Regulatory Toxicology and Pharmacology 126 (2021) 105044

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Inter-laboratory performance of ICE histopathology scoring to identify UN GHS Category 1 surfactants and non-extreme pH detergents☆

Chantra Eskes [a,*], Marcel V.W. Wijnands [b], Martina Hermann [c], Menk Prinsen [b,1], Caroline Bertein [d,1], Dominic Byrne [d,1], Robert Kreutzer [e], Vasanthi Mowat [f], Ian Taylor [f], Erwin van Vliet [a,1], Klaus Weber [e]

[a] SeCAM Services and Consultation on Alternative Methods, Magliaso, Switzerland
[b] TNO Triskelion, Zeist, the Netherlands
[c] Henkel AG & Co. KGaA, Duesseldorf, Germany
[d] A.I.S.E., Brussels, Belgium
[e] Anapath Services GmbH, Liestal, Switzerland
[f] Envigo, UK

ARTICLE INFO

Handling Editor: Dr. Martin Van den berg

Keywords:
Isolated chicken eye test
Histopathology
OECD TG 438
Detergent products

ABSTRACT

The inter-laboratory performance of Isolated Chicken Eye (ICE) histopathology scoring was assessed for predicting EU CLP/UN GHS Cat. 1 surfactants. Furthermore, the predictive capacity of ICE histopathology was evaluated for the combined dataset of surfactants and existing data for non-extreme pH ($2 < pH < 11.5$) detergents. Use of ICE histopathology led to increased sensitivity compared to the ICE test method alone for surfactants. When combined with the existing dataset of detergents, use of histopathology in addition to the standard ICE test method decreased the false negative rates from 64% (14/22) to 27% (6/22); increased accuracy from 53% (16/30) to 77% (23/30); and led to acceptable level of false positives (from 0/8 to 1/8 (12.5%). Moreover, good reproducibility of ICE histopathology predictions conducted on the same slides was found between pathologists and peer-reviewers from three independent laboratories (10/12 or 83%) and over time. Use of ICE histopathology was therefore found suitable to predict EU CLP/UN GHS Cat. 1 surfactants and non-extreme pH detergents. In addition, appropriate reproducibility of ICE histopathology was found, provided that i) an internal peer-review system was in place; ii) original slides were assessed to enable evaluation of three dimensional effects; and iii) appropriate training and proficiency appraisal were conducted.

1. Introduction

In the European Union's regulatory framework for classification and labelling of mixtures, the classification for potential hazards was governed until 2015 by the Dangerous Substances Directive (DSD, 67/548/EEC) and by the Dangerous Preparations Directive (DPD, 1999/45/EC). Since mid-2015 both directives have been replaced by the Regulation on Classification, Labelling and Packaging of Substances and Mixtures (EU CLP, Regulation No 1272/2008) which introduced the United Nations Globally Harmonised System (UN GHS). For non-extreme pH mixtures, the EU CLP give priority to the use of all available data for the mixture under evaluation; a second tier requires the use of bridging principles. If the first and second tiers are not conclusive, the additivity/calculation method might then be used.

An impact assessment carried out by the International Association for Soaps, Detergents and Maintenance Products (A.I.S.E.,[2]) showed that the use of the EU CLP/UN GHS additivity approach (EU CLP; EC, 2008) can result in over-conservative classification and labelling for eye effects

☆ Defined as: "a mixture (excluding dilutions of single surfactant) containing one or more surfactants at a final concentration of >3%, intended for washing and cleaning processes. Detergents may be in any form (liquid, powder, paste, bar, cake, moulded piece, shape, etc.) and marketed for or used in household, or institutional or industrial purposes." (OECD, 2018a).

* Corresponding author.
E-mail address: chantra.eskes@secam-ce.eu (C. Eskes).

[1] Affiliation at the time of the study.

[2] www.aise.eu. Accessed on 24.05.2021.

https://doi.org/10.1016/j.yrtph.2021.105044
Received 5 March 2021; Received in revised form 25 August 2021; Accepted 5 September 2021
Available online 8 September 2021

Table 1
UN GHS predicted classification based on the combination of categories obtained for corneal swelling, corneal opacity, and fluorescein retention as described in the OECD TG 438 (2018a).

UN GHS classification	Combinations of the 3 endpoints
No Category	3 x I
	2 x I, 1 x II
	2 x II, 1 x I
No Prediction Can be Made	Other combinations
Category 1	3 x IV
	2 x IV, 1 x III
	2 x IV, 1 x II[a]
	2 x IV, 1 x I[a]
	Corneal opacity ≥3 at 30 min (in at least 2 eyes)
	Corneal opacity = 4 at any time point (in at least 2 eyes)
	Severe loosening of the epithelium (in at least 1 eye)

[a] Combinations less likely to occur.

of many detergent and cleaning products not requiring such a classification according to consistent animal, *in vitro* and human experience data (A.I.S.E. unpublished data). As a result, daily use detergents such as hand dish wash liquids, may display a "corrosive" pictogram for severe eye effect (Category 1) that was previously (under the former DSD/DPD classification system) reserved for products truly being corrosive such as extreme pH drain cleaners. This resulting over-labelling can potentially confuse end-users and may lead to the underestimation of real risk when this is merited due to trivialization of labelling.

