が内部 peer review を通過し、外部 peer review に進むこととなった。現在、外部 peer review に向けた内部 peer reviewer への指摘に対応している。

C.1.2. TG の開発

1)皮膚感作性試験

皮膚感作性試験代替法 ADRA の TG 開発を、開発者の富士フイルムや国内外の専門家の協力を受け進めた。2 度に渡り、WNT から TG への意見を受け、修正して対応した。結果として、2019 年 4 月の OECD WNT 会議で TG 案（添付資料 1）の採択が内定した。

2)光安全性試験

ROS アッセイの TG 開発を国内外の専門家と密な連携をとり進めた。WNT から TG への意見を受け、修正して対応した。

結果として、2019 年 4 月の OECD WNT 会議で TG 案（添付資料 2）の採択が内定した。

3)腐食性試験

LabCyte EPI-MODEL24 を用いる腐食性試験代替法の TG 開発を、開発者の株式会社 J-TEC と密な連携をとりつつ進めた。WNT から TG への意見を受け、修正して対応した。結果として、2019 年 4 月の OECD WNT 会議で TG 案（添付資料 3）の採択が内定した。

4)免疫毒性試験

海外の専門家を招聘し、IL-2 Luc アッセイバリデーション報告書の peer review 会議を 2019 年 2 月 27 日から 28 日まで、東京にて開催した。外部評価委員として Henk van Loveren (Maastricht University, Netherland), Haley LaNef Ford (Seattle Genetics, Inc., USA), Barbara Kaplan (Mississippi State University, USA), Sang-Hyun Kim (Kyungpook National University, Korea), Fujio Kayama (Jichi Medical University, Japan), Takao Ashikaga (National Institute of Health Sciences, Japan)を招請した。また, Xingchao Geng (National Center for Safety Evaluation of Drugs (NCSED), China)は電話会議にて参加した。

Peer review会議において，バリデーション報告書に対し、別紙 1 に示す提案が届き、次年度に対応予定である。

また、今後、相場らの開発した *in vitro* 免疫毒性試験を円滑に TG に導くための準備として、本件に関する昨今の状況をまとめた DRP の SPSF を OECD に提案し（添付資料４）、2019 年 4 月の OECD WNT 会議で作業計画として採択された。

C.1.3. DA の開発協力

OECD 専門家会議（電話会議や対面会議）で皮膚感作性 DA の開発に寄与した。

現在、ヒト及び動物実験結果の再評価、適用限界の明確化、不確定要素の解決法について、それぞれサブワーキンググループが討議しており、その提案を待って DA の

最終的な議論が来年度になされる予定である。

D. 考察

D.1 . AOP の開発

本研究班から提案している免疫抑制の AOP154 案 ”Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response”は EAGMST における内部 peer review が終了し、外部 peer review が始まる。外部 peer review 終了まであと一息となった。来年度には必ず成立させたい。

D.2 . TG の開発

皮膚感作性試験代替法 ADRA、光安全性 ROS アッセイ及び LabCyte EPI-MODEL24 を用いる腐食性試験代替法に関しては、OECD　WNT にて、2019 年 4 月に TG として採択されることが内定した。来年度の正式採択に向け、OECD と調整していきたい。

一方、新たに *in vitro* 免疫毒性 DRP の開発のための SPSF を OECD に昨年 11 月に提出し、2019 年 4 月に OECD 作業計画に加えられることになった。この DRP の開発を基に、TG の開発を目指すために OECD へ働きかけを続けていきたい。

D.3 . IATA 及び DA の成立

感作性 DA に関する OECD 活動に対し、引き続き協力していく予定である。来年度は光安全性 IATA の開発を OECD に提案する予定であり、動物実験を用いない安全性評価の体系化を日本からも提案していく予定である。

E. 結論

昨年度からの OECD との継続した活動の中、本年度に TG や AOP が採択されたものはなかった。ただし、来年度に 3 試験法の TG、１件の AOP を成立させることができる目途がたった。

引き続き、OECD の活動の中で、日本が得意とする分野で主導権を握って、AOP や TG を公定化し、さらには IATA や DA の開発に協力していく予定である。

F. 研究発表

F.1. 論文発表

1. 小島　肇: 皮膚・粘膜毒性, トキシコロジー第３版, 朝倉書店, 279-286.
2. 小島　肇: 動物実験代替法, トキシコロジー第３版, 朝倉書店, 320-325.
3. 小島　肇: *in vitro* 実験の重要性と培養細胞の選択方法, 創薬のための細胞利用技術の最新動向と市場,（株）シーエムシー・リサーチ, 3-6.
4. Fujita M, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara Y, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H, Kasahara T： Cause of and countermeasures for oxidation of the cysteine-derived reagent used in the amino acid derivative reactivity assay, J Appl Toxicol. 2019 Feb;39(2):191-208. doi: 10.1002/jat.3707.
5. Mitachi T, Kouzui M, Maruyama R, Yamashita K, Ogata S, Kojima H, Itagaki H: Some non-sensitizers upregulate CD54 expression by activation of the NLRP3 inflammasome in THP-1 cells. J Toxicol Sci. 2019;44(3):213-224. doi: 10.2131/jts.44.213.
6. 小島　肇: 化学物質の毒性評価方法の現状と今後, 化学物質と環境, エコケミストリー研究会, 2019;154, 1-3.

7. 中村和昭, 諫田泰成, 山崎大樹, 片岡 健, 青井貴之, 中川誠人, 藤井万紀子, 阿久津英憲, 末盛博文, 浅香 勲, 中村幸夫, 小島 肇, 伊藤弓弦, 関野祐子, 古江－楠田美保：「培養細胞の観察の基本原則」の提案, 組織培養研究, 2018; 37(2), 123-131.
8. 小島　肇：化学物質や医薬品などの安全性評価に用いる動物実験代替法の技術開発の現状と展望, イルシー Japan, 2018;136, 23-31.
9. Kimura Y, Watanabe M, Suzuki N, Iwaki T, Yamakage K, Saito K, Nakajima Y, Fujimura C, Ohmiya Y, Omori T, Kojima H, Aiba S: The performance of an in vitro skin sensitisation test, IL-8 Luc assay (OECD442E), and the integrated approach with direct peptide reactive assay (DPRA), J Toxicol Sci. 2018; 43(12):741-749. doi: 10.2131/jts.43.741.

F.2. 学会発表

1. New trend on alternative to animal testing in Japan, Kojima H, OpenTox 2018, 2018/5/24, 国内.
2. Alternative Test Methods Developed in Japan and South Korea for Regulatory Use, Kojima H, 8th Conference of Alternative Methods, 2018/6/12, 国外.
3. In vitroモデルの創薬開発への活用, 小島肇, 日本組織培養学会第91回大会, 2018/6/16,国内.
4. Recent Activities for safety assessment, Kojima H, International Symposium on Cosmetic Regulation, 2018/7/13, 国内.
5. 教育講演：ガイドライン化を目指したin vitro試験系導入の具体的な留意点, 小島肇, 第45回日本毒性学会学術年会, 2018/7/20, 国内.
6. 皮膚・粘膜毒性、光毒性、代替試験法, 小島 肇, 第21回日本毒性学会基礎教育講習会, 2018/8/7, 国内.
7. New Approach on Alternative to Animal Test Methods in JaCVAM and Japanese projects, , Kojima H, The 15th Annual meeting of Korean Society of Alternative to Animal Experiments, 2018/8/24, 国外.
8. Introduction and research status of AAT in JaCVAM, Kojima H, The 2nd International Conference on Cosmetics Alternative Methods in NIFDC, 2018/9/20, 国外.
9. New methods validation and AAT regulatory acceptance in Japan, Kojima H, The 2nd International Conference on Cosmetics Alternative Methods in NIFDC, 2018/9/21, 国外.
10. Japanese Strategy on Alternative to Animal Test Methods for Systemic Toxicology, Kojima H, 20th International Congress on In Vitro Toxicology, 2018/10/16, 国外.
11. h-CLAT における NLRP3 インフラマソームの影響, 丸山 諒, 洪水麻衣, 三田地隆史, 小島 肇, 板垣 宏, 日本動物実験代替法学会第 31 回大会, 2018/11/24, 国内.
12. タンパク質のアレルギー性を評価する in vitro 試験法の開発 試薬中 LPS の影響除外に関する検討（第 1 報）, 小林（九十九）英恵, 生地加奈実, 山下邦彦, 小島肇, 板垣 宏, 日本動物実験代替法学会第 31 回大会, 2018/11/24, 国内.
13. タンパク質のアレルギー性を評価する in vitro 試験法の開発 薬中 LPS の影響除外に関する検討（第 2 報）, 生地加奈実, 小林英恵, 山下邦彦, 小島 肇, 板垣宏, 日本動物実験代替法学会第 31 回大会, 2018/11/24, 国内.

14. Multi-Immuno Tox Assay（MITA）：バリデーション研究の結果, 木村 裕, 安野理恵, 渡辺美香, 小林美和子, 岩城知子, 藤村千鶴, 近江谷克裕, 山影康次, 中島芳浩, 小林眞弓, 大森 崇, 足利太可雄, 小島 肇, 相場節也, 日本動物実験代替法学会第 31 回大会, 2018/11/24, 国内.
15. OECD AOPプロジェクト, 小島 肇, 第一回医薬品毒性機序研究会, 2019/1/10, 国内.
16. 毒性評価系の国際標準化に向けた戦略, 小島 肇, 毒性評価研究会, 2019/1/31, 国内.
17. ヒト健康影響を予測するための非動物実験の開発動向, 小島 肇, Translational and Regulatory Science Symposium, 2019/2/7, 国内.
18. 国際情勢から見た幹細胞ベースの毒性試験について, 小島 肇, 幹細胞を用いた化学物質リスク情報共有化コンソーシーアム scChemRISK, 2019/2/15, 国内.
19. AOP（Adverse Outcome Pathway; 有害性発現経路）に基づいた化学物質の安全性評価へ向けたチャレンジ, 山田隆志, 足利太可雄, 小島 肇, 広瀬明彦, 日本薬学会第 139 年会, 2019/3/23, 国内

G. 知的所有権の取得状況

G.1 特許取得

特になし

G.2. 実用新案登録

特になし

G.3 その他

特になし

H. 添付資料

1. OECD Draft Updated Test Guideline 442C for in chemico skin sensitisation assays addressing the Adverse Outcome Pathway Key Event on covalent binding to proteins
2. OECD GUIDELINE FOR THE TESTING OF CHEMICALS: ROS (REACTIVE OXYGEN SPECIES) ASSAY FOR PHOTOTOXICITY
3. OECD DRAFT UPDATED TEST GUIDELINE 431 ON INVITRO SKIN CORROSION, RECONSTRUCTED HUMAN EPIDERMIS TEST METHODS

別紙 1　Action Items to peer reviewers for the validation report on the IL-2 Luc assay

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.

PRP Comment: Together with a new title, the rationale needs to be stated clearly to be T-cell targeting.

Evaluation Criterion 2: 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.

PRP Comment: Needs to focus on IL-2, including the limitations described in the meeting minutes. The introduction needs to focus solely on IL-2 and the IL-2 Luc Assay. Discussion about its part in MITA should be left until the discussion section.

Evaluation Criterion 3: A detailed test method protocol should be available

PRP Comment: The commercial availability of the #2H4 cell line needs to be described.

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

PRP Comment: Acceptable

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

PRP Comment: We think only four or five negatives is not enough, so we suggest that some additional testing of negatives be performed.

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

PRP Comment: Predictive capacity needs to be reassessed based on today's proposed definition of T-cell–targeting chemicals.

Evaluation Criterion 7: All data should adequately support the assessment of the validity of the test method for peer review.

PRP Comment: A clear definition of the 35% threshold and a clear explanation of Criteria 5 and how it was developed is needed. Should the table in Appendix 8 include the test judgment? Also, delete DTH, tumor, infection, and NK activity but specify T-cell proliferation in the table in Appendix 8.

Evaluation Criterion 8: 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)

PRP Comment: The report needs to explain clearly and in detail what is meant by the phrase "in the spirit of GLP" and whether or not each laboratory performed their work in this spirit.

Evaluation Criterion 9: Applicability domain of the test method should be defined

PRP Comment: We recommend that the applicability domain be more clearly defined as noted in the PRP meeting minutes.

Evaluation Criterion 10: Proficiency chemicals should be set up in the proposed protocol

PRP Comment: None

Evaluation Criterion 11: Performance standards should be set up with the proposed protocol

PRP Comment: If performance standards are understood to be assay controls, then the use of three-fold stimulation of IL-2 Luc by PMA/IO and inhibition of stimulated IL-2 Luc by DEX and CYA are sufficient. We suggest that acceptance criteria for variability within test replicates be defined.

Evaluation Criterion 12: Advantages in terms of time, cost and animal welfare

PRP Comment: We suggest that the conclusion leave out mention of in vivo testing to confirm T-cell immunotoxicity and include discussion of the use of IL-2 Luc assay within MITA.

Evaluation Criterion 13: Completeness of all data and documents supporting the assessment of the validity of the test method.

PRP Comment: We suggest that data be redone to reassess predictive capacity based on today's proposed definition of T-cell–targeting chemicals. Also, a critical assessment of the 35% threshold in the context of the new definition of T-cell targeting is necessary.

Evaluation Criterion 14: Validation Study Management and Conduct

PRP Comment: None

Other considerations

PRP Comment: None

Conclusion

PRP Comment: We look forward to seeing a revised report based on our comments.

添付資料１

Section 4
Health effects

Test Guideline No. 442C

In Chemico Skin Sensitisation

Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins

18 June 2019

OECD Guidelines for the Testing of Chemicals

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Key–Event-Based Test Guideline For In Chemico Skin Sensitisation Assays Addressing The Adverse Outcome Pathway Key Event On Covalent Binding To Proteins

GENERAL INTRODUCTION

Covalent binding to proteins Key Event based Test Guideline.

1. A skin sensitiser refers to a substance that will lead to an allergic response following repeated skin contact as defined by the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS) (1). There is general agreement on the key biological events underlying skin sensitisation. The current knowledge of the chemical and biological mechanisms associated with skin sensitisation has been summarised as an Adverse Outcome Pathway (AOP) (2) starting with a molecular initiating event through intermediate events to the adverse effect, namely allergic contact dermatitis. This AOP focuses on chemicals that react with amino-acid residues (i.e. cysteine or lysine) such as organic chemicals. In this instance, the molecular initiating event (i.e. the first key event), is the covalent binding of electrophilic substances to nucleophilic centres in skin proteins. The second key event in this AOP takes place in the keratinocytes and includes inflammatory responses as well as changes in gene expression associated with specific cell signaling pathways such as the antioxidant/electrophile response element (ARE)-dependent pathways. The third key event is the activation of dendritic cells, typically assessed by expression of specific cell surface markers, chemokines and cytokines. The fourth key event is T-cell proliferation.

2. The assessment of skin sensitisation has typically involved the use of laboratory animals. The classical methods that use guinea-pigs, the Guinea Pig Maximisation Test (GPMT) of Magnusson and Kligman and the Buehler Test (OECD TG 406) (11) assess both the induction and elicitation phases of skin sensitisation. The murine tests, such as the LLNA (OECD TG 429) (12) and its three non-radioactive modifications — LLNA:DA (OECD TG 442A) (13), LLNA:BrdU-ELISA, and BrdU-FCM (OECD TG 442B) (14) — all assess the induction response exclusively and have gained acceptance, since they provide an advantage over the guinea pig tests in terms of animal welfare together with an objective measurement of the induction phase of skin sensitisation.

3. Mechanistically-based in chemico and in vitro test methods addressing the first three key events of the skin sensitisation AOP have been adopted for contributing to the evaluation of the skin sensitisation hazard potential of chemicals: the present Test Guideline assesses covalent binding to proteins, addressing the first key event; the OECD TG 442D assesses keratinocyte activation (15), the second key event and the OECD TG 442E addresses the activation of dendritic cells (16), the third key event of the skin sensitisation AOP. Finally, the fourth key event representing T-cell proliferation is indirectly assessed in the murine Local Lymph Node Assay (LLNA) (12).

Background and principles of the test methods included in the Key Event based Test Guideline

4. This Test Guideline (TG) describes in chemico assays that address mechanisms described under the first key event of the AOP for skin sensitisation, namely covalent binding to proteins (2). The Test Guideline comprises test methods to be used for supporting the discrimination between skin sensitisers and non-sensitisers in accordance with the UN GHS (1). The test methods currently described in this Test Guideline are:

- The Direct Peptide Reactivity Assay (DPRA) (Appendix I), and
- The Amino acid Derivative Reactivity Assay (ADRA) (Appendix II).

5. These two test methods are based on in chemico covalent binding to proteins and are considered to be scientifically valid. The DPRA has been evaluated in a European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)-lead validation study and subsequent independent peer review by the EURL ECVAM Scientific Advisory Committee (ESAC) (3) (4) (5). The ADRA underwent a validation study coordinated by the Japanese Center for the Validation of Alternative Methods (JaCVAM) (6) (7) (8) (9) followed by an independent peer-review (10).

6. The test methods included in this Test Guideline might differ with regard to the procedures used to generate the data but can each be used to address countries' requirements for test results on protein reactivity, while benefiting from the Mutual Acceptance of Data.

7. The correlation of protein reactivity with skin sensitisation potential is well established. (17) (18) (19). Nevertheless, since protein reactivity represents only one key event of the skin sensitisation AOP (2) (20), information generated with test methods developed to address this specific key event may not be sufficient as stand-alone methods to conclude on the presence or absence of skin sensitisation potential of chemicals. Therefore data generated with the test methods described in this Test Guideline are proposed to support the discrimination between skin sensitisers (i.e. UN GHS Category 1) and non-sensitisers when used within Integrated Approaches to Testing and Assessment (IATA), together with other relevant complementary information from in vitro assays addressing other key events of the skin sensitisation AOP as well as non-testing methods, including in silico modeling and read-across from chemical analogues (20). Examples on the use of data generated with these methods within Defined Approaches (DAs) i.e. approaches standardised both in relation to the set of information sources used and in the procedure applied to derive predictions—have been published (20) and can be employed as useful elements within IATA.

8. The test methods described in this Test Guideline do not allow either sub-categorisation of skin sensitisers into subcategories 1A and 1B (21), as defined by UN GHS (1) for authorities implementing these two optional subcategories, or potency prediction for safety assessment decisions. However, depending on the regulatory framework, positive results generated with these methods may be used on their own to classify a chemical into UN GHS Category 1.

9. Definitions are provided in the Annex.

Literature for introduction

(1) United Nations (UN) (2017), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Seventh revised edition, New York and Geneva, United Nations Publications. Available at: [https://www.unece.org/trans/danger/publi/ghs/ghs_rev07/07files_e0.html]

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

(3) GF Gerberick, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP (2004), Development of a peptide reactivity assay for screening contact allergens. Toxicol Sci. 81, 332-343.

(4) GF Gerberick, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP (2007), Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol Sci. 97, 417-427. .

(5) EC EURL-ECVAM (2013), Recommendation on the Direct Peptide Reactivity Assay (DPRA) for the skin sensitisation testing Available at: https://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendaion-on-the-directpeptide-reactivity-assay-dpra.

(6) M Fujita, Yamamoto Y, Tahara H, Kasahara T, Jimbo Y, Hioki T (2014), Development of a prediction method for skin sensitisation using novel cysteine and lysinederivatives. J PharmacolToxicol Methods. 70, 94-105.

(7) Y Yamamoto, Tahara H, Usami R, Kasahara T, Jimbo Y, Hioki T, Fujita M.(2015) A novel in chemico method to detect skin sensitisers in highly diluted reactionconditions. J Appl Toxicol. 35, 1348-1360.

(8) M Fujita, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara K, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H and Kasahara T (2019), Cause of and Countermeasures for Oxidation of the Cysteine-Derived Reagent Used in the Amino acid Derivative Reactivity Assay, J. Appl. Toxicology, Feb;39(2):191-208 (doi: 10.1002/jat.3707).

(9) OECD (2019), Draft validation report: Amino acid Derivative Reactivity Assay (ADRA) – JaCVAM Validation Study Report. Series on testing and Assessment n° xx. Organisation for Economic Cooperation and Development, Paris.

(10) OECD (2019), Amino acid Derivative Reactivity Assay (ADRA) – Report of the Peer Review Panel. Series on testing and Assessment n° xx. Organisation for Economic Cooperation and Development, Paris.

(11) OECD (1992), OECD Guidelines for the Testing of Chemicals No. 406. Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(12) OECD (2010), OECD Guidelines for Chemical Testing No. 429. Skin sensitisation: Local Lymph Node assay. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(13) OECD (2010), OECD Guidelines for Chemical Testing No. 442A.Skin sensitisation: Local Lymph Node assay: DA. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(14) OECD (2018), OECD Guidelines for Chemical Testing No. 442B. Skin sensitisation: Local Lymph Node assay: BrdU-ELISA or –FCM. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(15) OECD (2018), OECD Key Event based test Guideline 442D: In vitro Skin Sensitisation Assays Addressing AOP Key Event on Keratinocyte Activation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(16) OECD (2018), OECD Key event based test Guideline 442E: In Vitro Skin Sensitisation Assays Addressing the Key Event on Activation of Dendritic Cells on the Adverse Outcome Pathway for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(17) Landsteiner and Jacobs (1936), Studies on the sensitisation of animals with simple chemical compounds. Journal of Experimental Medicine 64:625-639.

(18) Dupuis and Benezra (1982), Allergic contact dermatitis to simple chemicals: a molecular approach. New York & Basel: Marcel Dekker Inc.

(19) JP Lepoittevin, Basketter DA, Goossens A, Karlberg AT (1998), Allergic contact dermatitis: the molecular basis, Springer, Berlin (doi: 10.1007/978-3-642-80331-4).

(20) OECD (2016), Series on Testing & Assessment No. 256: Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Annex 1 and Annex 2. ENV/JM/HA(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [https://community.oecd.org/community/iatass].

(21) B Wareing, Urbisch D, Kolle SN, Honarvar N, Sauer UG, Mehling A, Landsiedel R(2017) Prediction of skin sensitization potency sub-categories using peptide reactivity data, Toxicol In Vitro Dec;45(Pt 1):134-145 (doi: 10.1016/j.tiv.2017.08.015).

ANNEX - DEFINITIONS

Accuracy: The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance to mean the proportion of correct outcomes of a test method (1).

(Formula shown below.)

ADRA: Amino acid Derivative Reactivity Assay

AOP (Adverse Outcome Pathway): sequence of events from the chemical structure of a target chemical or group of similar chemicals through the molecular initiating event to an in vivo outcome of interest (2).

Calculation

Calculating depletion of either NAC or NAL

Depletion is calculated as follows:

Percent depletion of either NAC or NAL = {1- (NAC or NAL peak area in replicate injection ÷ mean NAC or NAL peak area in reference control C)} × 100

Calculating predictive capacity

There are several terms that are commonly used along with the description of sensitivity, specificity and accuracy. They are true positive (TP), true negative (TN), false negative (FN), and false positive (FP).

Sensitivity, specificity and accuracy are described in terms of TP, TN, FN, and FP.

Sensitivity: Number of true positives ÷ Number of all positive chemicals, TP ÷ (TP + FN)

Specificity: Number of true negatives ÷ Number of all negative chemicals, TN ÷ (TN + FP)

Accuracy: Number of correct predictions ÷ Number of all predictions, (TN + TP) ÷ (TN+TP+FN+FP)

Calibration curve: The relationship between the experimental response value and the analytical concentration (also called standard curve) of a known substance.

Coefficient of variation: a measure of variability that is calculated for a group of replicate data by dividing the standard deviation by the mean. It can be multiplied by 100 for expression as a percentage.

Defined Approach (DA): a DA consists of a fixed data interpretation procedure (e.g. statistical, mathematical models) applied to data (e.g. in silico predictions, in chemico, in vitro data) generated with a defined set of information sources to derive a prediction.

DPRA: Direct Peptide Reactivity Assay

EDTA: Ethylenediaminetetraacetic acid

EURL ECVAM: the European Union Reference Laboratory for Alternatives to Animal Testing

Hazard: Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

IATA (Integrated Approach to Testing and Assessment): A structured approach used for hazard identification (potential), hazard characterisation (potency), and/or safety assessment (potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weights all relevant data to inform regulatory decision regarding potential hazards, risks, and the need for further targeted and therefore minimal testing.

JaCVAM: Japanese Center for the Validation of Alternative Methods

LLNA: murine Local Lymph Node Assay issued as OECD TG 429 in 2010

Molecular Initiating Event: Chemical-induced perturbation of a biological system at the molecular level identified to be the starting event in the adverse outcome pathway.

Mixture: A solid or liquid comprising two or more substances which do not react chemically. (3)

Mono-constituent substance: A substance, defined by its quantitative composition, in which one main constituent comprises at least 80% (w/w) of the whole.

Multi-constituent substance: A substance, defined by its quantitative composition, in which two or more main constituents are present in concentrations ≥ 10% (w/w) and < 80% (w/w). Multi-constituent substances are the result of a manufacturing process. The difference between a mixture and a multi-constituent substance is that a mixture comprises two or more substances which do not react chemically, whereas a multi-constituent substance comprises two or more substances that do react chemically.

NAC: N-(2-(1-naphthyl)acetyl)-L-cysteine (4) (5) (6)

NAL: α-N-(2-(1-naphthyl)acetyl)-L-lysine (4) (5) (6)

Positive control: A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

Pre-haptens: chemicals which become sensitisers through abiotic transformation

Pro-haptens: chemicals requiring enzymatic activation to exert skin sensitisation potential

Reference control: An untreated sample containing all components of a test system, including the solvent or vehicle that is processed with the test chemical treated and other control samples to establish the baseline response for the samples treated with the test chemical dissolved in the same solvent or vehicle. When tested with a concurrent negative control, this sample also demonstrates whether the solvent or vehicle interacts with the test system.

Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method. (1)

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability. (1)

Reproducibility: The concordance of results obtained from testing the same substance using the same test protocol (see reliability). (1)

Sensitivity: The proportion of all positive/active chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method. (1) (Formula shown below.)

Specificity: The proportion of all negative/inactive chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method. (1) (Formula shown below.)

Substance: Chemical elements and their compounds in the natural state or resulting from a manufacturing process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process, but excluding solvents that may be separated without affecting the stability of the substance or changing its composition (3).

System suitability: Determination of instrument performance (e.g. sensitivity) by analysis of a reference standard prior to running the analytical batch (7).

Test chemical: The term test chemical is used to refer to the substance being tested.

TFA: Trifluoroacetic acid

United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS): A system proposing the classification of chemicals (substances and mixtures) according to standardised types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (3).

UVCB: substances of unknown or variable composition, complex reaction products or biological materials.

Valid test method: A test method considered to have sufficient relevance and reliability for a specific purpose and which is based on scientifically sound principles. A test method is never valid in an absolute sense, but only in relation to a defined purpose (1).

Literature for definitions

(1) 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. Organisation for Economic Cooperation and Development, Paris, France.

(2) OECD (2012), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Series on Testing and Assessment No. 168, OECD, Paris.

(3) United Nations (UN) (2013), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Fifth revised edition, UN New York and Geneva, 2013. Available at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html

(4) M Fujita, Yamamoto Y, Tahara H, Kasahara T, Jimbo Y and Hioki T (2014), Development of a prediction method for skin sensitisation using novel cysteine and lysine derivatives, Journal of pharmacological and toxicological methods, 70:94-105.

(5) Y Yamamoto, Tahara H, Usami R, Kasahara T, Jimbo Y, Hioki T and Fujita M (2015), A novel in chemico method to detect skin sensitisers in highly diluted reaction conditions, Journal of Applied Toxicology, 35(11):1348-60, (doi: 10.1002/jat.3139).

(6) M Fujita, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara K, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H and Kasahara T (2019), Cause of and Countermeasures for Oxidation of the Cysteine-Derived Reagent Used in the Amino acid Derivative Reactivity Assay, J. Appl. Toxicology, Feb;39(2):191-208 (doi: 10.1002/jat.3707).

(7) FDA (Food and Drug Administration) (2001), Guidance for Industry: Bioanalytical Method Validation 22pp. Accessible at:www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidance/ucm070 107.pdf - 138 (23)

APPENDIX I

In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA)

INITIAL CONSIDERATIONS, APPLICABILITY AND LIMITATIONS

10. The DPRA is proposed to address the molecular initiating event of the skin sensitisation AOP, namely protein reactivity, by quantifying the reactivity of test chemicals towards model synthetic peptides containing either lysine or cysteine (1). Cysteine and lysine percent peptide depletion values are then used to categorise a substance in one of four classes of reactivity for supporting the discrimination between skin sensitisers and non-sensitisers (2).

11. The DPRA test method proved to be transferable to laboratories experienced in high-performance liquid chromatography (HPLC) analysis. The level of reproducibility in predictions that can be expected from the test method is in the order of 85% within laboratories and 80% between laboratories (3). Results generated in the validation study (4) and published studies (5) overall indicate that the accuracy of the DPRA in discriminating sensitisers (i.e. UN GHS Cat. 1) from non-sensitisers is 80% (N=157) with a sensitivity of 80% (88/109) and specificity of 77% (37/48) when compared to LLNA results. The DPRA is more likely to under predict chemicals showing a low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals showing a high skin sensitisation potency (i.e. UN GHS subcategory 1A) (4) (5). However, the accuracy values given here for the DPRA as a stand-alone test method are only indicative since the test method should be considered in combination with other sources of information in the context of an IATA or a DA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore when evaluating non-animal methods for skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in the species of interest, i.e. humans. On the basis of the overall data available, the DPRA was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, skin sensitisation potency (as determined in *in vivo* studies) and physico-chemical properties (1) (2) (3) (5). Taken together, this information indicates the usefulness of the DPRA to contribute to the identification of skin sensitisation hazard.

12. The term "test chemical" is used in this Test Guideline to refer to what is being tested[1] and is not related to the applicability of the DPRA to the testing of substances and/or mixtures. This test method is not applicable for the testing of metal compounds since they are known to react with proteins with mechanisms other than covalent binding. A test chemical should be soluble in an appropriate solvent at a final concentration of 100 mM (see paragraph 10). However, test chemicals that are not soluble at this concentration may still be tested at lower soluble concentrations. In such a case, a positive result could still be used to support the identification of the test chemical as a skin sensitiser but no firm

[1] In June 2013, the Joint Meeting agreed that where possible, a more consistent use of the term “test chemical” describing what is being tested should now be applied in new and updated Test Guidelines.

conclusion on the lack of reactivity should be drawn from a negative result. Limited information is currently available on the applicability of the DPRA to mixtures of known composition (4) (5). The DPRA is nevertheless considered to be technically applicable to the testing of multi-constituent substances and mixtures of known composition (see paragraph 4 and 10). When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Appendix of the Test Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically. The current prediction model cannot be used for complex mixtures of unknown composition or for substances of unknown or variable composition, complex reaction products or biological materials (i.e. UVCB substances) due to the defined molar ratio of test chemical and peptide. For this purpose a new prediction model based on a gravimetric approach will need to be developed. In cases where evidence can be demonstrated on the non-applicability of the test method to other specific categories of chemicals, the test method should not be used for those specific categories of chemicals.

13. The test method described in this Appendix of the Test Guideline is an *in chemico* method that does not encompass a metabolic system. Chemicals that require enzymatic bioactivation to exert their skin sensitisation potential (i.e. pro-haptens) cannot be detected by the test method. Chemicals that become sensitisers after abiotic transformation (i.e. pre-haptens) are reported to be in most cases correctly detected by the test method (4) (9) (10). In the light of the above, negative results obtained with the test method should be interpreted in the context of the stated limitations and in the connection with other information sources within the framework of an IATA or a DA. Test chemicals that do not covalently bind to the peptide but promote its oxidation (i.e. cysteine dimerisation) could lead to a potential over estimation of peptide depletion, resulting in possible false positive predictions and/or assignment to a higher reactivity class (see paragraphs 21 and 22).

14. As described, the DPRA assay supports the discrimination between skin sensitisers and non-sensitisers. However, it may also potentially contribute to the assessment of sensitising potency (6) (11) when used in integrated approaches such as IATA or DA (12). However further work, preferably based on human data, is required to determine how DPRA results may possibly inform potency assessment.

PRINCIPLE OF THE TEST

15. The DPRA is an *in chemico* method which quantifies the remaining concentration of cysteine- or lysine-containing peptide following 24 hours incubation with the test chemical at 22.5-30°C. The synthetic peptides contain phenylalanine to aid in the detection. Relative peptide concentration is measured by high-performance liquid chromatography (HPLC) with gradient elution and UV detection at 220 nm. Cysteine- and lysine peptide percent depletion values are then calculated and used in a prediction model (see paragraph 21) which allows assigning the test chemical to one of four reactivity classes used to support the discrimination between sensitisers and non-sensitisers.

16. Prior to routine use of the method described in this Appenix, laboratories should demonstrate technical proficiency, using the ten proficiency substances listed in Annex 1.

PROCEDURE

17. This test method is based on the DPRA DB-ALM protocol nº 154 (7) which represents the protocol used for the EURL ECVAM-coordinated validation study. It is recommended that this protocol is used when implementing and using the method in the laboratory. The following is a description of the main components and procedures for the DPRA. If an alternative HPLC set-up is used, its equivalence to the validated set-up described in the DB-ALM protocol should be demonstrated (e.g. by testing the proficiency substances in Annex 1).

Preparation of the cysteine or lysine-containing peptides

18. Stock solutions of cysteine (Ac-RFAA**C**AA-COOH) and lysine (Ac-RFAA**K**AA-COOH) containing synthetic peptides of purity higher than 85% and preferably > 90%, should be freshly prepared just before their incubation with the test chemical. The final concentration of the cysteine peptide should be 0.667 mM in pH 7.5 phosphate buffer whereas the final concentration of the lysine peptide should be 0.667 mM in pH 10.2 ammonium acetate buffer. The HPLC run sequence should be set up in order to keep the HPLC analysis time less than 30 hours. For the HPLC set up used in the validation study and described in this test method, up to 26 analysis samples (which include the test chemical, the positive control and the appropriate number of solvent controls based on the number of individual solvents used in the test, each tested in triplicate), can be accommodated in a single HPLC run. All of the replicates analysed in the same run should use the identical cysteine and lysine peptide stock solutions. It is recommended to prove individual peptide batches for proper solubility prior to their use.

Preparation of the test chemical

19. Solubility of the test chemical in an appropriate solvent should be assessed before performing the assay following the solubilisation procedure described in the DPRA DB-ALM protocol (7). An appropriate solvent will dissolve the test chemical completely. Since in the DPRA the test chemical is incubated in large excess with either the cysteine or the lysine peptides, visual inspection of the forming of a clear solution is considered sufficient to ascertain that the test chemical (and all of its components in the case of testing a multi-constituent substance or a mixture) is dissolved. Suitable solvents are, acetonitrile, water, 1:1 mixture water:acetonitrile, isopropanol, acetone or 1:1 mixture acetone:acetonitrile. Other solvents can be used as long as they do not have an impact on the stability of the peptide as monitored with reference controls C (i.e. samples constituted by the peptide alone dissolved in the appropriate solvent; see Annex 2). If the test chemical is not soluble in any of the solvents mentioned above, DMSO can be used as a last resort and in minimal amounts. It is important to note that DMSO may lead to peptide dimerisation and as a result, it may be more difficult to meet the acceptance criteria. If DMSO is chosen, attempts should be made to first solubilise the test chemical in 300 µL of DMSO and dilute the resulting solution with 2700 µL of acetonitrile. If the test chemical is not soluble in this mixture, attempts should be made to solubilise the same amount of test chemicals in 1500 µL of DMSO and dilute the resulting solution with 1500 µL of acetonitrile. The test chemical should be pre-weighed into glass vials and dissolved immediately before testing in an appropriate solvent to prepare a 100 mM solution. For mixtures and multi-constituent substances of known composition, a single purity should be determined by the sum of the proportion of its constituents (excluding water), and a single apparent molecular weight should be determined by considering the individual molecular weights of each component

in the mixture (excluding water) and their individual proportions. The resulting purity and apparent molecular weight should then be used to calculate the weight of test chemical necessary to prepare a 100 mM solution. For polymers for which a predominant molecular weight cannot be determined, the molecular weight of the monomer (or the apparent molecular weight of the various monomers constituting the polymer) may be considered to prepare a 100 mM solution. However, when testing mixtures, multi-constituent substances or polymers of known composition, it should be considered to also test the neat chemical. For liquids, the neat chemical should be tested as such without any prior dilution by incubating it at 1:10 and 1:50 ratio with the cysteine and lysine peptides, respectively. For solids, the test chemical should be dissolved to its maximum soluble concentration in the same solvent used to prepare the apparent 100 mM solution. It should then be tested as such without any further dilution by incubating it at 1:10 and 1:50 ratio with the cysteine and lysine peptides, respectively. Concordant results (reactive or non-reactive) between the apparent 100 mM solution and the neat chemical should allow for a firm conclusion on the result.

Preparation of the positive control, reference controls and coelution controls

20. Cinnamic aldehyde (CAS 104-55-2; ≥95% food-grade purity) should be used as positive control (PC) at a concentration of 100 mM in acetonitrile. Other suitable positive controls providing mid-range depletion values may be used if historical data are available to derive comparable run acceptance criteria. In addition reference controls (i.e. samples containing only the peptide dissolved in the appropriate solvent) should also be included in the HPLC run sequence and these are used to verify the HPLC system suitability prior to the analysis (reference controls A), the stability of the reference controls over time (reference control B) and to verify that the solvent used to dissolve the test chemical does not impact the percent peptide depletion (reference control C) (see Annex 2). The appropriate reference control for each substance is used to calculate the percent peptide depletion for that substance (see paragraph 18). In addition, a co-elution control constituted by the test chemical alone for each of the test chemicals analysed should be included in the run sequence to detect possible co-elution of the test chemical with either the lysine or the cysteine peptide.

Incubation of the test chemical with the cysteine and lysine peptide solutions

21. Cysteine and lysine peptide solutions should be incubated in glass autosampler vials with the test chemical at 1:10 and 1:50 ratio respectively. If a precipitate is observed immediately upon addition of the test chemical solution to the peptide solution, due to low aqueous solubility of the test chemical, one cannot be sure how much test chemical remained in the solution to react with the peptide. Therefore, in such a case, a positive result could still be used, but a negative result is uncertain and should be interpreted with due care (see also provisions in paragraph 10 for the testing of chemicals not soluble up to a concentration of 100 mM). The reaction solution should be left in the dark at 22.5-30°C for 24±2 hours before running the HPLC analysis. Each test chemical should be analysed in triplicate for both peptides. Samples have to be visually inspected prior to HPLC analysis. If a precipitate or phase separation is observed, samples may be centrifuged at low speed (100-400xg) to force precipitate to the bottom of the vial as a precaution since large amounts of precipitate may clog the HPLC tubing or columns. If a precipitation or phase separation is observed after the incubation period, peptide depletion may be underestimated and a conclusion on the lack of reactivity cannot be drawn with sufficient confidence in case of a negative result.

Preparation of the HPLC standard calibration curve

22. A standard calibration curve should be generated for both the cysteine and the lysine peptides. Peptide standards should be prepared in a solution of 20% or 25% acetonitrile:buffer using phosphate buffer (pH 7.5) for the cysteine peptide and ammonium acetate buffer (pH 10.2) for the lysine peptide. Using serial dilution standards of the peptide stock solution (0.667 mM), 6 calibration solutions should be prepared to cover the range from 0.534 to 0.0167 mM. A blank of the dilution buffer should also be included in the standard calibration curve. Suitable calibration curves should have an $r^2>0.99$.

HPLC preparation and analysis

23. The suitability of the HPLC system should be verified before conducting the analysis. Peptide depletion is monitored by HPLC coupled with an UV detector (photodiode array detector or fixed wavelength absorbance detector with 220 nm signal). The appropriate column is installed in the HPLC system. The HPLC set-up described in the validated protocol uses a Zorbax SB-C-18 2.1 mm x 100 mm x 3.5 micron as preferred column. With this reversed-phase HPLC column, the entire system should be equilibrated at 30°C with 50% phase A (0.1% (v/v) trifluoroacetic acid in water) and 50% phase B (0.085% (v/v) trifluoroacetic acid in acetonitrile) for at least 2 hours before running. The HPLC analysis should be performed using a flow rate of 0.35 mL/min and a linear gradient from 10% to 25% acetonitrile over 10 minutes, followed by a rapid increase to 90% acetonitrile to remove other materials. Equal volumes of each standard, sample and control should be injected. The column should be re-equilibrated under initial conditions for 7 minutes between injections. If a different reversed-phase HPLC column is used, the set-up parameters described above may need to be adjusted to guarantee an appropriate elution and integration of the cysteine and lysine peptides, including the injection volume, which may vary according to the system used (typically in the range from 3-10 µL). Importantly, if an alternative HPLC set-up is used, its equivalence to the validated set-up described above should be demonstrated (e.g. by testing the proficiency substances in Annex 1). Absorbance is monitored at 220 nm. If a photodiode array detector is used, absorbance at 258 nm should also be recorded. It should be noted that some supplies of acetonitrile could have a negative impact on peptide stability and this has to be assessed when a new batch of acetonitrile is used. The ratio of the 220 peak area and the 258 peak area can be used as an indicator of co-elution. For each sample a ratio in the range of 90%<mean[2] area ratio of control samples<100% would give a good indication that co-elution has not occurred.

24. There may be test chemicals which could promote the oxidation of the cysteine peptide. The peak of the dimerised cysteine peptide may be visually monitored. If dimerisation appears to have occurred, this should be noted as percent peptide depletion may be over-estimated leading to false positive predictions and/or assignment to a higher reactivity class (see paragraphs 21 and 22).

25. The HPLC analysis should be timed to assure that the injection of the first sample starts 22 to 26 hours after the test chemical was mixed with the peptide solution. The HPLC run sequence should be set up in order to keep the HPLC analysis time less than 30 hours. For the HPLC set up used in the validation study and described in this test method, up to 26 analysis samples can be accommodated in a single HPLC run (see also paragraph 9). An example of HPLC analysis sequence is provided in Annex 2.

[2] For mean it is meant arithmetic mean throughout the document.

DATA AND REPORTING

Data evaluation

26. The concentration of cysteine or lysine peptide is photometrically determined at 220 nm in each sample by measuring the peak area (area under the curve, AUC) of the appropriate peaks and by calculating the concentration of peptide using the linear calibration curve derived from the standards.

27. The percent peptide depletion is determined in each sample by measuring the peak area and dividing it by the mean peak area of the relevant reference controls C (see Annex 2) according to the formula described below.

$$\text{Percent peptide depletion} = \left[1 - \left(\frac{\text{Peptide peak area in replicate injection}}{\text{Mean peptide peak area in reference controls } C}\right)\right] \times 100$$

Acceptance criteria

28. The following criteria should be met for a run to be considered valid:

a) the standard calibration curve should have an r^2>0.99,

b) the mean percent peptide depletion value of the three replicates for the positive control cinnamic aldehyde should be between 60.8% and 100% for the cysteine peptide and between 40.2% and 69.0% for the lysine peptide (for other positive controls a reference range needs to be established) and the maximum standard deviation (SD) for the positive control replicates should be <14.9% for the percent cysteine depletion and <11.6% for the percent lysine depletion and

c) the mean peptide concentration of reference controls A should be 0.50±0.05 mM and the coefficient of variation (CV) of peptide peak areas for the nine reference controls B and C in acetonitrile should be <15.0%.

If one or more of these criteria is not met the run should be repeated.

29. The following criteria should be met for a test chemical's results to be considered valid:

a) the maximum standard deviation for the test chemical replicates should be <14.9% for the percent cysteine depletion and <11.6% for the percent lysine depletion,

b) the mean peptide concentration of the three reference controls C in the appropriate solvent should be 0.50±0.05 mM.

If these criteria are not met the data should be rejected and the run should be repeated for that specific test chemical.

Prediction model

30. The mean percent cysteine and percent lysine depletion value is calculated for each test chemical. Negative depletion is considered as "0" when calculating the mean. By using the cysteine 1:10/lysine 1:50 prediction model shown in Table 1, the threshold of 6.38%

average peptide depletion should be used to support the discrimination between skin sensitisers and non-sensitisers in the framework of an IATA or DA. Application of the prediction model for assigning a test chemical to a reactivity class (i.e. low, moderate and high reactivity) may perhaps prove useful to inform potency assessment within the framework of an IATA or DA.

Table 1: Cysteine 1:10/lysine 1:50 prediction model[1]

Mean of cysteine and lysine % depletion	Reactivity Class	DPRA Prediction[2]
0% ≤ mean % depletion ≤ 6.38%	No or minimal reactivity	Negative
6.38% < mean % depletion ≤ 22.62%	Low reactivity	Positive
22.62% < mean % depletion ≤ 42.47%	Moderate reactivity	
42.47% < mean % depletion ≤ 100%	High reactivity	

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] A DPRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 4.

31. There might be cases where the test chemical (the substance or one or several of the components of a multi-constituent substance or a mixture) absorbs significantly at 220 nm and has the same retention time of the peptide (co-elution). Co-elution may be resolved by slightly adjusting the HPLC set-up in order to further separate the elution time of the test chemical and the peptide. If an alternative HPLC set-up is used to try to resolve co-elution, its equivalence to the validated set-up should be demonstrated (e.g. by testing the proficiency substances in Annex 1). When co-elution occurs the peak of the peptide cannot be integrated and the calculation of the percent peptide depletion is not possible. If co-elution of such test chemicals occurs with both the cysteine and the lysine peptides then the analysis should be reported as "inconclusive". In cases where co-elution occurs only with the lysine peptide, then the cysteine 1:10 prediction model reported in Table 2 can be used.

Table 2: Cysteine 1:10 prediction model[1]

Cysteine (Cys) % depletion	Reactivity class	DPRA prediction[2]
0% ≤ Cys % depletion ≤ 13.89%	No or minimal reactivity	Negative
13.89% < Cys % depletion ≤ 23.09%	Low reactivity	Positive
23.09% < Cys % depletion ≤ 98.24%	Moderate reactivity	
98.24% < Cys % depletion ≤ 100%	High reactivity	

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] A DPRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 4.

32. There might be other cases where the overlap in retention time between the test chemical and either of the peptides is incomplete. In such cases percent peptide depletion values can be estimated and used in the cysteine 1:10/lysine 1:50 prediction model, however assignment of the test chemical to a reactivity class cannot be made with accuracy.

33. A single HPLC analysis for both the cysteine and the lysine peptide should be sufficient for a test chemical when the result is unequivocal. However, in cases of results close to the threshold used to discriminate between positive and negative results (i.e. borderline results), additional testing may be necessary. If situations where the mean percent depletion falls in the range of 3% to 10% for the cysteine 1:10/lysine 1:50 prediction model or the cysteine percent depletion falls in the range of 9% to 17% for the cysteine 1:10 prediction model, a second run may be considered, as well as a third one in case of discordant results between the first two runs.

Test report

34. The test report should include the following information

Test chemical

- Mono-constituent substance
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
    - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
    - Treatment prior to testing, if applicable (e.g. warming, grinding);
    - Concentration(s) tested;
    - Storage conditions and stability to the extent available.
- Multi-constituent substance, UVCB and mixture:
    - Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
    - Physical appearance, water solubility and additional relevant physicochemical properties, to the extent available;
    - Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions or other information relevant for the conduct of the study;
    - Treatment prior to testing, if applicable (e.g. warming, grinding);
    - Concentration(s) tested;
    - Storage conditions and stability to the extent available.

Controls

- Positive control
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;

- Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
- Purity, chemical identity of impurities as appropriate and practically feasible, etc;
- Treatment prior to testing, if applicable (e.g. warming, grinding);
- Concentration(s) tested;
- Storage conditions and stability to the extent available;
- Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.

- Solvent/vehicle
  - Solvent/vehicle used and ratio of its constituents, if applicable;
  - Chemical identification(s), such as IUPAC or CAS name(s), CAS number(s), and/or other identifiers;
  - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
  - Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other solvents/vehicles than those mentioned in the test method are used and to the extent available;
  - Storage conditions and stability to the extent available;
  - Justification for choice of solvent for each test chemical;
  - For acetonitrile, results of test of impact on peptide stability.

Preparation of peptides, positive control and test chemical

- Characterisation of peptide solutions (supplier, lot, exact weight of peptide, volume added for the stock solution);
- Characterisation of positive control solution (exact weight of positive control substance, volume added for the test solution);
- Characterisation of test chemical solutions (exact weight of test chemical, volume added for the test solution).

HPLC instrument setting and analysis

- Type of HPLC instrument, HPLC and guard columns, detector, autosampler;
- Parameters relevant for the HPLC analysis such as column temperature, injection volumes, flow rate and gradient.

System suitability

- Peptide peak area at 220 nm of each standard and reference control A replicate;
- Linear calibration curve graphically represented and the r^2 reported;
- Peptide concentration of each reference control A replicate;

- Mean peptide concentration (mM) of the three reference controls A, SD and CV;
- Peptide concentration of reference controls A and C.

Analysis sequence

- For reference controls:
    - Peptide peak area at 220 nm of each B and C replicate;
    - Mean peptide peak area at 220 nm of the nine reference controls B and C in acetonitrile, SD an CV (for stability of reference controls over analysis time);
    - For each solvent used, the mean peptide peak area at 220 nm of the three appropriate reference controls C (for the calculation of percent peptide depletion);
    - For each solvent used, the peptide concentration (mM) of the three appropriate reference controls C;
    - For each solvent used, the mean peptide concentration (mM) of the three appropriate reference controls C, SD and CV.
- For positive control:
    - Peptide peak area at 220 nm of each replicate;
    - Percent peptide depletion of each replicate;
    - Mean percent peptide depletion of the three replicates, SD and CV.
- For each test chemical:
    - Appearance of precipitate in the reaction mixture at the end of the incubation time, if observed. If precipitate was re-solubilised or centrifuged;
    - Presence of co-elution;
    - Description of any other relevant observations, if applicable;
    - Peptide peak area at 220 nm of each replicate;
    - Percent peptide depletion of each replicate;
    - Mean of percent peptide depletion of the three replicate, SD and CV;
    - Mean of percent cysteine and percent lysine depletion values;
    - Prediction model used and DPRA prediction.

Proficiency testing

- If applicable, the procedure used to demonstrate proficiency of the laboratory in performing the test method (e.g. by testing of proficiency substances) or to demonstrate reproducible performance of the test method over time.

Discussion of the results

- Discussion of the results obtained with the DPRA test method;

- Discussion of the test method results in the context of an IATA if other relevant information is available.

Conclusion

LITERATURE FOR APPENDIX I

(1) Gerberick *et al.* (2004), Development of a peptide reactivity assay for screening contact allergens. Toxicological Sciences 81:332-343.

(2) Gerberick *et al.* (2007), Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach. Toxicological Sciences 97:417-427.

(3) EC EURL-ECVAM (2013), Recommendation on the Direct Peptide Reactivity Assay (DPRA) for skin sensitisation testing. Available at: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-direct-peptide-reactivity-assay-dpra.

(4) EC EURL ECVAM (2012), Direct Peptide Reactivity Assay (DPRA) Validation Study Report 74pp. Accessible at: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-direct-peptide-reactivity-assay-dpra

(5) Natsch *et al.* (2013), A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. Journal of Applied Toxicology, published online, 9 April 2013, DOI:10.1002/jat.2868.

