Toxicology in Vitro 34 (2016) 55–70

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

Multi-laboratory evaluation of SkinEthic HCE test method for testing serious eye damage/eye irritation using solid chemicals and overall performance of the test method with regard to solid and liquid chemicals testing

N. Alépée [a,*], E. Adriaens [b], M.H. Grandidier [a], M. Meloni [c], L. Nardelli [a], C.J. Vinall [d], F. Toner [d], C.S. Roper [d], A.R. Van Rompay [e], V. Leblanc [a], J. Cotovio [a]

[a] L'Oréal Research & Innovation, Aulnay Sous Bois, France
[b] Adriaens Consulting, Aalter, Belgium
[c] VitroScreen, Milano, Italy
[d] Charles River Laboratories Edinburgh, EH33 2NE, United Kingdom
[e] VITO NV (Flemish Institute for Technological Research), Mol, Belgium

ARTICLE INFO

Article history:
Received 14 October 2015
Received in revised form 19 February 2016
Accepted 21 February 2016
Available online 15 March 2016

Keywords:
Ocular toxicity
SkinEthic HCE
Reconstructed human corneal tissue
Validation
Reproducibility
Predictive capacity

ABSTRACT

A prospective multicentre study of the reconstructed human corneal epithelial tissue-based *in vitro* test method (SkinEthic™ HCE) was conducted to evaluate its usefulness to identify chemicals as either not classified for serious eye damage/eye irritation (No Cat.) or as classified (Cat. 1/Cat. 2) within UN GHS.

The aim of this study was to demonstrate the transferability and reproducibility of the SkinEthic™ HCE EITS protocol for solids and define its predictive capacity. Briefly, 60 chemicals were three times tested (double blinded) in 3 laboratories and 35 additional chemicals were tested three times in one laboratory. Good within laboratory reproducibility was achieved of at least 95% (57/60) and 96.8% (92/95) for the extended data set. Furthermore, the overall concordance between the laboratories was 96.7% (58/60). The accuracy of the SkinEthic™ HCE EITS for the extended dataset, based on bootstrap resampling, was 81.0% (95% CI: 78.9% to 83.2%) with a sensitivity of 90.5% (95% CI: 88.1% to 92.9%) and specificity of 73.6% (95% CI: 71.7% to 75.5%). Overall, 200 chemicals were tested (105 liquids (EITL protocol) and 95 solids (EITS protocol)) resulting in a sensitivity of 95.2%, specificity of 72.1% and accuracy of 83.7%, thereby meeting all acceptance criteria for predictive capacity.

1. Introduction

Considerable progress has been made in the partial replacement of the regulatory *in vivo* Draize rabbit eye test. Currently, four test methods are accepted by the Organisation for Economic Co-operation and Development (OECD) to classify chemicals as inducing serious eye damage according to the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS Category 1) (UN, 2013). Two test methods are organotypic assays, the Bovine Corneal Opacity and Permeability (BCOP) test method (OECD Test Guideline (TG) 437) and the Isolated Chicken Eye (ICE) test method (OECD TG 438) (OECD, 2013a, 2013b). Two test methods are performed on

Abbreviations: BCOP, Bovine Corneal Opacity and Permeability; BLR, between laboratory reproducibility; C, classified; CAS RN, Chemical Abstracts Service Registry Number; Cat 1, UN GHS classification for chemicals causing irreversible effects on the eye/serious damage to the eye; Cat 2, UN GHS classification for chemicals causing reversible effects on the eye/eye irritation, sub-categorised in 2A (irritant to eyes, eye effects are not fully reversible within 7 days of observation) and 2B (mildly irritant to eyes, eye effects fully reversible within 7 days of observation); CC, Conjunctival Chemosis; CI, confidence interval; CM, Cytosensor Microphysiometer; CO, Corneal Opacity; Conj, CR and/or CC; CR, Conjunctival Redness; CRL, Charles River Laboratories; DRD, Draize eye test Reference Database; EIT, Eye Irritation Test; EITL, Eye Irritation Testing of Liquids; EITS, Eye Irritation Testing of Solids; EURL ECVAM, European Union Reference Laboratory for Alternatives to Animal Testing; HCE, Human Corneal Epithelium; HPLC/UPLC, High/Ultra Performance Liquid Chromatography; ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; ICE, Isolated Chicken Eye; IR, Iritis; LO, L'Oréal; MTT, 3-[4,5-dimethylthiazole-2-yl]-2,5-diphenyltetrazolium bromide; NgC, negative control; NC, not classified; No Cat, chemicals not classified for serious eye damage/eye irritation under UN GHS/EU CLP; NSC, non-specific colour; NSC_{killed}, non-specific colour on killed tissues; NSMTT, non-specific reduction of MTT; OD, optical density; OECD, Organisation for Economic Co-operation and Development; PC, positive control; PBS, phosphate-buffered saline; pers, persistence; RhCE, Reconstructed human Cornea-like Epithelium; RhT, Reconstructed human Tissue; SCNM, study criteria not met; STE, short-time exposure; S.O.P., Standard Operating Procedure; TG, Test Guideline; UN GHS, United Nations Globally Harmonized System of Classification and Labelling of Chemicals; VITO, Flemish Institute for Technological Research; VMG, Validation Management Group; VRM, validation reference method; WLR, within laboratory reproducibility.

* Corresponding author at: L'Oréal Research & Innovation, 1 Avenue Eugène Schueller, 93600 Aulnay-sous-bois, France.
E-mail address: nalepee@rd.loreal.com (N. Alépée).

http://dx.doi.org/10.1016/j.tiv.2016.02.014
0887-2333/

Table 1
Overview of the chemicals tested in the multicentre study.

No.	Chemical	CAS RN	Generic chemical class	Functional group class	UN GHS[a]
1	1H-1,2,4-Triazole-1-ethanol, β-([1,1′-biphenyl]-4-yloxy)-α-(1,1-dimethylethyl)-	55179-31-2	Neutral organic	Alcohol, heterocyclic, aromatic	No Cat
2	2-Propanol, 1-(9H-carbazol-4-yloxy)-3-[[2-(2-methoxyphenoxy)ethyl]amino]-	72956-09-3	Organic base	Amine, alcohol, heterocyclic, aromatic	No Cat
3	Butane, 2,3-dimethyl-2,3-dinitro-	3964-18-9	Neutral organic	Alkane, nitro	No Cat
4	4-Pyrimidinol, 2,5,6-triamino-, 4-(hydrogen sulfate)	1603-02-7	Organic acid	Pyrimidine, sulfonic acid, heterocyclic	No Cat
5	2(1H)-Pyridinone, 6-hydroxy-4,5-dimethyl-	84540-47-6	Organic acid	Heterocyclic, phenol	No Cat
6	Phenol, 2-amino-	95-55-6	Neutral organic	Amine, alcohol, aromatic, phenol	No Cat
7	2(1H)-Pyrimidinethione	1450-85-7	Organic acid	Heterocyclic, thio-urea	No Cat
8	Phenol, 4,4′-methylenebis[2,6-bis(1,1-dimethylethyl)-	118-82-1	Neutral organic	Phenol, aromatic	No Cat
9	Benzenamine, 4,4′-sulfonylbis-	80-08-0	Neutral organic	Sulfoxide, amine, Aromatic (electrophile)	No Cat
10	Benzenesulfonic acid, 4-[2-(4-aminophenyl)diazenyl]-	104-23-4	Organic acid	Sulfuric acid, azoic, amine, aromatic	No Cat
11	Methanimidamide, N ′-(2,4-dimethylphenyl)-N-[[(2,4-dimethylphenyl)imino]methyl]-N-methyl-	33089-61-1	Neutral organic	Aromatic, amine	No Cat
12	Anthracene	120-12-7	Neutral organic	Aromatic	No Cat
13	Benzenamine, 4,4′-[1,4-phenylenebis(1-methylethylidene)]bis-	2716-10-1	Neutral organic	Aromatic	No Cat
14	5H-Dibenz[b,f]azepine, 10,11-dihydro-	494-19-9	Neutral organic	Heterocyclic, aromatic	No Cat
15	Magnesium carbonate hydroxide (Mg5(CO3)4(OH)2), hydrate (1:5)	56378-72-4	Inorganic salt	Magnesium, carbonate	No Cat
16	Phenol, 2,2′-methylenebis[6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)-	103597-45-1	Neutral organic	Phenol, heterocyclic	No Cat
17	Tetradecanoic acid, tetradecyl ester	3234-85-3	Neutral organic	Ester	No Cat
18	Guanidine, N,N-dimethyl-, sulfate (2:1)	598-65-2	Organic base	Amidine, salt, sulfate	No Cat
19	10H-Phenothiazine	92-84-2	Neutral organic	Amine, heterocyclic, thioether, aromatic	No Cat
20	3,5-Pyrazolidinedione, 4-butyl-1,2-diphenyl	50-33-9	Neutral organic	Amide, heterocyclic, aromatic	No Cat
21	Thiourea, N-phenyl-	103-85-5	Neutral organic	Amine, thioimine, aromatic, thio-urea	No Cat
22	Cellulose, ether with α-[2-hydroxy-3-(trimethylammonio)propyl]-ω-hydroxypoly(oxy-1,2-ethanediyl), chloride	68610-92-4	Cationic soap/surfactant	Amonium, polyether, polyglucosides	No Cat
23	Borate(1-), tetrafluoro-, potassium (1:1)	14075-53-7	Inorganic salt	Fluoroborate salts	No Cat
24	Benzoic acid, 4-hydroxy-, propyl ester	94-13-3	Neutral organic	Ester, phenol	No Cat
25	Silicic acid	1343-98-2	Inorganic acid	Silicium	No Cat
26	L-Ascorbic acid, 2-(dihydrogen phosphate), sodium salt (1:3)	66170-10-3	Organic acid	Phosphate, salt, alcohol	No Cat
27	Sulfurous acid, sodium salt (1:1)	7631-90-5	Inorganic acid	Sulfur compound, acid	No Cat
28	1H-Purine-2,6-dione, 3,7-dihydro-3,7-dimethyl-	83-67-0	Neutral organic	Purine, heterocyclic	No Cat
29	Urea, N-(4-chlorophenyl)-N′-(3,4-dichlorophenyl)-	101-20-2	Neutral organic	Urea, halogenated, aromatic, phenol	No Cat
30	Benzaldehyde, 3,4-dimethoxy-	120-14-9	Neutral organic	Aldehyde, ether, aromatic (electrophile)	No Cat
31	Bicyclo[2.2.1]heptane, 2,2-dimethyl-3-methylene	79-92-5	Neutral organic	Alkane cyclic	Cat 2B
32	1,4-Naphthalenedione, 2-hydroxy-	83-72-7	Neutral organic	Phenol, quinone (electrophile)	Cat 2B
33	Benzene, 1,3-dinitro-	99-65-0	Neutral organic	Nitro, aromatic	Cat 2B
34	Benzoic acid, 4-nitro-	62-23-7	Organic acid	Carboxylic acid, nitro	Cat 2B
35	Acetic acid, 2-chloro-, sodium salt (1:1)	3926-62-3	Neutral organic	Alkyl, halide, carboxylic acid (electrophile), acid chloride, salt	Cat 2B
36	2-Azetidinone, 4-(acetyloxy)-3-[(1R)-1-[[(1,1-dimethylethyl)dimethylsilyl]oxy]ethyl]-, (3R,4R)-	76855-69-1	Neutral organic	Ester, amide, siloxy	Cat 2A

Table 1 (*continued*)

No.	Chemical	CAS RN	Generic chemical class	Functional group class	UN GHS[a]
37	1,5-Naphthalenediol	83-56-7	Neutral organic	Phenol	Cat 2A
38	3-Pyridinol, 2-amino-	16867-03-1	Organic acid	Pyridine, amine, hydroxyl	Cat 2A
39	Propanoic acid, 3,3′-dithiobis-	1119-62-6	Organic acid	Carboxylic acid, disulfide	Cat 2A
40	Benzoic acid, 4-formyl-	619-66-9	Neutral organic	Carboxylic acid, aromatic (electrophile)	Cat 2A
41	Nitric acid ammonium salt (1:1)	6484-52-2	Inorganic base	Amine, nitrate, ammonium	Cat 2A
42	Phosphoric acid, bis(phenylmethyl) ester	1623-08-1	Organic acid	Organophosphorus compound, phosphate acid, aromatic	Cat 2A
43	Manganese, [N-[2-[(dithiocarboxy)amino]ethyl]carbamodithioato(2-)-κS,κS′]-	12427-38-2	Inorganic salt	Ester, thiocarbamate, Organometallic compound, amine, salt	Cat 2A[b]
44	Benzoic acid, sodium salt (1:1)	532-32-1	Organic salt	Carboxylic acid	Cat 2A
45	1H-1,2,4-Triazole, sodium salt (1:1)	41253-21-8	Organic base	Aryl triazole, triazole salt	Cat 1
46	1-Naphthaleneacetic acid	86-87-3	Organic acid	Carboxylic acid	Cat 1
47	Benzoic acid, 2,4-dihydroxy-	89-86-1	Organic acid	Acid, phenol	Cat 1
48	2,5-Hexanediol, 2,5-dimethyl-	110-03-2	Neutral organic	Alcohol, polyols	Cat 1
49	Phenol, 4-chloro-2-(phenylmethyl)-	120-32-1	Neutral organic	Phenol, halogen	Cat 1
50	1H-Pyrazole, 3,4-dimethyl-	2820-37-3	Neutral organic	Heterocyclic, pyrazole	Cat 1
51	Phenol, 4-(1,1,3,3-tetramethylbutyl)-	140-66-9	Neutral organic	Phenol	Cat 1
52	Pentanedioic acid, 2-oxo	328-50-7	Organic acid	Carboxylic acid, ketone	Cat 1
53	1,2-Benzisothiazol-3(2H)-one	2634-33-5	Neutral organic	Isothiazolinone, heterocyclic (electrophile)	Cat 1
54	Benzoic acid	65-85-0	Organic acid	Aromatic, carboxylic acid	Cat 1
55	2,4,11,13-Tetraazatetradecanediimidamide, N1,N14-bis(4-chlorophenyl)-3,12-diimino-	55-56-1	Organic base	Guanidine, halogenated, aromatic	Cat 1
56	Dodecanoic acid	143-07-7	Anionic soap/surfactant	Carboxylic acid, alcane	Cat 1
57	Methionine, N-acetyl-	1115-47-5	Organic acid	Acid, thioether, amide	Cat 1
58	10H-Phenothiazine-10-ethanamine, N,N,α-trimethyl-, hydrochloride (1:1)	58-33-3	Organic base	Amine, thioether, heterocyclic compound	Cat 1
59	Ethanedioic acid, sodium salt (1:2)	62-76-0	Organic acid	Carboxylic acid, salt	Cat 1
60	Phenol, 4,4′-(4,5,6,7-tetrabromo-1,1-dioxido-3H-2,1-benzoxathiol-3-ylidene)bis[2,6-dibromo-	4430-25-5	Neutral organic	Halogenated, phenol, sulfoxide	Cat 1

[a] Cat 1 is defined as causing irreversible effects on the eye/serious damage to the eye, (Cat 2) as causing reversible effects (fully reversible within 21 days) on the eye/eye irritation with Cat 2A (irritant to eyes) when the eye effects are not fully reversible within 7 days of observation and Cat 2B (mildly irritant to eyes) when the eye effects fully reverse within 7 days of observation. No Cat is defined as not requiring classification for serious eye damage/eye irritation.