A.I.S.E. believes it is critical to generate reliable data which can accurately predict the hazard potential to humans of detergent and cleaning products and ensure correct classification, but that no animal testing should be conducted on finished products. For this reason, A.I.S.E. conducted an *in vitro* testing program between 2010 and 2012 to investigate the usefulness of regulatory adopted *in vitro* test methods to reliably classify detergent and cleaning product formulations. The results showed that specific ICE histopathological effects were found to correlate with serious eye damage classification under EU CLP/UN GHS system induced by non-extreme pH ($2 < pH < 11.5$) detergents *in vivo* (Cazelle et al., 2014, 2015). Epithelial vacuolation (in mid and lower layers) and epithelial erosion (of at least moderate level) were found to be the most typical histopathology effects induced by the non-extreme pH detergents classified *in vivo* as EU CLP/UN GHS Category 1 (Cat. 1). Use of these histopathology criteria for non-extreme pH detergents substantially increased the sensitivity of the standard ICE prediction model for EU CLP/UN GHS Cat. 1 identification (from 0% to at least 75%, n = 8) whilst maintaining a good accuracy (73%, n = 30), and an acceptable specificity (from 100% to 73%, n = 22). In particular, it allowed the correct identification of 5 of 6 non-extreme pH detergents classified as EU CLP/UN GHS Cat. 1 based on *in vivo* persistence of effects i.e., having tissue effects that do not reverse 21 days after treatment and that do not lead to severity of effects that would warrant an EU CLP/UN GHS Cat. 1 classification (Cazelle et al., 2014). In contrast, for extreme pH detergents, 5 of the 6 tested *in vivo* EU CLP/UN GHS Cat. 1 were classified in *in vivo* studies due to severity of effects and not persistence. In this case, the A.I.S.E. histopathology criteria did not improve the sensitivity of the standard ICE test method (83%, n = 6), whilst it strongly decreased specificity (from 83% to 33%, n = 12), and accuracy (from 83% to 50%, n = 18) (Cazelle et al., 2015). These data indicate that there are specific applicability domains for the use of the ICE histopathology for non-extreme pH ($2 < pH < 11.5$) detergents that are likely based on the mode of action of the tested detergents.

In the present study, A.I.S.E. further investigated the suitability of ICE histopathology to identify EU CLP/UN GHS Cat. 1 surfactants in order to enlarge its dataset with chemicals having Draize *in vivo* eye hazard data. In particular, the reproducibility of ICE histopathology scoring between pathologists and peer-reviewers of multiple independent laboratories was assessed. Surfactants represent major ingredients for detergents often driving the ocular hazard classification of non-extreme pH ($2 < pH < 11.5$) detergents. Furthermore, both surfactants and non-extreme pH ($2 < pH < 11.5$) detergents have a similar proportion of substances classified *in vivo* based on persistence of effects (A.I.S.E., unpublished data). In addition, the predictive capacity of ICE histopathology was evaluated for the new dataset on surfactants combined with the existing dataset on non-extreme pH ($2 < pH < 11.5$) detergents. Detergents are identified here as a mixture (excluding dilutions of single surfactant) containing one or more surfactants at a final concentration of >3%, intended for washing and cleaning processes. Detergents may be in any form (liquid, powder, paste, bar, cake, moulded piece, shape, etc.) and marketed for or used in household, or institutional or industrial purposes (OECD TG 438, 2018a).

2. Material and methods

2.1. ICE test method

The ICE test method was performed according to the recently revised OECD Test Guideline 438 (OECD TG 438, 2018a) by TNO Triskelion. Briefly, chicken eyes were obtained from a food chain poultry slaughterhouse in the vicinity of the TNO Triskelion testing facility in the Netherlands. Sufficient suitable eyes (free of corneal opacity and fluorescein staining) were enucleated from the heads and placed in the ICE superfusion chambers generally within 1 h after slaughter. No eye or experiment was rejected because of pre-dose observations for eye quality during acclimatization or post-dose observations in the control eyes. In general, per testing day one test run of maximally 3 substances, tested simultaneously, was performed. The test run included a negative control eye treated with saline and positive control eyes treated with benzalkonium chloride 5% (w/v) for liquid substances or sodium hydroxide for powder substances.

Test substances were applied neat in one single dose of 30 μL (liquids) or 30 mg (solids) onto the cornea of isolated chicken eyes (3 eyes per test substance) for an exposure time of 10 s after which the cornea was rinsed with physiological saline. At defined intervals up to 4 h after treatment, assessment of corneal swelling, corneal opacity and fluorescein retention of the corneal epithelium was made by using a Haag-Streit slit-lamp microscope. On the basis of the severity (maximum mean score of the three eyes) of the observed findings for corneal swelling, corneal opacity and fluorescein retention, the effects were divided into four classes, i.e., I = no effect; II = slight effect; III = moderate effect; IV = severe effect (for detailed criteria refer to OECD TG 438, 2018a). The final irritation classification was determined using the highest mean score (of the three eyes) observed at any time point for corneal swelling and for corneal opacity, whereas for fluorescein retention the mean score obtained at 30 min post treatment was used. The ensuing classes obtained for the three endpoints were then combined to derive the *in vitro* classification based on the defined prediction model as described in Table 1.

2.2. ICE histopathology

The treated corneas (eyes) were collected by TNO in a neutral aqueous phosphate-buffered 4% solution of formaldehyde at termination of the ICE test, i.e. 4 h after treatment. For this purpose, the eyes were first cut in half with a scalpel just behind the level of the lens and through the vitreous body. The half with the cornea and lens was placed in a glass container with approximately 20 mL of formalin. After fixation for at least 24 h, the tissue was trimmed with scissors in such a way that a thin piece containing the entire cornea and the adjacent sclera were embedded in paraffin wax. Longitudinal serial slides (sectioned at 5 μM) were prepared from the central area of the cornea and further processed with the staining. In general, the directions given in the manual AFIP Laboratory Methods in Histotechnology (Prophet et al., 1992) were

Table 2
Semi-quantitative histopathology scoring system used for isolated chicken eyes that were fixed, trimmed, embedded in paraffin wax, sectioned and stained (extract from OECD TG 438, 2018a).