(6) Jaworska *et al.* (2013), Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice Journal of Applied Toxicology, published online, 14 May 2013, DOI: 10.1002/jat.2869.

(7) DB-ALM (INVITTOX) Protocol 154: Direct Peptide Reactivity assay (DPRA) for skin sensitisation testing 17pp. Accessible at: http://ecvam-dbalm.jrc.ec.europa.eu/

(8) ECETOC (2003), Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).

(9) Urbisch et al. (2016), Assessment of Pre- and Pro-haptens Using Nonanimal Test Methods for Skin Sensitization, Chem Res Toxicol. 29(5):901-13 doi: 10.1021/acs.chemrestox.6b00055

(10) Pattlewicz et al. (2016), Can currently available non-animal methods detect pre and pro-haptens relevant for skin sensitization? Regul Toxicol Pharmacol.; 82:147-155. doi: 10.1016/j.yrtph.2016.08.007

(11) Jaworska et al. (2015), Arch Toxicol. 2015 Dec;89(12):2355-83. doi: 10.1007/s00204-015-1634-2)

(12) OECD (2016). Series on Testing & Assessment No. 256: Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Annex 1 and Annex 2. ENV/JM/HA(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [https://community.oecd.org/community/iatass].

APPENDIX I, ANNEX 1

PROFICIENCY SUBSTANCES

In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay

Prior to routine use of the test method described in this test method, laboratories should demonstrate technical proficiency by correctly obtaining the expected DPRA prediction for the 10 proficiency substances recommended in Table 1 and by obtaining cysteine and lysine depletion values that fall within the respective reference range for 8 out of the 10 proficiency substances for each peptide. These proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that they are commercially available, that high quality *in vivo* reference data and high quality *in vitro* data generated with the DPRA are available, and that they were used in the EURL ECVAM-coordinated validation study to demonstrate successful implementation of the test method in the laboratories participating in the study.

Table 1: Recommended proficiency substances for demonstrating technical proficiency with the Direct Peptide Reactivity Assay

Proficiency substances	CASRN	Physical state	*In vivo* prediction[1]	DPRA prediction[2]	Range[3] of % cysteine peptide depletion	Range[3] of % lysine peptide depletion
2,4-Dinitrochlorobenzene	97-00-7	Solid	Sensitiser (extreme)	Positive	90-100	15-45
Oxazolone	15646-46-5	Solid	Sensitiser (extreme)	Positive	60-80	10-55
Formaldehyde	50-00-0	Liquid	Sensitiser (strong)	Positive	30-60	≤24
Benzylideneacetone	122-57-6	Solid	Sensitiser (moderate)	Positive	80-100	≤7
Farnesal	19317-11-4	Liquid	Sensitiser (weak)	Positive	15-55	≤25
2,3-Butanedione	431-03-8	Liquid	Sensitiser (weak)	Positive	60-100	10-45
1-Butanol	71-36-3	Liquid	Non-sensitiser	Negative	≤7	≤5.5
6-Methylcoumarin	92-48-8	Solid	Non-sensitiser	Negative	≤7	≤5.5
Lactic Acid	50-21-5	Liquid	Non-sensitiser	Negative	≤7	≤5.5
4-Methoxyacetophenone	100-06-1	Solid	Non-sensitiser	Negative	≤7	≤5.5

[1]The *in vivo* hazard and (potency) predictions are based on LLNA data (5). The *in vivo* potency is derived using the criteria proposed by ECETOC (8).

[2] A DPRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 4.

[3] Ranges determined on the basis of at least 10 depletion values generated by 6 independent laboratories.

APPENDIX I, ANNEX 2

EXAMPLES OF ANALYSIS SEQUENCE

Calibration standards and reference controls	STD1 STD2 STD3 STD4 STD5 STD6 Dilution buffer Reference control A, rep 1 Reference control A, rep 2 Reference control A, rep 3
Co-elution controls	Co-elution control 1 for test chemical 1 Co-elution control 2 for test chemical 2
Reference controls	Reference control B, rep 1 Reference control B, rep 2 Reference control B, rep 3
First set of replicates	Reference control C, rep 1 Cinnamic aldehyde, rep 1 Sample 1, rep 1 Sample 2, rep 1
Second set of replicates	Reference control C, rep 2 Cinnamic aldehyde, rep 2 Sample 1, rep 2 Sample 2, rep 2
Third set of replicates	Reference control C, rep 3 Cinnamic aldehyde, rep 3 Sample 1, rep 3 Sample 2, rep 3
Reference controls	Reference control B, rep 4 Reference control B, rep 5 Reference control B, rep 6

Three sets of reference controls (i.e. samples constituted only by the peptide dissolved in the appropriate solvent) should be included in the analysis sequence:
Reference control A: used to verify the suitability of the HPLC system.
Reference control B: included at the beginning and at the end of the analysis sequence to verify stability of reference controls over the analysis time.
Reference control C: included in the analysis sequence to verify that the solvent used to dissolve the test chemical does not impact the percent peptide depletion.

APPENDIX II

In Chemico Skin Sensitisation: Amino acid Derivative Reactivity Assay (ADRA)

INITIAL CONSIDERATIONS, APPLICABILITY AND LIMITATIONS

35. The ADRA is proposed to address the molecular initiating event of the skin sensitisation AOP—namely, protein reactivity—by quantifying the reactivity of test chemicals towards model synthetic amino acid derivatives containing either lysine or cysteine (1) (2) (3). Depletion values of cysteine and lysine derivatives are then used to support the discrimination between skin sensitisers and non-sensitisers (1) (2) (3).

36. The ADRA proved to be transferable to laboratories experienced in high-performance liquid chromatography (HPLC) analysis. ADRA's WLR was 100% (10/10), 100% (7/7), 90% (9/10), and 100% (10/10) in four participating laboratories. BLR for 40 test chemicals calculated based the results from three participating laboratories was 91.9% (4). For the 40 chemicals tested in the validation study in four laboratories, the cumulative accuracy was 86.9% (139/160), sensitivity was 81.5% (88/108), and specificity was 98.1% (51/52) (4) (5). Results from the validation study (4) (5) as well as from other published studies (3) indicate that ADRA identified sensitisers and non-sensitisers with an accuracy of 79% (98/124) (124 compounds that fall within ADRA's applicability domain), a sensitivity of 74% (65/88), and a specificity of 92% (33/36) relative to LLNA results (6). In addition, the prediction of human skin sensitisation for 73 compounds that fall within ADRA's applicability domain has an accuracy of 86% (63/73), a sensitivity of 85% (44/52), and a specificity of 90% (19/21) (6). However, the accuracy values given here for ADRA as a stand-alone test method are for reference only, since it is recommended that the test method be used in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General Introduction. Furthermore when evaluating non-animal methods for skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in the species of interest, which is humans. On the basis of the overall data available, ADRA's applicability domain was shown to include a variety of organic functional groups, reaction mechanisms, skin sensitisation potencies (as determined in *in vivo* studies), and physicochemical properties (1) (2) (3) (4). Following an independent peer review, the ADRA validation study was considered to demonstrate that this method should be acceptable as part of an integrated testing strategy for the predictive identification of skin sensitisation hazard (7).

37. The term "test chemical" is used in this Test Guideline to refer to what is being tested and is not related to the applicability of the ADRA to the testing of substances and/or mixtures. This test method is not applicable to the testing of metal compounds, which are known to react with proteins via mechanisms other than covalent binding. The test method described in this Appendix of the Test Guideline is an *in chemico* method that does not encompass a metabolic system. Chemicals that require enzymatic bioactivation to exert their skin sensitisation potential (i.e. pro-haptens) cannot be detected by the test method. Chemicals that become sensitisers after abiotic transformation (i.e. pre-haptens) are

reported to be in some cases correctly detected by the test method (1) (2) (3) (4). In the light of the above, negative results obtained with the test method should be interpreted in the context of the stated limitations and in the connection with other information sources within the framework of an IATA. Test chemicals that promote the oxidation of the *N*-(2-(1-naphthyl)acetyl)-L-cysteine (NAC) reagent (i.e. cysteine dimerisation) could lead to a potential over-estimation of NAC depletion, resulting in possible false positive predictions (see paragraphs 27 and 28); it may be possible to detect and quantify any NAC dimer formed by HPLC, thus confirming or ruling out that the NAC reagent has been depleted via oxidative dimerisation as opposed to reaction and covalent bonding to the test item substance(s).

38. The ADRA test method allows testing of poorly soluble chemicals. To be tested, a test chemical should be soluble in an appropriate solvent at a final concentration of 1 mM (see paragraph 15). Test chemicals that are not soluble at this concentration may still be tested at lower concentrations. In such cases, a positive result could still be used to support identification of the test chemical as a skin sensitiser but no firm conclusion on the lack of reactivity should be drawn from a negative result.

39. In general, many organic compounds absorb UV in the range of 220 nm. In the case of co-elution of the nucleophilic reagent and the test chemical, this might result in false negative prediction. This may happen with the DPRA which specifies that quantification of the peptide-based nucleophilic reagents has to be performed at 220 nm. In contrast to this, the nucleophilic reagents used in ADRA are quantified at 281 nm. The substances that absorb UV in this range of the spectrum are generally limited to those having conjugated double bonds, which significantly lowers the potential for co-elution (8).

40. The current prediction model cannot be used for complex mixtures of unknown composition or for substances of unknown or variable composition, complex reaction products, or biological materials (UVCB substances) due to the need for defined molar ratio of test chemical and nucleophilic reagents. Limited information is currently available on the applicability of the ADRA to mixtures (9) (10). A new protocol has to be developed for multi-constituent substances and mixtures to be used with test methods like ADRA, which utilise HPLC analysis to quantify the depletion of nucleophilic reagents (9) (10). Thus, although it is impossible to define fixed methods in this guideline, which can evaluate multi-constituent substances and mixtures, paragraph 16 describes an evaluation method that is considered to be applicable at the present time for multi-constituent substances or mixtures of known composition (9). Such substances were however not tested during the validation studies. When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically.

41. ADRA can be used to support the discrimination between skin sensitisers and non-sensitisers. Further work, preferably based on human data, is necessary to determine whether ADRA results can contribute to potency assessment when considered in combination with other information sources.

PRINCIPLE OF THE TEST

42. ADRA is an *in chemico* test method that quantifies residual concentrations of the cysteine derivative *N*-(2-(1-naphthyl)acetyl)-*L*-cysteine (CAS. 32668-00-1), which is known as NAC, and the lysine derivative α-*N*-(2-(1-naphthyl)acetyl)-*L*-lysine (CAS. 397841-92-8), known as NAL, following a 24±1 hour incubation at 25±1ºC in the presence of a test chemical. Both these derivatives include a naphthalene ring that is introduced to their *N*-terminal in order to facilitate UV detection. The relative concentrations of NAC and NAL are measured by high-performance liquid chromatography (HPLC) with gradient elution and UV detection at 281 nm. Percent depletion values are then calculated for both NAC and NAL and compared to a prediction model (see paragraph 26).

43. Prior to routine use of the method described in this test method, laboratories should demonstrate technical proficiency, using the ten proficiency substances listed in Annex 1 of this Appendix.

PROCEDURE

44. This test method is based on the protocol (11) used for the JaCVAM-coordinated ADRA validation study and is recommended for use when implementing ADRA at a laboratory. The main components and procedures for the ADRA are described below. Before using an alternative HPLC set-up, its equivalence to the validated set-up described in the protocol should be demonstrated, preferably by testing the proficiency substances in Annex 1 of this Appendix.

Quality of NAC and NAL

45. The Nucleophilic Reagents can be obtained as an ADRA Kit for Skin Sensitisation Test, from FUJIFILM Wako (FFWK) Pure Chemical Corporation, Catalog No. 296-80901. Manufacturing NAC/NAL is patented in Japan only, by Fujifilm Corporation. Therefore, manufacturers in other countries can produce NAC/NAL without permission. In case other NAC/NAL are used, these should satisfy three quality criteria described below. Quality checks can be obviated and ADRA testing performed without delay by purchasing NAC and NAL that have been manufactured specifically to satisfy these quality criteria.

Quality required for NAC and NAL:

1) Purity: Both NAC and NAL are to be at least 98% pure.

2) Stability: Using NAC and NAL stock solution, prepare a reference control free of any test chemical and quantify the residual levels of NAC and NAL both immediately after preparation (0 hours) and after a 24 hour incubation. Residual levels of NAC and NAL are to be a minimum of 90% in either case (11). The residual level of NAC is calculated as a percentage of the sum of NAC and the residual level of NAC dimers.

3) Reactivity: NAC and NAL are to be evaluated for reactivity with the ten proficiency substances given in Annex 1 and should satisfy the requirement given therein.

Preparation of the NAC and NAL stock solution

46. The solubility of individual NAC and NAL batches should be verified prior to use. NAC stock solution should be prepared to a concentration of 2 mM in 100 mM of pH 8.0 phosphate buffer, including 0.333 μM of EDTA, as well as NAL stock solution to a concentration of 2 mM in 100 mM of pH 10.2 phosphate buffer. These two stock solutions are then diluted in buffer to prepare 6.667 μM stock solutions. Both NAC and NAL stock solutions should be used as soon as possible after preparation (3). In the event that they are to be stored, these stock solutions may be frozen and stored for up to twelve months time at less than -75°C prior to use. The final concentration of the NAC solution is 5 μM in pH 8.0 phosphate buffer, and the final concentration of the NAL solution is 5 μM in pH 10.2 phosphate buffer.

Preparation of the test chemical solution

47. Solubility of the test chemical in an appropriate solvent should be assessed before performing the assay in accordance with the solubilisation procedure described in the ADRA JaCVAM protocol (11). An appropriate solvent should dissolve the test chemical completely. Since the ADRA protocol stipulates that the test chemical be incubated in an excess volume of both NAC and NAL, visual inspection of the clear test chemical solution is considered sufficient to confirm that the test chemical (and all its constituents, if testing a multi-constituent substance or a mixture) is dissolved. Suitable solvents are distilled water, acetonitrile and acetone. If the test chemical is not soluble in any of the solvents mentioned above, DMSO can be used as a last resort and in minimal amounts. It is important to note that DMSO may lead to dimerisation of the nucleophilic reagent NAC (12) and as a result, it may be more difficult to meet the acceptance criteria. If DMSO is chosen, attempts should be made to solubilise the test chemical in a 1:20 mixture of DMSO and acetonitrile (5% DMSO in acetonitrile). When using a DMSO-acetonitrile solvent, the test chemical should be dissolved in DMSO, and then this solution should be diluted 20-fold with acetonitrile to prepare a 1 mM test chemical solution. In case the use of DMSO leads to increased dimerisation of the NAC reagent, this can be checked analytically as the NAC dimer can be detected by HPLC. The test chemical should be pre-weighed into a disposable polypropylene tube and dissolved immediately before testing in an appropriate solvent to prepare a 1 mM solution.

48. Mono-constituent substances of unknown molecular weight may be tested in a test chemical solution at a concentration of 0.5 mg/mL rather than 1 mM (9). Polymers which are well characterised should also be tested at a concentration of 1 mM based on the mean number average molecular weight, in a manner analogous to the procedure for mono-constituent compounds.

49. Mixtures and multi constituent substances, of known composition are to be tested as follows:

> 1) Liquids: Generally, tested as an undiluted mixture. In cases where low solubility of the test item prevents formation of reaction solution, i.e. undissolved material, clouding, and/or precipitation is observed, a positive result may still be used in the assessment, whereas a negative result is uncertain and should be interpreted with due care. Insofar as results could be false positives, however, predictions should be interpreted with due care.
>
> 2) Solids: The test chemical should be dissolved to maximum soluble concentration in the same solvent used to prepare the 1 mM test chemical solution. The test

chemical solution of the highest concentration possible is then tested as an undiluted mixture. In cases where low solubility of the test item prevents formation of reaction solution, i.e. undissolved material, clouding, and/or precipitation is observed, a positive result may still be used in the assessment, whereas a negative result is uncertain and should be interpreted with due care. Insofar as results could be false positives, however, predictions should be interpreted with due care.

Preparation of the positive control, reference controls and co-elution controls

50. Phenylacetaldehyde (CAS 122-78-1, ≥90% purity) should be used as positive control (PC) at a concentration of 1 mM in acetonitrile. Other suitable positive controls that provide mid-range depletion values may be used if historical data are available to derive comparable run acceptance criteria. In addition, reference controls comprising only NAC or only NAL dissolved in the appropriate solvent should also be included in the HPLC run sequence, so they can be used to verify the HPLC system suitability prior to analysis (Reference Control A), the stability of the reference controls over time (Reference Control B), and any effects of the solvent used on depletion of NAC or NAL (Reference Control C) (See Annex 2). The percent NAC and NAL depletion for a test chemical is calculated using an appropriate reference control for that test chemical (see paragraph 23). Also, a co-elution control comprising only the test chemical should be included in the run sequence to detect possible co-elution of the test chemical with either the NAC or NAL.

Incubation of the test chemical with the NAC and NAL solutions

51. Both the NAC and the NAL solutions should be incubated with the test chemical at 1:50 ratio in a 96-well microplate. The observation of precipitate immediately upon addition of the test chemical solution to the NAC and the NAL solutions is an indication of poor solubility, which means that there is no way to know exactly how much test chemical is contained in the solution. Thus, although positive results can be used with confidence, negative results are uncertain and should be interpreted with due care (see also paragraph 4 regarding the testing of chemicals not soluble at concentrations as high as 1 mM). The reaction solution should be incubated in the dark at 25±1ºC for 24±1 hours before performing HPLC analysis. After incubation, trifluoroacetic acid (TFA) (≥ 98%) should be added as a fixing solution to stop the reaction (3).

HPLC preparation and analysis

52. Each test chemical should be analysed in triplicate to determine percent depletion for both NAC and NAL. Although adding the fixing solution does stop the reaction, measurement of the reaction solution is to be performed as soon as possible and in any case within three days after adding the fixing solution. For example, when HPLC analysis of NAC and NAL are performed separately using two 96-well microplates, up to 34 samples may be analysed at one time, including the test chemical, the positive control, and the appropriate number of solvent controls based on the number of individual solvents used in the test, each in triplicate. All of the replicates analysed in a single run should use identical batches of NAC and NAL stock solution. Test chemical and control solutions are to be visually inspected prior to HPLC analysis and may be centrifuged at low speed (100–400 × g) to force any precipitate to the bottom of the vial as a precaution against large amounts of precipitate clogging the HPLC tubing or columns. Observation of precipitation or phase separation after the incubation period is an indication that NAC and NAL depletion could be misleading, and negative results in that case are uncertain and should be interpreted with

due care, as well as for any precipitate observed at the beginning of the incubation period (see above).

53. A standard calibration curve should be generated for both NAC and NAL. Standard solutions of both NAC and NAL should be prepared in 20% acetonitrile in buffer and containing 0.5% trifluoroacetic acid. For NAC, a phosphate buffer at pH 8.0, and for NAL, a phosphate buffer at pH 10.2 should be used. Serial dilution of the NAC and NAL stock solutions (5.0 μM) will be used to prepare six calibration solutions in concentrations from 5.0 to 0.156 μM as well as a blank of the dilution buffer. Suitable calibration curves should have an $R^2 > 0.990$.

54. The suitability of the HPLC system should be verified before conducting the analysis. Both NAC and NAL depletion is monitored by HPLC coupled with an UV detector (photodiode array detector or fixed wavelength absorbance detector with 281 nm signal). The appropriate column is installed in the HPLC system. The recommended HPLC set-up described in the validated protocol uses a column (Base particle: core-shell type silica gel, Particle size: 2.5~2.7 μm, colomn size: 3.0 × 150 mm) as preferred column. With this reversed-phase HPLC column, the entire system should be equilibrated for at least 30 minutes at 40ºC with 50% phase A (0.1% (v/v) trifluoroacetic acid in water), 50% phase B (0.1% (v/v) trifluoroacetic acid in acetonitrile) before use. Then, the column is conditioned by running the gradient at least twice before actual use. The HPLC analysis should be performed using a flow rate of 0.30 mL/min and a linear gradient from 30% to 55% acetonitrile for NAC and from 25% to 45% acetonitrile for NAL within 10 minutes, followed by a rapid increase to 100% acetonitrile to remove other materials. Equal volumes of the standard solutions, test chemical solutions, and control solutions should be injected. The column should be re-equilibrated under initial conditions for 6.5 minutes between injections. If a different reversed-phase HPLC column is used, the set-up parameters described above may need to be adjusted to guarantee an appropriate elution and integration of the NAC and NAL, including the injection volume, which may vary according to the system used (typically in the range from 10–20 μL). Importantly, if an alternative HPLC set-up is used, its equivalence to the validated set-up described above should be demonstrated, preferably by testing the proficiency substances in Annex 1. Absorbance is monitored at 281 nm. If a photodiode array detector is used, absorbance at 291 nm should also be recorded. It should be noted that some batches of acetonitrile could have a negative impact on NAC and NAL stability and this has to be assessed when a new batch of acetonitrile is used. The ratio of the 281 nm peak area and the 291 nm peak area can be used as an indicator of co-elution. For each sample a ratio in the range of 90% < mean area ratio of control samples < 100% would give a good indication that co-elution has not occurred. An example of HPLC analysis sequence is provided in Annex 2.

55. There are some test chemicals that could potentially promote oxidation of NAC. The peak of the dimerised NAC may be monitored visually. Any apparent dimerisation should be noted, since overestimation of NAC depletion could result in false-positive predictions (See paragraphs 26 and 27).

DATA AND REPORTING

Data evaluation

56. The concentration of both NAC and NAL is photometrically determined at 281 nm in each sample by measuring the peak area (area under the curve, AUC) of the appropriate peaks and by calculating the concentration of both NAC and NAL using the linear calibration curve derived from the standards.

57. The percent depletion for both NAC and NAL is determined in each sample by measuring the peak area and dividing it by the mean peak area of the relevant Reference Controls C (See Annex 2) according to the formula described below.

$$\text{Percent NAC or NAL depletion} = \left[1-\left[\frac{\text{NAC or NAL peak area in replicate injection}}{\text{Mean NAC or NAL peak area in reference controls C}}\right]\right]\text{x}100$$

Acceptance criteria

58. The following criteria should be met:

a) the standard calibration curve should have an $R^2 > 0.990$,

b) the mean percent NAC and NAL depletion value of the three replicates for the positive control phenylacetaldehyde should be between 6% and 30% for NAC and between 75% and 100% for NAL, while the maximum standard deviation (SD) for the positive control replicates should be < 10% for both NAC and NAL depletion, and

c) the mean NAC and NAL concentration of both Reference Controls A and Reference Control C should be 3.2–4.4 μM and the coefficient of variation (CV) of NAC and NAL peak areas for the nine Reference Controls B and C in acetonitrile should be < 10%.

If one or more of these criteria is not satisfied, the data should be rejected and the run should be repeated.

59. The following criteria should be satisfied for a test chemical's results to be accepted as valid:

a) the maximum standard deviation for the test chemical replicates should be < 10% for the percent depletion of both NAC and NAL,

b) the mean NAC and NAL concentration of the three Reference Controls C in the appropriate solvent should be 3.2–4.4 μM.

If one or more of these criteria is not satisfied, the data should be rejected and the run should be repeated.

Prediction model

60. The mean percent depletion of NAC and NAL is calculated for each test chemical. Negative depletion is considered to be "0" when calculating the mean. By using the NAC/NAL prediction model shown in Table 1, the threshold of 4.9% mean percent depletion should be used to support the discrimination between skin sensitisers and non-sensitiser in the framework of an IATA or a DA.

Table 1: NAC/NAL prediction model[1]

Mean NAC and NAL percent depletion	ADRA prediction[2]
Less than 4.9%	Negative
4.9% or higher	Positive

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] An ADRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 3.

61. Co-elution occurs when the test chemical (the substance or one or several of the constituents of a multi-constituent substance or a mixture) absorbs significantly at 281 nm and has the same retention time as NAC or NAL. Co-elution may be resolved by slightly adjusting the HPLC set-up in order to further separate the elution time of the test chemical and NAC or NAL. If an alternative HPLC set-up is used to try to resolve co-elution, its equivalence to the validated set-up should be demonstrated, preferably by testing the proficiency substances in Annex 1. When co-elution occurs, it is not possible to integrate the peak of the NAC or NAL, thereby preventing calculation of the percent depletion of NAC or NAL. If co-elution of test chemicals occurs with both the NAC and NAL and separation of elution time is not feasible, then the analysis should be reported to be inconclusive. In cases where co-elution occurs only with NAL and separation of elution time is not feasible, the NAC-only prediction model (See Table 2) can be used to make a prediction.

Table 2: NAC-only prediction model[1]

Mean NAC percent depletion	ADRA prediction[2]
Less than 5.6%	Negative
5.6% or higher	Positive

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] An ADRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 3.

62. When a result is unequivocal, a single HPLC analysis for both NAC and NAL should be sufficient for a test chemical. Additional testing is sometimes necessary, however, when the results lie close to the threshold value used to discriminate between positive and negative results (borderline results). If the mean percent depletion falls between 3.0% and 10.0% when using the NAC/NAL prediction model or the NAC percent depletion falls between 4.0% and 11.0% when using the NAC-only prediction model, a second run is advisable, as is a third run in the event of discordant results between the first two runs.

Test report

63. The test report should include the following information:

Test chemical

- Mono-constituent substance
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available
    - Purity, chemical identity of impurities as appropriate and practically feasible, etc.
    - Treatment prior to testing, if applicable (warming, grinding)
    - Concentration(s) tested
    - Storage conditions and stability to the extent available
- Multi-constituent substance, UVCB, and mixtures
    - Characterisation by chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available
    - Physical appearance, water solubility, and additional relevant physicochemical properties, to the extent available
    - Molecular weight (or apparent molecular weight) for mixtures or polymers of known composition, or other information relevant to the study
    - Treatment prior to testing, if applicable (warming, grinding)
    - Concentration(s) tested
    - Storage conditions and stability, to the extent available.