[b] Study criteria are not met (SCNM) for the *in vivo* Draize rabbit eye test, the UN GHS classification corresponds with at least a Cat 2A, the results of this study are published in the Eye Irritation Reference Chemicals data bank published by the European Centre for Toxicology and Ecotoxicology of Chemicals (ECETOC, 1998). The study was terminated on day 7, in four animals all tissue scores fully reversed to 0 by day 7, one animal had CR = 1 and another animal had CR and CC = 1.

confluent cell monolayers; the Fluorescein Leakage (FL) test method (OECD TG 460) is a cytotoxicity and cell function-based assay and the Short-Time Exposure (STE) test method (OECD TG 491) is a cytotoxicity based assay (OECD, 2012; OECD, 2015a). Furthermore, at this time, three methods are adopted by the OECD for the identification of chemicals not requiring a classification for serious eye damage/eye irritation (UN GHS No Category). The organotypic BCOP (OECD TG 437) and ICE (OECD TG 438) test methods were adopted for this purpose (OECD, 2013a, 2013b). The third test method uses a three-dimensional Reconstructed human Cornea-like Epithelium (RhCE) (OECD TG 492) and measures cytotoxicity (OECD, 2015b). OECD TG 492 covers at this time only the commercially available EpiOcular™ Eye Irritation Test (EIT) validated by the European Union Reference Laboratory for Alternatives to Animal testing (EURL ECVAM) as a result of the EURL ECVAM/Cosmetics Europe prospective validation study (Freeman et al., 2010). In addition, the Cytosensor Microphysiometer (CM) (Hartung et al., 2010), a cytotoxicity and cell function-based method, has been endorsed as scientifically valid for the identification Cat 1 and No Cat chemicals for limited applicability domains (ESAC, 2009; ICCVAM, 2010). This method is currently in the process of review by the OECD. A more detailed description of the principle and background of the methods listed above is presented in the review published by Wilson et al. (2015).

Despite all these efforts, full replacement of the *in vivo* Draize rabbit eye test is, however, yet to be achieved. During a workshop held in 2005 by the European Centre for the Validation of Alternative Methods (ECVAM), a testing scheme was proposed using a Bottom-Up or Top-Down progression of *in vitro* tests (Scott et al., 2010). The outcome of this expert meeting identified that test methods using RhCE could be considered for incorporation into a testing strategy as an initial step in a Bottom-Up approach or the second step in a Top-Down approach. Currently, there are two such test methods available namely, the EpiOcular™EIT (Kaluzhny et al., 2011; Pfannenbecker et al., 2013; OECD, 2015b) and the SkinEthic™ Human Corneal Epithelium (HCE) test (Van Goethem et al., 2006; Cotovio et al., 2007, 2010; Alépée et al., 2013). Recently, the SkinEthic™ HCE model has been optimized for the evaluation of liquids (Eye Irritation Testing of Liquids, EITL) in a multicentre prospective study demonstrating an overall predictive capacity of 84.4% with 99.0% sensitivity and 68.5% specificity (Alépée et al., 2016).

The current paper presents an optimization of the SkinEthic™ HCE test method for the Eye Irritation Testing of Solids (EITS protocol). The primary aim of this multicentre study was to assess the reliability and relevance of the test method to discriminate solid chemicals not requiring classification for serious eye damage/eye irritancy (No Cat) from chemicals requiring classification and labelling (Cat 1 and Cat 2). In addition, the discrimination between No Cat and Cat 1/Cat 2 chemicals is based on measuring tissue viability in the RhCE constructs and this is usually determined by enzymatic reduction of yellow 3-[4,5-

dimethylthiazole-2-yl]2,5-diphenyltetrazolium bromide (MTT) to purple reduced MTT (formazan) (Mosmann, 1983). Formazan is quantified photometrically and viability is calculated from the percentage MTT conversion in the treated cultures relative to the corresponding negative controls (100% viability). A known limitation of the photometric MTT-reduction assay is the possible interference of strongly coloured test chemicals with the absorbance measurement of formazan. Recently, Cosmetics Europe demonstrated the usefulness of High/Ultra Performance Liquid Chromatography (HPLC/UPLC)–spectrophotometry for detection of formazan in *in vitro* Reconstructed human Tissue (RhT)-based test methods (Alépée et al., 2015). Test chemicals that interfere too strongly with the MTT-reduction assay may still be assessed using HPLC/UPLC photometry instead of standard absorbance (OD), in fact as described in OECD TG 492, the extracted MTT formazan may be quantified using either standard absorbance (OD) or HPLC/UPLC spectrophotometry (OECD, 2015b). Therefore, in the current study the applicability of the HPLC/UPLC–spectophotometry assay for the determination of tissue viability in the SkinEthic™ HCE test method was also used. The second aim of this paper was to evaluate the reliability and relevance of the SkinEthic™ HCE EITS and EITL protocol using a total of 200 chemicals. In two multicentre studies, 120 chemicals (60 liquids and 60 solids) were tested in three laboratories, detailed results of the 60 solids will be presented in the current paper and the results of the liquids were published recently by Alépée and co-workers (Alépée et al., 2016). An additional 80 chemicals (45 liquids and 35 solids) were tested by the test method developer L'Oréal only to increase the range of chemical classes.

2. Materials and methods

2.1. Tissues, media and reagents

The SkinEthic™ HCE tissues and tissue maintenance medium were purchased from Episkin SA (Lyon, France). Tissues were shipped in agarose semi-solid culture medium. Upon receipt, the tissue cultures were placed onto 1 mL fresh maintenance medium (6-well plate) and incubated overnight in standard culture conditions (37 °C, 5% CO_2, ≥95% humidity). Following this equilibration period, the cultures were then transferred into a 24-well plate containing 300 μL fresh maintenance medium per well.

3-[4,5-Dimethylthiazole-2-yl]-2,5-diphenyltetrazolium bromide reagent (MTT, CAS RN ??), Ca^{2+}- and Mg^{2+}-free Dulbecco's phosphate-buffered saline (PBS), and methyl acetate were purchased from Sigma (CAS RN 79-20-9, Sigma-Aldrich, France).

2.2. Chemicals

2.2.1. Validation set

A total of 60 solid chemicals representing different chemical classes were selected and are listed in Table 1. The chemicals were chosen to provide a balanced representation of chemicals not requiring classification (n = 30) and chemicals inducing serious eye damage/eye irritation (Cat 1, n = 16/Cat 2, n = 14). MTT and/or colour interfering chemicals were also selected. All chemicals were sourced and blind coded independently for each laboratory and distributed to the testing laboratories by VitroScreen (Milano, Italy). Fig. 1 shows a scheme of the management structure of the study. PBS and methyl acetate were used as negative control (NgC) and positive control (PC), respectively. Each laboratory tested each chemical in at least three independent runs performed with different tissue batches. During each run, a maximum of 13 solid test chemicals, NgC and PC were all tested concurrently on two tissue replicates. A test chemical concurrently tested on two tissue replicates is called a test herein after.

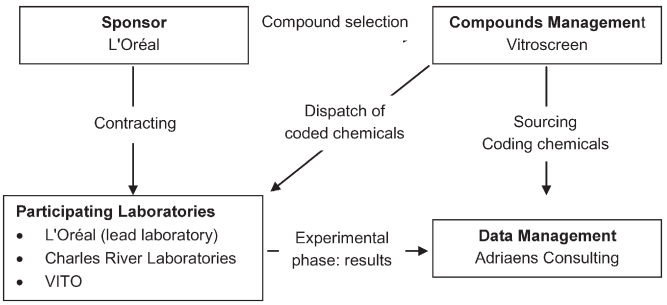

Fig. 1. Management structure of the SkinEthic™ HCE EITS validation study.

2.2.2. Additional chemicals

In order to enlarge the chemical diversity and to increase the dataset for evaluating the predictive capacity of the SkinEthic™ HCE EITS protocol, 35 additional chemicals were evaluated unblinded by L'Oréal (LO) in three independent runs. The chemicals represented 23 non-classified and 12 classified chemicals (Table 2). In total, 95 chemicals (53 non-classified and 42 classified chemicals, consisting of 24 Cat 1 and 18 Cat 2 chemicals) were evaluated on SkinEthic™ HCE test method for the Eye Irritation Testing of Solids.

2.3. Participating laboratories

The within and between laboratory reproducibility (WLR and BLR, respectively) of the SkinEthic™ HCE EITS method was assessed in three laboratories: L'Oréal (L'Oréal Research & Innovation, Aulnay sous Bois, France), Charles River Laboratories (CRL, Edinburgh, United Kingdom) and VITO NV (Flemish Institute for Technological Research, Mol, Belgium).

2.4. Technology transfer

Identical protocols and Excel templates for data collection were transferred to each laboratory. CRL and VITO received formal hands-on training in assay methodology and analysis from L'Oréal Research & Innovation, using the EITS protocol. The laboratory assistants tested 9 commercially available chemicals in at least three independent runs. This set of chemicals contained a colourant (quinacrine dihydrochloride), an MTT reducer (dihydroxy-2.6-toluene) and an MTT reducing colourant (tetrabromophenol blue). The colourants were selected in order to evaluate the crucial rinsing step procedure and the additional controls which are needed for tissue colouring chemicals. The MTT interacting chemicals were chosen with the intention to perform the specific controls for direct MTT reduction by chemicals.

2.5. SkinEthic™ HCE tissue

2.5.1. Principle of the test system

The SkinEthic™ HCE model uses immortalized human corneal epithelial cells cultured in a chemically defined medium. When cultured at the air–liquid interface on a permeable synthetic membrane insert, the epithelial cells stratify and differentiate into a 3-dimensional tissue which bears close resemblance to normal human corneal epithelium. The tissue construct contains at least four viable layers including columnar basal cells, transitional wing cells and superficial squamous cells. Other structural features of corneal tissue, such as the presence of mature desmosomes and intermediate filaments, as well as the expression of corneal specific cytokeratin 64 kD (K.3) similar to that of the normal human corneal epithelium, have been described (Nguyen et al., 2003).

2.5.2. Eye Irritation Test Solid protocol

SkinEthic™ HCE tissues (0.5 cm^2) were topically exposed to 30 mg ± 2 mg of solid chemical for 4 h ± 5 min at 37 °C at 5% CO_2 in a humidified incubator (standard culture conditions). If necessary,

Table 2
Overview of the extended solid chemical dataset.

No.	Test substance	CAS RN	Generic chemical class	Functional group class	UN GHS[a]
61	[1,2,4]Triazolo[1,5-c]pyrimidine-2-sulfonamide, N-(2,6-difluorophenyl)-8-fluoro-5-methoxy-	145701-23-1	Organic acid	Heterocyclic, sulfonamide, aromatic, fluor, ether	No Cat
62	1-Propanaminium, N,N,N-trimethyl-3-[(1-oxo-2-propen-1-yl)amino]-, chloride (1:1), polymer with 2-propenamide	75150-29-7	Organic cationic	Acrylate, ammonium (electrophile)	No Cat
63	1,3,5,7-Tetraazatricyclo[3.3.1.13,7]decane	100-97-0	Organic base	Amine, hydrocarbon, cyclic	No Cat
64	2-Pyrimidinamine, 4,6-dimethyl-N-phenyl-	53112-28-0	Neutral organic	Aromatic, amine, heterocyclic, pyrimidine	No Cat
65	2H-1,2,4-Benzothiadiazine-7-sulfonamide, 6-chloro-3-[[(phenylmethyl)thio]methyl]-, 1,1-dioxide	91-33-8	Neutral organic	Heterocyclic, thioether, sulfonyl	No Cat
66	3-Pyrazolidinone, 1-phenyl-	92-43-3	Neutral organic	Heterocyclic	No Cat
67	3H-Pyrazol-3-one, 2-(4-aminophenyl)-2,4-dihydro-5-(1-pyrrolidinyl)-	30707-77-8	Organic base	Heterocyclic, amine, aromatic	No Cat
68	4-Quinazolinamine, N-[3-chloro-4-[(3-fluorophenyl)methoxy]phenyl]-6-iodo-	231278-20-9	Neutral organic	Heterocyclic, quinazoline, halogenated, aromatic, ether	No Cat
69	4H-1,3,5-Oxadiazin-4-imine, 3-[(2-chloro-5-thiazolyl)methyl]tetrahydro-5-methyl-N-nitro-	153719-23-4	Neutral organic	Guanidine, heterocyclic, ether	No Cat
70	Aluminium hydroxide (Al(OH)3)	21645-51-2	Inorganic acid	Hydroxide, aluminium	No Cat
71	Benzene, 2-chloro-1-fluoro-4-nitro-	350-30-1	Neutral organic	Halogenated, nitro, aromatic	No Cat
72	Benzoic acid, 2-[4-(diethylamino)-2-hydroxybenzoyl]-, hexyl ester	302776-68-7	Organic acid	Aromatic, amine, ester, ketone, phenol	No Cat
73	Benzoic acid, 4-iodo-2-[[[[(4-methoxy-6-methyl-1,3,5-triazin-2-yl)amino]carbonyl]amino]sulfonyl]-, methyl ester, sodium salt (1:1)	144550-36-7	Organic acid	Sulfonamide, urea, triazine, ester, aromatic, heterocyclic, salt	No Cat
74	Benzoic acid, 4,4′,4″-(1,3,5-triazine-2,4,6-triyltriimino)tris-, 1,1′,1″-tris(2-ethylhexyl) ester	88122-99-0	Neutral organic	Ester, heterocyclic	No Cat
75	Benzoic acid, 5-(aminosulfonyl)-2,4-dichloro-	2736-23-4	Organic acid	Halogenated, carboxylic acid, sulfonyl	No Cat
76	Ethanol, 2,2′-[[3-methyl-4-[2-(4-nitrophenyl)diazenyl]phenyl]imino]bis-	3179-89-3	Neutral organic	Alcohol, aromatic, amine, azo, nitro	No Cat
77	Glutamic acid, hydrate (1:1)	19285-83-7	Organic acid, organic base	Amine, carboxylic acid	No Cat
78	Glycine, N,N′-1,2-ethanediylbis[N-(carboxymethyl)-, potassium salt, hydrate (1:2:2)	25102-12-9	Organic acid, organic base	Salt, amine, carboxylic acid, aminoacid salt	No Cat
79	Methanaminium, 1-carboxy-N,N,N-trimethyl-, hydroxide (1:1)	590-47-6	Organic acid	Carboxylic acid, ammonium	No Cat
80	Pentanenitrile, 4,4-dimethyl-3-oxo-	59997-51-2	Neutral organic	Nitrile, ketone	No Cat
81	Perylo[3,4-cd:9,10-c'd']dipyran-1,3,8,10-tetrone, 5,6,12,13-tetrachloro-	156028-26-1	Neutral organic	Halogenated, anhydride, aromatic (electrophile)	No Cat
82	Phenol, 2,2′-[6-(4-methoxyphenyl)-1,3,5-triazine-2,4-diyl]bis[5-[(2-ethylhexyl)oxy]-	187393-00-6	Neutral organic	Ether, phenol, triazine, aromatic, heterocyclic	No Cat
83	Phosphorothioic acid, O,O-diethyl O-(3,5,6-trichloro-2-pyridinyl) ester	2921-88-2	Neutral organic	Halogenated, heterocyclic, ester, organophosphates	No Cat
84	2,6-dichloro-5-fluoro-beta-oxo-3-pyridinepropanoate	96568-04-6	Neutral organic	Pyridine, halogen, ketone, ester (electrophile)	Cat 2B
85	Benzene, 1,4-dibutoxy-	104-36-9	Neutral organic	Ether, aromatic	Cat 2B
86	Cyclopropanecarboxylic acid, 3-[(1Z)-2-chloro-3,3,3-trifluoro-1-propen-1-yl]-2,2-dimethyl-, (2-methyl[1,1′-biphenyl]-3-yl)methyl ester, (1R,3R)-rel-	82657-04-3	Neutral organic	Ester, halogenated (electrophile)	Cat 2B
87	1,3-Benzenediamine, 4,4′-[1,3-propanediylbis(oxy)]bis-, hydrochloride (1:4)	74918-21-1	Organic base	Ether, aromatic, amide	Cat 2A
88	1,3-Benzenediol, 2-methyl-	608-25-3	Neutral organic	Phenol	Cat 1
89	3-Piperidinemethanol, 4-(4-fluorophenyl)-1-methyl-, (3S,4R)-	105812-81-5	Organic base	Amine, alcohol, fluor, aromatic	Cat 1
90	4,7-Methanoisobenzofuran-1,3-dione, 4,5,6,7,8,8-hexachloro-3a,4,7,7a-tetrahydro-	115-27-5	Neutral organic	Hydrocarbon cycle, halogenated, anhydrous (electrophile)	Cat 1[b]
91	Benzene, 1,2-dichloro-4-isocyanato-	102-36-3	Neutral organic	Halogenated, aromatic, isocyanate (electrophile),	Cat 1
92	Benzenesulfonic acid, 2,2′-([1,1′-biphenyl]-4,4′-diyldi-2,1-ethenediyl)bis-, sodium salt (1:2)	27344-41-8	Organic acid	Sulfonic acid, aromatic	Cat 1
93	Quinacrine dihydrochloride	69-05-6	Organic base	Amine, halogen, aromatic, heterocyclic	Cat 1
94	Spiro[isobenzofuran-1(3H),9′-[9H]xanthen]-3-one, 2′,4′,5′,7′-tetrabromo-4,5,6,7-tetrachloro-3′,6′-dihydroxy-, sodium salt (1:2)	18472-87-2	Organic acid	Phenol, halogenated aromatic, salt, ester	Cat 1
95	Sulfuric acid, sodium salt (1:1)	7681-38-1	Inorganic acid	Sulfonic acid, salt	Cat 1

[a] Cat 1 is defined as causing irreversible effects on the eye/serious damage to the eye, (Cat 2) as causing reversible effects (fully reversible within 21 days) on the eye/eye irritation with Cat 2A (irritant to eyes) when the eye effects are not fully reversible within 7 days of observation and Cat 2B (mildly irritant to eyes) when the eye effects fully reverse within 7 days of observation. No Cat is defined as not requiring classification for serious eye damage/eye irritation.