Parameter	Observation	Score	Description[a]
Epithelium: erosion	Very slight	½	Few single cells up to the entire single superficial layer
	Slight	1	Up to 3 layers are gone
	Moderate	2	Up to 50% of the epithelial layer is gone[a]
	Severe	3	Epithelial layer is gone up to the basement membrane
Epithelium: vacuolation	Very slight	½	Single to few scattered cells
Separately scored for the top, mid, and lower parts of the epithelium[b]	Slight	1	Groups of vacuolated cells or single string of cells with small vacuoles
	Moderate	2	Up to 50% of the epithelium consists of vacuolated cells[a]
	Severe	3	50–100% of the epithelium consists of vacuolated cells
Epithelium: necrosis[c]	Normal	-	<10 necrotic cells[d]
	Very slight	½	10–20 necrotic cells[d]
	Slight	1	20–40 necrotic cells[d]
	Moderate	2	Many necrotic cells but <50% of the epithelial layer
	Severe	3	50–100% of the epithelial layer is necrotic.
Stroma: pyknotic nuclei [e; f]	Normal	-	<5 pyknotic nuclei
	Slight	1	5–10 pyknotic nuclei
In top or bottom region	Moderate	2	>10 pyknotic nuclei
Stromal disorder of fibres [f]	Present	P	Irregular appearance of the fibres.
Endothelium: necrosis	Present	P	The endothelium consists of only one layer, so a grade is not relevant

Note: Annex II of the OECD GD 160 (12) displays an Atlas with typical photomicrographs of untreated as well as treated Isolated Chicken Eyes illustrating the various possible histopathological effects described above.

[a] Over the entire cornea except in case of test chemicals (e.g. some solid chemicals) causing localized effects despite of the homogenous application of the test chemical as required within the OECD TG 438. In this case the evaluation should be based on the localized effects at the site(s) of exposure.

[b] Top, mid and lower parts represent equal one third parts of the epithelial layer each. If the top layer is missing, the mid layer does not become the 'new' top layer, but is still the mid layer (see Annex II of the OECD GD 160 for more details (12)).

[c] Only necrosis of attached cells/tissues.

[d] Necrotic cells are counted across the entire length of the cornea (there is no need for a specific fixed length to report cell counts because the entire length of the cornea is consistent on each slide as there is almost no variation in the size of the chicken eyes used and in the size of the samples evaluated microscopically). The scoring system uses absolute cell counts from 'normal' to 'slight', versus a percentage for 'moderate' and 'severe'. This is due to the way the evaluation is performed by the examiner: necrotic cells are seen as individual items. If there are more, they are usually scattered. Therefore, the examiner counts them to get an impression of the amount of necrosis. This is in contrast to erosion, for which the first effect the examiner notices is that a part of the epithelium is missing, so it makes sense to use an estimated percentage of loss.

[e] The ICE test method already includes a precise measurement of the thickness of the cornea using a slit lamp microscope. Therefore, swelling of the stroma is not separately scored during the subsequent histopathological evaluation.

[f] The stromal effects that are scored consist of (1) pyknotic nuclei, which originate from the scoring system used by Maurer (2001) based on his observations in corneas of rabbits after *in vivo* exposure (described as keratocyte loss/necrosis), and of (2) disorder of fibres. Regarding (1), the presence of pyknotic nuclei is observed only occasionally and the development of pyknotic nuclei is proposed to be dependent on the depth of injury and/or the inflammation process of the cornea *(in vivo)*. Furthermore, due to the elongated form of the stromal fibroblasts, normal nuclei could be misleadingly considered as pyknotic nuclei depending on the section orientation of cells. Regarding (2), the observation and scoring of disorder of fibres may be difficult because the stromal fibres already show a "natural" disorder. The processing of the cornea for microscopy can also contribute to an artificial disorder of stromal fibres. In both cases (pyknotic nuclei and disorder of fibres), these observations coincide with severe corneal effects already observed by the slit-lamp microscope observations, and with effects observed in the mid and/or lower epithelial layer.

Table 3
Histopathology decision criteria to be used in addition to the standard validated ICE test method for the identification of EU CLP/UN GHS Cat. 1 non-extreme pH ($2 < pH < 11.5$) detergents and surfactants (extract from OECD, 2018a).

Tissue layer	Effects triggering eye serious damage (EU CLP/UN GHS Category 1) identification
Epithelium	- erosion ≥ moderate (score 2), in at least 2 out of 3 eyes - and/or any vacuolation (≥very slight, score ½) observed in the mid and/or lower parts, in at least 2 out of 3 eyes - or if erosion ≥ moderate (score 2) in 1 out of 3 eyes + vacuolation ≥ very slight in mid and/or low part (score ½) is observed in at least another eye out of the 3 eyes - and/or necrosis ≥ moderate (score 2), observed in at least 2 out of 3 eyes

Table 4
Prediction model used for the identification of non-extreme pH ($2 < pH < 11.5$) detergents and surfactants based on ICE histopathology evaluations (extract from OECD, 2018a).

Standard ICE	ICE histopathology criteria described in Table 3	EU CLP/UN GHS Classification
No Prediction Can be Made	Criteria met	EU CLP/UN GHS Category 1
No Prediction Can be Made	Criteria not met	No Prediction Can be Made

followed using Periodic Acid-Schiff (PAS) staining as described previously (Prinsen et al., 2011). Semi-quantitative microscopic evaluation of PAS stained corneas was performed by TNO Triskelion and by the participating laboratories according to the criteria described in Table 2.