Controls

- Positive control
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
    - Purity, chemical identity of impurities, as appropriate or feasible
    - Treatment prior to testing, if applicable (warming, grinding)
    - Concentration(s) tested
    - Storage conditions and stability, to the extent available;
    - Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.

- Solvent
    - Solvent used and ratio of its constituents, if applicable
    - Chemical identification(s), such as IUPAC or CAS name(s), CAS number(s), and/or other identifiers
    - Purity, chemical identity of impurities, as appropriate and feasible
    - Physical appearance, molecular weight, and additional relevant physicochemical properties when solvents other than those mentioned in the test method are used
    - Storage conditions and stability, to the extent available
    - Justification for choice of solvent for each test chemical
    - Impact on NAC and NAL stability when using acetonitrile

Preparation of NAC and NAL, positive control and test chemical solution

- Characterisation of NAC and NAL solutions (supplier, lot, exact weight of NAC and NAL, volume added for the stock solution)
- Characterisation of positive control solutions (exact weight of positive control reagent, volume added for the control solution)
- Characterisation of test chemical solutions (exact weight of test chemical, volume added for the test chemical solution)

HPLC instrument setting and analysis

- Type of HPLC instrument, HPLC and guard columns, detector, autosampler
- Parameters relevant for the HPLC analysis such as column temperature, injection volumes, flow rate and gradient

System suitability

- NAC and NAL peak area at 281 nm of each standard and reference control A replicate
- Linear calibration curve graphically represented and the R2 reported
- NAC and NAL concentration of each Reference Control A replicate
- Mean NAC and NAL concentration (µM) of the three reference controls A, SD and CV
- NAC and NAL concentration of Reference Controls A and C.

Analysis sequence

- For Reference Controls
    - NAC and NAL peak area at 281 nm of each replicate of Reference Controls B and C
    - Mean NAC and NAL peak area at 281 nm of the nine Reference Controls B and C in acetonitrile, SD and CV (for stability of reference controls over analysis time)

- For each solvent used, the mean NAC and NAL peak area at 281 nm of the three appropriate Reference Controls C (for the calculation of percent NAC and NAL depletion)
- For each solvent used, the NAC and NAL concentration (μM) of the three appropriate Reference Controls C
- For each solvent used, the mean NAC and NAL concentration (μM) of the three appropriate Reference Controls C, SD and CV.

- For positive controls
  - NAC and NAL peak area at 281 nm of each replicate
  - Percent NAC and NAL depletion of each replicate
  - Mean percent NAC and NAL depletion of the three replicates, SD and CV.

- For each test chemical
  - Appearance of precipitate in the reaction mixture at the end of the incubation time, if observed. If precipitate was re-solubilised or centrifuged;
  - Presence of co-elution
  - Description of any other relevant observations, if applicable
  - NAC and NAL peak area at 281 nm of each replicate
  - Percent NAC and NAL depletion of each replicate
  - Mean of percent NAC and NAL depletion of the three replicate, SD and CV
  - Mean of percent NAC and percent NAL depletion values
  - Prediction model used and ADRA prediction

Proficiency testing

- If applicable, the procedure used to demonstrate proficiency of the laboratory in performing the test method (testing of proficiency substances, etc.) or to demonstrate reproducible performance of the test method over time.

Discussion of the results

- Discussion of the results obtained with the ADRA test method
- Discussion of the test method results in the context of an IATA if other relevant information is available

Conclusion

Literature for Appendix II

(1) Fujita M, Yamamoto Y, Tahara H, Kasahara T, Jimbo Y and Hioki T (2014), Development of a prediction method for skin sensitisation using novel cysteine and lysine derivatives, Journal of Pharmacological and Toxicological Methods, 70:94-105. DOI: 10.1016/j.vascn.2014.06.001.

(2) Yamamoto Y, Tahara H, Usami R, Kasahara T, Jimbo Y, Hioki T and Fujita M (2015), A novel *in chemico* method to detect skin sensitisers in highly diluted reaction conditions, Journal of Applied Toxicology, 35:1348-60. DOI: 10.1002/jat.3139.

(3) Fujita M, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara K, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H and Kasahara T (2019), Cause of and Countermeasures for Oxidation of the Cysteine-Derived Reagent Used in the Amino acid Derivative Reactivity Assay, Journal of Applied Toxicology, 39,191-208. DOI: 10.1002/jat.3707.

(4) ADRA validation study, *in preparation for submission*

(5) OECD (2019), Draft validation report: Amino acid Derivative Reactivity Assay (ADRA) – JaCVAM Validation Study Report. Series on testing and Assessment n° xx. Organisation for Economic Cooperation and Development, Paris.

(6) Yamamoto Y, Fujita M, Wanibuchi S, Sato A, Katsuoka Y, Ono A, Kasahara T. (2019), Availability of ADRA for prediction of skin sensitization by combining multiple alternative methods. in preparation for submission

(7) OECD (2019), Amino acid Derivative Reactivity Assay (ADRA) – Report of the Peer Review Panel. Series on testing and Assessment n° xx. Organisation for Economic Cooperation and Development, Paris.

(8) Fujira M, Yamamoto Y, Wanibuchi S, Katsuoka Y, Kasahara T, (2019), The underlying factors that explain why nucleophilic reagents rarely co-elute with test chemicals in the ADRA. Journal of Pharmacological and Toxicological Methods. DOI: 10.1016/j.vascn.2019.02.004

(9) YamamotoY, Fujita M, Manibuchi S, Katsuoka Y, Ono A, Kasahara T.. (2019), Expanding the applicability of the Amino acid Derivative Reactivity Assay (ADRA): Determining a weight concentration for preparation of test chemical solutions that yields a predictive capacity identical to the conventional method using molar concentration and demonstrating the capacity to detect sensitizers in liquid mixtures. Journal of Pharmacological and Toxicological Methods. DOI: 10.1016/j.vascn.2019.01.001

(10) Fujita M, Yamamoto Y, Wanibuchi S, Katsuoka Y, Kasahara T A Newly Developed Means of HPLC-Fluorescence Analysis for Predicting the Skin Sensitization Potential of Multi-Constituent Substances Using ADRA. Toxicology In Vitro, under revision

(11) ADRA protocol: JaCVAM Statements. Available at: http://www.jacvam.jp/en_effort/effort02.html

(12) James P. Tam JP, Cui Rong Wu CR, Wen Liu W, Jing Wen Zhang JW. (1991), Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. Journal of the American Chemical Society, 113, 6657–6662. DOI: 10.1021/ja00017a044

(13) Natsch A, Ryan CA, Foertsch L, Emter R, Jaworska J, Gerberick F, Kern P. (2013). A dataset on 145 chemicals tested in alternative assays for skin sensitisation undergoing prevalidation. *Journal of Applied Toxicology*, **33**(11):1337-52, DOI:10.1002/jat.2868.

(14) Gerberick GF, Vassalo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin J-P, (2007), Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach, Toxicological Sciences, 97, 417-427. DOI: 10.1093/toxsci/kfm064.

(15) Basketter DA, Scholes EW, (1992), Comparison of the local lymph node assay with the guinea-pig maximization test for the detection of a range of contact allergens, Food and Chemical Toxicology, 30, 65-69.

(16) ECETOC (2003). Contact sensitisation: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).

APPENDIX II, ANNEX 1

Proficiency Substances

In Chemico Skin Sensitisation: Amino acid Derivative Reactivity Assay (ADRA)

Prior to routine use of the test method, laboratories should demonstrate technical proficiency by correctly obtaining the expected ADRA prediction for the 10 proficiency substances recommended in Table 1 and by obtaining NAC and NAL depletion values that fall within the respective reference ranges for 8 out of the 10 proficiency substances. These proficiency substances were selected to represent the full range of responses for skin sensitisation hazards. Other selection criteria were that they are commercially available, that high quality *in vivo* reference data and high quality ADRA data are available, and that they were used during the JaCVAM-coordinated validation study to demonstrate successful implementation.

Table 1. Recommended chemicals for demonstrating technical proficiency with ADRA

No.	Test chemicals	CAS No.	Physical state	Molecular weight	*In vivo* Prediction[1]	ADRA prediction[2]	Range of % depletion	
							NAC[3]	NAL[3]
1	*p*-Benzoquinone	106-51-4	Solid	108.09	Sensitiser (extreme)	Positive	90-100	40-70
2	Chloramine T trihydrate	7080-50-4	Solid	281.69	Sensitiser (strong)	Positive	90-100	90-100

3	Trans-Cinnamaldehyde	14371-10-9	Liquid	132.16	Sensitiser (moderate)	Positive	40-100	≤20
4	Palmitoyl Chloride	112-67-4	Liquid	274.87	Sensitiser (moderate)	Positive	≤10	50-100
5	Imidazolidinyl urea	39236-46-9	Solid	388.29	Sensitiser (weak)	Positive	10-45	≤10
6	Farnesol	4602-84-0	Liquid	222.37	Sensitiser (weak)	Positive	20-40	≤15
7	Glycerol	56-81-5	Liquid	92.09	Non-sensitiser	Negative	≤7	≤7
8	Benzyl alcohol	100-51-6	Liquid	108.14	Non-sensitiser	Negative	≤7	≤7
9	Dimethyl isophthalate	1459-93-4	Solid	194.19	Non-sensitiser	Negative	≤7	≤7
10	Propyl paraben	94-13-3	Solid	110.11	Non-sensitiser	Negative	≤7	≤7

[1]The *in vivo* hazard (and potency) predictions are based on LLNA data. (13) (14) (15). The *in vivo* potency is derived using the criteria proposed by ECETOC (16).

[2] An ADRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 3.

[3] Ranges determined on the basis of at least 10 depletion values generated by 5 independent laboratories.

APPENDIX II, ANNEX 2

EXAMPLES OF ANALYSIS SEQUENCE

Each sample of HPLC analysis should be analysed in number order below. Refer to the table showing Examples of HPLC Sample Analysis Sequences for more practical sequences about HPLC analysis.

1. Start to analyse calibration standards and Reference Control A (N = 3).

2. The co-elution Control does not need to be analysed by turns if it is analysed after analysis of standard solution and Reference Control A.

3. Reference Control B should be analysed three times (total six times) before and after the analysis of sample, Reference Control C and Positive Control.

4. The Reference Control C, Positive Control and Test chemical solutions are analysed. (After the first set of replicates of each sample is analysed, the second set of replicates of each should be analysed).

Calibration standards and reference controls	STD1 STD2 STD3 STD4 STD5 STD6 Dilution buffer Reference control A, rep 1 Reference control A, rep 2 Reference control A, rep 3
Co-elution controls	Co-elution control 1 for test chemical 1 Co-elution control 2 for test chemical 2
Reference controls	Reference control B, rep 1 Reference control B, rep 2 Reference control B, rep 3
First set of replicates	Reference control C, rep 1 Phenylacetaldehyde, rep 1 Sample 1, rep 1 Sample 2, rep 1
Second set of replicates	Reference control C, rep 2 Phenylacetaldehyde, rep 2 Sample 1, rep 2 Sample 2, rep 2
Third set of replicates	Reference control C, rep 3 Phenylacetaldehyde, rep 3 Sample 1, rep 3 Sample 2, rep 3
Reference controls	Reference control B, rep 4 Reference control B, rep 5 Reference control B, rep 6

Three sets of reference controls (NAC or NAL dissolved in the appropriate solvent) should be included in the analysis sequence:

Reference control A: Control for verifying validity of the HPLC system. Reference Control A is used to verify concentration of NAC and NAL from each calibration curve after addition of acetonitrile rather than test chemical.

Reference control B: Control for verifying stability of reaction solution under analysis. Reference Control B is used to verify variability (CV) of each three NAC/NAL peak areas in the solution after addition of acetonitrile rather than test chemical at the start of analysis and at the end of analysis.

Reference control C:

Control for calculating NAC/NAL depletion of each test chemical solution. To calculate depletion of NAC/NAL, measure three Reference Controls C after addition of solvent instead of test chemical. Prepare reference Control C for all solvents used to dissolve the test chemicals.

添付資料２

Section 4
Health effects

Test Guideline No. 495

Reactive Oxygen Species (ROS) Assay for Photoreactivity

18 June 2019

OECD Guidelines for the Testing of Chemicals

OECD GUIDELINE FOR TESTING OF CHEMICALS

Ros (Reactive Oxygen Species) Assay For Photoreactivity

INTRODUCTION

1. Phototoxicity is defined as a toxic response is elicited by topically or systemically administered photoreactive chemicals after the exposure of the body to environmental light. Several classes of photoreactive chemicals could cause phototoxic reactions when activated by light at otherwise non-toxic doses. Phototoxicity can be categorized as photoirritation, photoallergy, and photogenotoxicity (1). Photoirritation is characterized as an acute light-induced skin response to a photoreactive chemical. Photoallergy is an immune-mediated reaction in which light may cause a structural change in a drug so that it acts as a hapten, possibly by binding to proteins in the skin (2). Photogenotoxicity is a genotoxic response after exposure to a chemical by two mechanisms: either directly by photoexcitation of DNA or indirectly by excitation of photoreactive chemicals.

2. In 2002, regulatory agencies in the US (US Food and Drug Administration, FDA) and EU (European Medicines Agency, EMA) published guidelines for photosafety assessments of drug candidates (3)(4). In 2004, the Organisation for Economic Co-operation and Development (OECD) adopted Test Guideline 432: *In vitro* 3T3 Neutral Red Uptake (NRU) Phototoxicity Test as a validated methodology for evaluating the phototoxic potential of chemicals (5). The EMA also published a concept paper in 2008 (6), which proposes a testing strategy that merges the testing proposals recommended by FDA and EMA. Considering these documents, the International Council of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) published ICH S10 guideline, "Photosafety Evaluation of Pharmaceuticals" in 2014 (7).

3. According to above referenced guidelines, chemicals or drug candidates need to be examined for their phototoxic potential. Since light must be absorbed by a compound in order for photochemical reactions to take place (8), the phototoxic potential of chemicals is related to the photochemical properties of compounds, especially light absorption properties within 290–700 nm. The guidelines suggested the need for measurement of the light absorption properties of chemicals as a first round of screening (3)(4). The ICH S10 guideline recommends UV-visible light absorption spectral analysis as a criterion for evaluating the phototoxic potentials of drugs (7); however, UV-visible light absorption of chemicals would not always correlate directly with their phototoxic potential, so a

combination of UV data (molar extinction coefficient, MEC) with other appropriate screening systems might be advantageous in avoiding false predictions.

4. In addition to light absorption and distribution to light-exposed tissue, the generation of a reactive species from chemicals following absorption of UV-visible light is described as a key determinant of chemicals for causing direct phototoxic reactions in an older guidance document (7)(9). Thus, the Reactive Oxygen Species (ROS) assay (10)(11) has been also included by the ICH S10 guideline as an optional initial *in chemico* screening tool for evaluating the photoreactivity of pharmaceuticals (7).

5. As an alternative method for *in vivo* phototoxicity testing, the OECD TG432 (5) describes an *in vitro* 3T3 NRU Phototoxicity Test and sets specific criteria for evaluating phototoxic hazard. The 3T3 NRU Phototoxicity Test evaluates photo-cytotoxicity by the relative reduction in viability of cells exposed to the chemical in the presence versus absence of light. Chemicals identified by this test are likely to be photoreactive, following systemic application and distribution to the skin, or after topical application. Although most of the photoirritant chemicals were correctly identified by the 3T3 NRU Phototoxicity Test, it provided false predictions for almost half of the chemicals in the photoallergens group. However, the 3T3 NRU Phototoxicity Test was not originally designed for specific prediction of chemical photoallergenicity (2). The 3T3 NRU Phototoxicity Test seems to be less reliable for photoallergenicity prediction. The photochemical assays such as ROS assay and UV/VIS spectral analysis can predict photoallergenic potential of tested chemicals, although there is still a substantial risk of false positive predictions (12).

6. Definitions used are provided in Annex A.

INITIAL CONSIDERATION AND LIMITATIONS

7. Before photosafety assessments are considered, a UV-visible light absorption spectrum of the test chemical should be determined according to OECD Test Guideline 101(13). Based on an analysis of data, the ICH S10 guideline has suggested that no further photosafety testing is needed if the MEC of a chemical is less than 1,000 $L \cdot mol^{-1} \cdot cm^{-1}$ (7). Few phototoxic chemicals showed a MEC less than 1,000 $L \cdot mol^{-1} \cdot cm^{-1}$ and these chemicals may not need to be tested in the ROS assay or any other photosafety assessments (9) (14) (15). Data collected for the limits of photoreactivity are discussed in Henry *et al.* (16) and Bauer *et al.* (17). It should be noted that phototoxicity by indirect mechanisms (e.g., pseudoporphyria or porphyria), although rare, could still occur. For compounds with MEC values of 1000 $L \cdot mol^{-1} \cdot cm^{-1}$ or higher, if the drug developer chooses to conduct a test for photoreactivity a negative result could support a decision that no further photosafety assessment is warranted.

8. The reliability and relevance of the ROS assay was recently evaluated in a multi-laboratory validation study using two different solar simulators (18)(19)(20)(21). In both solar simulators, the intra- and inter-day precisions for quinine, a positive control, were found to be above 90%, and the data suggested high inter-laboratory reproducibility (19). In a multi-laboratory validation study, the ROS assay on 2 standards and 42 coded chemicals, including 23 phototoxins and 19 non-phototoxic drugs/chemicals, provided no false negative predictions upon defined criteria as compared with the *in vitro/in vivo* phototoxicity. The, sensitivity, individual specificity, positive and negative predictivities of the ROS assay on the 42 tested chemical were calculated to be 100%, 42–82%, 75–92% and 100%, respectively. The ROS assay was designed for qualitative photoreactivity assessment of chemicals, the principle of which is

monitoring of type I (an electron or hydrogen transfer, resulting in the formation of free radical species) and type II (an energy transfer from excited triplet photosensitizer to the oxygen) photochemical reactions in test chemicals exposed to simulated sunlight (10), possibly leading to photodegradation and various phototoxic reactions, including photoirritation, photoallergy, and photogenotoxicity. Further, this assay has been optimised for detecting positive test chemicals. Test chemicals found to be negative in the ROS assay are likely to be negative in *in vivo* test systems; however, additional data may be required to determine if chemicals that are photoreactive in the ROS assay are likely to be positive *in vivo*. The test has not been designed to address indirect mechanisms of phototoxicity, such as effects of metabolites of a test chemical.

9. The applicability domain of the ROS assay is currently restricted to only those chemicals that meet the solubility criteria outlined in the protocol (see paragraph 22). Insoluble chemicals in the reaction mixtures are not suitable for testing with the ROS assay using this protocol (DMSO or NaPB solvent) but might be tested in the ROS assay with addition of solubility enhancers in the reaction mixtures (22)(23)(24). However, further characterization and standardization of procedures using these alternative vehicles should be performed by testing proficiency chemicals before incorporation into routine use. In the ROS assay, superoxide anion (SA) can be measured upon the reduction of nitroblue tetrazolium, and the determination of singlet oxygen (SO) can be made on the basis of bleaching of *p*-nitrosodimethylaniline by oxidized imidazole (11). Test chemicals that interfere with these reactions are sometimes best considered outside of the applicability domain of the ROS assay. For example, ascorbic acid and other reducing chemicals reduce the tetrazolium salt to formazan directly (25). Some skin-lightening cosmetics may also have potent reducing properties that interfere with ROS determinations. Ascorbic acid also accelerates the oxidation of imidazole derivatives (26), providing false positive prediction in the ROS assay.

10. The term "test chemical" is used in this Test Guideline to refer to what is being tested and is not related to the applicability of the ROS assay to the testing of mono-constituent chemicals, multi-constituent chemicals and/or mixtures. Based on the data currently available, the ROS assay was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, phototoxic potency (as determined in *in vivo* studies) and physicochemical properties. Limited information is currently available on the applicability of the ROS assay to multi-constituent chemicals/mixtures (27). When considering testing of mixtures, difficult to test chemical (e.g. unstable) or chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically.

PRINCIPLE OF THE TEST

11. Chemical phototoxicity can be caused by topical and systemic application of chemicals in combination with exposure to environmental light. There are several classes of chemicals that are nontoxic by themselves but could become reactive in the skin or eyes when exposed to environmental light and thereby result in toxicity. The primary event in any phototoxic reaction is the absorption of photons of a wavelength that induces excitation of the chromophore. The excitation energy is often transferred to oxygen molecules, followed by generation of ROS, including SA through type I photochemical reactions and SO through type II photochemical reactions by photo-excited molecules. These appear to be the principal intermediate species in many phototoxic responses. Direct reaction of excited chromophores with cellular constituents may also lead to phototoxicity. Therefore, while the ROS assay may not detect all ultimate mechanisms of phototoxicity, the

determination of ROS generation from chemicals irradiated with simulated sunlight is indicative of phototoxic potential.

12. In the ROS assay, SO generation is detected by spectrophotometric measurement of *p*-nitrosodimethylaniline (RNO) bleaching, followed by decreased absorbance of RNO at 440 nm (28). Although SO does not react chemically with RNO, the RNO bleaching is a consequence of SO capture by the imidazole ring, which results in the formation of a trans-annular peroxide intermediate capable of inducing the bleaching of RNO, as follows:

SO + Imidazole → [Peroxide intermediate] → Oxidized imidazole

[Peroxide intermediate] + RNO → RNO + Products

13. SA generation is detected by observing the reduction of nitroblue tetrazolium (NBT). As indicated below; NBT can be reduced by SA via a one-electron transfer reaction, yielding partially reduced (2 e^-) monoformazan (NBT^+) as a stable intermediate (29). Thus, SA can reduce NBT to NBT^+, the formation of which can be monitored spectrophotometrically at 560 nm.

$SA + NBT \rightarrow O_2 + NBT^+$

Demonstration of Proficiency

14. Prior to routine use of the test method described in this Test Guideline, laboratories should demonstrate technical proficiency, using the proficiency chemicals listed and described in Annex C. The 9 proficiency chemicals (Nos. 1–9) for the two recommended solar simulators (Suntest CPS+ or CPS and SXL-2500V2) or the 17 proficiency chemicals (Nos. 1–17) for a solar simulator other than the two recommended models are to be tested to ensure that measured values of SO and SA on all proficiency chemicals are within the range described in Annex C.

PROCEDURE

Solar simulator

15. Typically calibrated solar simulators are used because photoreactivity in the presence of natural sunlight is of concern, due to the spectral differences of global positioning and the time of day. For other circumstances where photoreactivity in response to artificial light is of interest, other sources of light may be considered. An appropriate solar simulator is to be used for irradiation of UV and visible light. The irradiation power distribution is to be kept as close to that of outdoor daylight as possible by using an appropriate filter to reduce UVC wavelengths. Recommended test conditions are as follows:

Solar simulator with filter to reduce UV wavelengths <290 nm (See Annex B)

- 1.8 to 2.2 mW/cm^2 (e.g. the indicator setting value of 250 W/m^2 for CPS+) for 1 hour,
- 6.5 to 7.9 J/cm^2 of UVA intensity (Annex B).

SXL-2500V2 (Seric) with UV filter (to reduce wavelengths <300 nm)

- 3.0 to 5.0 mW/cm^2 for 1 hour,
- 11 to 18 J/cm^2 of UVA intensity (Annex B).

16. The solar simulator is to be equipped with an appropriate temperature control or fan to stabilize the temperature during irradiation, because ROS production is affected by

temperature. Standard temperature for a solar simulator with temperature control is 25°C. The acceptable temperature range during irradiation is 20 to 29°C (20)(21).

Quartz reaction container

17. A quartz reaction container is used to avoid loss of UV due to passing through a plastic lid and vaporization of the reaction mixture (20)(21)(30). Specifications for the recommended container are provided in Annex D. If a different container is used, a lid or seal with high UV transmittance should be used. In this case, a feasibility study using the reference chemicals (Nos. 1–17) is to be conducted to determine an appropriate level of exposure to UV and visible light.

Reagents

18. All reagents should be used within 1 month after preparation and should be sonicated immediately prior to use (20)(21). Representative preparation methods are shown as follows:

20 mM sodium phosphate buffer (NaPB), pH 7.4

- Weigh 593 mg of $NaH_2PO_4 \cdot 2H_2O$ (CAS No. 13472-35-0) and 5.8 g of $Na_2HPO_4 \cdot 12H_2O$ (CAS No. 10039-32-4), add 900 mL of purified water, adjust with HCl to a pH of 7.4, dilute with purified water up to 1 L, and mix.
- Store in a refrigerator or at room temperature.

0.2 mM *p*-nitrosodimethylaniline (RNO, CAS No. 138-89-6)

- Dissolve 3 mg of RNO in 100 mL of 20 mM NaPB.
- Store in a refrigerator and protect from light.

0.2 mM imidazole (CAS No. 288-32-4)

- Dissolve 13.6 mg of imidazole in 10 mL of 20 mM NaPB.
- Dilute the 20 mM imidazole solution 100 times with 20 mM NaPB.
- Store in a refrigerator and protect from light.

0.4 mM nitroblue tetrazolium chloride (NBT, CAS No. 298-83-9)

- Dissolve 32.7 mg of NBT in 100 mL of 20 mM NaPB.
- Store in a refrigerator and protect from light.

Solvents

19. Use analytical grade DMSO at first. For chemicals that are not soluble in DMSO, 20 mM NaPB is to be used as a solvent. Some chemicals react with DMSO and test chemical stability in DMSO should be determined. If the test chemical is not soluble or stable in DMSO or NaPB, other solvents may be used. However the test chemical must be demonstrated to be stable in the selected solvent, and SO and SA ranges for proficiency chemicals must fall within the ranges defined in Annex C.