[b] SCNM for the *in vivo* Draize rabbit eye test, the UN GHS classification corresponds with a Cat 1. The summary results of this study are published in the DRD (Barroso et al., 2016). The study was terminated on day 14 with CO = 2 in 3/6, CO = 1 in 2/6, IR = 1 in 4/6, CR = 1 in 5/6, CC = 2 in 1/6 and CC = 1 in 4/6 animals.

the test substance was first crushed to a very fine powder and before applying the chemical on the tissue, 30 μL PBS was pipetted onto the epithelium to improve optimal contact of the powder with the epithelium. Two tissues were used per test substance (NgC, PC, or chemical). After 4 hour treatment, tissues were rinsed with 25 mL of PBS to remove the residual test chemical from the tissue surface. After rinsing, the tissues were immersed into 4 mL fresh maintenance medium at room temperature for 30 min ± 2 min. At the end of the post-soak immersion, tissues were transferred to a new 6-well plate containing 1 mL of maintenance medium and were incubated for 18 h ± 30 min at standard culture conditions. After the incubation period, duplicate tissues were assessed for tissue viability.

2.5.3. Tissue viability assessment

Following the 18 hour post-incubation period, tissues were carefully rinsed with PBS. Each tissue was transferred to a new well containing 300 μL of freshly prepared MTT (1 mg/mL) solution for a 3 h ± 15 min incubation period under standard culture conditions. Then the tissue inserts were rinsed with 300 μL PBS and transferred into new plates containing 1.5 mL of isopropanol per well for at least 2 h to extract the reduced MTT (formazan crystals) out of the tissues. 200 μL aliquots of formazan solution extracts were transferred to 96-well flat bottom microtiter plates for optical density (OD) measurement using a spectrophotometer equipped with a 570 nm filter (filter band pass ±30 nm). Isopropanol was used as a blank. The percentage viability of each of the treated cultures was calculated from the percentage MTT conversion in the treated cultures relative to the corresponding negative controls (100% viability). Results were expressed as mean OD and mean % viability and the difference of viability between the two replicate tissues.

2.5.4. Assessment for direct MTT reduction by the chemical

Possible interference of each chemical, *e.g.* ability to reduce MTT in absence of tissue, was verified before the start of the experiment. Therefore, 30 mg of the chemical was added to 300 μL of MTT solution (1 mg/mL), the mixture was incubated at 37 °C protected from light for 3 h. If the MTT solution colour turns blue or purple, the chemical interacts with the MTT. In case of MTT interaction, non-specific reduction of the MTT by the chemical needs to be determined in a separate experiment by using killed epithelial tissues. The killed tissues were treated with the chemical, rinsed and exposed to MTT according to the standard protocol. In addition, two killed tissues were treated with PBS as control. For the determination of the final viability, this non-specific reduction of MTT (%NSMTT) was taken into account and the viability was 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).

2.5.5. Adapted controls for colouring chemicals

Coloured chemicals or chemicals able to develop a colour after contact with the tissue can generate a remaining non-specific colour (NSC). Therefore, each chemical was checked on a single occasion, for its colourant properties. In order to determine non-specific colouring, all steps of the EITS protocol were followed except the MTT incubation since 300 μL of maintenance medium was dispensed instead of MTT medium. The % NCS was determined after isopropanol extraction and OD reading in similar conditions. For the determination of the final viability, the % NCS was taken into account and the viability was calculated as: the percent tissue viability obtained with living

Table 3
Characteristics of the chemicals tested with photometric MTT-reduction and HPLC/UPLC–spectrophotometry.

No.	Chemicals	CAS RN	MTT reducer (Y/N)	Colour interference (Y/N)	Physical state
45[c]	1,2-Ethanediamine, N1-[3-(trimethoxysilyl)propyl]-	1760-24-3	Y	N	Liquid
20[c]	1,3-Benzodioxole, 5-[[2-(2-butoxyethoxy)ethoxy]methyl]-6-propyl-	51-03-6	Y	N	Liquid
96	1H-1,2,4-Triazole, 1-[[2-(2,4-dichlorophenyl)-4-propyl-1,3-dioxolan-2-yl]methyl]-	60207-90-1	N	N	Liquid
8[c]	2-Propenoic acid, 2-methyl-, 2-ethoxyethyl ester	2370-63-0	N	N	Liquid
97	Basazol C Blue pr 8056	1686090-84-5	Y[a]	Y	Liquid
32[c]	Benzamide, N,N-diethyl-3-methyl-	134-62-3	N	N	Liquid
98	Benzenamine, 4-[(2,6-dichlorophenyl)(4-imino-3,5-dimethyl-2,5-cyclohexadien-1-ylidene)methyl]-2,6-dimethyl-, phosphate (1:1) 1%	74578-10-2	Y[a]	Y	Liquid
99	Benzenamine, 4,4′-[(4-imino-3-methyl-2,5-cyclohexadien-1-ylidene)methylene]bis[2-methyl-, hydrochloride (1:1) 1% (w/v) aqueous	3248-91-7	Y	Y	Liquid
100	Benzene, (1-methylpropyl)-	135-98-8	N	N	Liquid
40[c]	D-Gluconic acid, compd. with N1,N14-bis(4-chlorophenyl)-3,12-diimino-2,4,11,13-tetraazatetradecanediimidamide (2:1) (20% aq)	18472-51-0	N	N	Liquid
101	Poly(oxy-1,2-ethanediyl), α-sulfo-ω-hydroxy-, C12-14-alkyl ethers, sodium salts (30% aq)	68891-38-3	N	N	Liquid
88	1,3-Benzenediol, 2-methyl-	608-25-3	Y	N	Solid
102	1,4-Benzenediamine, 2-nitro-N1-2-propen-1-yl-	160219-76-1	Y	Y	Solid
103	2-Anthracenesulfonic acid, 4-[[4-(acetylamino)phenyl]amino]-1-amino-9,10-dihydro-9,10-dioxo-, sodium salt (1:1)	6424-85-7	Y	Y	Solid
35	Acetic acid, 2-chloro-, sodium salt (1:1)	3926-62-3	N	N	Solid
104	Benzenamine, 4-[(2,6-dichlorophenyl)(4-imino-3,5-dimethyl-2,5-cyclohexadien-1-ylidene)methyl]-2,6-dimethyl-, phosphate (1:1)	74578-10-2	Y[a]	Y	Solid
105	Benzenamine, 4,4′-[(4-imino-3-methyl-2,5-cyclohexadien-1-ylidene)methylene]bis[2-methyl-, hydrochloride (1:1)	3248-91-7	Y[a]	Y	Solid
106	Ethanol, 2,2′-[[4-[(2-methoxyethyl)amino]-3-nitrophenyl]imino]bis-	23920-15-2	Y	Y	Solid
26	L-Ascorbic acid, 2-(dihydrogen phosphate), sodium salt (1:3)	66170-10-3	Y	N	Solid
41	Nitric acid ammonium salt (1:1)	6484-52-2	N	N	Solid
60	Phenol, 4,4′-(4,5,6,7-tetrabromo-1,1-dioxido-3H-2,1-benzoxathiol-3-ylidene)bis[2,6-dibromo-	4430-25-5	Y	Y	Solid
107	Quinoxaline, 1,2,3,4-tetrahydro-6-nitro-	41959-35-7	Y	Y	Solid
29	Urea, N-(4-chlorophenyl)-N′-(3,4-dichlorophenyl)-	101-20-2	N	N	Solid
108	Xanthylium, 3,6-bis(diethylamino)-9-[2-(methoxycarbonyl)phenyl]-, tetrafluoroborate(1-) (1:1)	134429-57-5	Y[b]	Y	Solid

[a] Not applicable due to incompatibility or overflow results obtained using optical density endpoint parameter.
[b] Difference > 20% for the adapted controls.
[c] Detailed information on the irritation potency of the liquid chemicals is presented in Alépée et al., 2016. As such these No correspond to the sequence No for the liquid chemicals as defined in Alépée et al., 2016.

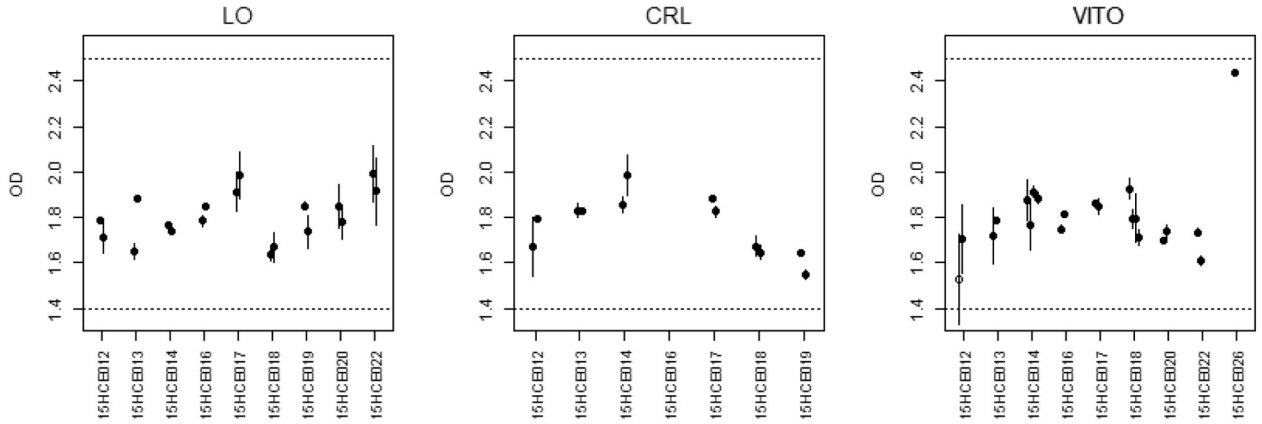

Fig. 2. OD values for the negative control for the different batches. Dots correspond with the mean of two tissues, bars correspond with the single values. Up to 4 runs were performed with the same batch (15HCE014 and 15HCE018) for VITO. (●) Qualified test and (○) unqualified test.

tissues exposed to the colour interfering chemical and incubated with MTT solution minus the percent non-specific colour obtained with living tissues exposed to the colour interfering test chemical and incubated with medium without MTT, run concurrently to the test being corrected (%NSC).

Therefore a coloured chemical can, in some cases, interfere with the MTT pre-check. In that case, each colouring chemical was applied onto two killed tissues and incubated in maintenance medium instead of MTT solution to determine the non-specific colour on killed tissues (NSC_{killed}). The final viability was calculated as: the percent tissue viability obtained with living tissues exposed to the test chemical minus %NSMTT minus %NSC 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 ran concurrently to the test being corrected ($\%NSC_{killed}$).

2.6. Prediction model

Based on the relative viability, the SkinEthic™ HCE EITS protocol can distinguish between chemicals not requiring classification for serious eye damage/eye irritancy (No Cat) from chemicals requiring classification and labelling (Cat 1 and Cat 2) according to UN GHS. No Cat (NC; UN GHS No Cat) corresponding with chemicals that result in a mean tissue viability >50%, classified (C; UN GHS Cat 1/Cat 2) correspond with chemicals that result in a mean tissue viability ≤50%.

2.7. Acceptance criteria

A run was considered qualified if the following criteria were met, mean OD of the NgC was ≥1.4 and ≤2.5 and mean % viability of the PC was ≤30. In addition, the difference of viability between the two replicate tissues of a single test chemical was ≤20 in the same run whatever the test item (for PC, NgC, test chemical and all adapted controls) (SkinEthic™ HCE EIT SOP, 2015). As mentioned before, each chemical was tested three times in three independent runs in each laboratory. If a test did not meet the acceptance criteria in a run, a maximum of two additional independent tests was performed for each chemical in each laboratory.

2.8. Comparison between photometric MTT-reduction and HPLC/UPLC–spectrophotometry for the assessment of tissue viability reword

A total of 24 chemicals, 11 liquids and 13 solids, representing non-coloured and coloured chemicals were selected and are listed in Table 3. The majority of the chemicals (22/24) were already tested in a previous study performed by Cosmetics Europe where the usefulness of the HPLC/UPLC–spectrophotometry to measure formazan in RhT systems for eye/skin irritation and skin corrosion was demonstrated (Alépée et al., 2015, 2016). Prior to the testing, each chemical was checked for its colourant and/or MTT reducing properties as described in Sections 2.5.4 and 2.5.5 in order to determine the use of adapted controls for the determination of non-specific colouration and/or MTT reduction. The viability of the chemicals was assessed

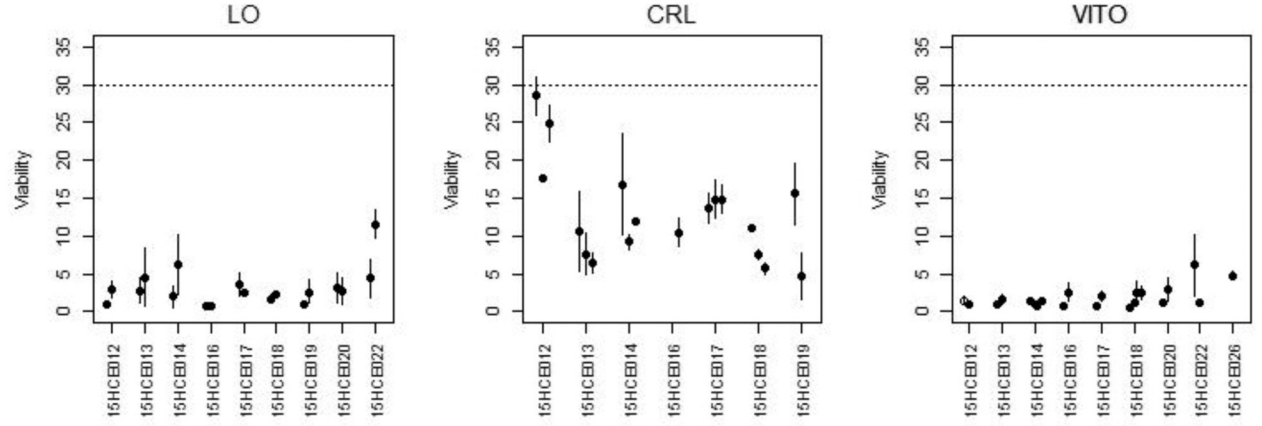

Fig. 3. Viability for the positive control for the different batches. Dots correspond with the mean of two tissues, bars correspond with the single values. Up to 4 runs were performed with the same batch (15HCE014 and 15HCE018). (●) Qualified test and (○) unqualified test.