Due to the standard ICE exposure, test chemicals usually cause homogenous effects in the cornea of the isolated chicken eyes, so that the mean of histopathological effects over the entire slide was scored. In cases where test chemicals caused focal or multifocal effects confined to certain spots despite their homogenous application (e.g., as for some solid test chemicals), the histopathological scoring was conducted based on the localized adverse effects observed where exposure to the test chemical occurred.

Only effects that are observed were scored. No assumptions were made (e.g., if the top layer of the epithelium was missing it was considered not possible to score for vacuolation in that layer). Attention was also given to distinguishing treatment-related effects from histopathological artefacts and/or background morphology. For this purpose, the recommendations given in the Atlas presented within the revised OECD GD 160 (2018b) were used. Furthermore, consolidated training, transferability and proficiency appraisal were conducted as described below. A predicted EU CLP/UN GHS Cat. 1 classification or a 'No Prediction Can Be Made' result was derived from the observed histopathological effects based on the decision criteria and prediction model shown in Tables 3 and 4. Effects on stroma such as pyknotic nuclei and endothelium effects were also reported as observations.

2.3. Preliminary training and transferability studies for ICE histopathology

Due to the fact that, at the time of its performance, no other studies investigating the reproducibility of histopathology sections for regulatory purposes were known to A.I.S.E., this study represented a 'learning by doing' process. As a consequence, preliminary trials took place that

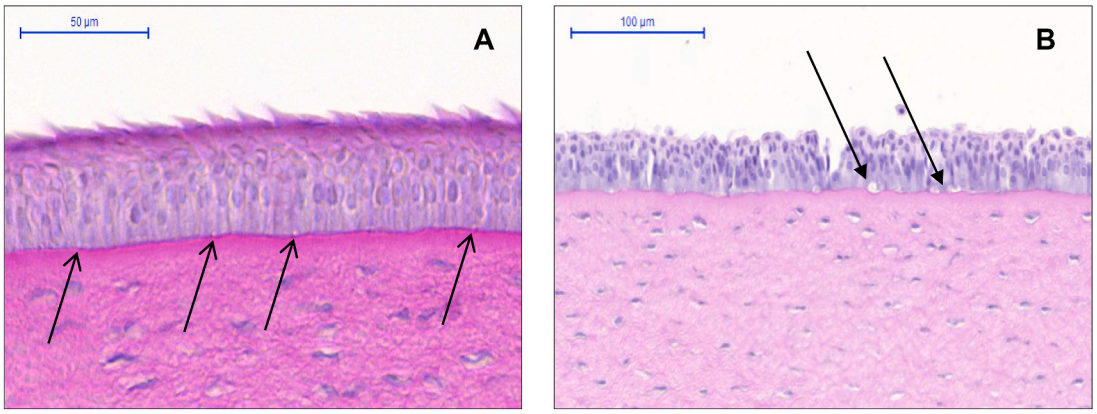

Fig. 1. Example of the differences between: A) 'small vacuoles' observed close to the basement membrane which are part of the background morphology alterations observed for saline controls. Arrows show examples of small vacuoles (i.e., < 1/3 the size of the nuclei of adjacent epithelial cells) close to the basement membrane observed as part of background morphology. B) 'true vacuoles' to be scored as an adverse effect in a treated eye. Arrows show examples of true vacuoles (≥ 1/3 the size of the nuclei of adjacent epithelial cells) that should be scored as a histopathological effect related to test chemical exposure and which is not part of background morphology.

Table 5
Clarification of histological effects that allow distinguishing 'true vacuoles' to be scored from the 'small vacuoles' observed as background morphology alterations.

Vacuole size	Vacuole location	Score for vacuolation? (Yes/No)
Equal or bigger than 1/3 the size of the nuclei of adjacent epithelial cells	Anywhere in epithelium	Yes
Smaller than '1/3 the size of the nuclei of adjacent epithelial cells'	Attached to the basement membrane	No
Smaller than '1/3 the size of the nuclei of adjacent epithelial cells'	Above the first row of nuclei (from bottom to top) in the epithelium	Yes
Smaller than '1/3 the size of the nuclei of adjacent epithelial cells'	Between basement membrane and first row of nuclei in the epithelium (from bottom to top)	Consult additional information[a]

[a] Make use of a step-wise approach: Step 1: Observe the other two isolated chicken eyes. If no effects are observed in the two other eyes vacuolation is not scored. If unequivocal effects are observed in one of the two eyes, vacuolation is scored. Step 2: If the results are still unclear based on the other two eyes (e.g., unclear vacuolation), consult observations from the standard ICE test method to take into account e.g. possible localized effects by solid materials.

helped to define the elements needed for appropriate conduct of a reproducibility assessment of ICE histopathology involving multiple laboratories. These trials are described below, whereas the final evaluation of the ICE histopathology reproducibility is discussed in a separate section.

A meeting between pathologists, from three laboratories that conduct the standard ICE testing was organized to discuss the best approach for assessing the reproducibility of ICE histopathology. From this meeting the following recommendations were made:

- Need to perform histopathology training followed by a transferability assessment based on the evaluation of blinded slides before reproducibility is assessed;
- Need for laboratories to have an internal system of peer-review of histopathology observations in place as recommended by e.g. Morton et al. (2010) and by the OECD Advisory document n. 16 on GLP requirements for peer review of histopathology (OECD, 2014);
- Need to have clear guidance on the performance of histopathology observations such as provided within the revised OECD GD 160 (2018b), including the Atlas on ICE histopathology observations and the need for pathologists to build a bandwidth of background morphological effects based on negative controls.