Test chemicals

20. Test chemicals must be prepared fresh, immediately prior to use unless data demonstrate their stability in storage. It is recommended that all chemical handling and the initial treatment of cells be performed under light conditions that would avoid photoactivation or degradation of the test chemical prior to irradiation. Chemicals should be tested at 200 µM (final concentration). A 20-µM concentration can be used if precipitation occurs before light exposure, coloration, or other interference is observed in the reaction mixture at 200 µM. A positive result at 20 µM can be used to indicate

photoreactivity; however, a negative result at the lower 20 µM concentration is not indicative of absence of photoreactivity. The molecular weight of the test chemical must be available.

21. The test chemical solutions are to be prepared immediately before use in a solvent as described in paragraph 19. Each test chemical is to be weighed in a tube, and solvent added to achieve a 10 mM concentration of the test chemical (20)(21). The tube is to be mixed with a vortex mixer and sonicated for 5 to 10 minutes. All preparations are to be protected from strong UV and intense visible light (e.g. direct overhead light, working near windows exposed to natural light) at all times during preparation. When precipitation before light exposure or other interference is observed in the reaction mixture at 200 µM, a 1-mM solution (20 µM as the final concentration) is to be prepared by dilution of the stock solution of chemicals at 10 mM using DMSO. For chemicals that are not soluble in DMSO, 20 µL of DMSO (2 v/v%) is to be contained in the reaction mixture.

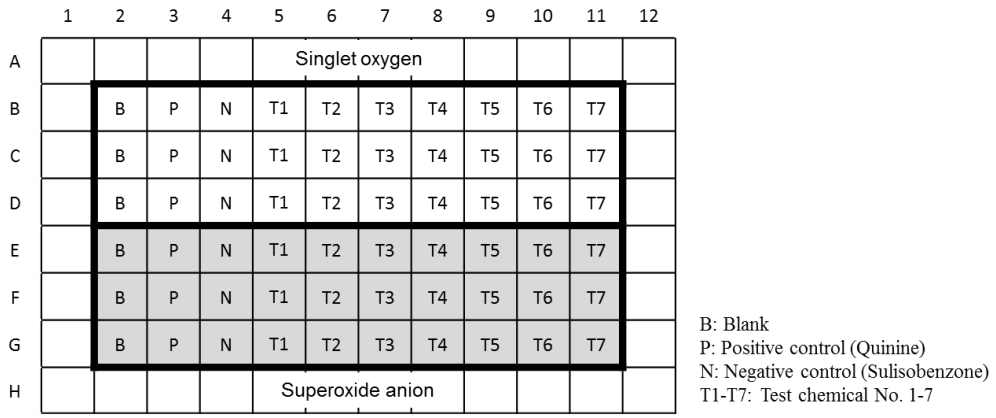
Positive and negative controls

22. Stock solutions of quinine hydrochloride (a positive control, CAS No. 6119-47-7) and sulisobenzone (a negative control, CAS No. 4065-45-6) are to be prepared at 10 mM each in DMSO (final concentration of 200 µM) according to the above procedure, divided into tubes, and stored in a freezer (generally below -20°C) for up to 1 month. The stock solution is to be thawed just before the experiment and used within the day.

Test procedure

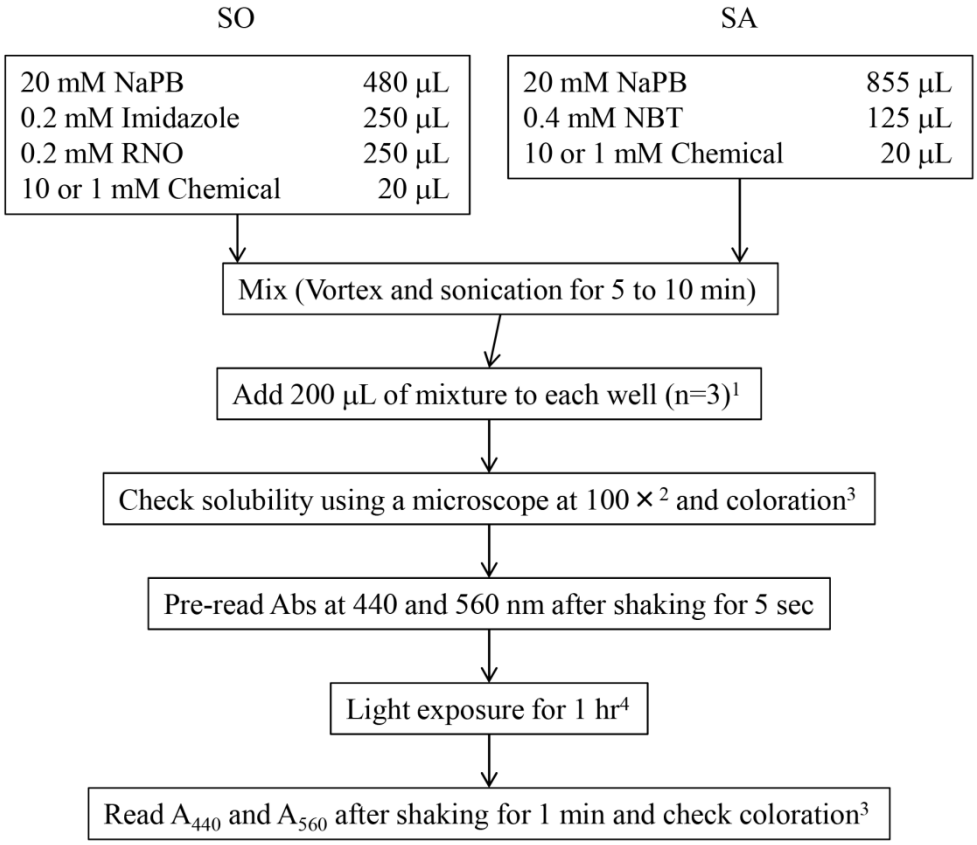
23. A typical 96-well plate configuration is as follows, but other configurations are also acceptable:

Figure 1. An example of a typical plate configuration



24. A tube (e.g. 1.5 mL micro tube) and a plastic clear flat bottomed 96-well microplate are to be used. The reaction mixture is to be prepared by vortex mixing and/or sonication under UV-cut illumination or shade. The same volume of DMSO, 20 µL, is to be added in a vehicle control instead of test chemical solution.

Figure 2. Workflow diagram if the stock solution of the test chemical is prepared in DMSO.



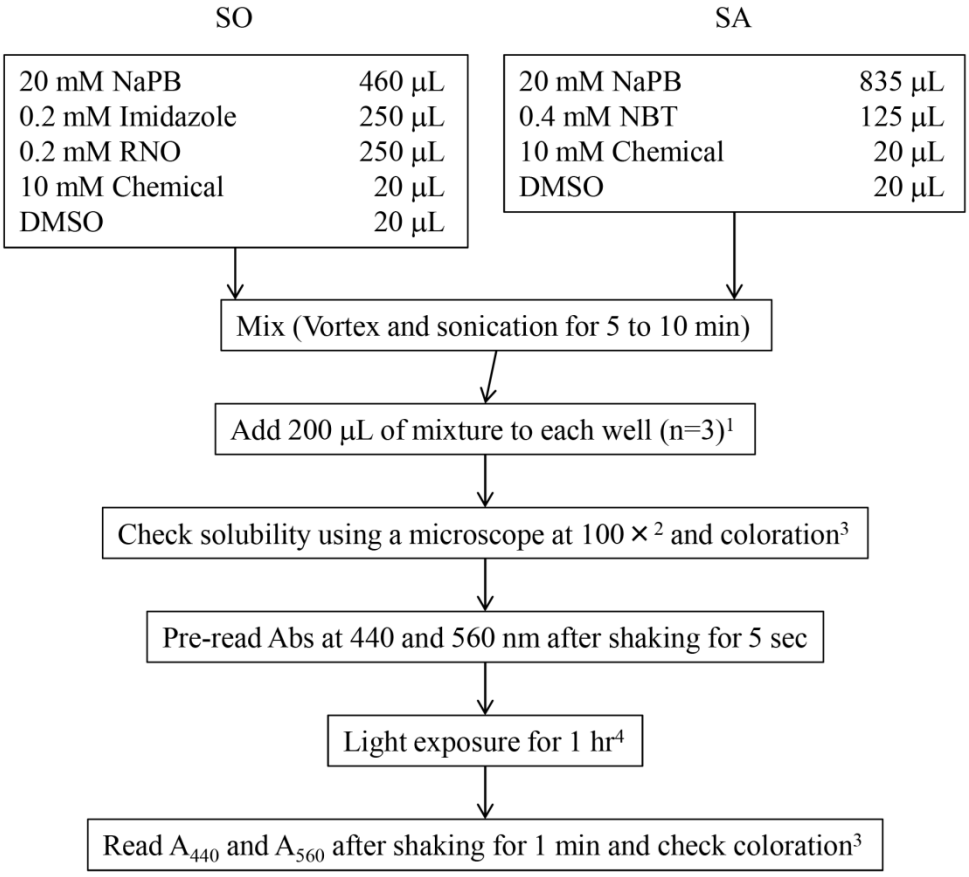
[1] Avoid using peripheral wells. More than one test chemical can be tested on a plate.

[2] Some chemicals might precipitate in the reaction mixture. It is therefore important to check solubility prior to irradiation. Solubility of each reaction mixture in its well is to be observed with a microscope prior to irradiation. Test chemical concentrations are to be selected so as to avoid precipitation or cloudy solutions.

[3] The reaction mixture is to be checked for coloration with the naked eye.

[4] The 96-well plate is to be placed in the quartz reaction container. A quartz cover is to be set on the plate and fastened with bolts. Ensure that temperature and other ambient conditions are stable when using the solar simulator. Measure UVA intensity and temperature at the plate position using a UVA detector and thermometer both before and after irradiation.

Figure 3. Workflow diagram if the stock solution of the test chemical is prepared in 20 mM NaPB.



[1] Avoid using peripheral wells. More than one test chemical can be tested on a plate.

[2] Some chemicals might precipitate in the reaction mixture. It is therefore important to check solubility prior to irradiation. Solubility of each reaction mixture in its well is to be observed with a microscope prior to irradiation. Test chemical concentrations are to be selected so as to avoid precipitation or cloudy solutions.

[3] The reaction mixture is to be checked for coloration with the naked eye.

[4] The 96-well plate is to be placed in the quartz reaction container. A quartz cover is to be set on the plate and fastened with bolts. Ensure that temperature and other ambient conditions are stable when using the solar simulator. Measure UVA intensity and temperature at the plate position using a UVA detector and thermometer both before and after irradiation.

DATA AND REPORTING

Data analysis

25. Data from three wells for each chemical concentration is used to calculate mean and standard deviation.

SO

Decrease of A_{440} ×1000 = [A_{440} (–) – A_{440} (+) – (a – b)] ×1000
A_{440} (–): Absorbance before light exposure at 440 nm
A_{440} (+): Absorbance after light exposure at 440 nm
a: Vehicle control before light exposure (mean)
b: Vehicle control after exposure (mean)

SA

Increase of A_{560} ×1000 = [A_{560} (+) – A_{560} (–) – (b – a)] ×1000
A_{560} (–): Absorbance before light exposure at 560 nm
A_{560} (+): Absorbance after light exposure at 560 nm
a: Vehicle control before light exposure (mean)
b: Vehicle control after exposure (mean)

Criteria for data acceptance

26. The following criteria are to be satisfied in each experiment.

- No precipitation of test chemical in the reaction mixture before light exposure.
- No color interference by test chemical in the reaction mixture before or after light exposure.
- No technical problems, including temperature range (20–29°C), when collecting data set.
- The ranges of raw A_{440} and A_{560} values: 0.02 to 1.5.
- Historical positive and negative control values are to be developed by each laboratory based on a mean +/-2 SD. The following range was defined based on the 95% confidence interval (mean +/- 1.96SD) obtained from the validation data. When a solar simulator other than a recommended model is used, establish modified criteria based on 95% confidence interval.
-

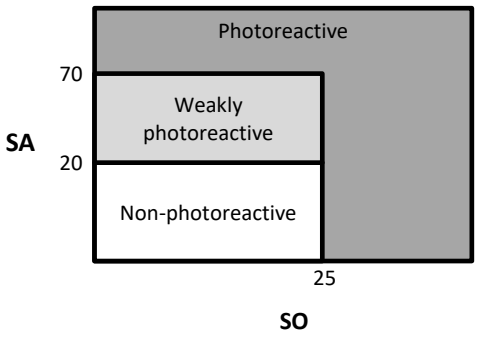
  Positive control (quinine hydrochloride) value at 200 µM (mean of 3 wells)
  SO: 319 to 583
  SA: 193 to 385
  Negative control (sulisobenzone) value at 200 µM (mean of 3 wells)
  SO: -9 to 11
  SA: -20 to 2
- Laboratories should demonstrate technical proficiency, as described in Annex C, prior to routine use of the test method described in this Test Guideline.

Criteria for judgment

27. Each test chemical is to be judged as follows:

ROS assay prediction model

Judgment[1, 2]	Concentration[3]	SO (mean of 3 wells)[6]		SA (mean of 3 wells)[6]
Photoreactive	200 μM	≥25	and	≥70
		<25 and/or I[4]	and	≥70
		≥25	and	<70 and/or I[4]
Weakly photoreactive	200 μM	<25	and	≥20, <70
Photoreactive	20 μM	≥25	and	≥20
Non-photoreactive	200 μM	<25	and	<20
Inconclusive	The results do not meet any of the above-mentioned criteria.[5]			



[1] A single experiment is sufficient for judging results, because the ROS assay shows good intra- and inter-laboratory reproducibility in the validation studies.

[2] If precipitation, coloration, or other interference is observed at both 20 and 200 μM, the chemical is considered incompatible with the ROS assay and judged as inconclusive.

[3] 20 μM can be used for judgment when precipitation or coloration is observed at 200 μM. A positive results at 20 μM can be used to indicate photoreactivity; however, a negative result at the lower 20 μM concentration is not indicative of absence of photoreactivity.

[4] Interference such as precipitation or coloration.

[5] Positive prediction can be made on the basis of SO only, SA only, or both; however, both SO and SA values should be obtained for reliable negative prediction.

[6] Classification criteria defined in published manuscripts. (11)(20)(21)

Data quality

28. Studies for regulatory purposes are to be conducted to the highest of quality standards, with data collection records readily available, in compliance with GLP regulations whenever possible, and all documents checked by the Quality Assurance Unit of the laboratory.

Test report

29. The test report should include the following information:

Test chemical:

- identification data, common generic names and IUPAC and CAS number, if known;
- physical nature and purity;
- physicochemical properties relevant to conduct of the study;
- UV/vis absorption spectrum;
- stability and photostability, if known.

Control chemicals:

- name, manufacturer, and lot No.;
- physical nature and purity;
- storage condition;
- preparation of control chemical solutions;
- final concentrations tested.

Solvent:

- name, manufacturer, and lot No.;
- justification for choice of solvent;
- solubility of the test chemical in solvent.

Irradiation condition:

- manufacturer and type of the solar simulator used;
- rationale for selection of the solar simulator used;
- UVA detector used;
- UVA irradiance, expressed in mW/cm^2
- UVA dose, expressed in J/cm^2;
- temperature before and after irradiation.

ROS assay procedure.

Acceptance and decision criteria.

Results.

Discussion.

Conclusions.

LITERATURE

(1) Moore, D. E. (2002). Drug-induced cutaneous photosensitivity: incidence, mechanism, prevention and management. *Drug Saf*, **25**, 345-72

(2) Onoue, S., Seto, Y., Sato, H., Nishida, H., Hirota, M., Ashikaga, T., Api, A. M., Basketter, D. and Tokura, Y. (2017). Chemical photoallergy: photobiochemical mechanisms, classification, and risk assessments. *J Dermatol Sci*, **85**, 4-11

(3) The European Agency for the Evaluation of Medicinal Products, Evaluation of Medicines for Human Use, Committee for Proprietary Medicinal Products (EMEA/CPMP) (2002). Note for Guidance on Photosafety Testing, CPMP/SWP/398/01.

(4) The Food and Drug Administration, Center for Drug Evaluation and Research (FDA/CDER) (2002). Guidance for Industry, Photosafety Testing.

(5) The Organisation for Economic Co-operation and Development (OECD) (2004). OECD guideline for testing of chemicals, 432, *In vitro* 3T3 NRU phototoxicity test.

(6) The European Agency for the Evaluation of Medicinal Products, Evaluation of Medicines for Human Use, Committee for Proprietary Medicinal Products (EMEA/CPMP) (2008). Concept Paper on the Need for Revision of the Note for Guidance on Photosafety testing, CPMP/SWP/398/01.

(7) International Council on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) (2014). ICH guideline S10 Guidance on photosafety evaluation of pharmaceuticals.

(8) *Radiation. Encyclopedia Britannica Online. Retrieved 2009-11-09.*

(9) The Organisation for Economic Co-operation and Development (OECD) (1997). Environmnental Health and Safety Publications, Series on Testing and Assessment No.7 "Guidance Document on Direct Phototransformation of Chemicals in Water" Environment Directorate.

(10) Onoue, S. and Tsuda, Y. (2006). Analytical studies on the prediction of photosensitive/phototoxic potential of pharmaceutical substances. *Pharm Res*, **23**, 156-64

(11) Onoue, S., Kawamura, K., Igarashi, N., Zhou, Y., Fujikawa, M., Yamada, H., Tsuda, Y., Seto, Y. and Yamada, S. (2008). Reactive oxygen species assay-based risk assessment of drug-induced phototoxicity: classification criteria and application to drug candidates. *J Pharm Biomed Anal*, **47**, 967-72

(12) Onoue, S., Ohtake, H., Suzuki, G., Seto, Y., Nishida, H., Hirota, M., Ashikaga, T. and Kouzuki, H. (2016). Comparative study on prediction performance of photosafety testing tools on photoallergens. *Toxicol In Vitro*, **33**, 147-52

(13) The Organisation for Economic Co-operation and Development (OECD) (1981). OECD guideline for testing of chemicals, 101, UV-VIS absorption spectra.

(14) Onoue, S., Suzuki, G., Kato, M., Hirota, M., Nishida, H., Kitagaki, M., Kouzuki, H. and Yamada, S. (2013). Non-animal photosafety assessment approaches for cosmetics based on the photochemical and photobiochemical properties. *Toxicol In Vitro*, **27**, 2316-24

(15) Lovell, W. W. (1993). A scheme for in vitro screening of substances for photoallergenic potential. *Toxicol In Vitro*, **7**, 95-102

(16) Henry, B., Foti, C. and Alsante, K. (2009). Can light absorption and photostability data be used to assess the photosafety risks in patients for a new drug molecule? *J Photochem Photobiol B*, **96**, 57-62

(17) Bauer, D., Averett, L. A., De Smedt, A., Kleinman, M. H., Muster, W., Pettersen, B. A. and Robles, C. (2014). Standardized UV-vis spectra as the foundation for a threshold-based, integrated photosafety evaluation. *Regul Toxicol Pharmacol*, **68**, 70-5

(18) Onoue, S., Hosoi, K., Wakuri, S., Iwase, Y., Yamamoto, T., Matsuoka, N., Nakamura, K., Toda, T., Takagi, H., Osaki, N., Matsumoto, Y., Kawakami, S., Seto, Y., Kato, M., Yamada, S., Ohno, Y. and Kojima, H. (2013). Establishment and intra-/inter-laboratory validation of a standard

protocol of reactive oxygen species assay for chemical photosafety evaluation. *J Appl Toxicol*, **33**, 1241-1250

(19) Onoue, S., Hosoi, K., Toda, T., Takagi, H., Osaki, N., Matsumoto, Y., Kawakami, S., Wakuri, S., Iwase, Y., Yamamoto, T., Nakamura, K., Ohno, Y. and Kojima, H. (2014). Intra-/inter-laboratory validation study on reactive oxygen species assay for chemical photosafety evaluation using two different solar simulators. *Toxicol In Vitro*, **28**, 515-23

(20) ROS assay validation management team (2013). Validation report for the international validation study on ROS (Reactive Oxygen Species) assay as a test evaluating phototoxic potential of chemicals (Atlas Suntest version), http://www.jacvam.jp/files/news/ROS_Assay_full%2020130920%20atlas_fourth%20data.pdf (accessed May, 2015).

(21) ROS assay validation management team (2013). Validation report for the international validation study on ROS (Reactive Oxygen Species) assay as a test evaluating phototoxic potential of chemicals (Seric version), http://www.jacvam.jp/files/news/ROS_Assay_full%2020131009Seric%20fourth.pdf (accessed May, 2015).

(22) Onoue, S., Yamauchi, Y., Kojima, T., Igarashi, N. and Tsuda, Y. (2008). Analytical studies on photochemical behavior of phototoxic substances; effect of detergent additives on singlet oxygen generation. *Pharm Res*, **25**, 861-8

(23) Seto, Y., Kato, M., Yamada, S. and Onoue, S. (2013). Development of micellar reactive oxygen species assay for photosafety evaluation of poorly water-soluble chemicals. *Toxicol In Vitro*, **27**, 1838-46

(24) Onoue, S., Kato, M. and Yamada, S. (2014). Development of an albuminous reactive oxygen species assay for photosafety evaluation under experimental biomimetic conditions. *J Appl Toxicol*, **34**, 158-65

(25) Bruggisser, R., von Daeniken, K., Jundt, G., Schaffner, W. and Tullberg-Reinert, H. (2002). Interference of plant extracts, phytoestrogens and antioxidants with the MTT tetrazolium assay. *Planta Med*, **68**, 445-8

(26) Imanaga, Y. (1955). Autooxidation of L-ascorbic acid and imidazole nucleus. *J Biochem*, **42**, 657-67

(27) Nishida, H., Hirota, M., Seto, Y., Suzuki, G., Kato, M., Kitagaki, M., Sugiyama, M., Kouzuki, H. and Onoue, S. (2015). Non-animal photosafety screening for complex cosmetic ingredients with photochemical and photobiochemical assessment tools. *Regul Toxicol Pharmacol*, **72**, 578-85

(28) Kraljic, I. and Mohsni, S. E. (1978). A new method for the detection of singlet oxygen in aqueous solutions. *Photochem Photobiol*, **28**, 577-81

(29) Pathak, M. A. and Joshi, P. C. (1984). Production of active oxygen species ($^1O_2$ and O_2^-.) by psoralens and ultraviolet radiation (320-400 nm). *Biochim Biophys Acta*, **798**, 115-26

(30) Onoue, S., Igarashi, N., Yamada, S. and Tsuda, Y. (2008). High-throughput reactive oxygen species (ROS) assay: an enabling technology for screening the phototoxic potential of pharmaceutical substances. *J Pharm Biomed Anal*, **46**, 187-93

Annex A. Definitions

3T3 NRU Phototoxicity Test: *In vitro* 3T3 neutral red uptake phototoxicity test.

Irradiance: The intensity of UV or visible light incident on a surface, measured in W/m^2 or mW/cm^2.

Dose of light: The quantity [= intensity × time (seconds)] of UV or visible light incident on a surface, expressed in J/m^2 or J/cm^2.

MEC: Molar Extinction Coefficient (also called molar absorptivity) is a constant for any given molecule under a specific set of conditions (e.g. solvent, temperature, and wavelength) and reflects the efficiency with which a molecule can absorb a photon (typically expressed as $L \cdot mol^{-1} \cdot cm^{-1}$).

Photoreactivity: The property of chemicals that react with another molecule as a consequence of absorption of photons.

Phototoxicity: Toxic responses that can be elicited after the exposure of skin to certain chemicals and subsequent exposure to light, or that is induced similarly by skin irradiation after systemic administration of a chemical.

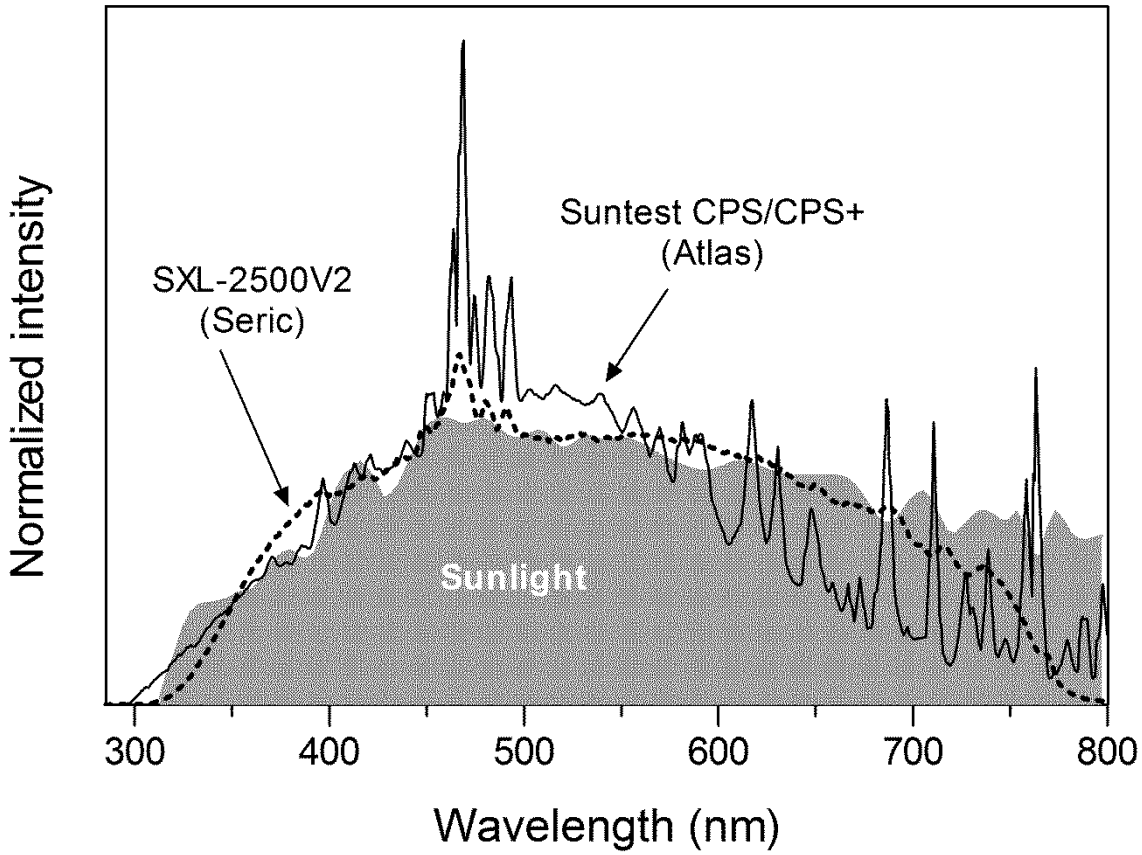
ROS: Reactive Oxygen Species, including superoxide anion (SA) and singlet oxygen (SO).