Table 4
WLR for SkinEthic™ HCE EITS protocol. Mean cell viability (n = 2) for 60 chemicals and concordance of prediction within a laboratory.

No.	*In vivo* UN GHS	LO				CRL				VITO			
		Run 1	Run 2	Run 3	WLR concord.	Run 1	Run 2	Run 3	WLR concord.	Run 1	Run 2	Run 3	WLR concord.
1	No Cat	96.0	87.3	103.9	Yes	99.3	101.7	90.1	Yes	98	101.3	81.3	Yes
2	No Cat	114.4	102.8	102.8	Yes	89.1	103.3	110.5	Yes	82.4	87.0	-	Yes
3	No Cat	105.4	96.3	99.2	Yes	90.1	91.7	83.8	Yes	93.6	108.1	103.8	Yes
4	No Cat	51.8 (51.8)	41.9 (41.9)	59.5 (59.5)	No	52.8 (54.8)	53.6 (55.9)	43.7 (46.1)	No	79.3 (80.2)	46.1 (47.0)	47.1 (48.0)	No
5[a,b]	No Cat	2.2 (3.0)	2.5 (3.3)	2.6 (3.3)	Yes	1.6 (2.3)	1.7 (2.5)	1.9 (2.7)	Yes	2.5 (3.2)	2.4 (3.3)	2.5 (3.4)	Yes
6[a,b]	No Cat	100.3 (102.0)	98.1 (99.7)	102.6 (104.4)	Yes	83.5 (84.3)	94.1 (95.0)	95 (95.8)	Yes	115.1 (115.3)	75.7 (75.9)	101.7 (101.9)	Yes
7[a,b]	No Cat	96.8 (96.8)	97.6 (97.6)	77.3 (77.3)	Yes	89.1 (89.0)	89.2 (89.0)	85.5 (85.5)	Yes	74.7 (74.4)	91.5 (91.4)	98.3 (98.2)	Yes
8	No Cat	101.5	112.1	99.1	Yes	80.3	101.1	99.7	Yes	88.6	99.1	97.2	Yes
9	No Cat	105.6	112.1	98.7	Yes	97.9	93.1	99.6	Yes	105.4	107.1	107.7	Yes
10[a,b]	No Cat	105.8 (105.9)	91.2 (91.4)	104.3 (104.5)	Yes	87.7 (87.9)	88.2 (88.4)	97.1 (97.3)	Yes	83.6 (83.8)	101.1 (101.2)	100.4 (100.7)	Yes
11[a]	No Cat	28.7	15.3	16.7	Yes	79.1 (79.1)	92.4 (92.4)	100.1 (100.1)	Yes	36.6	81.5	37.4	No
12	No Cat	94.4	105.7	103.1	Yes	91.5	98.5	99.4	Yes	95.7	104.9	88.9	Yes
13	No Cat	106.9	105.4	92.0	Yes	91.5	96.6	79.2	Yes	91.1	105	102.2	Yes
14	No Cat	123.7	121.3	112.6	Yes	98.7	113.2	128.6	Yes	137.5	116.1	106.9	Yes
15	No Cat	91.7	85.8	91.9	Yes	87.8	98.6	88.7	Yes	99.2	85.6	100.5	Yes
16	No Cat	101.9	109.7	100.9	Yes	91.1	92.8	97.9	Yes	110.8	102.6	71.8	Yes
17	No Cat	91.2	87.9	98.4	Yes	88.4	93.6	106.9	Yes	112.8	93.3	91.4	Yes
18	No Cat	0.5	0.5	0.7	Yes	0.5	0.6	0.5	Yes	0.8	0.8	0.8	Yes
19	No Cat	109.3	108.8	117.5	Yes	103.8	111.3	103.4	Yes	121.9	114.7	111.8	Yes
20	No Cat	102.7	92.8	105	Yes	92.9	93.8	93.4	Yes	108.0	101.6	101.9	Yes
21[a]	No Cat	111.5 (111.5)	117.6 (117.6)	106.5 (106.5)	Yes	98.0 (98.0)	91 .0 (91.0)	89.5 (89.5)	Yes	108.0	109.5	76.7	Yes
22	No Cat	79.6	71.1	73.6	Yes	73.5	64.9	63.6	Yes	74.2	72.1	78.5	Yes
23	No Cat	84.6	95.2	102.6	Yes	81.8	90.3	90.1	Yes	101.8	97.6	92.5	Yes
24[a]	No Cat	2.2 (2.2)	0.7 (0.7)	0.7 (0.7)	Yes	0.3 (0.3)	0.5 (0.5)	0.1 (0.1)	Yes	0.7 (0.7)	0.6 (0.6)	0.5 (0.5)	Yes
25	No Cat	101.3	92.9	100.1	Yes	82	83.4	90.4	Yes	94.5	93.4	88.2	Yes
26[a]	No Cat	1.0 (1.0)	3.0 (3.0)	8.1 (8.1)	Yes	0.5 (0.5)	0.9 (0.9)	0.9 (0.9)	Yes	8.3 (8.3)	3.6 (3.6)	8.8 (8.8)	Yes
27[a]	No Cat	1.7 (3.6)	0.0 (0.7)	1.2 (2.9)	Yes	1.9 (3.0)	1.7 (2.8)	1.9 (3.2)	Yes	1.6 (2.9)	1.7 (2.9)	2.0 (3.3)	Yes
28	No Cat	93.6	97.6	86.8	Yes	85.7	88.8	101.9	Yes	90.1	88	101.4	Yes
29	No Cat	110.5	114.6	103.4	Yes	97.6	97.6	100.7	Yes	101.2	109.3	82.5	Yes
30[a]	No Cat	0.6	0.9	0.8	Yes	0.1 (0.2)	0.4 (0.5)	0.5 (0.6)	Yes	0.0	0.6	0.8	Yes
31	Cat 2B	20.5	22.8	6.8	Yes	5.1	7.4	31.3	Yes	23.3	9.9	15.4	Yes
32[a,b]	Cat 2B	0.9 (1.0)	1.0 (1.1)	1.6 (1.8)	Yes	0.8 (0.9)	1.0 (1.0)	0.6 (0.7)	Yes	0.6 (0.7)	1.2 (1.4)	0.8 (1.0)	Yes
33[a]	Cat 2B	91.1 (91.1)	87.4 (87.4)	98.3 (98.3)	Yes	33.8	79.7	87.5	No	92.4	73.1	93.6	Yes
34	Cat 2B	3.6	2.9	3.4	Yes	5	2.3	1.4	Yes	0.6	5.9	5.1	Yes
35	Cat 2B	0.5	0.5	0.6	Yes	0.7	0.5	0.3	Yes	0.7	0.6	1.0	Yes
36	Cat 2A	68.7	60.6	58.8	Yes	61.7	60.9	71.2	Yes	56.3	69.0	63.6	Yes
37[a,b]	Cat 2A	24.7 (27.4)	56.5 (59.2)	47.2 (49.7)	No	63.2 (64.8)	45.4 (46.8)	57.8 (59.2)	No	41.4 (43.1)	71.0 (72.6)	43.3 (44.5)	No
38[a,b]	Cat 2A	19.2 (20.8)	18.5 (20.0)	27.5 (29.3)	Yes	35.7 (36.6)	9.5 (10.2)	30.6 (31.3)	Yes	12.8 (14.0)	48.0 (49.0)	39.1 (40.5)	Yes
39	Cat 2A	0.6	1.6	0.6	Yes	1	2.1	0.5	Yes	22.4	0.7	1.6	Yes
40	Cat 2A	1.5	0.5	1.2	Yes	0.7	0.6	1.6	Yes	0.5	5.6	0.8	Yes
41	Cat 2A	0.9	0.7	0.7	Yes	0.7	0.5	0.6	Yes	0.6	0.6	1.0	Yes
42	Cat 2A	0.6	0.6	0.6	Yes	6.6	0.4	0.4	Yes	0.5	0.5	0.8	Yes
43[a,b]	Cat 2A[c]	1.3 (2.3)	0.9 (2.0)	0.9 (2.0)	Yes	0.5 (1.5)	1.1 (2.1)	1 (1.9)	Yes	0.8 (2.1)	2.8 (4.0)	1.7 (3.0)	Yes
44	Cat 2A	0.5	0.5	0.7	Yes	0.3	0.8	0.4	Yes	0.7	0.9	0.6	Yes
45[a]	Cat 1	0.3 (0.3)	0.3 (0.3)	0.4 (0.4)	Yes	0.5	0.3	0	Yes	0.5 (0.5)	0.4 (0.4)	0.4 (0.4)	Yes
46	Cat 1	9.4	2	15.2	Yes	4.1	5.4	1.9	Yes	2.8	8.9	1.7	Yes
47	Cat 1	0.6	0.7	0.8	Yes	0.5	0.5	0.1	Yes	0.6	0.4	0.8	Yes
48	Cat 1	0.2	0.4	0.3	Yes	0.1	0.1	0.5	Yes	0.0	0.2	0.4	Yes
49[a,b]	Cat 1	1.1 (19.5)	0.6 (20.9)	4.1 (23.4)	Yes	8.2 (19.8)	3.1 (13.4)	3.8 (14.5)	Yes	3.6 (20.1)	11.9 (28.2)	3.5 (20.7)	Yes
50	Cat 1	0.4	0.5	0.5	Yes	0.4	0.3	0.2	Yes	0.7	0.4	0.5	Yes
51[a]	Cat 1	0.0 (0.5)	2.1 (4.5)	7.1 (9.8)	Yes	2.8 (4.0)	2.6 (4.0)	0.3 (1.8)	Yes	0.0 (0.8)	0.0 (0.3)	4.7 (6.8)	Yes
52	Cat 1	0.7	0.6	0.7	Yes	0.5	0.5	0.2	Yes	0.7	0.6	0.8	Yes
53	Cat 1	0.7	0.7	0.7	Yes	0.5	0.4	0.3	Yes	0.7	0.6	0.9	Yes
54	Cat 1	0.5	0.5	0.8	Yes	0.4	0.5	0.5	Yes	0.6	0.6	0.7	Yes
55[a]	Cat 1	0.5 (0.5)	0.5 (0.5)	0.6 (0.6)	Yes	0.6 (0.8)	0.9 (1.0)	0.7 (0.9)	Yes	0.6 (0.8)	0.5 (0.7)	0.8 (1.1)	Yes
56	Cat 1	0.3	0.4	0.9	Yes	0.6	0.6	0.5	Yes	1.1	0.4	0.8	Yes
57	Cat 1	0.6	0.7	0.6	Yes	0.5	0.4	0.6	Yes	0.1	0.5	0.5	Yes
58[a]	Cat 1	1.2 (1.2)	1.1 (1.1)	1.2 (1.2)	Yes	1.0 (1.0)	1.0 (1.0)	0.9 (0.9)	Yes	1.0 (1.0)	1.1 (1.1)	1.0 (1.0)	Yes
59	Cat 1	4.1	1.3	0.8	Yes	4.4	13.3	12	Yes	2.4	2.0	7.3	Yes
60[a,b]	Cat 1	1.2 (28.5)	0.0 (22.5)	2.1 (34.6)	Yes	0.0 (24.7)	0.0 (24.2)	0 (24.1)	Yes	3.0 (29.5)	4.2 (30.6)	0.3 (30.0)	Yes

using photometric MTT-reduction and HPLC/UPLC-spectophotometry. The SkinEthic™ HCE EITL (Alépée et al., 2016) or EITS protocol (Section 2.5.2) were performed for liquids and solids, respectively. The resulting formazan tissue extracts were analysed by photometry (OD) and HPLC/UPLC-spectrophotometric analysis. The agreement in viability between the MTT and HPLC/UPLC-spectrophotometric method was assessed with a scatter plot. The line of equality was used as a visual tool for agreement. A dot that falls on the line or that is close to the line corresponds with a chemical with equal viability values or values close to each for the different endpoint detection systems.

Coloured test chemicals or test chemicals that become coloured in contact with water or isopropanol that interfere too strongly with the MTT-reduction assay may still be assessed using HPLC/UPLC photometry instead of standard absorbance (OD). This is because the HPLC/UPLC system allows for the separation of the MTT formazan from the chemical before its quantification (18).

Viability was summarized as mean and the difference. Simple linear regression was used for assessing the agreement between the OD and HPLC/UPLC–spectrophotometry detection methods. The assumption of linearity was verified with a scatter plot of the standardized residuals *versus* the viability and the normality of the residuals was verified with a QQ-plot. An alpha level of 0.05 was used as significance level. All analyses were performed in R version 3.2.0 (R Core Team, 2015).

2.9. Statistical data analyses

The frequency of non-qualified runs and non-qualified tests per laboratory was reported. The WLR, BLR, accuracy, sensitivity and specificity of the SkinEthic™ HCE EITS test method were calculated according to the rules described in the OECD Performance Standards for the Assessment of Proposed Similar or Modified In Vitro Reconstructed human Cornea-like Epithelium (RhCE) Test Methods for Eye Hazard (ENV/JM/MONO(2015)23, 2015). Furthermore, the Validation Management Group (VMG) specified the following minimum values for WLR (≥85%), BLR (≥80%), accuracy (≥75%), sensitivity (≥90%) and specificity (≥60%) for the prospective Eye Irritation Validation Study (EIVS) of RhCE-based test methods conducted by EURL ECVAM and Cosmetics Europe (EC EURL ECVAM, 2014; Kaluzhny et al., 2015; Barroso et al., 2015).

2.9.1. Within laboratory reproducibility (WLR)

For each laboratory, the mean viability of each run for each chemical was calculated. The WLR of the independent runs was evaluated based on the concordance of predictions (C or NC) of the qualified tests. WLR was reported with the Wilson's 95% confidence intervals (CI) for proportions. The Wilson CI's, based on the score test, provides more reliable values for small samples and estimates close to 1.0 (Agresti and Coull, 1998).

2.9.2. Between laboratory reproducibility (BLR)

For each laboratory, the mean viability and standard deviation of the independent qualified tests was calculated to obtain a final classification for each chemical. The evaluation of the BLR was on the concordance of the final predictions classified (C) or No Cat (NC). BLR was reported with the Wilson 95% CI.

2.9.3. Predictive capacity

The predictive capacity of the assay was evaluated by comparing the prediction results, on the basis of the individual laboratory results using all qualified tests obtained for each chemical (as recommended by the OECD guidance document ENV/JM/MONO(2015)23, 2015), with the existing proposed classification. Therefore, 2 × 2 contingency tables (C versus NC) were constructed and sensitivity (probability of predicting C given the true state is serious eye damage/eye irritancy (Category 1 or Category 2), specificity (probability of predicting NC given the true state is No Category), and accuracy were calculated.

In addition to the calculation of predictive capacity as recommended by the OECD guidance document, the uncertainty of the point estimates for accuracy, sensitivity and specificity was assessed with the bootstrap resampling method. Bootstrap resampling (10.000 times with sample size = 1) was used to obtain 95% CI's for accuracy, sensitivity and specificity. The rationale for performing bootstrap resampling with size n = 1 is that in reality a chemical will be tested only once. Therefore, it was opted to calculate sensitivity, specificity, and accuracy on 10.000 simulated sets of 60 chemicals, based on observed predictions (9 predictions per chemical). Briefly, random sampling with sample size n = 1 was performed per chemical (pool of 9 predictions, being 3 runs for each of the 3 laboratories) for the set of 60 chemicals. Next, the accuracy, sensitivity and specificity were calculated for each of the 10.000 resampling sets. The mean of the bootstrap sample and 95% CI applying the percentile method was calculated for the three performance parameters.