During the preliminary trials it was noticed that use of photomicrographs did not allow the observation of effects for which a three-dimensional assessment is required such as 'ghost cells' (i.e., cells that lost their nuclei during technical processing and so should not be scored as vacuolation) and 'foamy cells' (i.e., cells presenting very fine vacuolation which causes a foamy appearance of the cytoplasm, an effect which should be scored for vacuolation). To address this, further training, transferability and reproducibility assessment were conducted based on original ICE histopathology sections (and not on photomicrographs or scanned histopathology sections).

Furthermore, during the preliminary trials, the reference laboratory (i.e., TNO Triskelion) noted the need to better clarify 'small vacuoles' that are part of the background morphology of the ICE sections, from 'true vacuoles' that should be scored as histopathology effects related to test chemical exposure (Fig. 1). In order to thoroughly characterize the 'small vacuoles,' A.I.S.E. and TNO Triskelion assessed a total of 112 slides, including 38 slides with sections of saline negative controls obtained from at least 8 independent studies and 74 slides with sections of treated samples coming from at least three independent studies (representing 8 slides with sections of positive controls and triplicate slides from both 12 EU CLP/UN GHS Cat. 1 test chemicals, and 10 test chemicals not inducing EU CLP/UN GHS Cat. 1 effects *in vivo*). Based on this evaluation, the most relevant aspects to allow the discrimination of a 'true vacuole' (derived from exposure to a test chemical) from a 'small vacuole' observed as background morphology were identified to be the vacuole size and location. On this basis, clarifications on scoring criteria have been defined as described in Table 5. When using these criteria, no true vacuoles were observed in the saline negative controls evaluated.

In addition, it was noted that necrosis of cells/tissue that detach but are not washed away should not be scored. Only necrosis of cells/tissue that remain attached to the tissue are scored. This is due to the fact that some detached cells may remain in the prepared slides, whereas depending on the severity of effects some detached cells may be removed from the slides due to e.g. technical issues that are not related to chemical exposure. Finally, it was noted that effects/changes close to the limbus should be scored if the tissue architecture is preserved. However, effects/changes occurring within the limbus should not be scored as it might represent effects not linked to the chemical exposure.

Table 6
Individual scoring of ICE histopathology effects obtained from the pathologists and peer-reviewers of three independent laboratories. Each cell shows the scores given for the three observed eyes according to the criteria described in Table 2. In case no effects were observed a '-' is reported. Text in bold represent scores having met the criteria for identification of EU CLP/UN GHS Cat. 1.

Part 1 Chemical name	*In vivo*EU CLP/UN GHS Cat.	Epithelial Erosion			Epithelial Necrosis			Epithelial Vacuolation (top)			Epithelial Vacuolation (mid/low)		
		Ref. lab.	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2
Benzensulphonylchloride	**Draize Cat. 1**	0.5	–	–	**2**	**3**	–	1	2	1	**1/1**	**2/2**	**1/1**
		0.5	0.5	–	**2**	**3**	–	1	1	2	**2/1**	**3/3**	**2/2**
		–	–	–	**2**	**2**	–	1	2	2	**1/1**	**2/1**	**2/1**
Cetyl pyridinium bromide 100%	**Draize Cat. 1**	1	**1**	**1**	–	–	–	0.5	2	–	**0.5/**	**0.5/1**	**1/2**
		2	**2**	**2**	–	–	–	–	–	–	**0.5**	**0.5/2**	**-/2**
		1	**2**	**2**	–	–	0.5	–	–	–	**0.5/ 0.5**	**1/2**	**1/2**
											0.5/ 0.5		
Cetyl pyridinium bromide 10%	**Draize Cat. 1**	1	1	1	0.5	–	–	0.5	–	–	**-/1**	**0.5/**	**-/0.5**
		1	1	1	–	–	–	0.5	–	–	**-/1**	**0.5**	**-/0.5**
		1	1	2	0.5	–	–	0.5	–	–	**-/0.5**	**0.5/ 0.5**	**–**
												–	
Domiphen bromide 10%	**Draize Cat. 1**	1	**2**	0.5	–	0.5	0.5	–	–	1	**0.5/**	**1/0.5**	**0.5/-**
		1	**2**	1	–	–	–	–	–	–	**0.5**	**1/1**	**1/0.5**
		0.5	**–**	–	–	–	–	–	–	–	**0.5/ 0.5**	**0.5/ 0.5**	**-/-**
											–		
Stearyltrimethylammonium chloride - 10%	**Draize Cat. 1**	1	0.5	1	–	–	–	2	1	1	**0.5/**	**0.5/1**	**0.5/**
		1	1	0.5	0.5	–	–	3	1	1	**0.5**	**1/1**	**0.5**
		1	1	0.5	–	–	0.5	3	2	1	**1/1**	**1/0.5**	**1/-**
											1/0.5		**0.5/ 0.5**
Benzalkonium chloride 5%	**Draize Cat. 1**	**3**	**3**	**3**	–	–	–	–	–	–	**-/0.5**	-/-	**-/0.5**
		2	**2**	**2**	–	2	–	1	–	–	**-/-**	-/-	**0.5/-**
		3	**3**	**2**	–	–	–	–	–	–	**-/0.5**	-/-	**0.5/-**
Cetyl pyridinium bromide 1% - study 1 (warmed)	Draize Cat. 2A	1	1	1	–	–	0.5	1	1	3	-/-	-/-	-/-
		1	1	1	–	–	–	0.5	–	1	-/-	-/-	-/-
		1	1	1	–	0.5	0.5	0.5	1	1	-/-	-/-	-/0.5
Cetyl pyridinium bromide 1% - study 2 (unwarmed)	Draize Cat. 2A	1	1	0.5	–	–	–	2	0.5	0.5	-/-	-/-	-/-
		1	0.5	1	–	–	–	3	–	0.5	-/-	-/-	-/-
		1	1	1	–	0.5	0.5	3	0.5	1	-/-	-/-	0.5/-
N-Lauroyl sarcosine Na salt 10%	Draize Cat. 2A	0.5	0.5	1	0.5	0.5	0.5	0.5	–	0.5	-/-	-/-	-/-
		0.5	0.5	0.5	0.5	1	0.5	0.5	–	–	-/-	-/-	-/-
		0.5	0.5	0.5	1	1	–	0.5	–	–	-/-	-/-	-/0.5
1-Ethyl-3-methylimidazolium ethyl sulphate	Draize No Cat.	0.5	0.5	0.5	0.5	0.5	–	0.5	–	0.5	-/-	-/-	**-/0.5**
		0.5	–	–	–	–	–	0.5	0.5	0.5	-/0.5	-/0.5	**-/-**
		0.5	0.5	–	0.5	0.5	–	0.5	–	1	-/-	-/-	**-/0.5**
Cetyl pyridinium bromide −0.1%	Draize No Cat.	0.5	0.5	0.5	–	–	–	1	–	1	-/-	-/-	-/-
		0.5	0.5	0.5	–	–	–	1	–	0.5	-/0.5	-/0.5	-/-
		0.5	0.5	–	–	–	–	1	–	1	-/-	-/-	-/-
Polysorbate 20 (Tween 20)	Draize No Cat.	0.5	0.5	–	0.5	0.5	–	1	–	–	-/-	**-/0.5**	-/-
		0.5	0.5	0.5	1	1	–	1	–	–	-/-	**-/0.5**	-/0.5
		–	0.5	0.5	–	0.5	–	1	–	–	-/-	**-/-**	-/-