SA: Superoxide anion is one of radical species, generated from photo-irradiated chemicals through type I photochemical reaction.

SO: Singlet oxygen is one of radical species, generated from photo-irradiated chemicals through type II photochemical reaction.

UV light wavebands: The designations recommended by the CIE (Commission Internationale de L'Eclairage) are: UVA (315–400 nm) UVB (280–315 nm) and UVC (100–280 nm). Other designations are also used; the division between UVB and UVA is often placed at 320 nm, and the UVA may be divided into UV-A1 and UV-A2 with a division made at about 340 nm.

Annex B. Spectrum of solar stimulators used in the validation studies.



Annex C. Proficiency Chemicals

Prior to routine use of the test method described in this Test Guideline, laboratories should demonstrate technical proficiency by correctly obtaining the expected ROS prediction for proficiency chemicals recommended in the Table. For Suntest CPS/CPS+ (Atlas) or SXL-2500V2 (Seric) solar simulators, nine chemicals (Nos. 1–9) are to be tested. For other solar simulators, all 17 chemicals (Nos. 1–17) are to be tested. These proficiency chemicals were selected to represent the range of responses for phototoxic potential. Other selection criteria were that they are commercially available, that high quality *in vivo* reference data and high quality *in vitro* data generated with the ROS assay are available, and that they were used in the JaCVAM-coordinated validation study to demonstrate successful implementation of the test method in the laboratories participating in the study (20)(21).

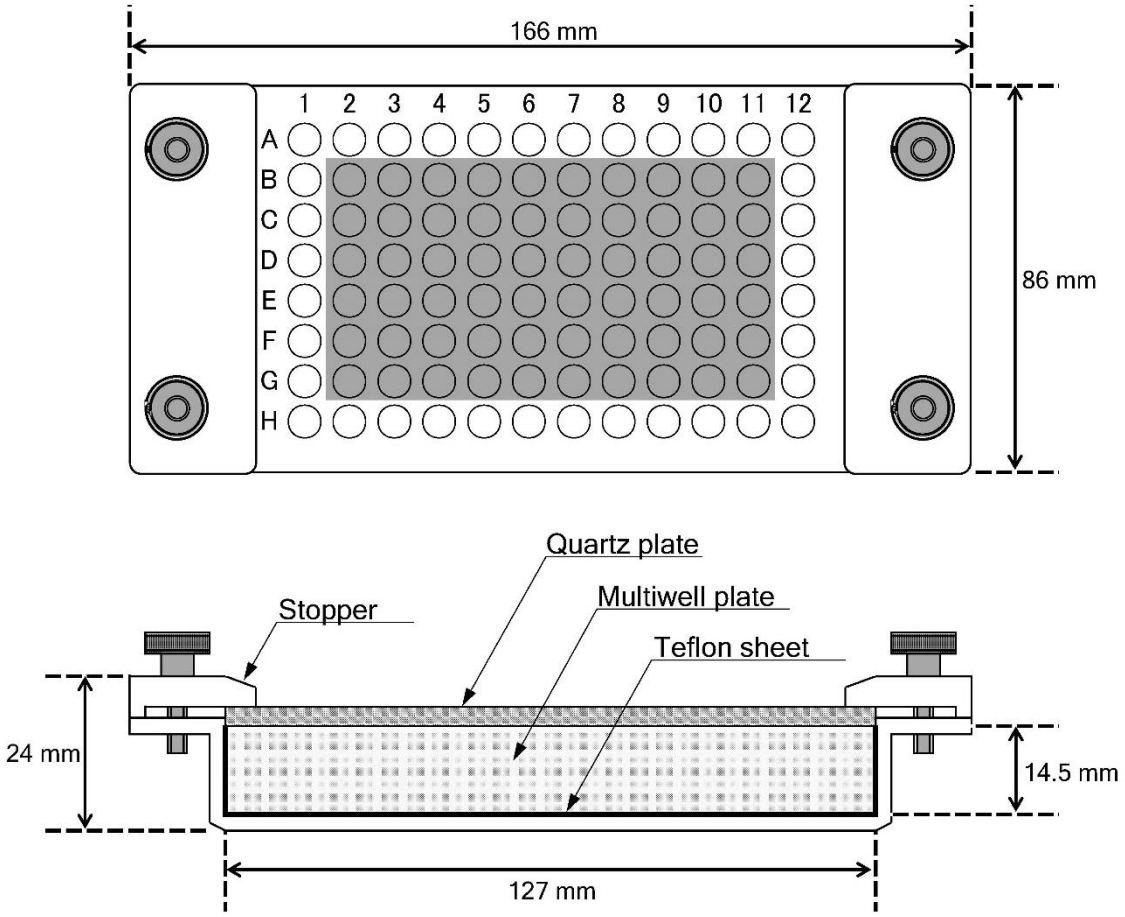
Table A C.1. Table of proficiency chemicals.

The expected ROS prediction for proficiency chemicals and the acceptable range..

No.	Chemical[1]	CAS No.	SO [2]	SA [2]	Solvent	Concentration
1	*p*-Aminobenzoic acid	150-13-0	-8 to 12	-11 to 7	DMSO	200 µM
2	Benzocaine	94-09-7	-7 to 9	-7 to 17	DMSO	200 µM
3	Doxycycline hydrochloride	10592-13-9	115 to 429	230 to 468	DMSO	200 µM
4	Erythromycin	114-07-8	-15 to 11	-9 to 21	DMSO	200 µM
5	Fenofibrate	49562-28-9	77 to 203	-31 to 11	DMSO	20 µM
6	L-Histidine	71-00-1	-8 to 12	8 to 120	NaPB	200 µM
7	Norfloxacin	70458-96-7	131 to 271	57 to 161	DMSO	200 µM
8	8-Methoxy psoralen	298-81-7	31 to 137	0 to 126	DMSO	200 µM
9	Octyl salicylate	118-60-5	-5 to 11	-8 to 20	DMSO	20 µM
10	Acridine	260-94-6	182 to 328	121 to 243	DMSO	200 µM
11	Chlorpromazine hydrochloride	69-09-0	-56 to 70	66 to 106	DMSO	200 µM
12	Diclofenac	15307-79-6	34 to 416	47 to 437	DMSO	200 µM
13	Furosemide	54-31-9	31 to 225	-7 to 109	DMSO	200 µM
14	Ketoprofen	22071-15-4	120 to 346	77 to 151	DMSO	200 µM
15	Nalidixic acid	389-08-2	54 to 246	88 to 470	DMSO	200 µM
16	Omeprazole	73590-58-6	-221 to 103	30 to 216	DMSO	200 µM
17	Promethazine hydrochloride	58-33-3	20 to 168	-3 to 77	DMSO	200 µM

[1]All chemicals are solid
[2]The values were calculated as means +/- 1.96 SD from the validation data..

Annex D. Quartz reaction container used in the validation studies.



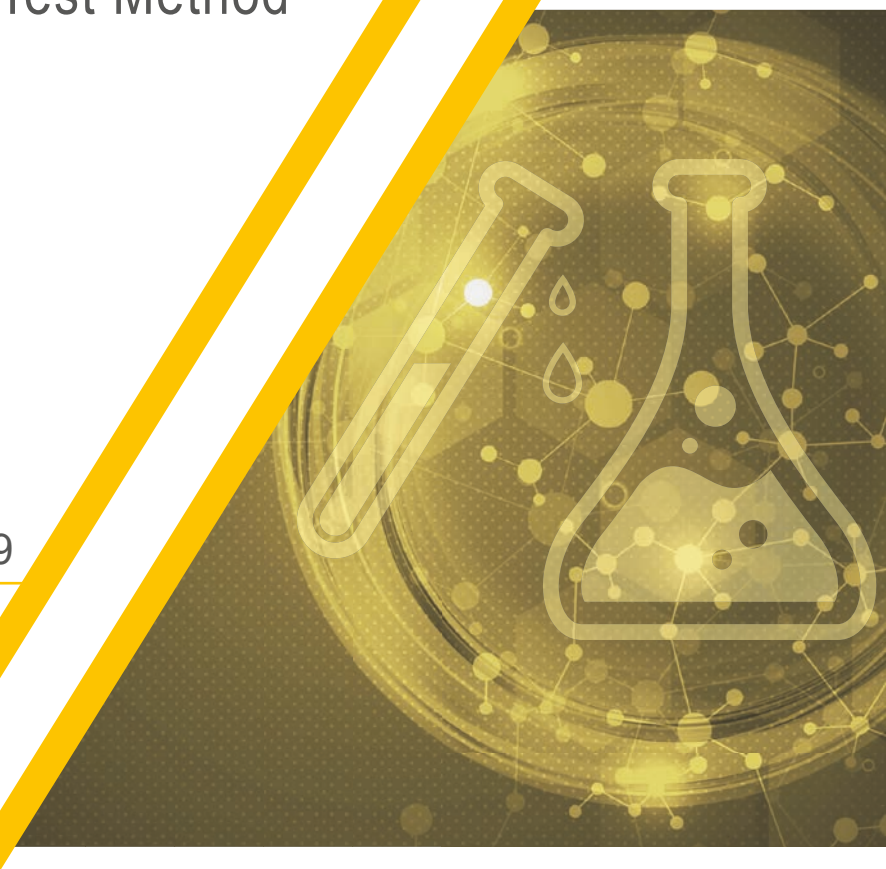
Recommended thickness of quartz plate: ca. 3 mm.

添付資料３

Section 4
Health effects

Test Guideline No. 431

In Vitro Skin Corrosion: Reconstructed Human Epidermis (RhE)Test Method



18 June 2019

OECD Guidelines for the Testing of Chemicals

OECD GUIDELINE FOR TESTING OF CHEMICALS

In Vitro Skin Corrosion: Reconstructed Human Epidermis (RhE) Test Method

INTRODUCTION

1. Skin corrosion refers to the production of irreversible damage to the skin manifested as visible necrosis through the epidermis and into the dermis, following the application of a test chemical [as defined by the United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS)] (1). This updated Test Guideline 431 provides an in vitro procedure allowing the identification of non-corrosive and corrosive substances and mixtures in accordance with UN GHS (1). It also allows a partial sub-categorisation of corrosives.

2. The assessment of skin corrosion potential of chemicals has typically involved the use of laboratory animals (OECD Test Guideline 404 (TG 404); originally adopted in 1981 and revised in 1992, 2002 and 2015) (2). In addition to the present TG 431, two other in vitro test methods for testing corrosion potential of chemicals have been validated and adopted as OECD Test Guidelines 430 (3) and 435 (4). Furthermore the in vitro OECD TG 439 (5) has been adopted for testing skin irritation potential. A document on Integrated Approaches to Testing and Assessment (IATA) for Skin Corrosion and Irritation describes several modules which group information sources and analysis tools, and provides guidance on (i) how to integrate and use existing testing and non-testing data for the assessment of skin irritation and skin corrosion potentials of chemicals and (ii) proposes an approach when further testing is needed (6).

3. This Test Guideline addresses the human health endpoint skin corrosion. It makes use of reconstructed human epidermis (RhE) (obtained from human derived non-transformed epidermal keratinocytes) which closely mimics the histological, morphological, biochemical and physiological properties of the upper parts of the human skin, i.e. the epidermis. This Test Guideline was originally adopted in 2004 and updated in 2013, 2016 and 2019 to include additional test methods using the RhE models. The Test Guideline was also updated in 2015 to introduce the possibility to use the methods to support the sub-categorisation of corrosive chemicals, and to refer to the IATA guidance document, and introduce the use of an alternative procedure to measure viability.

4. Five validated test methods using commercially available RhE models are included in this Test Guideline, as described below. Prevalidation studies (7), followed by a formal validation study for assessing skin corrosion (8) (9) (10) have been conducted (11) (12) for two of these commercially available test methods, EpiSkin™ Standard Model (SM), and EpiDerm™ Skin Corrosivity Test (SCT) (EPI-200) (referred to in the following text as the Validated Reference Methods – VRMs, EpiSkin^{TM}=VRM1, EpiDerm^{TM}= VRM2). The outcome of these studies led to the recommendation that the two VRMs mentioned above could be used for regulatory purposes for distinguishing corrosive (C) from non-corrosive (NC) substances, and that the EpiSkin™ could moreover be used to support sub-categorisation of corrosive substances (13) (14) (15). Two other commercially available in vitro skin corrosion RhE test methods have subsequently shown similar results to the EpiDerm™ SCT according to PS-based Validation (16) (17) (18). These are the SkinEthicTM RHE1 and epiCS® (previously named EST-1000) that can also be used for regulatory purposes for distinguishing corrosive from non-corrosive substances (19) (20). Post validation studies performed by the RhE model producers in the years 2012 to 2014 with a refined protocol correcting interferences of unspecific MTT reduction by the test chemicals improved the performance of both discrimination of C/NC as well as supporting sub-categorization of corrosives (21) (22). Further statistical analyses of the post-validation data generated with Epiderm™ SCT, SkinEthic™ RHE and epiCS® have been performed to identify alternative predictions models that improved the predictive capacity for sub-categorisation (23). Finally, the LabCyte EPI-MODEL24 is another commercially available in vitro skin corrosion RhE test that was shown to be scientific similar to the VRMs and can therefore be used for regulatory purposes to distinguish corrosive from non-corrosive substances as well as support sub-categorization of corrosives (40) (41) (42)(43).

5. Before a proposed similar or modified in vitro RhE test method for skin corrosion other than the VRMs can be used for regulatory purposes, its reliability, relevance (accuracy), and limitations for its proposed use should be determined to ensure its similarity to the VRMs, in accordance with the requirements of the Performance Standards (PS) (24) set out in accordance with the principles of Guidance Document No.34 (25). The Mutual Acceptance of Data will only be guaranteed after any proposed new or updated test method following the PS have been reviewed and included in this Test Guideline. The test methods included in this Test Guideline can be used to address countries' requirements for test results on in vitro test method for skin corrosion, while benefiting from the Mutual Acceptance of Data.

DEFINITIONS

6. Definitions used are provided in Annex I.

INITIAL CONSIDERATIONS

7. This Test Guideline allows the identification of non-corrosive and corrosive substances and mixtures in accordance with the UN GHS (1). This Test Guideline further supports the sub-categorisation of corrosive substances and mixtures into optional Sub-category 1A, in accordance with the UN GHS (1), as well as a combination of Sub-categories 1B and 1C (21) (22) (23). A limitation of this Test Guideline is that it does not allow discriminating between skin corrosive Sub-category 1B and Sub-category 1C in accordance with the UN GHS (1) due to the limited set of well-known in vivo corrosive Sub-category 1C chemicals. The five test methods under this test guideline are able to discriminate sub-categories 1A versus 1B-and-1C versus NC.

8. A wide range of chemicals representing mainly individual substances has been tested in the validation studies supporting the test methods included in this Test Guideline. The original database of the validation study conducted for identification of non-corrosives versus corrosives amounted to 60 chemicals covering a wide range of chemical classes (8) (9) (10). Testing to demonstrate sensitivity, specificity, accuracy and within-laboratory-reproducibility of the assay for sub-categorisation was further performed by the test method developers using 79 to 80 chemicals also covering a wide range of chemical classes, and results were reviewed by the OECD (21) (22) (23). On the basis of the overall data available, the Test Guideline is applicable to a wide range of chemical classes and physical states including liquids, semi-solids, solids and waxes. The liquids may be aqueous or non-aqueous; solids may be soluble or insoluble in water. Whenever possible, solids should be ground to a fine powder before application; no other prior treatment of the sample is required. In cases where evidence can be demonstrated on the non-applicability of test methods included in the Test Guideline to a specific category of test chemicals, these test methods should not be used for that specific category of test chemicals. In addition, this Test Guideline is assumed to be applicable to mixtures as an extension of its applicability to substances. However, due to the fact that mixtures cover a wide spectrum of categories and composition, and that only limited information is currently available on the testing of mixtures, in cases where evidence can be demonstrated on the non-applicability of the Test Guideline to a specific category of mixtures (e.g. following a strategy as proposed in (26)), the Test Guideline should not be used for that specific category of mixtures. When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically. Such considerations are not needed, when there is a regulatory requirement for testing of the mixture. Gases and aerosols have not been assessed yet in validation studies (8) (9) (10). While it is conceivable that these can be tested using RhE technology, the current Test Guideline does not allow testing of gases and aerosols.

9. Test chemicals absorbing light in the same range as MTT formazan and test chemicals able to directly reduce the vital dye MTT (to MTT formazan) may interfere with the tissue viability measurements and need the use of adapted controls for corrections. The type of adapted controls that may be required will vary depending on the type of interference produced by the test chemical and the procedure used to measure MTT formazan (see paragraphs 25-31).

10. While this Test Guideline does not provide adequate information on skin irritation, it should be noted that OECD TG 439 specifically addresses the health effect skin irritation in vitro and is based on the same RhE test system, though using another protocol (5). For a full evaluation of local skin effects after a single dermal exposure, the Guidance Document No. 203 on Integrated Approaches for Testing Assessment should be consulted (6). This IATA approach includes the conduct of in vitro tests for skin corrosion (such as described in this Test Guideline) and skin irritation before considering testing in living animals. It is recognized that the use of human skin is subject to national and international ethical considerations and conditions.

PRINCIPLE OF THE TEST

11. The test chemical is applied topically to a three-dimensional RhE model, comprised of non-transformed, human-derived epidermal keratinocytes, which have been cultured to

form a multi-layered, highly differentiated model of the human epidermis. It consists of organized basal, spinous and granular layers, and a multi-layered stratum corneum containing intercellular lamellar lipid layers representing main lipid classes analogous to those found in vivo.

12. The RhE test method is based on the premise that corrosive chemicals are able to penetrate the stratum corneum by diffusion or erosion, and are cytotoxic to the cells in the underlying layers. Cell viability is measured by enzymatic conversion of the vital dye MTT [3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue tetrazolium bromide; CAS number 298-93-1], into a blue formazan salt that is quantitatively measured after extraction from tissues (27). Corrosive chemicals are identified by their ability to decrease cell viability below defined threshold levels (see paragraphs 35 and 36). The RhE-based skin corrosion test methods have shown to be predictive of in vivo skin corrosion effects assessed in rabbits according to the OECD guideline 404 (2).

DEMONSTRATION OF PROFICIENCY

13. Prior to routine use of any of the five validated RhE test methods that adhere to this Test Guideline, laboratories should demonstrate technical proficiency by correctly classifying the twelve Proficiency Substances listed in Table 1. In case of the use of a method for sub-classification, also the correct sub-categorisation should be demonstrated. In situations where a listed substance is unavailable or where justifiable, another substance for which adequate in vivo and in vitro reference data are available may be used (e.g. from the list of reference chemicals (24)) provided that the same selection criteria as described in Table 1 are applied.

Table 1. List of Proficiency Substances[1]

Substance	CASRN	Chemical Class[2]	UN GHS Cat. Based on *In Vivo* results [3]	Cat. Based on *In Vitro* results[4]	Mean cell viability for VRMs				Physical State
					VRM1		VRM2		
					3 min	60 min.	3 min.	60 min	
Sub-category 1A *In Vivo* Corrosives									
Bromoacetic acid	79-08-3	Organic acid	1A	(3) 1A	3	2.8	3.2	2.8	S
Boron trifluoride dihydrate	13319-75-01	Inorganic acid	1A	(3) 1A	2.4	4.2	4.4	10.1	L
Phenol	108-95-2	Phenol	1A	(3) 1A	29.8	21.8	22.6	13.5	S
Dichloroacetyl chloride	79-36-7	Electrophile	1A	(3) 1A	5.6	6.3	1.3	1.4	L
Combination of sub-categories 1B-and-1C *In Vivo* Corrosives									
Glyoxylic acid monohydrate	563-96-2	Organic acid	1B-and-1C	(3) 1B-and-1C	110.4	22.5	90.4	3.1	S
Lactic acid	598-82-3	Organic acid	1B-and-1C	(3) 1B-and-1C	80.2	9.4	90	3.5	L
Ethanolamine	141-43-5	Organic base	1B	(3) 1B-and-1C	66.2	40.3	69.7	9.3	Viscous
Hydrochloric acid (14.4%)	7647-01-0	Inorganic acid	1B-and-1C	(3) 1B-and-1C	69.3	5.7	80.8	9	L

***In Vivo* Non Corrosives**									
Phenethyl bromide	103-63-9	Electrophile	NC	(3) NC	141	117.2	112.5	71.2	N
4-Amino-1,2,4-triazole	584-13-4	Organic base	NC	(3) NC	116.8	120.6	105.7	88.2	N
4-(methylthio)-benzaldehyde	3446-89-7	Electrophile	NC	(3) NC	136.7	150.4	85.4	81.6	N
Lauric acid	143-07-7	Organic acid	NC	(3) NC	102	117.4	90.7	64.4	N

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; UN GHS = United Nations Globally Harmonized System (1); VRM = Validated Reference Method, EpiSkin™=VRM1, EpiDerm™= VRM2; NC = Not Corrosive

[1]The proficiency substances, sorted first by corrosives versus non-corrosives, then by corrosive sub-category and then by chemical class, were selected from the substances used in the ECVAM validation studies EpiSkin™ and EpiDerm™ (8) (9) (10) and from post-validation studies based on data provided by EpiSkin™ (22), EpiDerm™, SkinEthic™ and epiCS® developers (23). Unless otherwise indicated, the substances were tested at the purity level obtained when purchased from a commercial source (8) (10). The selection includes, to the extent possible, substances that: (i) are representative of the range of corrosivity responses (e.g. non-corrosives; weak to strong corrosives) that the VRMs are capable of measuring or predicting; (ii) are representative of the chemical classes used in the validation studies; (iii) have chemical structures that are well-defined; (iv) induce reproducible results in the VRM; (v) induce definitive results in the in vivo reference test method; (vi) are commercially available; and (vii) are not associated with prohibitive disposal costs.

[2]Chemical class assigned by Barratt et al. (8).

[3]The corresponding UN Packing groups are I, II and III, respectively, for the UN GHS 1A, 1B and 1C.

[4]The in vitro predictions reported in this table were obtained with all five test methods covered in TG 431; for phenol though the LabCyte EPI-MODEL24 had slightly discordant results across runs, i.e. 1A-1BC-1BC; other methods achieved these classifications in validation or post-validation testing performed by the test method developers.

[5]The viability values obtained in the ECVAM Skin Corrosion Validation Studies were not corrected for direct MTT reduction (killed controls were not performed in the validation studies). However, the post-validation data generated by the test method developers that are presented in this table were acquired with adapted controls (23).

14. As part of the proficiency exercise, it is recommended that the user verifies the barrier properties of the tissues after receipt as specified by the RhE model manufacturer. This is particularly important if tissues are shipped over long distance/time periods. Once a test method has been successfully established and proficiency in its use has been demonstrated, such verification will not be necessary on a routine basis. However, when using a test method routinely, it is recommended to continue to assess the barrier properties in regular intervals.

PROCEDURE

15. The following is a generic description of the components and procedures of the RhE test methods for skin corrosion assessment covered by this Test Guideline. The RhE models endorsed as scientifically valid for use within this Test Guideline, i.e. the EpiSkin™ (SM), EpiDerm™ (EPI-200), SkinEthic™ RHE, epiCS® and LabCyte EPI-MODEL24 (16) (17) (19) (28) (29) (30) (31) (32) (33) (40) (41), can be obtained from commercial sources. Standard Operating Procedures (SOPs) for these five RhE models are available (34) (35) (36) (37) (42), and their main test method components are summarised in Annex 2. It is recommended that the relevant SOP be consulted when implementing and using one of these methods in the laboratory. Testing with the five RhE test methods covered by this Test Guideline should comply with the following:

RHE TEST METHOD COMPONENTS

General conditions

16. Non-transformed human keratinocytes should be used to reconstruct the epithelium. Multiple layers of viable epithelial cells (basal layer, stratum spinosum, stratum granulosum) should be present under a functional stratum corneum. The stratum corneum should be multi-layered containing the essential lipid profile to produce a functional barrier with robustness to resist rapid penetration of cytotoxic benchmark chemicals, e.g. sodium dodecyl sulphate (SDS) or Triton X-100. The barrier function should be demonstrated and may be assessed either by determination of the concentration at which a benchmark chemical reduces the viability of the tissues by 50% (IC50) after a fixed exposure time, or by determination of the exposure time required to reduce cell viability by 50% (ET50) upon application of the benchmark chemical at a specified, fixed concentration (see paragraph 18). The containment properties of the RhE model should prevent the passage of material around the stratum corneum to the viable tissue, which would lead to poor modelling of skin exposure. The RhE model should be free of contamination by bacteria, viruses, mycoplasma, or fungi.