All analyses were performed with R version 3.1.1. (R Core Team, 2015).

3. Results and discussions

3.1. Multicentre study: SkinEthic™ HCE EITS test method

Before starting the multicentre study, the effectiveness of the training phase was assessed in a transferability study. During this study, the laboratory technicians from CRL and VITO tested each 9 solids under blinded conditions in at least three independent runs. These data were submitted to the test method developer (L'Oréal) for quality check. The results (data not shown) demonstrated effectiveness of the training.

Next, a total of 60 chemicals were tested at least three times in three laboratories. Overall, L'Oréal produced two unqualified tests (chemicals No. 11 and No. 25) over the 18 series that were performed. Charles River Laboratories (CRL) performed 18 series, all resulting in qualified tests. In total 21 series were performed by two operators at VITO. One series was excluded due to high deviation between the viability of the replicate negative control tissues (difference 25.6%, batch 15HCE012, Fig. 2). Furthermore, 10 tests were unqualified, the difference of viability of the replicate tissues was >20% for chemicals No. 1, 2 (3 times unqualified tests), 6, 11, 13, 37, and 46. According to the Standard Operating Procedure (SkinEthic™ HCE EIT SOP, 2015), if the difference of viability between the two related replicate tissues exceeds 20, the test substance should be retested in an additional run (up to 5 tests, including retesting). Chemical No. 2 was five times tested by VITO and resulted three times in unqualified results. Therefore only two qualified tests are available for this chemical.

Notes to Table 4:
LO: L'Oréal; CRL: Charles River Laboratories; concord.: concordance.
Cells with a grey background correspond to classified prediction (mean cell viability ≤50%).
[a]Chemical was identified as an MTT reducer by at least one laboratory. Values in brackets correspond to uncorrected viabilities, i.e. before subtraction of viability measured in killed epithelial tissue.
[b]Chemical corresponds with a colourant. Values in brackets correspond to uncorrected viabilities, i.e. before subtraction of non-specific colouring.
[c]SCNM met for the *in vivo* Draize rabbit eye test, the UN GHS classification corresponds with at least a Cat 2A, the results of this study are published in the Eye Irritation Reference Chemicals data bank published by the European Centre for Toxicology and Ecotoxicology of Chemicals (ECETOC, 1998). The study was terminated on day 7, in four animals all tissue scores fully reversed to 0 by day 7, one animal had CR = 1 and another animal had CR and CC = 1.

In total, 56 independent qualified series were performed by the three laboratories, the mean OD of the NgC was within the acceptance limit (between 1.4 and 2.5, Fig. 2). The mean viability of the PC (methyl acetate) was below the acceptance threshold of 30% (range: 0.7% to 28.6%, Fig. 3).

Among the chemicals, seven colourants that were also MTT reducers, were identified by the three laboratories (No. 5, 6, 7, 37, 38, 43, and 60, Table 4), requiring the use of adapted controls for the determination of non-specific colouration and MTT reduction. Six chemicals (No. 24, 26, 27, 51, 55, and 58) were identified as MTT reducers by the three laboratories. Two chemicals (No. 4 and 49) were identified as MTT reducers by the three laboratories and as colourants by CRL only. Two chemicals (No. 10 and 32) were identified as MTT reducers by L'Oréal and CRL and as colourants by the

Table 5
BLR for SkinEthic™ HCE EITS protocol. Mean cell viability of 3 independent runs for 60 chemicals and agreement of prediction between the laboratories.

No	*In vivo* UN GHS/EU CLP	LO		CRL		VITO		BLR concord.
		Cell viability (%)	*In vitro* class	Cell viability (%)	*In vitro* class	Cell viability (%)	*In vitro* class	
1	No Cat	95.7 ± 8.3	NC	97.0 ± 6.1	NC	93.5 ± 10.7	NC	Yes
2	No Cat	106.7 ± 6.7	NC	101.0 ± 10.9	NC	84.7 ± 3.3[a]	NC	Yes
3	No Cat	100.3 ± 4.7	NC	88.5 ± 4.2	NC	101.8 ± 7.5	NC	Yes
4	No Cat	51.1 ± 8.8	NC	50.0 ± 5.5	NC	57.5 ± 18.9	NC	Yes
5	No Cat	2.4 ± 0.2	C	1.7 ± 0.2	C	2.5 ± 0.1	C	Yes
6	No Cat	100.3 ± 2.3	NC	90.9 ± 6.4	NC	97.5 ± 20.0	NC	Yes
7	No Cat	90.6 ± 11.5	NC	87.9 ± 2.1	NC	88.2 ± 12.2	NC	Yes
8	No Cat	104.2 ± 6.9	NC	93.7 ± 11.6	NC	95.0 ± 5.6	NC	Yes
9	No Cat	105.5 ± 6.7	NC	96.9 ± 3.4	NC	106.7 ± 1.2	NC	Yes
10	No Cat	100.4 ± 8.0	NC	91.0 ± 5.3	NC	95.0 ± 9.9	NC	Yes
11	No Cat	20.2 ± 7.4	C	90.5 ± 10.6	NC	51.8 ± 25.7	NC	No
12	No Cat	101.1 ± 5.9	NC	96.5 ± 4.3	NC	96.5 ± 8.0	NC	Yes
13	No Cat	101.4 ± 8.2	NC	89.1 ± 8.9	NC	99.4 ± 7.4	NC	Yes
14	No Cat	119.2 ± 5.8	NC	113.5 ± 15.0	NC	120.2 ± 15.7	NC	Yes
15	No Cat	89.8 ± 3.5	NC	91.7 ± 6.0	NC	95.1 ± 8.3	NC	Yes
16	No Cat	104.2 ± 4.8	NC	93.9 ± 3.5	NC	95.1 ± 20.6	NC	Yes
17	No Cat	92.5 ± 5.4	NC	96.3 ± 9.5	NC	99.2 ± 11.8	NC	Yes
18	No Cat	0.6 ± 0.1	C	0.5 ± 0.1	C	0.8 ± 0.0	C	Yes
19	No Cat	111.9 ± 4.9	NC	106.2 ± 4.5	NC	116.1 ± 5.2	NC	Yes
20	No Cat	100.2 ± 6.5	NC	93.4 ± 0.5	NC	103.8 ± 3.6	NC	Yes
21	No Cat	111.9 ± 5.6	NC	92.8 ± 4.5	NC	98.1 ± 18.5	NC	Yes
22	No Cat	74.8 ± 4.4	NC	67.3 ± 5.4	NC	74.9 ± 3.3	NC	Yes
23	No Cat	94.1 ± 9.1	NC	87.4 ± 4.9	NC	97.3 ± 4.7	NC	Yes
24	No Cat	1.2 ± 0.9	C	0.3 ± 0.2	C	0.6 ± 0.1	C	Yes
25	No Cat	98.1 ± 4.5	NC	85.3 ± 4.5	NC	92.0 ± 3.4	NC	Yes
26	No Cat	4.0 ± 3.7	C	0.8 ± 0.2	C	6.9 ± 2.9	C	Yes
27	No Cat	1.0 ± 0.9	C	1.8 ± 0.1	C	1.8 ± 0.2	C	Yes
28	No Cat	92.7 ± 5.5	NC	92.1 ± 8.6	NC	93.2 ± 7.2	NC	Yes
29	No Cat	109.5 ± 5.7	NC	98.6 ± 1.8	NC	97.7 ± 13.7	NC	Yes
30	No Cat	0.8 ± 0.2	C	0.3 ± 0.2	C	0.5 ± 0.4	C	Yes
31	Cat 2B	16.7 ± 8.7	C	14.6 ± 14.5	C	16.2 ± 6.7	C	Yes
32	Cat 2B	1.2 ± 0.4	C	0.8 ± 0.2	C	0.9 ± 0.3	C	Yes
33	Cat 2B	92.3 ± 5.5	NC	67.0 ± 29.0	NC	86.4 ± 11.5	NC	Yes
34	Cat 2B	3.3 ± 0.4	C	2.9 ± 1.9	C	3.9 ± 2.9	C	Yes
35	Cat 2B	0.5 ± 0.1	C	0.5 ± 0.2	C	0.8 ± 0.2	C	Yes
36	Cat 2A	62.7 ± 5.3	NC	64.6 ± 5.7	NC	63.0 ± 6.4	NC	Yes
37	Cat 2A	42.8 ± 16.4	C	55.5 ± 9.1	NC	51.9 ± 16.6	NC	No
38	Cat 2A	21.7 ± 5.0	C	25.3 ± 13.9	C	33.3 ± 18.3	C	Yes
39	Cat 2A	0.9 ± 0.6	C	1.2 ± 0.8	C	8.2 ± 12.3	C	Yes
40	Cat 2A	1.1 ± 0.5	C	1.0 ± 0.6	C	2.3 ± 2.9	C	Yes
41	Cat 2A	0.8 ± 0.1	C	0.6 ± 0.1	C	0.7 ± 0.2	C	Yes
42	Cat 2A	0.6 ± 0.0	C	2.5 ± 3.6	C	0.6 ± 0.2	C	Yes
43	Cat 2A[b]	1.0 ± 0.2	C	0.9 ± 0.3	C	1.8 ± 1.0	C	Yes
44	Cat 2A	0.6 ± 0.1	C	0.5 ± 0.3	C	0.7 ± 0.2	C	Yes
45	Cat 1	0.3 ± 0.1	C	0.3 ± 0.3	C	0.4 ± 0.1	C	Yes
46	Cat 1	8.9 ± 6.6	C	3.8 ± 1.8	C	4.5 ± 3.9	C	Yes
47	Cat 1	0.7 ± 0.1	C	0.4 ± 0.2	C	0.6 ± 0.2	C	Yes
48	Cat 1	0.3 ± 0.1	C	0.2 ± 0.2	C	0.2 ± 0.2	C	Yes
49	Cat 1	1.9 ± 1.9	C	5.0 ± 2.8	C	6.3 ± 4.8	C	Yes
50	Cat 1	0.5 ± 0.1	C	0.3 ± 0.1	C	0.5 ± 0.2	C	Yes
51	Cat 1	3.1 ± 3.7	C	1.9 ± 1.4	C	1.6 ± 2.7	C	Yes
52	Cat 1	0.7 ± 0.1	C	0.4 ± 0.2	C	0.7 ± 0.1	C	Yes
53	Cat 1	0.7 ± 0.0	C	0.4 ± 0.1	C	0.7 ± 0.2	C	Yes
54	Cat 1	0.6 ± 0.2	C	0.5 ± 0.1	C	0.6 ± 0.1	C	Yes
55	Cat 1	0.5 ± 0.1	C	0.7 ± 0.2	C	0.6 ± 0.2	C	Yes
56	Cat 1	0.5 ± 0.3	C	0.6 ± 0.1	C	0.8 ± 0.4	C	Yes
57	Cat 1	0.6 ± 0.1	C	0.5 ± 0.1	C	0.4 ± 0.2	C	Yes
58	Cat 1	1.2 ± 0.1	C	1.0 ± 0.1	C	1.0 ± 0.1	C	Yes
59	Cat 1	2.1 ± 1.8	C	9.9 ± 4.8	C	3.9 ± 3.0	C	Yes
60	Cat 1	1.1 ± 1.1	C	0.0 ± 0.0	C	2.5 ± 2.0	C	Yes

LO: L'Oréal; CRL: Charles River Laboratories; concord.: concordance.
Values correspond with mean ± SD of 3 independent runs; [a]Mean of two independent valid runs. [b]SCNM for the *in vivo* Draize rabbit eye test, the UN GHS classification corresponds with at least a Cat 2A, the results of this study are published in the Eye Irritation Reference Chemicals data bank published by the European Centre for Toxicology and Ecotoxicology of Chemicals (ECETOC, 1998). The study was terminated on day 7, in four animals all tissue scores fully reversed to 0 by day 7, one animal had CR = 1 and another animal had CR and CC = 1.

Table 6
Predictive capacity for the set of 60 chemicals based on individual laboratory predictions: overall and for each laboratory.

In vivo UN GHS	Cumulative		L'Oréal		CRL		VITO	
	C	NC	C	NC	C	NC	C	NC
Classified (n)	**249**	21	**83**	7	**83**	7	**83**	7
No Category (n)	63	**206**	22	**68**	19	**71**	22	**67**
Total (n)	539		180		180		179[a]	
Sensitivity (%)	92.2		92.2		92.2		92.2	
Specificity (%)	76.6		75.6		78.9		75.3	
Accuracy (%)	84.4		83.9		85.6		83.3	

Bold values are concordant *in vivo-in vitro* prediction.
[a] For chemical No. 2 only two valid runs were obtained over the five runs.

three laboratories. Three chemicals were identified as MTT reducers by one laboratory only (CRL No. 11 and 30, and L'Oréal Nr 33). Chemical No. 21 was identified as an MTT reducer by L'Oréal and CRL and Chemical No. 45 was identified as MTT reducer by L'Oréal and VITO. Both uncorrected and corrected (final) viabilities were reported in the Table 4.

3.1.1. Within laboratory reproducibility

The reliability of the SkinEthic™ HCE EITS protocol was assessed in terms of concordance in predictions for the independent valid runs. The results for each laboratory are presented in Table 4. The WLR was 96.7% (95% CI: 88.6%–99.1%) for L'Oréal and 95.0% (95% CI: 86.3%–98.3%) for VITO and CRL. Chemicals No. 4 and 37 resulted in discordant results in the three laboratories. The discordant predictions obtained for chemical No. 4 and No. 37 can be attributed to the viability which was in the middle range. Chemicals No. 11 and No. 33 resulted in a discordant prediction in one laboratory. At VITO, one result (viability: 81.5%) obtained for chemical No. 11, deviated clearly from the other two runs (36.6% and 37.4%). CRL obtained a disagreement in prediction for chemical No. 33, with a lower viability (33.8%) in the first run in comparison with the other two runs (79.7% and 87.5%).

In conclusion, low variation (WLR ≥ 95%) between the independent runs was observed within the laboratories, indicating that the SkinEthic™ HCE EITS protocol is robust. This means that the WLR is higher than 95%, which is the minimum value set by the VMG (Barroso et al., 2015).

3.1.2. Between laboratory reproducibility

In order to assess the transferability of the method, mean viability of the independent qualified tests within each laboratory was calculated to determine the final classification for each laboratory. The results are presented in Table 5. Fifty eight of the 60 chemicals were consistently classified (NC/C) by the three laboratories resulting in a BLR of 96.7% (95% CI: 88.6%–99.1%). The BLR for the pair-wise comparisons was 96.7% (58/60 chemicals) for L'Oréal and CRL and for L'Oréal and VITO, a 100% concordance was obtained between CRL and VITO. Chemicals No. 11, and 37 resulted in discordant predictions. The BLR of the SkinEthic™ HCE EITS test method was higher than the defined minimum value of 80% set by the VMG (Barroso et al., 2015).