(*continued on next page*)

2.4. ICE histopathology reproducibility study

Based on the clarification of histopathology effects as described above, the reproducibility of ICE histopathology scoring by multiple laboratories was assessed. The following laboratories took part in this ICE histopathology reproducibility assessment:

- TNO Triskelion (Netherlands) as the reference laboratory;
- Envigo (UK) as an experienced laboratory with ICE histopathology;
- Anapath (Switzerland) as a laboratory experienced with histopathology but naïve with regard to ICE histopathology.

The reproducibility of the ICE histopathology was determined independently from the reproducibility of the standard ICE test method in order to understand the contribution from histopathology independently from the known reproducibility of the standard ICE test method (OECD, 2018c). Although the overall reproducibility of the standard ICE together with the ICE histopathology has not been characterized in the present study, such characterization was not considered necessary on the basis that:

- ICE histopathology is used only to further identify EU CLP/UN GHS Cat. 1 non-extreme pH detergents and surfactants, whereas test chemicals predicted as No Prediction Can Be Made remain as such if ICE histopathology criteria are not met,
- Histopathology may be conducted by a different laboratory than the laboratory conducting the standard ICE test method (i.e., without histopathology),
- The approach undertaken allows the evaluation of the reproducibility of the standard ICE test method and the reproducibility of the ICE histopathology separately.

Table 6 (*continued*)

Part 1 Chemical name	*In vivo*EU CLP/UN GHS Cat.	Epithelial Erosion			Epithelial Necrosis			Epithelial Vacuolation (top)			Epithelial Vacuolation (mid/low)		
		Ref. lab.	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2
Part 2 Chemical name	*In vivo*EU CLP/UN GHS Cat.	Stromal disorder of fibres			Stromal pyknotic nuclei			Endothelium necrosis			Notes		
		Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2	Ref. lab	Lab. 1	Lab. 2
Benzensulphonylchloride	**Draize Cat. 1**	–	–	–	–	–	–	–	–	–	A	A	A
		–	–	–	–	–	–	–	–	–	A	A	–
		–	–	–	–	–	–	–	–	–	A	A	–
Cetyl pyridinium bromide 100%	**Draize Cat. 1**	–	–	–	–	–	–	–	–	–	B	B	–
		–	–	–	–	–	–	–	–	–	B	B	–
		–	–	–	–	–	–	–	–	–	B	B	D
Cetyl pyridinium bromide 10%	**Draize Cat. 1**	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
Domiphen bromide 10%	**Draize Cat. 1**	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	C	–
Stearyltrimethylammonium chloride - 10%	**Draize Cat. 1**	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
Benzalkonium chloride 5%	**Draize Cat. 1**	–	–	–	–	–	–	Present	–	Present	B	–	D
		–	–	–	–	–	–	Present	Present	–	B	–	D
		–	–	–	–	–	–	Present	Present	–	B	–	D
Cetyl pyridinium bromide 1% - study 1 (warmed)	Draize Cat. 2A	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
Cetyl pyridinium bromide 1% - study 2 (unwarmed)	Draize Cat. 2A	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
N-Lauroyl sarcosine Na salt 10%	Draize Cat. 2A	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
1-Ethyl-3-methylimidazolium ethyl sulphate	Draize No Cat.	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
Cetyl pyridinium bromide −0.1%	Draize No Cat.	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
Polysorbate 20 (Tween 20)	Draize No Cat.	–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–
		–	–	–	–	–	–	–	–	–	–	–	–

A: Multifocal; B: Epithelial layer (partly) detached from basal membrane; C: Clefts in epithelium but no detachment; D: Parts of the mid layer still present.