Functional conditions

Viability

17. The assay used for quantifying tissue viability is the MTT-assay (27). The viable cells of the RhE tissue construct reduce the vital dye MTT into a blue MTT formazan precipitate, which is then extracted from the tissue using isopropanol (or a similar solvent). The OD of the extraction solvent alone should be sufficiently small, i.e., OD < 0.1. The extracted MTT formazan may be quantified using either a standard absorbance (OD) measurement or an HPLC/UPLC-spectrophotometry procedure (38). The RhE model users should ensure that each batch of the RhE model used meets defined criteria for the negative control. An acceptability range (upper and lower limit) for the negative control OD values should be established by the RhE model developer/supplier. Acceptability ranges for the negative control OD values for the five validated RhE test methods included in this Test Guideline are given in Table 2. An HPLC/UPLC-Spectrophotometry user should use the negative control OD ranges provided in Table 2 as the acceptance criterion for the negative control. It should be documented that the tissues treated with negative control are stable in culture (provide similar OD measurements) for the duration of the exposure period.

Table 2. Acceptability ranges for negative control OD values to control batch quality

	Lower acceptance limit	Upper acceptance limit
EpiSkin™ (SM)	= 0.6	= 1.5
EpiDerm™ SCT (EPI-200)	= 0.8	= 2.8
SkinEthic™ RHE	= 0.8	= 3.0
epiCS	= 0.8	= 2.8
LabCyte EPI-MODEL24 SCT	= 0.7	= 2.5

Barrier function

18. The stratum corneum and its lipid composition should be sufficient to resist the rapid penetration of certain cytotoxic benchmark chemicals (e.g. SDS or Triton X-100), as estimated by IC50 or ET50 (Table 3). The barrier function of each batch of the RhE model used should be demonstrated by the RhE model developer/vendor upon supply of the tissues to the end user (see paragraph 21).

Morphology

19. Histological examination of the RhE model should be performed demonstrating multi-layered human epidermis-like structure containing stratum basale, stratum spinosum, stratum granulosum and stratum corneum and exhibits lipid profile similar to lipid profile of human epidermis. Histological examination of each batch of the RhE model used demonstrating appropriate morphology of the tissues should be provided by the RhE model developer/vendor upon supply of the tissues to the end user (see paragraph 21).

Reproducibility

20. Test method users should demonstrate reproducibility of the test methods over time with the positive and negative controls. Furthermore, the test method should only be used if the RhE model developer/supplier provides data demonstrating reproducibility over time with corrosive and non-corrosive chemicals from e.g. the list of Proficiency Substances (Table 1). In case of the use of a test method for sub-categorisation, the reproducibility with respect to sub-categorisation should also be demonstrated.

Quality control (QC)

21. The RhE model should only be used if the developer/supplier demonstrates that each batch of the RhE model used meets defined production release criteria, among which those for viability (paragraph 17), barrier function (paragraph 18) and morphology (paragraph 19) are the most relevant. These data are provided to the test method users, so that they are able to include this information in the test report. Only results produced with QC accepted tissue batches can be accepted for reliable prediction of corrosive classification. An acceptability range (upper and lower limit) for the IC50 or the ET50 is established by the RhE model developer/supplier. The acceptability ranges for the five validated test methods are given in Table 3.

Table 3. **QC batch release criterion**

	Lower acceptance limit	Upper acceptance limit
EpiSkin™ (SM) (18 hours treatment with SDS)(33)	IC_{50} = 1.0 mg/mL	IC_{50} = 3.0 mg/mL
EpiDerm™SCT (EPI-200) (1% Triton X-100)(34)	ET_{50} = 4.0 hours	ET_{50} = 8.7 hours
SkinEthic™ RHE (1% Triton X-100)(35)	ET_{50} = 4.0 hours	ET_{50} = 10.0 hours
epiCS (1% Triton X-100)(36)	ET_{50} = 2.0 hours	ET_{50} = 7.0 hours
LabCyte EPI-MODEL24 SCT (18 hours treatment with SDS) (42)	IC_{50} = 1.4 mg/mL	IC_{50} = 4.0 mg/mL

Application of the Test Chemical and Control Substances

22. At least two tissue replicates should be used for each test chemical and controls for each exposure time. For liquid as well as solid chemicals, sufficient amount of test chemical should be applied to uniformly cover the epidermis surface while avoiding an infinite dose, i.e. a minimum of 70 µL/cm2 or 30 mg/cm2 should be used. Depending on the methods, the epidermis surface should be moistened with deionized or distilled water before application of solid chemicals, to improve contact between the test chemical and the epidermis surface (34) (35) (36) (37) (42). Whenever possible, solids should be tested as a fine powder. The application method should be appropriate for the test chemical (see e.g. references (34-37). At the end of the exposure period, the test chemical should be carefully washed from the epidermis with an aqueous buffer, or 0.9% NaCl. Depending on which of the five validated RhE test methods is used, two or three exposure periods are used per test chemical (for all five valid RhE models: 3 min and 1 hour; for EpiSkin™ an additional exposure time of 4 hours). Depending on the RhE test method used and the exposure period assessed, the incubation temperature during exposure may vary between room temperature and 37ºC.

23. Concurrent negative and positive controls (PC) should be used in each run to demonstrate that viability (with negative controls), barrier function and resulting tissue sensitivity (with the PC) of the tissues are within a defined historical acceptance range. The suggested PC chemicals are glacial acetic acid or 8N KOH depending upon the RhE model used (see Annex 2 and relevant SOP for details). It should be noted that 8N KOH is a direct MTT reducer that might require adapted controls as described in paragraphs 25 and 26. The suggested negative controls are 0.9% (w/v) NaCl or water.

Cell Viability Measurements

24. The MTT assay, which is a quantitative assay, should be used to measure cell viability under this Test Guideline (27). The tissue sample is placed in MTT solution of appropriate concentration (0.3, 0.5 or 1 mg/mL, see Annex 2 and relevant SOP for details) for 3 hours. The precipitated blue formazan product is then extracted from the tissue using a solvent (e.g. isopropanol, acidic isopropanol), and the concentration of formazan is

measured by determining the OD at 570 nm using a filter band pass of maximum ± 30 nm, or by an HPLC/UPLC spectrophotometry procedure (see paragraphs 30 and 31) (38).

25. Test chemicals may interfere with the MTT assay, either by direct reduction of the MTT into blue formazan, and/or by colour interference if the test chemical absorbs, naturally or due to treatment procedures, in the same OD range of formazan (570 ± 30 nm, mainly blue and purple chemicals). Additional controls should be used to detect and correct for a potential interference from these test chemicals such as the non-specific MTT reduction (NSMTT) control and the non-specific colour (NSC) control (see paragraphs 26 to 30). This is especially important when a specific test chemical is not completely removed from the tissue by rinsing or when it penetrates the epidermis, and is therefore present in the tissues when the MTT viability test is performed. Detailed description of how to correct direct MTT reduction and interferences by colouring agents is available in the SOPs for the test methods (34) (35) (36) (37) (42).

26. To identify direct MTT reducers, each test chemical should be added to freshly prepared MTT medium (34) (35) (36) (37) (42). If the MTT mixture containing the test chemical turns blue/purple, the test chemical is presumed to directly reduce the MTT, and further functional check on non-viable epidermis should be performed, independently of using the standard absorbance (OD) measurement or an HPLC/UPLC-spectrophotometry procedure. This additional functional check employs killed tissues that possess only residual metabolic activity but absorb the test chemical in similar amount as viable tissues. Each MTT reducing chemical is applied on at least two killed tissue replicates per exposure time, which undergo the whole skin corrosion test. The true tissue viability is then calculated as the percent tissue viability obtained with living tissues exposed to the MTT reducer minus the percent non-specific MTT reduction obtained with the killed tissues exposed to the same MTT reducer, calculated relative to the negative control run concurrently to the test being corrected (%NSMTT).

27. To identify potential interference by coloured test chemicals or test chemicals that become coloured when in contact with water or isopropanol and decide on the need for additional controls, spectral analysis of the test chemical in water (environment during exposure) and/or isopropanol (extracting solution) should be performed. If the test chemical in water and/or isopropanol absorbs light in the range of 570 ± 30 nm, further colorant controls should be performed or, alternatively, an HPLC/UPLC-spectrophotometry procedure should be used in which case these controls are not required (see paragraphs 30 and 31). When performing the standard absorbance (OD) measurement, each interfering coloured test chemical is applied on at least two viable tissue replicates per exposure time, which undergo the entire skin corrosion test but are incubated with medium instead of MTT solution during the MTT incubation step to generate a non-specific colour (NSCliving) control. The NSCliving control needs to be performed concurrently per exposure time per coloured test chemical (in each run) due to the inherent biological variability of living tissues. The true tissue viability is then calculated as the percent tissue viability obtained with living tissues exposed to the interfering test chemical and incubated with MTT solution minus the percent non-specific colour obtained with living tissues exposed to the interfering test chemical and incubated with medium without MTT, run concurrently to the test being corrected (%NSCliving).

28. Test chemicals that are identified as producing both direct MTT reduction (see paragraph 26) and colour interference (see paragraph 27) will also require a third set of controls, apart from the NSMTT and NSCliving controls described in the previous

paragraphs, when performing the standard absorbance (OD) measurement. This is usually the case with darkly coloured test chemicals interfering with the MTT assay (e.g., blue, purple, black) because their intrinsic colour impedes the assessment of their capacity to directly reduce MTT as described in paragraph 26. These test chemicals may bind to both living and killed tissues and therefore the NSMTT control may not only correct for potential direct MTT reduction by the test chemical, but also for colour interference arising from the binding of the test chemical to killed tissues. This could lead to a double correction for colour interference since the NSCliving control already corrects for colour interference arising from the binding of the test chemical to living tissues. To avoid a possible double correction for colour interference, a third control for non-specific colour in killed tissues (NSCkilled) needs to be performed. In this additional control, the test chemical is applied on at least two killed tissue replicates per exposure time, which undergo the entire testing procedure but are incubated with medium instead of MTT solution during the MTT incubation step. A single NSCkilled control is sufficient per test chemical regardless of the number of independent tests/runs performed, but should be performed concurrently to the NSMTT control and, where possible, with the same tissue batch. The true tissue viability is then calculated as the percent tissue viability obtained with living tissues exposed to the test chemical minus %NSMTT minus %NSCliving plus the percent non-specific colour obtained with killed tissues exposed to the interfering test chemical and incubated with medium without MTT, calculated relative to the negative control run concurrently to the test being corrected (%NSCkilled).

29. It is important to note that non-specific MTT reduction and non-specific colour interferences may increase the readouts of the tissue extract above the linearity range of the spectrophotometer. On this basis, each laboratory should determine the linearity range of their spectrophotometer with MTT formazan (CAS # 57360-69-7) from a commercial source before initiating the testing of test chemicals for regulatory purposes. In particular, the standard absorbance (OD) measurement using a spectrophotometer is appropriate to assess direct MTT-reducers and colour interfering test chemicals when the ODs of the tissue extracts obtained with the test chemical without any correction for direct MTT reduction and/or colour interference are within the linear range of the spectrophotometer or when the uncorrected percent viability obtained with the test chemical already defined it as a corrosive (see paragraphs 35 and 36). Nevertheless, results for test chemicals producing %NSMTT and/or %NSCliving ≥ 50% of the negative control should be taken with caution.

30. For coloured test chemicals which are not compatible with the standard absorbance (OD) measurement due to too strong interference with the MTT assay, the alternative HPLC/UPLC-spectrophotometry procedure to measure MTT formazan may be employed (see paragraph 31) (37). The HPLC/UPLC-spectrophotometry system allows for the separation of the MTT formazan from the test chemical before its quantification (38). For this reason, NSCliving or NSCkilled controls are never required when using HPLC/UPLC-spectrophotometry, independently of the chemical being tested. NSMTT controls should nevertheless be used if the test chemical is suspected to directly reduce MTT or has a colour that impedes the assessment of the capacity to directly reduce MTT (as described in paragraph 26). When using HPLC/UPLC-spectrophotometry to measure MTT formazan, the percent tissue viability is calculated as percent MTT formazan peak area obtained with living tissues exposed to the test chemical relative to the MTT formazan peak obtained with the concurrent negative control. For test chemicals able to directly reduce MTT, true tissue viability is calculated as the percent tissue viability obtained with living tissues exposed to the test chemical minus %NSMTT. Finally, it should be noted that direct MTT-reducers

that may also be colour interfering, which are retained in the tissues after treatment and reduce MTT so strongly that they lead to ODs (using standard OD measurement) or peak areas (using UPLC/HPLC-spectrophotometry) of the tested tissue extracts that fall outside of the linearity range of the spectrophotometer cannot be assessed, although these are expected to occur in only very rare situations.

31. HPLC/UPLC-spectrophotometry may be used also with all types of test chemicals (coloured, non-coloured, MTT-reducers and non-MTT reducers) for measurement of MTT formazan (38). Due to the diversity of HPLC/UPLC-spectrophotometry systems, qualification of the HPLC/UPLC-spectrophotometry system should be demonstrated before its use to quantify MTT formazan from tissue extracts by meeting the acceptance criteria for a set of standard qualification parameters based on those described in the U.S. Food and Drug Administration guidance for industry on bio-analytical method validation (38) (39). These key parameters and their acceptance criteria are shown in Annex 4. Once the acceptance criteria defined in Annex 4 have been met, the HPLC/UPLC-spectrophotometry system is considered qualified and ready to measure MTT formazan under the experimental conditions described in this Test Guideline.

Acceptance Criteria

32. For each test method using valid RhE models, tissues treated with the negative control should exhibit OD reflecting the quality of the tissues as described in table 2 and should not be below historically established boundaries. Tissues treated with the PC, i.e. glacial acetic acid or 8N KOH, should reflect the ability of the tissues to respond to a corrosive chemical under the conditions of the test method (see Annex 2 and relevant SOP for details). The variability between tissue replicates of test chemical and/or control substances should fall within the accepted limits for each valid RhE model requirements (see Annex 2 and relevant SOP for details) (e.g. the difference of viability between the two tissue replicates should not exceed 30%). If either the negative control or PC included in a run fall out of the accepted ranges, the run is considered as not qualified and should be repeated. If the variability of test chemicals falls outside of the defined range, its testing should be repeated.

Interpretation of Results and Prediction Model

33. The OD values obtained for each test chemical should be used to calculate percentage of viability relative to the negative control, which is set at 100%. In case HPLC/UPLC-spectrophotometry is used, the percent tissue viability is calculated as percent MTT formazan peak area obtained with living tissues exposed to the test chemical relative to the MTT formazan peak obtained with the concurrent negative control. The cut-off percentage cell viability values distinguishing corrosive from non-corrosive test chemical (or discriminating between different corrosive sub-categories) are defined below in paragraphs 35 and 36 for each of the test methods covered by this Test Guideline and should be used for interpreting the results.

34. A single testing run composed of at least two tissue replicates should be sufficient for a test chemical when the resulting classification is unequivocal. However, in cases of borderline results, such as non-concordant replicate measurements, a second run may be considered, as well as a third one in case of discordant results between the first two runs.

35. The prediction model for the EpiSkin™ skin corrosion test method (9) (34) (22), associated with the UN GHS (1) classification system, is shown in Table 4:

Table 4. EpiSkin™ prediction model

Viability measured after exposure time points (t='3,' 60 and 240 minutes)	Prediction to consider
35% after 3 min exposure	**Corrosive:** Optional Sub-category 1A *
= 35% after 3 min exposure **AND** < 35% after 60 min exposure **OR** = 35% after 60 min exposure **AND** < 35% after 240 min exposure	**Corrosive:** A combination of optional Sub-categories 1B-and-1C
= 35% after 3 min exposure	**Non-corrosive**

*) According to the data generated in view of assessing the usefulness of the RhE test methods for supporting sub-categorisation, it was shown that around 22 % of the Sub-category 1A results of the EpiSkin™ test method may actually constitute Sub-category 1B or Sub-category 1C substances/mixtures (i.e. over classifications) (see Annex 3).

36. The prediction models for the EpiDerm™ SCT (10) (23) (35), the SkinEthic™ RHE (17) (18) (23) (36), the epiCS® (16) (23) (37) and LabCyte EPI-MODEL24 (41) (42) skin corrosion test methods, associated with the UN GHS (1) classification system, are shown in Table 5:

Table 5. EpiDerm™ SCT, SkinEthic™ RHE epiCS® and LabCyte EPI-MODEL24 SCT

Viability measured after exposure time points (t=3 and 60 minutes)	Prediction to be considered
STEP 1 for EpiDerm™ SCT, SkinEthic™ RHE, epiCS® and LabCyte EPI-MODEL24 SCT	
< 50% after 3 min exposure	Corrosive
= 50% after 3 min exposure AND < 15% after 60 min exposure	Corrosive
= 50% after 3 min exposure AND = 15% after 60 min exposure	Non-corrosive
STEP 2 for EpiDerm™ SCT - for substances/mixtures identified as Corrosive in step 1	
< 25% after 3 min exposure	Optional Sub-category 1A *
= 25% after 3 min exposure	A combination of optional Sub-categories 1B-and-1C
STEP 2 for SkinEthic™ RHE - for substances/mixtures identified as Corrosive in step 1	
< 18% after 3 min exposure	Optional Sub-category 1A *
= 18% after 3 min exposure	A combination of optional Sub-categories 1B-and-1C
STEP 2 for epiCS® - for substances/mixtures identified as Corrosive in step 1	
< 15% after 3 min exposure	Optional Sub-category 1A *
= 15% after 3 min exposure	A combination of optional Sub-categories 1B-and-1C
STEP 2 for LabCyte EPI-MODEL24 SCT - for substances/mixtures identified as Corrosive in step 1	
< 15% after 3 min exposure	Optional Sub-category 1A *
= 15% after 3 min exposure	A combination of optional Sub-categories 1B-and-1C

* According to the data generated in view of assessing the usefulness of the RhE test methods for supporting sub-categorisation, it was shown that around 29%, 31%, 33% and 30% of the Sub-category 1A results of the EpiDerm™ SCT, SkinEthic™ RHE epiCS® and LabCyte EPI-MODEL24 SCT, respectively, may actually constitute Sub-category 1B or Sub-category 1C substances/mixtures (i.e. over-classifications) (see Annex 3).

DATA AND REPORTING

Data

37. For each test, data from individual tissue replicates (e.g. OD values and calculated percentage cell viability for each test chemical, including classification) should be reported in tabular form, including data from repeat experiments as appropriate. In addition, means and ranges of viability and CVs between tissue replicates for each test should be reported. Observed interactions with MTT reagent by direct MTT reducers or coloured test chemicals should be reported for each tested chemical.

Test Report

38. The test report should include the following information:

Test Chemical and Control Substances:

- Mono-constituent substance: chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc;
- Multi-constituent substance, UVCB and mixture: characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents;
- Physical appearance, water solubility, and any additional relevant physicochemical properties;
- Source, lot number if available;
- Treatment of the test chemical/control substance prior to testing, if applicable (e.g. warming, grinding);
- Stability of the test chemical, limit date for use, or date for re-analysis if known;
- Storage conditions.

RhE model and protocol used and rationale for it (if applicable)

Test Conditions:

- RhE model used (including batch number);
- Calibration information for measuring device (e.g. spectrophotometer), wavelength and band
- pass (if applicable) used for quantifying MTT formazan, and linearity range of measuring device;
- Description of the method used to quantify MTT formazan;
- Description of the qualification of the HPLC/UPLC-spectrophotometry system, if applicable;
- Complete supporting information for the specific RhE model used including its performance. This should include, but is not limited to:
    - i) Viability;
    - ii) Barrier function;
    - iii) Morphology;
    - iv) Quality controls (QC) of the model;
- Reference to historical data of the model. This should include, but is not limited to acceptability of the QC data with reference to historical batch data;
- Demonstration of proficiency in performing the test method before routine use by testing of the proficiency substances.

Test Procedure:

- Details of the test procedure used (including washing procedures used after exposure period);
- Doses of test chemical and control substances used;
- Duration of exposure period(s) and temperature(s) of exposure;
- Indication of controls used for direct MTT-reducers and/or colouring test chemicals, if applicable;
- Number of tissue replicates used per test chemical and controls (PC, negative control, and NSMTT, NSCliving and NSCkilled, if applicable), per exposure time;
- Description of decision criteria/prediction model applied based on the RhE model used;
- Description of any modifications of the test procedure (including washing procedures).
- Run and Test Acceptance Criteria:
- Positive and negative control mean values and acceptance ranges based on historical data;
- Acceptable variability between tissue replicates for positive and negative controls;
- Acceptable variability between tissue replicates for test chemical.

Results:

- Tabulation of data for individual test chemicals and controls, for each exposure period, each run and each replicate measurement including OD or MTT formazan peak area, percent tissue viability, mean percent tissue viability, differences between replicates, SDs and/or CVs if applicable;
- If applicable, results of controls used for direct MTT-reducers and/or colouring test chemicals including OD or MTT formazan peak area, %NSMTT, %NSCliving, %NSCkilled, differences between tissue replicates, SDs and/or CVs (if applicable), and final correct percent tissue viability;
- Results obtained with the test chemical(s) and control substances in relation to the defined run and test acceptance criteria;
- Description of other effects observed;
- The derived classification with reference to the prediction model/decision criteria used.

Discussion of the results:

Conclusions:

LITERATURE

1. UN. (2013). United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Fifth Revised Edition, UN New York and Geneva. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html]
2. OECD. (2015). Guideline for Testing of Chemicals. (No. 404.): Acute Dermal Irritation, Corrosion, Organisation for Economic Cooperation and Development, Paris.
3. OECD. (2015). Guideline for the Testing of Chemicals (No. 430.): In Vitro Skin Corrosion: Transcutaneous Electrical Resistance (TER). Organisation for Economic Cooperation and Development, Paris.
4. OECD. (2015). Guideline for the Testing of Chemicals (No. 435.): In Vitro Membrane Barrier Test Method. Organisation for Economic Cooperation and Development, Paris. (5) OECD. (2015). Guideline for the Testing of Chemicals (No. 439.): In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method. Organisation for Economic Cooperation and Development, Paris.
5. OECD. (2015). Guideline for the Testing of Chemicals (No. 439.): In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method. Organisation for Economic Cooperation and Development, Paris.
6. OECD. (2014). Guidance Document on Integrated Approaches to Testing and Assessment of Skin Irritation/Corrosion. Environment, Health and Safety Publications, Series on Testing and Assessment, (No. 203.) Organisation for Economic Cooperation and Development, Paris.
7. Botham P.A., Chamberlain M., Barratt M.D., Curren R.D., Esdaile D.J., Gardner J.R., Gordon V.C., Hildebrand B., Lewis R.W., Liebsch M., Logemann P., Osborne R., Ponec M., Regnier J.F., Steiling W., Walker A.P., and Balls M. (1995). A Prevalidation Study on In Vitro Skin Corrosivity Testing. The report and Recommendations of ECVAM Workshop 6.ATLA 23, 219-255.
8. Barratt M.D., Brantom P.G., Fentem J.H., Gerner I., Walker A.P., and Worth A.P. (1998). The ECVAM International Validation Study on In Vitro Tests for Skin Corrosivity. 1. Selection and distribution of the Test Chemicals. Toxicol.In Vitro 12, 471-482.
9. Fentem J.H., Archer G.E.B., Balls M., Botham P.A., Curren R.D., Earl L.K., Esdaile D.J., Holzhutter H.- G., and Liebsch M. (1998). The ECVAM International Validation Study on In Vitro Tests for Skin Corrosivity. 2. Results and Evaluation by the Management Team. Toxicol.in Vitro 12, 483-524.
10. Liebsch M., Traue D., Barrabas C., Spielmann H., Uphill, P., Wilkins S., Wiemann C., Kaufmann T., Remmele M. and Holzhütter H. G. (2000). The ECVAM Prevalidation Study on the Use of EpiDerm for Skin Corrosivity Testing, ATLA 28, pp. 371-401.
11. Balls M., Blaauboer B.J., Fentem J.H., Bruner L., Combes R.D., Ekwall B., Fielder R.J., Guillouzo A., Lewis R.W., Lovell D.P., Reinhardt C.A., Repetto G., Sladowski D., Spielmann H. et Zucco F. (1995). Practical Aspects of the Validation

of Toxicity Test Procedures. The Report and Recommendations of ECVAM Workshops, ATLA 23, 129-147.

12. ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). (1997). Validation and Regulatory Acceptance of Toxicological Test Methods. NIH Publication No. 97-3981. National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. Available at: [http://www.iccvam.niehs.nih.gov/docs/guidelines/validate.pdf].

13. ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). (2002). ICCVAM evaluation of EpiDerm™ (EPI-200), EPISKIN™ (SM), and the Rat Skin Transcutaneous Electrical Resistance (TER) Assay: In Vitro Test Methods for Assessing Dermal Corrosivity Potential of Chemicals. NIH Publication No. 02-4502. National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.Available at:[http://www.iccvam.niehs.nih.gov/methods/epiddocs/epis_brd.pdf]

14. EC-ECVAM. (1998). Statement on the Scientific Validity of the EpiSkin™ Test (an In Vitro Test for Skin Corrosivity), Issued by the ECVAM Scientific Advisory Committee (ESAC10), 3 April 1998. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

15. EC-ECVAM. (2000). Statement on the Application of the EpiDerm™ Human Skin Model for Skin Corrosivity Testing, Issued by the ECVAM Scientific Advisory Committee (ESAC14), 21 March 2000. Available at: [http://ecvam.jrc.ec.europa.eu].