3.1.3. Predictive capacity

The predictive capacity was calculated for each laboratory and for the cumulative results of the three laboratories using the cut-off of 50% viability to distinguish between chemicals not requiring classification for serious eye damage/eye irritancy (No Cat) from chemicals requiring classification and labelling (Cat 1 and Cat 2) according to UN GHS (Table 6). The calculations were based on the individual predictions derived from the qualified tests for each chemical in each laboratory. The three laboratories obtained a sensitivity of 92.2%. The specificity varied between 75.3% (VITO), 75.6% (L'Oréal), and 78.9% (CRL). An accuracy of 83.9%, 85.6%, and 83.3% was obtained by L'Oréal, CRL, and VITO, respectively. In order to estimate the uncertainty of the sensitivity, specificity and accuracy estimates, the bootstrap resampling method was used. The bootstrap sample consisted of 10.000 resamplings of size 1 per chemical for the set of 60 chemicals. The distribution of the bootstrap samples is presented in Fig. 4. This resulted in an overall sensitivity of 91.9% (95% CI: 90.0% to 93.3%), a specificity of 76.6% (95% CI: 73.3% to 80.0%), and an accuracy of 84.3% (95% CI: 81.7% to 86.7%). In conclusion, the SkinEthic™ HCE EITS test method exceeds the defined values for sensitivity (≥90%), specificity (≥60%) and accuracy (≥75%) that were set by the VMG (Barroso et al., 2015).

3.2. Additional data: SkinEthic™ HCE EITS test method

The lead laboratory (L'Oréal) tested 35 additional chemicals (Table 2) in three independent runs. Twenty three chemicals did not require classification *in vivo* and 12 chemicals were classified. A concordant prediction was obtained for 34 of the 35 chemicals, resulting in a WLR of 97.1% (Table 7). The predictive capacity to distinguish chemicals not requiring classification from classified chemicals was determined for the extended dataset (60 chemicals of the multicentre study and 35 additional chemicals). This resulted in an accuracy of 80.7% with a 89.7% sensitivity and a 73.6% specificity for L'Oréal only (Table 8). Overall, the sensitivity, specificity, and accuracy based on the individual predictions of the three laboratories were 91.2%, 75.4%, and 82.9% respectively. The bootstrap estimates for this extended dataset of 95 chemicals, correspond with an overall sensitivity of 90.5% (95% CI: 88.1% to 92.9%), a specificity of 73.6% (95% CI: 71.7% to 75.5%), and an accuracy of 81.0% (95% CI: 78.9% to 83.2%). The distribution of the bootstrap samples is presented in Fig. 5. In conclusion, also for the extended set of 95 solid chemicals, the SkinEthic™ HCE EITS test method met all the acceptance criteria set by the VMG (Barroso et al., 2015).

3.3. HPLC/UPLC–spectrophotometry

Since it is known that the photometric MTT-reduction assay can interfere with strongly coloured and/or strong MTT-reducing test

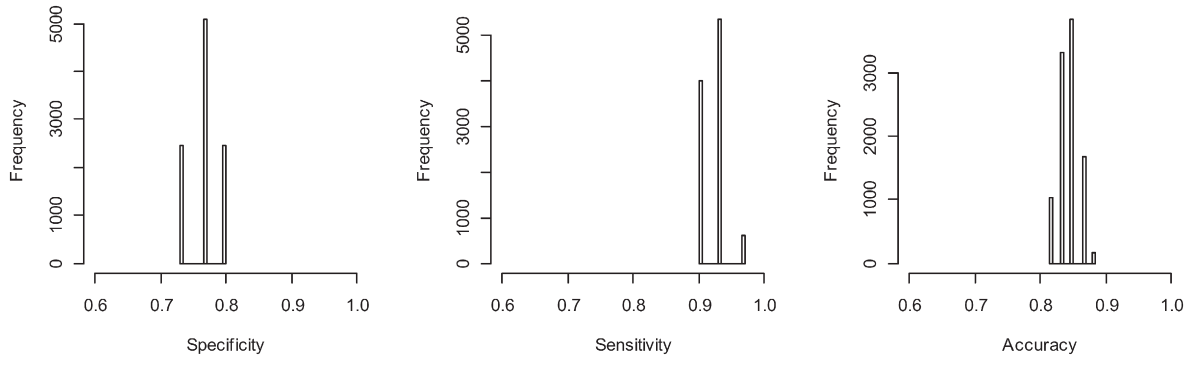

Fig. 4. Distribution of the bootstrap sample representing 10,000 resamplings of size 1 per chemical for the set of 60 chemical (multicentre study).

Table 7
WLR for SkinEthic™ HCE EITS protocol. Mean cell viability (n = 2) for 35 additional chemicals and agreement of prediction within L'Oréal (LO).

No.	*In vivo* UN GHS/ EU CLP	L'Oréal			
		Run 1	Run 2	Run 3	WLR concordance
61	No Cat	96.5	100.7	101.4	Yes
62	No Cat	83.1	81.6	72.9	Yes
63	No Cat	43.3	45.5	39.0	Yes
64	No Cat	97.1	97.5	95.7	Yes
65	No Cat	111.1	105.9	102.8	Yes
66	No Cat	3.2	3.7	0.5	Yes
66[a]		(5.2)	(5.7)	(2.5)	
67	No Cat	106.2	109	98	Yes
67[a, b]		(106.4)	(109.2)	(98.1)	
68	No Cat	104.7	100.5	87.2	Yes
68[b]		(104.9)	(100.6)	(87.7)	
69	No Cat	101	96.5	96.1	Yes
70	No Cat	103.8	103.4	101.2	Yes
71	No Cat	0.6	0.4	0.4	Yes
72	No Cat	105.9	100.5	101.4	Yes
73	No Cat	78.5	87.6	87.7	Yes
74	No Cat	105.1	95.8	104.2	Yes
75	No Cat	71.5	61.4	68	Yes
76	No Cat	113.3	107.3	118.2	Yes
76[a, b]		(127.4)	(131.6)	(128.4)	
77	No Cat	3.8	2.1	0.8	Yes
78	No Cat	0.6	0.6	0.7	Yes
79	No Cat	63.3	66.9	62.2	Yes
80	No Cat	58.4	44.7	28.5	No
80[a]		(58.5)	(44.8)	(28.6)	
81	No Cat	33.4	29.1	22.5	Yes
81[b]		(33.4)	(29.1)	(22.6)	
82	No Cat	103	97.7	102.9	Yes
83	No Cat	96.9	105.5	104.2	Yes
84	Cat 2B	34.6	41.7	28.7	Yes
85	Cat 2B	99.2	105.9	120.2	Yes
86	Cat 2B	96.5	99.5	97.6	Yes
87	Cat 2A	2.3	2.2	2.8	Yes
87[a, b]		(8.9)	(9.6)	(9.8)	
88	Cat 1	2.8	3.1	3.4	Yes
88[a]		(4.7)	(5.2)	(5.4)	
89	Cat 1	0.4	0.5	0.5	Yes
89[a]		(0.4)	(0.5)	(0.5)	
90	Cat 1[c]	0.6	0.7	0.6	Yes
91	Cat 1	26.4	20	13.5	Yes
92	Cat 1	15.4	7.5	20.8	Yes
92[b]		(15.5)	(7.5)	(20.9)	
93	Cat 1	0.5	0.5	0.4	Yes
93[b]		(0.5)	(0.5)	(0.5)	
94	Cat 1	0.1	0.2	0.3	Yes
94[b]		(0.2)	(0.3)	(0.4)	
95	Cat 1	0.4	0.5	0.5	Yes

Cells with a grey background correspond to classified prediction (mean cell viability ≤50%).
[a]Chemical was identified as an MTT reducer by at least one laboratory. Values in brackets correspond to uncorrected viabilities, i.e. before subtraction of viability measured in killed epithelial tissue.
[b]Chemical corresponds with a colourant. Values in brackets correspond to uncorrected viabilities, i.e. before subtraction of non-specific colouring.
[c]Study criteria were not met for the *in vivo* Draize rabbit eye test, the UN GHS classification corresponds with a Cat 1. The summary results of this study are published in the DRD (Barroso et al., 2016). The study was terminated on day 14 with CO = 2 in 3/6, CO = 1 in 2/6, IR = 1 in 4/6, CR = 1 in 5/6, CC = 2 in 1/6 and CC = 1 in 4/6 animals.

chemicals, the viability of 11 liquid and 13 solid chemicals was assessed using photometric MTT-reduction and HPLC/UPLC-spectophotometry. The MTT reducing and colour interfering properties are presented in Table 3. Photometric measurements (MTT) could not be obtained for 5 out of 24 chemicals. The remaining 19 chemicals covered 9 chemicals without MTT-reducing and colouring properties, and 10 chemicals with MTT-reducing and/or colouring properties.

A simple linear relationship between OD and HPLC/UPLC–spectrophotometry seems reasonable. From the residuals plot (Fig. 6A), it can be observed that the viability for 1,2-ethanediamine, N1-[3-(trimethoxysilyl)propyl]- (chemical treated tissue 2) (#45c) and for 2: 2-Anthracenesulfonic acid, 4-[[4-(acetylamino)phenyl]amino]-1-amino-9,10-dihydro-9,10-dioxo-, sodium salt (1:1) (chemical treated tissue 2) (#103) obtained with HPLC/UPLC–spectrophotometry was higher and lower in comparison with OD, respectively. The mean difference in viability between both methods was 24.9% and 18.0%, respectively, whereas for all other chemicals the absolute difference was on average ≤ 6.7% (data not shown). The high fit ($R^2 = 0.99$) and the slope of the regression model which is close to 1 (slope = 0.98 with 95% CI: 0.95; 1.02) confirms that, for chemicals that are compatible with use of OD, high agreement is observed between measurement of tissue viability by OD and HPLC/UPLC–spectrophotometry (Fig. 6B). This supports the findings of the study published by Alépée and co-workers that HPLC/UPLC–spectrophotometry and OD measurements gave similar results in terms of tissue viability for the EpiOcular™ EIT test method (2015). Out of the 24 chemicals tested in the current study with the SkinEthic™ HCE EIT test method, 19 were also evaluated by the EpiOcular™ EIT test method. Therefore, HPLC/UPLC–spectrophotometry can be used to measure formazan irrespective of the tissue model used.

3.4. Misclassified chemicals

The misclassified chemicals were investigated in more detail by taking into account the functional group and the UH GHS category. In total, out of 42 classified chemicals that were tested, five false negative results were obtained. Chemical No. 37 (*in vivo* Cat 2A) was predicted NC in 4 out of 9 runs, chemical No. 33 (*in vivo* Cat 2B) was predicted NC in 8 out of 9 runs and chemicals No. 36, 85 and 86 (*in vivo* Cat 2A, 2B and 2B, respectively) were always predicted NC (Table 4 and Table 7). The five false negatives correspond with four different functional groups (nitro-compound, two esters, phenol and ether), therefore it is unlikely that the under-predictions are related with the functional group. Of the 53 *in vivo* No Cat chemicals, 38 were correctly identified as NC. Twelve *in vivo* UN GHS No Cat chemicals (No. 5, 18, 24, 26, 27, 30, 63, 66, 71, 77, 78, and 81) were consistently predicted C (mean viability < 50%) in all runs. Three additional *in vivo* No Cat chemicals (No. 4, 11, and 80) resulted in a false positive prediction in at least two runs. No relation was observed between the false positives and the functional group of the chemical (12 different functional groups).

3.5. Overall performance of the SkinEthic™ HCE EIT method

Any new test method proposed for use under OECD TG 492 should be evaluated prior to their use for regulatory purposes to establish their similarity to the Validated Reference Method (VRM) and to determine the reliability and relevance to identify chemicals not requiring classification for serious eye damage/eye irritation according to UN GHS. The reliability and relevance of SkinEthic™ HCE EIT test method was determined using 200 commercially available chemicals with different physical states (105 liquids and 95 solids) representing different organic functional groups. The overall set contained several colour interfering chemicals (1 liquid and 7 solids), MTT reducers (7 liquids and 12 solids) and MTT reducing colourants (1 liquid and 10 solids). In total, 120 chemicals (60 liquids and 60 solids) covering 16 different functional groups were evaluated in the SkinEthic™ HCE EITL (Alépée et al., 2016) and SkinEthic™ HCE EITS (current study) validation study. In terms of physical state and UN GHS Categories, the 120

Table 8
Predictive capacity for the set of 95 chemicals based on individual laboratory predictions: overall and for each laboratory.

In vivo UN GHS	Cumulative		L'Oréala		Charles River Laboratoriesb		VITOb	
	C	NC	C	NC	C	NC	C	NC
Classified (n)	**279**	27	**113**	13	**83**	7	**83**	7
No Category (n)	83	**255**	42	**117**	19	**71**	22	**67**
Total (n)	644		285		180		179c	
Sensitivity (%)	91.2		89.7		92.2		92.2	
Specificity (%)	75.4		73.6		78.9		75.3	
Accuracy (%)	82.9		80.7		85.6		83.3	

Bold values are concordant *in vivo-in vitro* prediction.
[a] Predictions based on all chemicals (60 from the multicentre study and 35 additional chemicals).
[b] Predictions based on the 60 chemicals from the multicentre study.
[c] For chemical No. 2 only two valid runs were obtained over the five runs.

chemicals, were distributed as follows: 32 Cat 1 (16 liquids and 16 solids), 17 Cat 2A (8 liquids and 9 solids), 13 Cat 2B (8 liquids and 5 solids) and 58 No Cat (28 liquids and 30 solids) chemicals. Furthermore, the lead laboratory (L'Oréal) tested 80 additional chemicals (45 liquids and 35 solids) in three independent runs enlarging the number of functional groups with one additional group. The chemicals covered 19 Cat 1 (11 liquids and 8 solids), 12 Cat 2A (11 liquids and 1 solid), 4 Cat 2B (1 liquid and 3 solids) and 45 No Cat (22 liquids and 23 solids) chemicals.

The reliability (WLR and BLR) of the SkinEthic™ HCE EIT method was assessed in terms of concordance in predictions. The WLR, based on the set of 120 chemicals, was 91.7% (EITL 88.3% and EITS 95.0%) for CRL, 94.2% for VITO (EITL 93.3% and EITS 95.0%) and 95.8% for LO (EITL 95.0% and EITS 96.7%). The WLR for the extended set of 200 chemicals that were tested by LO only was 95.0% (EITL 93.3% and EITS 96.8%). The WLR of the SkinEthic™ HCE EITL method was slightly less in comparison to the VRM, the EpiOcular™ EIT test method obtained a WLR of 96.3%, 98.1% and 98.1% in three laboratories for the Liquids protocol. For the EpiOcular™ EIT Solids protocol, a WLR of 96.6% was obtained in one laboratory, this is comparable with the WLR for solids obtained in the current study (EC EURL ECVAM, 2014; Barroso et al., 2015). The overall BLR for the HCE EIT method, based on the set of 120 chemicals, was 95.0% (EITL 93.3% and EITS 96.7%). The EpiOcular™ EIT test method resulted in a BLR of 94.4% for the Liquids protocol and 92.0% for the Solids original protocol (EC EURL ECVAM, 2014; Barroso et al., 2015). It is important to note that a strict comparison of the WLR and BLR should not be made since the chemical sets were

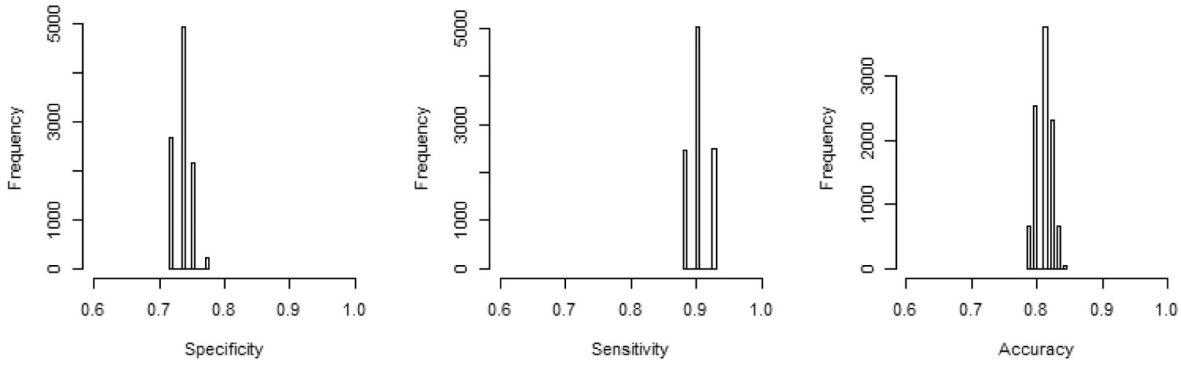

Fig. 5. Distribution of the bootstrap sample representing 10,000 resamplings of size 1 per chemical for the extended data set of 95 chemicals.