The current study addressed the reproducibility of histopathological semi-quantitative scoring and its resulting classification predictions. For this purpose, a meeting was organized between the participating laboratories at the premises of the reference laboratory (TNO Triskelion). In order to ensure appropriate training and proficiency appraisal, the following procedures were undertaken before assessing the ICE histopathology reproducibility by the pathologists and peer-reviewers from two additional laboratories:

- Clarifications were given to the pathologists and peer-review pathologists of the participating laboratories on the ICE histopathology criteria as defined by the reference laboratory and explained above;
- Training was conducted on a first preliminary set of blinded slides covering a range of histopathology effects (n = 17 slides), in which scoring and observations from the pathologists and peer-review pathologists of the two participating laboratories were discussed with the reference laboratory to ensure ICE histological clarifications were well understood.
- Proficiency in scoring ICE histopathology effects was then appraised using a second set of blinded slides covering a range of histopathology effects (n = 23 slides) by the pathologists and peer-reviewers of the two participating laboratories. At the end of this proficiency appraisal, any remaining discrepancies as compared to the scoring from the reference laboratory were reviewed by the participating laboratories and discussed independently with the reference laboratory.
- Finally, the reproducibility of ICE histopathology was assessed by the pathologists and peer-reviewers from the three independent laboratories.

The reproducibility of ICE histopathology was assessed based on surfactants having existing *in vivo* Draize data sought thorough collaboration with Cosmetics Europe (CosEU[3]) and Environment and Health - Risk Assessment & Management (ERASM[4]). Based on this collaboration, and on surfactants evaluated by A.I.S.E., a total of 68 slides were selected covering 6 EU CLP/UN GHS Cat.1 & 6 EU CLP/UN GHS Non-Cat. 1 (representing three EU CLP/UN GHS Cat. 2 and three EU CLP/UN GHS No Cat.) surfactants having existing *in vivo* Draize data as well

[3] www.cosmeticseurope.eu. Accessed on 24.05.2021.
[4] www.erasm.org. Accessed on 24.05.2021.

Table 7
ICE histopathology-based predictions obtained from the pathologists and peer-reviewers of three independent laboratories.

Chemical name	CAS n.	Physical state	Type of surfactant	*In vivo* EU CLP/ UN GHS Cat.	Ref. lab	Lab. 1	Lab. 2
Benzensulphonylchloride	98-09-9	L	Anionic	**Draize Cat. 1 (severity & persistence)**	**Cat.1 (vacuolation & necrosis)**	**Cat.1 (vacuolation & necrosis)**	**Cat.1 (vacuolation)**
Cetyl pyridinium bromide 100%	140-72-7	S	Cationic	**Cat. 1 (by extrapolation)**	**Cat.1 (vacuolation)**	**Cat.1 (vacuolation & erosion)**	**Cat.1 (vacuolation & erosion)**
Cetyl pyridinium bromide 10%	140-72-7	L	Cationic	**Draize Cat. 1 (severity & persistence)**	**Cat.1 (vacuolation)**	**Cat.1 (vacuolation)**	**Cat.1 (vacuolation)**
Domiphen bromide 10%	538-71-6	L	Cationic	**Draize Cat. 1 (severity & persistence)**	**Cat.1 (vacuolation)**	**Cat.1 (vacuolation & erosion)**	**Cat.1 (vacuolation)**
Stearyltrimethylammonium chloride - 10% unwarmed	112-03-8	L	Cationic	**Draize Cat. 1 (persistence only)**	**Cat.1 (vacuolation)**	**Cat.1 (vacuolation)**	**Cat.1 (vacuolation)**
Benzalkonium chloride 5%	63449-41-2	L	Cationic	**Draize Cat. 1 (severity & persistence)**	**Cat.1 (erosion & vacuolation)**	**Cat.1 (erosion)**	**Cat.1 (erosion & vacuolation)**
Cetyl pyridinium bromide 1% warmed - study 1	140-72-7	L	Cationic	Draize Cat. 2A	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
Cetyl pyridinium bromide 1% unwarmed - study 2	140-72-7	L	Cationic	Draize Cat. 2A	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
N-Lauroyl sarcosine Na salt 10%	137-16-6	L	Anionic	Draize Cat. 2A	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
1-Ethyl-3-methylimidazolium ethyl sulphate	342573-75-5	L	Cationic	Draize No Category	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction	**Cat.1 (vacuolation)**
Cetyl pyridinium bromide −0.1%	140-72-7	L	Cationic	Draize No Category	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
Polysorbate 20 (Tween 20)	9005-64-5	L	Non-ionic	Draize No Category	No triggers for Cat. 1 prediction	**Cat.1 (vacuolation)**	No triggers for Cat. 1 prediction

as the corresponding positive and negative controls of each independent ICE histopathology study. Among the six EU CLP/UN GHS Non-Cat. 1 surfactants, two represented the same surfactant prepared in two different manners, i.e. warmed and unwarmed dilution, as such procedure was found in previous studies (data not shown) to have a potential impact on the final outcome.

Scoring of ICE histopathology sections was conducted separately over time by each participating laboratory during the meeting to ensure independence. Sections were scored by the pathologists and respective peer-review pathologists from each laboratory who evaluated at least 1/3 of the slides as recommended in the recently adopted revised OECD GD 160 (2018b). Finally, A.I.S.E. provided a continuous presence during the evaluation to ensure that no communication occurred between the pathologists and peer-reviewers of the participating laboratories and the reference laboratory.