16. Hoffmann J., Heisler E., Karpinski S., Losse J., Thomas D., Siefken W., Ahr H.J., Vohr H.W. and Fuchs H.W. (2005). Epidermal-Skin-Test 1000 (EST-1000)-A New Reconstructed Epidermis for In Vitro Skin Corrosivity Testing. Toxicol.In Vitro 19, 925-929.

17. Kandárová H., Liebsch M., Spielmann,H., Genschow E., Schmidt E., Traue D., Guest R., Whittingham A., Warren N, Gamer A.O., Remmele M., Kaufmann T., Wittmer E., De Wever B., and Rosdy M. (2006). Assessment of the Human Epidermis Model SkinEthic RHE for In Vitro Skin Corrosion Testing of Chemicals According to New OECD TG 431. Toxicol.In Vitro 20, 547 559.

18. Tornier C., Roquet M. and Fraissinette A.B. (2010). Adaptation of the Validated SkinEthic™ Reconstructed Human Epidermis (RHE) Skin Corrosion Test Method to 0.5 cm2 Tissue Sample. Toxicol. In Vitro 24, 1379-1385.

19. EC-ECVAM. (2006). Statement on the Application of the SkinEthic™ Human Skin Model for Skin Corrosivity Testing, Issued by the ECVAM Scientific Advisory Committee (ESAC25), 17 November 2006. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

20. EC-ECVAM. (2009). ESAC Statement on the Scientific Validity of an In-Vitro Test Method for Skin Corrosivity Testing: the EST-1000, Issued by the ECVAM Scientific Advisory Committee (ESAC30), 12 June 2009. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

21. OECD. (2013). Summary Document on the Statistical Performance of Methods in OECD Test Guideline 431 for Sub-categorisation. Environment, Health, and Safety

Publications, Series on Testing and Assessment (No. 190.). Organisation for Economic Cooperation and Development, Paris.

22. Alépée N., Grandidier M.H., and Cotovio J. (2014). Sub-Categorisation of Skin Corrosive Chemicals by the EpiSkin™ Reconstructed Human Epidermis Skin Corrosion Test Method According to UN GHS: Revision of OECD Test Guideline 431. Toxicol.In Vitro 28, 131-145.
23. Desprez B., Barroso J., Griesinger C., Kandárová H., Alépée N., and Fuchs, H. (2015). Two Novel Prediction Models Improve Predictions of Skin Corrosive Sub-categories by Test Methods of OECD Test Guideline No. 431. Toxicol. In Vitro 29, 2055 2080.
24. OECD. (2015). Performance Standards for the Assessment of Proposed Similar or Modified In Vitro Reconstructed Human Epidermis (RHE) Test Methods For Skin Corrosion in Relation to OECD TG 431. Environmental Health and Safety Publications, Series on Testing and Assessment (No. 219). Organisation for Economic Cooperation and Development, Paris
25. OECD. (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. .Environment, Health and Safety Publications, Series on Testing and Assessment (No. 34.), Organisation for Economic Cooperation and Development, Paris.
26. Eskes C. et al. (2012). Regulatory Assessment of In Vitro Skin Corrosion and Irritation Data Within the European Framework: Workshop Recommendations. Regul.Toxicol.Pharmacol. 62, 393-403.
27. Mosmann T. (1983). Rapid Colorimetric Assay for Cellular Growth and Survival: Application to Proliferation and Cytotoxicity Assays. J. Immunol. Methods 65, 55-63.
28. Tinois E., et al. (1994). The Episkin Model: Successful Reconstruction of Human Epidermis In Vitro. In: In Vitro Skin Toxicology. Rougier A.,. Goldberg A.M and Maibach H.I. (Eds): 133-140.
29. Cannon C. L., Neal P.J., Southee J.A., Kubilus J. and Klausner M. (1994), New Epidermal Model for Dermal Irritancy Testing. Toxicol.in Vitro 8, 889 - 891.
30. Ponec M., Boelsma E, Weerheim A, Mulder A, Bouwstra J and Mommaas M. (2000). Lipid and Ultrastructural Characterization of Reconstructed Skin Models. Inter. J. Pharmaceu. 203, 211 - 225.
31. Tinois E., Tillier, J., Gaucherand, M., Dumas, H., Tardy, M. and Thivolet J. (1991). In Vitro and Post Transplantation Differentiation of Human Keratinocytes Grown on the Human Type IV Collagen Film of a Bilayered Dermal Substitute. Exp. Cell Res. 193: 310-319.
32. Parenteau N.L., Bilbo P, Nolte CJ, Mason VS and Rosenberg M. (1992). The Organotypic Culture of Human Skin Keratinocytes and Fibroblasts to Achieve Form and Function. Cytotech. 9, 163-171.
33. Wilkins L.M., Watson SR, Prosky SJ, Meunier SF and Parenteau N.L. (1994). Development of a Bilayered Living Skin Construct for Clinical Applications. Biotech.Bioeng.43/8, 747-756.

34. EpiSkin™ (December 2011)SOP,. INVITTOX Protocol (No. 118.). EpiSkin™ Skin Corrosivity Test.. Available at: [http://www.ecvam.jrc.ec.europa.eu.hm;].

35. EpiDremTM SOP. (February 2012). Version MK-24-007-0024 Protocol for : In vitro EpiDerm™ Skin Corrosion Test (EPI-200-SCT), for Use with MatTek Corporation's Reconstructed Human Epidermal Model EpiDerm. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

36. SkinEthic™ RHE SOP,INVITTOX Protocol (January 2012). SkinEthic™ Skin Corrosivity Test. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

37. EpiCS® SOP, Version 4.1 (January 2012). In Vitro Skin Corrosion: Human Skin Model Test Epidermal Skin Test 1000 (epiCS®) CellSystems. Available at: [http://www.ecvam.jrc.ec.europa.eu.html].

38. Alépée N., Barroso J., De Smedt A., De Wever B., Hibatallah J., Klaric M., Mewes K.R., Millet M., Pfannenbecker U., Tailhardat M., Templier M., and McNamee P. Use of HPLC/UPLCspectrophotometry for Detection of MTT Formazan in In Vitro Reconstructed Human Tissue (RhT)- based Test Methods Employing the MTT Assay to Expand their Applicability to Strongly Coloured Test Chemicals. Manuscript in Preparation.

39. US FDA. (2001). Guidance for Industry: Bioanalytical Method Validation. U.S. Department of Health and Human Services, Food and Drug Administration. (May 2001). Available at: [http://www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf].

40. Katoh M., Hamajima F., Ogasawara T., and Ken-ichiro Hata. (2010). Assessment of the human epidermal model LabCyte EPI-MODEL for in vitro skin corrosion testing according to the OECD test Guideline 431. J Toxicol. Sci. 35, 411-417.

41. LabCyte Validation Management Team (2018). Validation Study for in vitro skin corrosion test method using reconstructed human epidermal tissue LabCyte EPI-MODEL24. Available at: [http://www.jacvam.jp/files/doc/01_05/01_05_Z1.pdf].

42. LabCyte EPI-MODEL24 SCT SOP, Version 1.6. (May, 2017). Skin corrosion test using the reconstructed human model “LabCyte EPI-MODEL24”. Available at: [http://www.jacvam.jp/files/doc/01_05/01_05_Z2.pdf].

43. Report of the Peer-review of the validation study for LabCyte EPI-MODEL24 In Vitro Skin Corrosion Test Method. Available at [http://www.jacvam.jp/files/doc/01_05/01_05_Z3.pdf].

ANNEX 1- DEFINITIONS

Accuracy: The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with "concordance" to mean the proportion of correct outcomes of a test method (25).

Cell viability: Parameter measuring total activity of a cell population e.g. as ability of cellular mitochondrial dehydrogenases to reduce the vital dye MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5- diphenyltetrazolium bromide, Thiazolyl blue), which depending on the endpoint measured and the test design used, correlates with the total number and/or vitality of living cells.

Chemical: means a substance or a mixture.

Concordance: This is a measure of test method performance for test methods that give a categorical result, and is one aspect of relevance. The term is sometimes used interchangeably with accuracy, and is defined as the proportion of all chemicals tested that are correctly classified as positive or negative. Concordance is highly dependent on the prevalence of positives in the types of test chemical being examined (25).

ET50: Can be estimated by determination of the exposure time required to reduce cell viability by 50% upon application of the benchmark chemical at a specified, fixed concentration, see also IC50.

GHS (Globally Harmonized System of Classification and Labelling of Chemicals): A system proposing the classification of chemicals (substances and mixtures) according to standardized types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (1).

HPLC: High Performance Liquid Chromatography.

IATA: Integrated Approach on Testing and Assessment.

IC50: Can be estimated by determination of the concentration at which a benchmark chemical reduces the viability of the tissues by 50% (IC50) after a fixed exposure time, see also ET50.

ET50. Infinite dose: Amount of test chemical applied to the epidermis exceeding the amount required to completely and uniformly cover the epidermis surface.

Mixture: means a mixture or solution composed of two or more substances in which they do not react.

Mono-constituent substance: A substance, defined by its quantitative composition, in which one main constituent is present to at least 80% (w/w).

MTT: 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; Thiazolyl blue tetrazolium bromide.

Multi-constituent substance: A substance, defined by its quantitative composition, in which more than one main constituent is present in a concentration ≥ 10% (w/w) and <

80% (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

NC: Non corrosive.

NSCkilled control: Non-Specific Colour control in killed tissues.

NSCliving control : Non-Specific Colour control in living tissues.

NSMTT: Non-Specific MTT reduction.

OD: Optical Density

PC: Positive Control, a replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

Performance standards (PS): Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are; (i) essential test method components; (ii) a minimum list of Reference Chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (iii) the similar levels of reliability and accuracy, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of Reference Chemicals (25).

Relevance: Description of relationship of the test method to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test method correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (25).

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility (25).

Run: A run consists of one or more test chemicals tested concurrently with a negative control and with a PC.

Sensitivity: The proportion of all positive/active chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (25).

Skin corrosion in vivo: The production of irreversible damage of the skin; namely, visible necrosis through the epidermis and into the dermis, following the application of a test chemical for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of observation at 14 days, by discoloration due to blanching of the skin, complete areas of alopecia, and scars. Histopathology should be considered to evaluate questionable lesions.

Specificity: The proportion of all negative/inactive chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (25).

Substance: means chemical elements and their compounds in the natural state or obtained by any production process, including any additive necessary to preserve the stability of the

product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing its composition.

Test chemical: means what is being tested.

UPLC: Ultra-High Performance Liquid Chromatography.

UVCB: substances of unknown or variable composition, complex reaction products or biological materials.

ANNEX 2 - MAIN TEST METHOD COMPONENTS OF THE RhE TEST METHODS VALIDATED FOR SKIN CORROSION TESTING

Nr.	1	2	3	4	5
Test Method Component	**EpiSkin™**	**EpiDerm™ SCT**	**SkinEthic™ RHE**	**epiCS®**	**LabCyte EPI-MODEL24 SCT**
Model surface	0.38 cm^2	0.63 cm^2	0.5 cm^2	0.6 cm^2	0.3 cm^2
Number of tissue replicates	At least 2 per exposure time	2-3 per exposure time	At least 2 per exposure time	At least 2 per exposure time	At least 2 per exposure time
Treatment doses and application	Liquids and viscous: 50 ± 3 µL (131.6 µL/cm^2) Solids: 20± 2 mg (52.6 mg/cm^2) +100 µL±5µL NaCl solution (9 g/L) Waxy/sticky: 50 ± 2 mg (131.6 mg/cm^2) with a nylon mesh	Liquids: 50 µL (79.4 µL/cm^2) with or without a nylon mesh *Pre-test compatibility of test chemical with nylon mesh* Semi solids: 50 µL (79.4 µL/cm^2) Solids: 25 µL H_2O (or necessary) + 25 mg (39.7 mg/cm^2) Waxes: flat "disc like" piece of ca. 8 mm diameter placed atop the tissue wetted with 15µL H_2O.	Liquids and viscous:40 ± 3 µL (80µL/cm^2) using nylonmesh *Pre-test compatibility of test chemical with nylon mesh* Solids: 20 µL ± 2µl H_2O + 20± 3 mg (40 mg/cm2) Waxy/sticky: 20 ± 3 mg (40 mg/cm^2) with a nylon mesh	Liquids and viscous:50 µL (83.3µL/cm^2) using nylonmesh *Pre-test compatibility of test chemical with nylon mesh* Semi solids: 50 µL (83.3 µL/cm^2) Solids: 25 mg (41.7 mg/cm^2) + 25 µL H_2O (or more if necessary) Waxy/sticky: flat "cookie like" piece of ca. 8 mm diameter placed atop the tissue wetted with 15µL H_2O	Liquids and viscous:50 µL (166.7µL/cm^2) Solids: 50± 2 mg (166.7 mg/cm^2) + 50 µL H_2O Waxy: Use a positive displacement pipette and tip as liquid and viscous substance.

Nr.	1	2	3	4	5
Test Method Component	**EpiSkin™**	**EpiDerm™ SCT**	**SkinEthic™ RHE**	**epiCS®**	**LabCyte EPI-MODEL24 SCT**
Pre-check for direct MTT reduction	50 µL (liquid) or 20 mg (solid) + 2 mL MTT 0.3 mg/mL solution for 180±5 min at 37°C, 5% CO_2, 95% RH ➔ if solution turns blue/purple, water-killed adapted controls should be performed	50 µL (liquid) or 25 mg (solid) + 1 mL MTT 1 mg/mL solution for 60 min at 37°C, 5% CO_2, 95% RH ➔ if solution turns blue/purple, freeze-killed adapted controls should be performed	40 µL (liquid) or 20 mg (solid) + 1 mL MTT 1 mg/mL solution for 180±15 min at 37°C, 5% CO_2, 95% RH ➔ if solution turns blue/purple, freeze-killed adapted controls should be performed	50 µL (liquid) or 25 mg (solid) + 1 mL MTT 1 mg/mL solution for 60 min at 37°C, 5% CO_2, 95% RH ➔ if solution turns blue/purple, freeze-killed adapted controls should be performed	50 µL (liquid) or 50 mg (solid) + 500 µL MTT 0.5 mg/mL solution for 60 min at 37°C, 5% CO_2, 95% RH ➔ if solution turns blue/purple, freeze-killed adapted controls should be performed
Pre-check for colour interference	10 µL (liquid) or 10 mg (solid) + 90µL H2O mixed for 15 min at RT ➔ if solution becomes coloured, living adapted controls should be performed	50 µL (liquid) or 25 mg (solid) + 300 µL H2O mixed for 60 min at 37oC, 5% CO2, 95% RH ➔ if solution becomes coloured, living adapted controls should be performed	40 µL (liquid) or 20 mg (solid) + 300 µL H2O mixed for 60 min at RT ➔ if solution becomes coloured, living adapted controls should be performed	50 µL (liquid) or 25 mg (solid) + 300 µL H2O mixed for 60 min at 37oC, 5% CO2, 95% RH ➔ if solution becomes coloured, living adapted controls should be performed	50 µL (liquid) or 50 mg (solid) + 500 µL H2O mixed for 60 min at 37oC, 5% CO2, 95% RH ➔ if solution becomes coloured, living adapted controls should be performed
Exposure time and temperature	3 min, 60 min (±5 min) and 240 min (±10 min) In ventilated cabinet Room Temperature (RT, 18-28oC)	3 min at RT, and 60 min at 37oC, 5% CO2, 95% RH	3 min at RT, and 60 min at 37oC, 5% CO2, 95% RH	3 min at RT, and 60 min at 37oC, 5% CO2, 95% RH	3 min at RT, and 60 min at 37oC, 5% CO2, 95% RH
Rinsing	25 mL 1x PBS (2 mL/throwing)	20 times with a constant soft stream of 1x PBS	20 times with a constant soft stream of 1x PBS	20 times with a constant soft stream of 1x PBS	10 times or more with a constant strong stream of 1x PBS
Negative control	50 µL NaCl solution (9 g/L) Tested with every exposure time	50 µL H2O Tested with every exposure time	40 µL H2O Tested with every exposure time	50 µL H2O Tested with every exposure time	50 µL H2O Tested with every exposure time
Positive control	50 µL Glacial acetic acid Tested only for 4 hours	50 µL 8N KOH Tested with every exposure time	40 µL 8N KOH Tested only for 1 hour	50 µL 8N KOH Tested with every exposure time	50 µL 8N KOH Tested only for 1 hour
MTT solution	2 mL 0.3 mg/mL	300 µL 1 mg/mL	300 µL 1 mg/mL	300 µL 1 mg/mL	500 µL 0.5 mg/mL
MTT incubation	180 min (±15 min) at 37oC, 5%	180 min at 37oC, 5% CO2, 95%	180 min (±15 min) at 37oC, 5%	180 min at 37oC, 5% CO2, 95%	180 min (±5 min) at 37oC, 5%

Nr.	1	2	3	4	5
Test Method Component	**EpiSkin™**	**EpiDerm™ SCT**	**SkinEthic™ RHE**	**epiCS®**	**LabCyte EPI-MODEL24 SCT**
time and temperature	CO2, 95% RH	RH	CO2, 95% RH	RH	CO2, 95% RH
Test Method Component	EpiSkin™ EIT	EpiDerm™ SCT	SkinEthic™ RHE EIT	epiCS®	LabCyte EPI-MODEL24 SCT
Extraction solvent	500 µL acidified isopropanol (0.04 N HCl in isopropanol) (isolated tissue fully immersed)	2 mL isopropanol (extraction from top and bottom of insert)	1.5 mL isopropanol (extraction from top and bottom of insert)	2 mL isopropanol (extraction from top and bottom of insert)	300 µL isopropanol (isolated tissue fully immersed)
Extraction time And temperature	Overnight at RT, protected from light	Overnight without shaking at RT or for 120 min with shaking (~120 rpm) at RT	Overnight without shaking at RT or for 120 min with shaking (~120 rpm) at RT	Overnight without shaking at RT or for 120 min with shaking (~120 rpm) at RT	Overnight at RT, protected from light
OD reading	570 nm (545 - 595 nm) without reference filter	570 nm (or 540 nm) without reference filter	570 nm (540 - 600 nm) without reference filter	540 - 570 nm without reference filter	570 nm with reference filter 650 nm
Tissue Quality Control	18 hours treatment with SDS 1.0mg/mL ≤ IC_{50} ≤ 3.0mg/mL	Treatment with 1% Triton X-100 4.08 hours ≤ ET_{50} ≤ 8.7 hours	Treatment with 1% Triton X-100 4.0 hours ≤ ET_{50} ≤ 10.0 hours	Treatment with 1% Triton X-100 2.0 hours ≤ ET_{50} ≤ 7.0 hours	18 hours treatment with SDS 1.4mg/mL ≤ IC_{50} ≤ 4.0 mg/mL
Acceptability Criteria	1. Mean OD of the tissue replicates treated with the negative control (NaCl) should be ≥ 0.6 and ≤ 1.5 for every exposure time 2. Mean viability of the tissue replicates exposed for 4 hours with the positive control (glacial acetic acid), expressed as % of the negative control, should be ≤ 20% 3. In the range 20-100% viability and for ODs ≥ 0.3, difference of viability between the two tissue replicates should not	1. Mean OD of the tissue replicates treated with the negative control (H_2O) should be ≥ 0.8 and ≤ 2.8 for every exposure time 2. Mean viability of the tissue replicates exposed for 1 hour with the positive control (8N KOH), expressed as % of the negative control, should be ≤ 15% 3. In the range 20 - 100% viability, the Coefficient of Variation (CV) between tissue replicates should be ≤ 30%	1. Mean OD of the tissue replicates treated with the negative control (H_2O) should be ≥ 0.8 and ≤ 3.0 for every exposure time 2. Mean viability of the tissue replicates exposed for 1 hour (and 4 hours, if applicable) with the positive control (8N KOH), expressed as % of the negative control, should be ≤ 15% 3. In the range 20-100% viability and for ODs ≥ 0.3, difference of viability between the two	1. Mean OD of the tissue replicates treated with the negative control (H_2O) should be ≥ 0.8 and ≤ 2.8 for every exposure time 2. Mean viability of the tissue replicates exposed for 1 hour with the positive control (8N KOH), expressed as % of the negative control, should be ≤ 15%. 3. In the range 20-100% viability and for ODs ≥ 0.3, difference of viability between the two tissue replicates should not	1. Mean OD of the tissue replicates treated with the negative control (H_2O) should be ≥ 0.7 and ≤ 2.5 for every exposure time 2. Mean viability of the tissue replicates exposed for 1 hour with the positive control (8N KOH), expressed as % of the negative control, should be ≤ 15%. 3. In the range 20-100% viability and for ODs ≥ 0.3, difference of viability between the two tissue replicates should not

Nr.	1	2	3	4	5
Test Method Component	**EpiSkin™**	**EpiDerm™ SCT**	**SkinEthic™ RHE**	**epiCS®**	**LabCyte EPI-MODEL24 SCT**
	exceed 30%.		tissue replicates should not exceed 30%.	exceed 30%.	exceed 30%.

ANNEX 3 - PERFORMANCE OF TEST METHODS FOR SUB-CATEGORISATION

The table below provides the performances of the five test methods calculated based on a set of 79 or 80 chemicals tested by the five test developers. Calculations of four test methods (EpiSkin™, EpiDerm™ SCT, SkinEthic™ RHE and epiCS®) were performed by the OECD Secretariat, reviewed and agreed by an expert subgroup (21) (23). Calculation of LabCyte EPI-MODEL24 SCT was performed by the test developer, reviewed and agreed by the validation management group and a peer review panel (41) (43).

STATISTICS ON PREDICTIONS OBTAINED ON THE ENTIRE SET OF CHEMICALS (n= 80 chemicals tested over 2 independent runs for epiCS® or 3 independent runs for EpiDerm™ SCT, EpiSkin™ and SkinEthic™RHE *i.e.* respectively 159* or 240 classifications.
n= 79 chemicals tested over 3 independent runs for LabCyte EPI-MODEL24 SCT, *i.e.* 237 classification.)**
*one chemical was tested once in epiCS® because of no availability (23).
** one chemical was not tested in LabCyte EPI-MODEL24 SCT because of no availability.

	EpiSkin	**EpiDerm**	**SkinEthic**	**epiCS**	**LabCyte EPI-MODEL24**
Overclassifications:					
1B-and-1C overclassified 1A	21.5%	29.0%	31.2%	32.8%	30.0%
NC overclassified 1B-and-1C	20.7%	23.4%	27.0%	28.4%	18.9%
NC overclassified 1A	0.0%	2.7%	0.0%	0.0%	2.7%
Overclassified as Corrosive	20.7%	26.1%	27.0%	28.4%	21.6%
Global overclassification rate (all categories)	17.9%	23.3%	24.5%	25.8%	21.5%
Underclassifications:					
1A underclassified 1B-and-1C	16.7%	16.7%	16.7%	12.5%	13.9%
1A underclassified NC	0.0%	0.0%	0.0%	0.0%	0.0%
1B-and-1C underclassified NC	2.2%	0.0%	7.5%	6.6%	0.0%
Global underclassification rate (all categories)	3.3%	2.5%	5.4%	4.4%	2.1%
Correct Classifications:					
1A correctly classified	83.3%	83.3%	83.3%	87.5%	86.1%
1B-and-/1C correctly classified	76.3%	71.0%	61.3%	60.7%	70.0%
NC correctly classified	79.3%	73.9%	73.0%	71.62%	78.4%
Overall Accuracy	78.8%	74.2%	70.0%	69.8%	76.4%

ANNEX 4 - Key parameters and acceptance criteria for qualification of an HPLC/UPLC-spectrophotometry system for measurement of MTT formazan extracted from RhE tissues

Parameter	Protocol Derived from FDA Guidance (36)(38)	Acceptance Criteria
Selectivity	Analysis of isopropanol, living blank (isopropanol extract from living RhE tissues without any treatment), dead blank (isopropanol extract from killed RhE tissues without any treatment)	$Area_{interference}$ = 20% of $Area_{LLOQ}$[1]
Precision	Quality Controls (i.e., MTT formazan at 1.6 g/mL, 16 g/mL and 160 g/mL) in isopropanol (n=5)	CV = 15% or = 20% for the LLOQ
Accuracy	Quality Controls in isopropanol (n=5)	%Dev = 15% or = 20% for LLOQ
Matrix Effect	Quality Controls in living blank (n=5)	85% = %Matrix Effect= 115%
Carryover	Analysis of isopropanol after an ULOQ[2] standard	$Area_{interference}$ = 20% of $Area_{LLOQ}$
Reproducibility (intra-day)	3 independent calibration curves (based on 6 consecutive 1/3 dilutions of MTT formazan in isopropanol starting at ULOQ, i.e., 200 g/mL); Quality Controls in isopropanol (n=5)	Calibration Curves:%Dev = 15% or = 20% for LLOQ Quality Controls:
Reproducibility (inter-day)	Day 1: 1 calibration curve and Quality Controls in isopropanol (n=3) Day 2: 1 calibration curve and Quality Controls in isopropanol (n=3) Day 3: 1 calibration curve and Quality Controls in isopropanol (n=3)	%Dev= 15% and CV = 15%
Short Term Stability of MTT Formazan in RhE Tissue Extract	Quality Controls in living blank (n='3)' analysed the day of the preparation and after 24 hours of storage at room temperature	%Dev = 15%
Long Term Stability of MTT Formazan in RhE Tissue Extract, if required	Quality Controls in living blank (n='3)' analysed theday of the preparation and after several days of storageat a specified temperature (e.g., 4ºC, -20ºC, -80ºC)	%Dev = 15%

Note:
[1]LLOQ: Lower Limit of Quantification, defined to cover 1-2% tissue viability, i.e., 0.8 µg/mL.
[2]ULOQ: Upper Limit of Quantification, defined to be at least two times higher than the highest expected MTT formazan concentration in isopropanol extracts from negative controls i.e., 200 µg/mL.