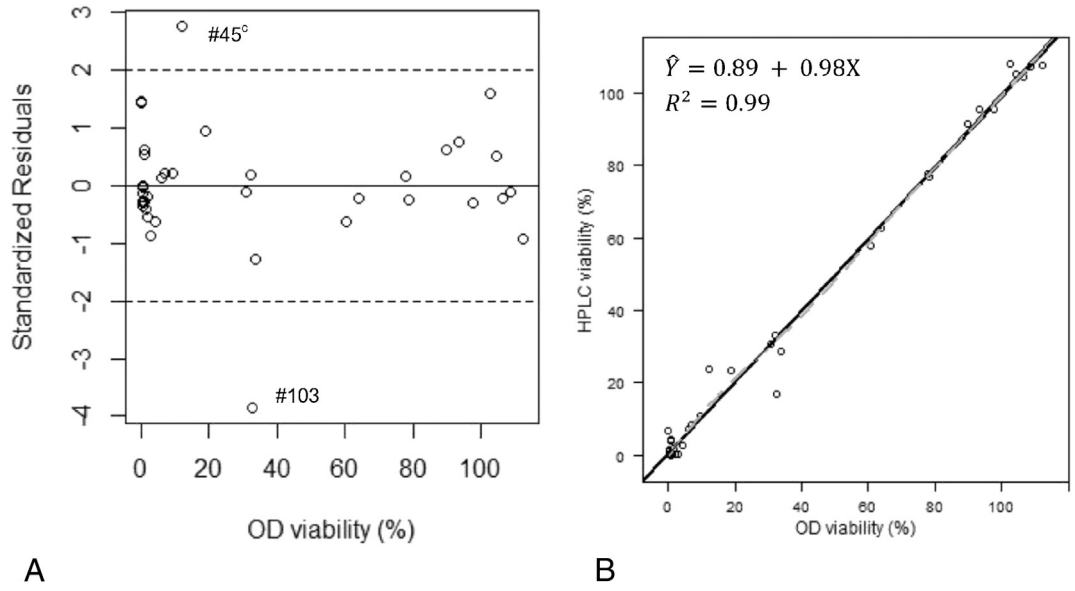

Fig. 6. (A) Plot of the standardized residuals against HPLC/UPLC–spectrophotometry viability. (B) Correlation of individual tissue viability values (%) measured by OD and HPLC/UPLC–spectrophotometry in SkinEthic™ HCE EIT test method for 19 chemicals. The black line corresponds with the linear regression line and the dotted grey curve corresponds with the LOESS curve or lowess fit, a locally weighted smooth fit. Linearity is confirmed since the LOESS fit line is very close to the linear regression line. #45[c]: 1,2-ethanediamine, N1-[3-(trimethoxysilyl)propyl]- CAS No. 1760-24-3 (tissue 2) and #103: 2-Anthracenesulfonic acid, 4-[[4-(acetylamino)phenyl]amino]-1-amino-9,10-dihydro-9,10-dioxo-, sodium salt (1:1) CAS No. 6424-85-7 (tissue 2).

different; the non-commercially available proprietary chemicals evaluated using EpiOcular™ EIT test method during the EURL ECVAM/Cosmetics Europe study (EC EURL ECVAM, 2014) were not evaluated in the current SkinEthic™ HCE EIT method. Overall, the SkinEthic™ HCE EIT method met the performance acceptance criteria, the minimum values for WLR and BLR set by the VMG were 85% and 80%, respectively (Barroso et al., 2015).

The overall predictive capacity of the SkinEthic™ HCE EIT method to distinguish between chemicals not requiring classification for serious eye damage/eye irritancy (No Cat) from chemicals requiring classification and labelling (Cat 1 and Cat 2) according to UN GHS was evaluated for the validation set. The accuracy based on the individual predictions obtained in the three laboratories for the set of 120 chemicals was 84.6% with a specificity of 73.1% and sensitivity of 95.3%. For the liquids (EITL) and solids (EITS) the accuracy was 84.8% and 84.4%, with a specificity of 69.4% and 76.6%, and sensitivity of 98.3% and 92.2%, respectively. Considering all the data of the SkinEthic™ HCE EIT method (120 chemicals of the validation study and 80 additional chemicals), an accuracy of 83.7% was obtained with a specificity of 72.1% (based on 103 chemicals) and sensitivity of 95.2% (based on 97 chemicals). The EpiOcular™ EIT has an overall accuracy of 80% (based on 112 chemicals), sensitivity of 96% (based on 57 chemicals), specificity of 63% (based on 55 chemicals) when compared to reference *in vivo* rabbit eye test data (EC EURL ECVAM, 2014; OECD, 2015b). Again, it is important to note that a strict comparison of the predictive capacity values should not be made since the number of chemicals tested and sets are different. The SkinEthic™ HCE EIT method met the performance acceptance criteria set by the VMG in terms of sensitivity, specificity and accuracy that should be equal to or higher than 90%, 60% and 75%, respectively.

In what follows, a more in depth analyses of misclassified liquids and solids will be presented with respect to the drivers of *in vivo* classification. Recent papers have shown the importance of understanding these effects for the evaluation of alternative methods (Barroso et al., 2013; Adriaens et al., 2014). A full description of all the ocular effects that drive classification is available for a large set of reference chemicals, the so-called Draize eye test Reference Database (DRD) published by Barroso et al. (2016). In order to evaluate the predictive capacity of the SkinEthic™ HCE EIT test method and its limitations, the misclassified chemicals were correlated with the *in vivo* drivers of

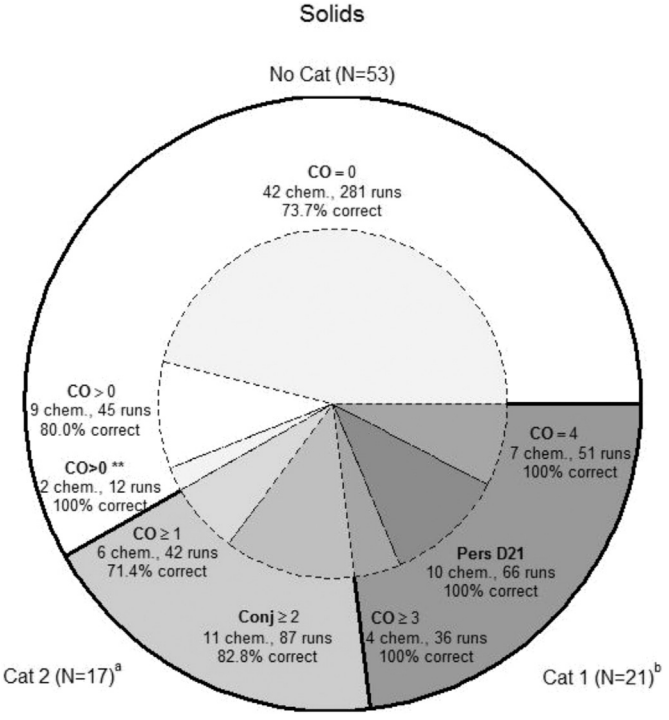

Fig. 7. Distribution of the solid chemicals (chem.) according to the drivers of classification (UN GHS Cat 1 and Cat 2) and according to the subgroups (UN GHS No Cat) as defined by Barroso et al. (2016). The proportion correct predictions corresponds with the number of runs that were correctly predicted over the total number of runs that were performed. [a]The data of 17/18 Cat 2 chemicals are included in the chart, chemical No. 43 was excluded since the driver could not be identified (SCNM). [b]The data of 21/24 Cat 1 chemicals are included in the chart, chemicals No. 49 and No. 93 were excluded since multiple studies were available that resulted in a different driver and No. 90 was excluded since the driver could not be identified (SCNM). "*CO = 0*" Corneal Opacity (CO) scores equal to 0 in all animals and all observed time points in the Draize eye test; "*CO > 0*" CO scores greater than 0 in at least one animal for at least one observed time point, **correspond with No Cat studies for which at least one animal had a mean of the scores of days 1–3 above the classification cut-off for at least one endpoint but not enough animals to generate a classification; "*CO mean ≥ 1*" mean CO scores of days 1–3 ≥ 1 in ≥60% of the animals; "*Conj mean ≥ 2*" mean Conjunctival Redness (CR) and/or Conjunctival Chemosis (CC) during the first three observation days ≥2 in ≥60% of the animals in absence of "*CO mean ≥ 1*"; "*CO mean ≥ 3*" mean CO scores of days 1–3 ≥ 3 in ≥60% of the animals; "*IR mean > 1.5*" mean Iritis (IR) scores of days 1–3 > 1.5 in ≥60% of the animals in absence of "*CO mean ≥ 3*"; "*Pers D21*" persistence of any ocular effect on day 21 in the absence of severity ("*CO mean ≥ 3*" and "*IR mean > 1.5*"); "*CO = 4*" at any observation time during the study in the absence of both severity and persistence.

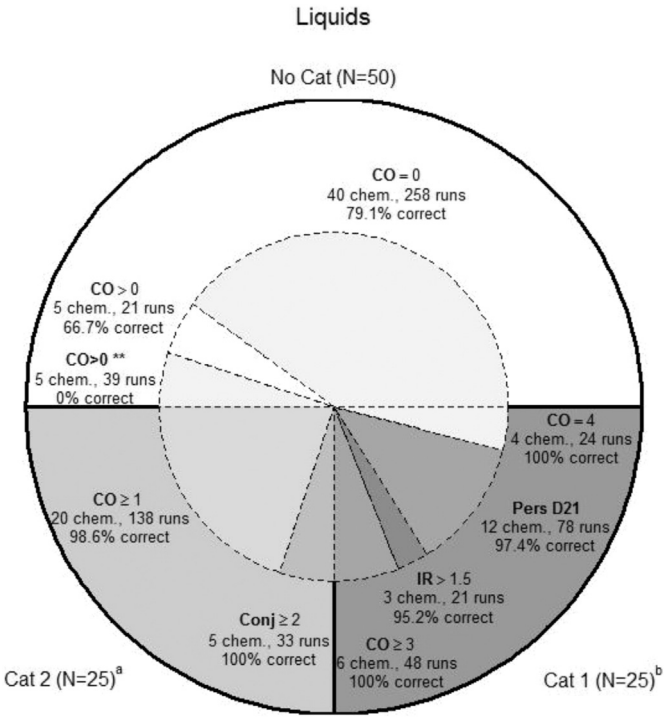

Fig. 8. Distribution of the liquid chemicals (chem.) according to the drivers of classification (UN GHS Cat 1 and Cat 2) and according to the subgroups (UN GHS No Cat) as defined by Barroso et al. (2016). The proportion correct predictions corresponds with the number of runs that were correctly predicted over the total number of runs that were performed. [a]The data of 25/28 Cat 2 chemicals are included in the chart, 1-decanol and poly(ethylene glycol) butyl ether were excluded since the driver could not be identified (SCNM), ethanol was excluded since multiple studies were available resulting in different drivers/classifications. [b]The data of 25/27 Cat 1 chemicals are included in the chart, anisole was excluded since the driver could not be identified (SCNM) and n-butanol was excluded since multiple studies were available that resulted in a different driver. "*CO = 0*" Corneal Opacity (CO) scores equal to 0 in all animals and all observed time points in the Draize eye test; "*CO > 0*" CO scores greater than 0 in at least one animal for at least one observed time point, **correspond with No Cat studies for which at least one animal had a mean of the scores of days 1–3 above the classification cut-off for at least one endpoint but not enough animals to generate a classification; "*CO mean ≥ 1*" mean CO scores of days 1–3 ≥ 1 in ≥60% of the animals; "*Conj mean ≥ 2*" mean Conjunctival Redness (CR) and/or Conjunctival Chemosis (CC) during the first three observation days ≥2 in ≥60% of the animals in absence of "*CO mean ≥ 1*"; "*CO mean ≥ 3*" mean CO scores of days 1–3 ≥ 3 in ≥60% of the animals; "*IR mean > 1.5*" mean Iritis (IR) scores of days 1–3 > 1.5 in ≥60% of the animals in absence of "*CO mean ≥ 3*"; "*Pers D21*" persistence of any ocular effect on day 21 in the absence of severity ("*CO mean ≥ 3*" and "*IR mean > 1.5*"); "*CO = 4*" at any observation time during the study in the absence of both severity and persistence.

classification as presented in the DRD (Barroso et al., 2016). Extensive analyses of *in vivo* UN GHS Cat 1 studies, presented in the DRD, showed that the most important drivers for Cat 1 classification are CO mean ≥ 3 (mean scores calculated from grading at day 1, 2 and 3 after instillation of the chemical in the eye) and CO persistence (pers) on day 21 in the absence of severity (CO mean ≥ 3) (Adriaens et al., 2014; Barroso et al., 2016). The most important drivers for a Cat 2 classification are CO mean ≥ 1 and Conjunctival Redness (CR) mean ≥ 2. Barroso et al. (2016) also suggested a critical revision of the current UN GHS decision criteria, one of the key conclusions of this analysis was that all classifiable Cat 1 effects should be present in more than 60% of the animals. The most important drivers of Cat 1 and Cat 2 classification named above were well represented in the solids and liquids chemicals set evaluated with the SkinEthic™ HCE EITS and EITL protocol (Figs. 7 and 8). The results of 9 chemicals (4 solids: No. 43, 49, 90, and 93 and 5 liquids: 1-decanol, ethanol, poly(ethylene glycol) butyl ether, anisole, and n-butanol) were not included in the pie charts since the driver could not be identified or because multiple studies were available for the same chemical and the driver differed between the repeat studies. From the regulatory perspective, it is important to note that out of the 51 *in vivo* UN GHS Cat 1 chemicals (24 solids and 27 liquids) that were tested with the optimized SkinEthic™ HCE EITL or EITS protocols, 49 (96.1%) chemicals were always correctly identified as C. Of note, Figs. 7 and 8 includes the results of 21 Cat 1 solids and 25 Cat 1 liquids. The false negative result obtained by VITO for the liquid 3-(2-aminoethylamino)propyl]-trimethoxysilane, was probably related to instability of the chemical (Alépée et al., 2016). The second chemical with a false negative prediction (1 out of 9 runs) was the liquid tetraethylene glycol diacrylate, this chemical was classified Cat 1 based on Iritis (IR) in the Draize eye test, an endpoint know to be of minor importance in driving Cat 1 classification (Barroso et al., 2016). With respect to the most important drivers of Cat 1 classification, an excellent predictive capacity was obtained for the solids with 100% correct predictions (Fig. 7) and a very high predictive capacity was obtained for the liquids with 100% correct predictions for the driver CO mean ≥ 3 and 97.4% correct predictions for the liquids that were classified Cat 1 based on persistence (Fig. 8). Overall 84.8% (39/46) of the *in vivo* UN GHS Cat 2 chemicals (18 solids and 28 liquids) were always predicted C. Of note, Figs. 7 and 8 show the results of 17 solids and 25 liquids, as mentioned before, the results of 1 solid and 3 liquids were omitted. For two *in vivo* UN GHS Cat 2B liquid chemicals (2-methyl-1-pentanol and ethyl-2-methylacetoacetate), only 1 out of 9 runs resulted in an under-prediction. The viability for 2-methyl-1-pentanol was 83.6% in one run whereas in 8 other runs, the viability was below 2% (Alépée et al., 2016). For ethyl-2-methylacetoacetate a borderline classification was obtained in one run (62%), whereas the viability in the other 8 runs was <60%. The solid, 1,5-Znaphthalenediol was under-predicted in 4 out of 9 runs, this chemical was predicted Cat 2A based on conjunctival effects only (Conj, corresponding with CR and CC effects). Four additional solid chemicals (No. 33, 36, 85, and 86) were under-predicted in the majority of the runs. Chemical No. 36 and No. 86 were both classified Cat 2 based on Corneal Opacity and No. 33 and No. 85 were classified Cat 2 based on conjunctival effects only. The performance of the SkinEthic™ HCE EIT method in terms of most important drivers of Cat 2 classification was high (Figs. 7 and 8). Cat 2 chemicals (6 solids and 20 liquids) that were classified based on Corneal Opacity were correctly predicted in 71.4% (solids) and 96.5% (liquids) of the runs. Furthermore, although the SkinEthic™ HCE EIT method models the cornea, a substantial proportion of the Cat 2 chemicals (11 solids and 5 liquids) that were classified based on conjunctival effects only were also identified correctly (82.8% and 100% of the runs for solids and liquids, respectively). This provides evidence that the SkinEthic™ HCE EIT method can also identify chemicals which result in *in vivo* conjunctival effects only. Concerning the false positives, it is interesting to note that the liquids of the subgroups CO > 0 and CO > 0** were often over-predicted by the SkinEthic™ HCE EITL method (44% and 100% of the runs, respectively). In the Draize eye test, those liquid chemicals induced CO scores greater than 0 in at least one animal for at least one observed time point. Moreover, for five liquid chemicals (CO > 0**), CO mean over the first three days was equal to or greater than 1 in one animal. This means that the method is very sensitive in detecting such *in vivo* effects. This relationship was not observed for the solids. In fact, the subgroups CO > 0 and CO > 0** were correctly predicted in 80% and 100% of the runs, respectively.

4. Conclusions

The present work assessed the reliability (WLR and BLR) and relevance (predictive capacity) of the SkinEthic™ HCE EIT test method to discriminate chemicals not requiring classification for serious eye damage/eye irritancy (No Cat) from chemicals requiring classification and labelling (Cat 1 and Cat 2). The SkinEthic™ HCE EIT test method is however not intended to differentiate between UN GHS Cat 1 (serious eye damage) and UN GHS Cat 2 (eye irritation). In fact, a chemical that is identified as requiring classification for eye irritation/serious eye damage with SkinEthic™ HCE EIT will require additional testing by another tier of a test strategy (Scott et al., 2010). A definitive classification can be established using e.g., OECD TG 437, 438, 460 or 492 (OECD TG 492, 2015).

The present study demonstrated that the SkinEthic™ HCE EIT method met the performance acceptance criteria set by the VMG in terms of WLR, BLR, sensitivity, specificity and accuracy that should be equal to or higher than 85%, 80%, 90%, 60% and 75%, respectively. The SkinEthic™ HCE EIT test method is currently in the work plan 2015 programme of the OECD for identifying No Cat chemicals.

Conflict of interest statement

The authors declare there are no conflicts of interest.

Funding

No external funding was received for this study.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

The authors gratefully acknowledge Tadeja Kerin and Iwona Wilinska (CRL, Edinburgh, United Kingdom), and An Jacobs and Karen Hollanders (VITO NV, Mol, Belgium) for their technical assistance and fruitful discussions. Likewise, the authors wish to thank Grégoire Sawka and Maeva Talleur (L'Oréal R&I, Aulnay sous Bois, France) for their in-depth technical support in the conduct of this study.

References

Adriaens, E., Barroso, J., Eskes, C., Hoffmann, S., McNamee, P., Alépée, N., Bessou-Touya, S., de Smedt, A., De Wever, B., Pfannenbecker, U., Tailhardat, M., Zuang, V., 2014. Retrospective analysis of the Draize test for serious eye damage/eye irritation: importance of understanding the *in vivo* endpoints under UN GHS/EU CLP for the development and evaluation of *in vitro* test methods. Arch. Toxicol. 88, 701–723. http://dx.doi.org/10.1007/s00204-013-1156-8.

Agresti, A., Coull, B.A., 1998. Approximate is better than exact for interval estimation of binomial proportions. Am. Stat. 52, 119126.

Alépée, N., Bessou-Touya, S., Cotovio, J., De Smedt, A., De Wever, B., Faller, C., Jones, P., Le Varlet, B., Marrec-Fairley, M., Pfannenbecker, U., Tailhardat, M., Van Goethem, F., McNamee, P., 2013. Cosmetics Europe multi-laboratory pre-validation of the SkinEthic™ reconstituted human corneal epithelium test method for the prediction of eye irritation. Toxicol. in Vitro 27, 1476–1488. http://dx.doi.org/10.1016/j.tiv.2013.02.009.

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., McNamee, P., 2015. Use of HPLC/UPLC–spectrophotometry for detection of formazan in in vitro Reconstructed

human Tissue (RhT)-based test methods employing the MTT-reduction assay to expand their applicability to strongly coloured test chemicals. Toxicol. in Vitro 29, 741–761. http://dx.doi.org/10.1016/j.tiv.2015.02.005.

Alépée, N., Leblanc, V., Adriaens, E., Grandidier, M.H., Lelièvre, D., Meloni, M., Nardelli, L., Roper, C.S., Santirocco, E., Toner, F., Van Rompay, A., Vinall, J., Cotovio, J., 2016. Multi-laboratory validation of SkinEthic HCE test method for testing serious eye damage/eye irritation using liquid chemicals. Toxicol. in Vitro 31, 43–53. http://dx.doi.org/10.1016/j.tiv.2015.11.012.

Barroso, J., Alépée, N., De Smedt, A., De Wever, B., Hibatallah, J., McNamee, P., Mewes, K.R., Millet, M., Pfannenbecker, U., Tailhardat, M., Templier, M., 2013. The importance of understanding drivers of irritation *in vivo* for selection of chemicals used in the development and evaluation of *in vitro* eye irritation assays: Cosmetics Europe analysis. Poster Presented at the In Vitro Testing Industrial Platform (IVTIP) Spring 2013 Meeting, Southampton, UK, May 14–16, 2013 (Available from: https://www.cosmeticseurope.eu/downloads/6057.html. Accessed on 05/10/2015).

Barroso, J., Alépée, N., Cole, T., Eskes, C., Freeman, S.J., Liska, R., McNamee, P., Pfannenbecker, U., Reus, A.A., Rubingh, C.M., Schaeffer, M.W., Zuang, V., 2015. EURL ECVAM — Cosmetics Europe prospective validation study of Reconstructed human Tissue-based test methods for serious eye damage/eye irritation testing. Poster Presented at the Congress of the European Societies of Toxicology (EUROTOX), Porto, Portugal, September, 13–16, 2015 (Available from: http://dx.doi.org/10.1016/j.toxlet.2015.08.507. Accessed on 05/10/2015).

Barroso, J., Pfannenbecker, U., Adriaens, E., Alépée, N., Cluzel, M., De Smedt, A., Hibatallah, J., Klaric, M., Mewes, K.R., Millet, M., Templier, M., McNamee, P., 2016. Cosmetics Europe compilation of historical serious eye damage/eye irritation *in vivo* data analysed by drivers of classification to support the selection of chemicals for development and evaluation of alternative methods/strategies: the Draize eye test Reference Database (DRD). Arch. Toxicol. (in press).

Core Team, R., 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (URL http://www.R-project.org/).

Cotovio, J., Grandidier, M.-H., Lelièvre, D., Bremond, C., Flamand, N., Loisel-Joubert, S., Van Der Lee, A., Capallere, C., Meunier, J.-R., Leclaire, J., 2007. The use of the reconstructed human corneal model (HCE) to assess in vitro eye irritancy of chemicals. ALTEX 14 (special issue), 343–350.

Cotovio, J., Grandidier, M.H., Lelièvre, D., Bremond, C., Amsellem, C., Maloug, S., Ovigne, J.M., Loisel-Joubert, S., Van Der Lee, A., Minondo, A.M., Capallere, C., Bertino, C., Alépée, N., Tinois-Tessonneaud, E., De Brugerolle de Fraissinette, A., Meuniera, J.R., Leclaire, J., 2010. In vitro assessment of eye irritancy using the reconstructed human corneal epithelial SkinEthic™ HCE model: application to 435 substances from consumer products industry. Toxicol. in Vitro 24, 523–537. http://dx.doi.org/10.1016/j.tiv.2009.11.010.

EC EURL ECVAM, 2014. Validation Study Report on the EURL ECVAM — Cosmetics Europe Prospective Validation Study of Reconstructed Human Corneal Epithelium-based Test Methods for Identifying Chemicals Not Requiring Classification and Labelling for Serious Eye Damage/Eye Irritation Testing (Available at: in publication).

ECETOC, 1998. Eye irritation reference chemicals data bank. Technical Report n. 48 (2), Brussels, Belgium.

ENV/JM/MONO(2015)23, 2015. OECD performance standards for the assessment of proposed similar or modified in vitro Reconstructed human Cornea-like Epithelium (RhCE) test methods for eye hazard, OECD environment, health and safety publications. Series on Testing & Assessment No. 216, p. 21.

ESAC, 2009. Statement on the Scientific Validity of Cytotoxicity/Cell-function Based *In Vitro* Assay for Eye Irritation Testing. Issued by the ECVAM Scientific Advisory Committee (ESAC) at its 31st Meeting, 10th July 2009. European Commission Joint Research Centre. Institute for Health and Consumer Protection. European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) (Available from: https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/docs-eye-irritation/ESAC31_CBA_eye-irritation_20091005.pdf. accessed 31/07/2015).

Freeman, S.J., Alépée, N., Barroso, J., Cole, T., Compagnoni, A., Rubingh, C., Eskes, C., Lammers, J., McNamee, P., Pfannenbecker, U., Zuang, V., 2010. Prospective validation study of reconstructed human tissue models for eye irritation testing. ALTEX 27 (special issue), 261–266.

Hartung, T., Bruner, L., Curren, R., Eskes, C., Goldberg, A., McNamee, P., Scott, L., Zuang, V., 2010. First alternative method validated by a retrospective weight-of-evidence approach to replace the Draize eye test for the identification of non-irritant substances for a defined applicability domain. ALTEX 27, 43–51.

ICCVAM, 2010. ICCVAM Test Method Evaluation Report: Current Validation Status of *In Vitro* Test Methods Proposed for Identifying Eye Injury Hazard Potential of Chemicals and Products. NIH Publication No. 10-7553. National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA (Available from: http://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/ocular/in-vitro-test-methods/tmer/index.html. (accessed 31.07.2015)).

Kaluzhny, Y., Kandarova, H., Handa, P., Kunilus, J., d'Argembeau-Thornton, L., Klausner, M., 2011. Development of the EpiOcular™ Eye Irritation Test (EpiOcular-EIT) for hazard identification and labelling of eye irritating chemicals in response to the requirements of the cosmetics directive and REACH legislation. ATLA 39, 339–364.

Kaluzhny, Y., Kandarova, H., Handa, Y., DeLuca, J., Truong, T., Hunter, A., Kearney, P., d'Argembeau-Thornton, L., Klausner, M., 2015. The EpiOcular Eye Irritation Test (EIT) for hazard identification and labelling of eye irritating chemicals: protocol optimisation for solid materials and the results after extended shipment. ATLA 43, 101–127.

Mosmann, T., 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J. Immunol. Methods 65, 55–63.

Nguyen, D.H., Beuerman, R.W., De Wever, B., Rosdy, M., 2003. Three-dimensional construct of the human corneal epithelium for in vitro toxicology. In: Salem, H., Katz, S.A. (Eds.), Alternative Toxicological Methods. CRC Press, pp. 147–159.

OECD, 2012. Test No. 460: Fluorescein Leakage test method for identifying ocular corrosives and severe irritants. OECD Guideline for the Testing of Chemicals, Section 4. OECD Publishing, Paris http://dx.doi.org/10.1787/9789264185401-en (16 pp. OECD Publishing).

OECD, 2013a. Test No. 437: Bovine Corneal Opacity and Permeability test method for identifying i) chemicals inducing serious eye damage and ii) chemicals not requiring classification for eye irritation or serious eye damage. OECD Guideline for the Testing of Chemicals, Section 4. OECD Publishing, Paris http://dx.doi.org/10.1787/9789264203846-en (27 pp. OECD Publishing).

OECD, 2013b. Test No. 438: Isolated Chicken Eye test method for identifying i) chemicals inducing serious eye damage and ii) chemicals not requiring classification for eye irritation or serious eye damage. OECD Guideline for the Testing of Chemicals, Section 4. OECD Publishing, Paris http://dx.doi.org/10.1787/9789264203860-en (20 pp. OECD Publishing).

OECD, 2015a. Test No. 491: The Short Time Exposure *in vitro* test method for identifying i) chemicals inducing serious eye damage and ii) chemicals not requiring classification for eye irritation or serious eye damage. OECD Guideline for the Testing of Chemicals, Section 4. OECD publishing http://dx.doi.org/10.1787/9789264229822-en 13 pp.

OECD, 2015b. Test No. 492: Reconstructed Human Cornea-like Epithelium (RhCE) test method for identifying chemicals not requiring classification and labelling for eye irritation or serious eye damage. OECD Guideline for the Testing of Chemicals, Section 4. OECD Publishing http://dx.doi.org/10.1787/9789264242548-en 26 pp.

Pfannenbecker, U., Bessou-Touya, S., Faller, C., Harbell, J., Jacob, T., Raabe, H., Tailhardat, M., Alépée, N., De Smedt, A., De Wever, B., Jones, P., Kaluzhny, Y., Le Varlet, B., McNamee, P., Marrec-Fairly, M., Van Goethem, F., 2013. Cosmetics Europe multi-laboratory pre-validation of the EpiOcular™ reconstituted human tissue test method for the prediction of eye irritation. Toxicol. in Vitro http://dx.doi.org/10.1016/j.tiv.2012.11.007.

Scott, L., Eskes, C., Hoffmann, S., Adriaens, E., Alépée, N., Bufo, M., Clothier, R., Facchini, D., Faller, C., Guest, R., Harbell, J., Hartung, T., Kamp, H., Varlet, B.L., Meloni, M., McNamee, P., Osborne, R., Pape, W., Pfannenbecker, U., Prinsen, M., Seaman, C., Spielmann, H., Stokes, W., Trouba, K., Berghe, C.V., Van Goethem, F., Vassallo, M., Vinardell, P., Zuang, V., 2010. A proposed eye irritation testing strategy to reduce and replace *in vivo* studies using Bottom-Up and Top-Down approaches. Toxicol. in Vitro 24, 1–9. http://dx.doi.org/10.1016/j.tiv.2009.05.019.

SkinEthic™ HCE EIT SOP, 2015. SkinEthic™ HCE EIT for the prediction of acute ocular irritation of chemicals. Available at: [http://ecvam-dbalm.jrc.ec.europa.eu/].

UN, 2013. United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS). ST/SG/AC.10/30/Rev.5, Part 3 Health Hazards—Chapter 3.3 Serious Eye Damage/Eye Irritation. United Nations, New York, USA, and Geneva, Switzerland, pp. 135–147 (Available from: http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev05/English/ST-SG-AC10-30-Rev5e.pdf. accessed 31/07/2015).

Van Goethem, F., Adriaens, E., Alépée, N., Straube, F., De Wever, B., Cappadoro, M., Catoire, S., Hansen, E., Wolf, A., Vanparys, P., 2006. Prevalidation of a new in vitro reconstituted human cornea model to assess the eye irritating potential of chemicals. Toxicol. in Vitro 20, 1–17. http://dx.doi.org/10.1016/j.tiv.2005.05.002.

Wilson, S.L., Ahearne, M., Hopkinson, A., 2015. An overview of current techniques for ocular toxicity testing. Toxicol. in Vitro 327, 32–46. http://dx.doi.org/10.1016/j.tox.2014.11.003.