3. Results and discussions

3.1. ICE histopathology reproducibility between pathologists and peer-reviewers from three independent laboratories

The detailed scoring given by each laboratory are shown in Table 6, and the EU CLP/UN GHS predicted Categories are shown in Table 7. Based on the obtained scores the following concordance in prediction between the pathologists and peer-reviewers from the three participating laboratories was found:

- 83% (10/12) of the assessed surfactants had concordant predictions between the pathologists and peer-reviewers of the three independent laboratories;
- There was concordance on the main histopathological effects triggering EU CLP/UN GHS Cat. 1 prediction among the pathologist and peer-reviewer from each of the three laboratories;
- All six Draize Cat. 1 surfactants were correctly predicted as positives. Furthermore, all six non-Cat.1 Draize surfactants were correctly predicted as negatives when based on the majority of outcomes from all labs.
- The same results were observed for the reference laboratory (TNO Triskelion) as an individual laboratory, whereas for the two other laboratories, all six Draize Cat. 1 surfactants were correctly predicted as positives whereas five of the six non-Cat.1 Draize surfactants were correctly predicted as negatives.

3.2. ICE histopathology reproducibility over time

Reproducibility over time for both non-extreme pH ($2 < pH < 11.5$) detergents and surfactants, was also investigated based on the existing dataset and using the same scoring system over time. Results are shown in Tables 8 and 9 respectively. An overall concordance of 17/18 and 6/6 was found respectively for non-extreme pH detergents and for surfactants, indicating an appropriate reproducibility over a time-frame of 2–4 years.

3.3. ICE histopathology predictive capacity

The results obtained with the new dataset on surfactants and with the existing A.I.S.E. dataset for detergents evaluated with the same histopathology criteria as used here, are shown in Table 10. The ensuing predictive capacity is shown in Table 11 for detergents and surfactants having Draize *in vivo* data (Cat. 1 and non-Cat. 1), and LVET *in vivo* Cat. 1 data. Use of histopathology in addition to the standard ICE test method according to the OECD TG 438, decreased the rate of false negatives from 64% (14/22) to 27% (6/22), increased the overall accuracy from 53% (16/30) to 77% (23/30) and maintained an acceptable level of false positives (from 0/8 to 1/8 (12.5%)). Furthermore, if an epithelial necrosis threshold of 1 instead of 2 was used for the identification of EU CLP/UN GHS Cat. 1 surfactants, the rate of false negatives would be

Table 8
Reproducibility ICE histopathology blinded slides over time for non-extreme pH (2 < pH < 11.5) detergents.

A.I.S.E. non-extreme pH detergents	Physical state	Approximate surfactant content (%)	*In vivo* EU CLP/ UN GHS [a]	ICE histopathology [b]	
				TNO Triskelion - 2013	TNO Triskelion - 2017
HDWL 1	L	35	**LVET – Cat. 1**	**Cat. 1 (vacuolation)**	**Cat. 1 (vacuolation)**
HDWL 2	L	40	**LVET – Cat. 1**	**Cat. 1 (erosion)**	**Cat. 1 (erosion)**
HDWL 3	L	45	**LVET – Cat. 1**	**Cat. 1 (vacuolation)**	**Cat. 1 (vacuolation)**
HDWL 4	L	35	**LVET – Cat. 1**	**Cat. 1 (vacuolation & erosion)**	**Cat. 1 (vacuolation & erosion)**
HDWL 5	L	30	**LVET – Cat. 1**	**Cat. 1 (vacuolation)**	**Cat. 1 (vacuolation)**
Laundry powder 1	S	25	**LVET – Cat. 1**	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
Laundry liquid 1	L	50	**LVET – Cat. 1**	**Cat. 1 (erosion)**	**Cat. 1 (erosion)**
Laundry liquid 2	L	55	**LVET – Cat. 1**	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
HDWL 13	L	20	Draize – Cat. 2A	**Cat. 1 (vacuolation)**	No triggers for Cat. 1 prediction
HDWL 14	L	20	Draize – Cat. 2A	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
HDWL 17	L	30	Draize – NC	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
APC 3	L	10	LVET Non-Cat. 1 [c]	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
APC 4	L	5	LVET Non-Cat. 1 [c]	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
HDWL 15	L	25	LVET Non-Cat. 1 [c]	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
HDWL 16	L	45	LVET Non-Cat. 1 [c]	**Cat. 1 (erosion)**	**Cat. 1 (erosion)**
Laundry liquid 4	L	45	LVET Non-Cat. 1 [c]	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
Laundry powder 4	S	20	LVET Non-Cat. 1 [c]	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction
Laundry powder 5	S	15	LVET Non-Cat. 1 [c]	No triggers for Cat. 1 prediction	No triggers for Cat. 1 prediction

APC: All purposes cleaner; Cat.: Category; HDWL: Hand dishwash liquid; L: liquid; LVET; Low Volume Eye Test NC: No Category; S: solid.
[a] LVET classification derived using the Draize criteria for classification.
[b] Based on histopathological criteria as described in the revised OECD GD 160 (2017).
[c] Representing a No Category if the Draize criteria for classification was used.

Table 9
Reproducibility of ICE histopathology blinded slides over time for surfactants.

Chemical name	CAS	*In vivo* EU CLP/ UN GHS Cat.	TNO Triskelion 2014–2015	Three labs June 2017
			ICE histopathology [a]	ICE histopathology [a]
Benzalkonium chloride (5%)	63449-41-2	**Draize – Cat. 1 (severity & persistence)**	**Cat. 1 (erosion) - Sept 2014**	**Cat. 1 (erosion)**
Cetylpyridinium bromide (10%)	140-72-7	**Draize – Cat. 1 (severity & persistence)**	**Cat.1 (vacuolation) - Sept 2014**	**Cat.1 (vacuolation)**
Stearyltrimethylammonium chloride - 10% unwarmed	112-03-8	**Draize – Cat. 1 (persistence)**	**Cat.1 (vacuolation) - Sept. 2015**	**Cat.1 (vacuolation)**
N-Lauroyl sarcosine Na salt (10%)	137-16-6	Draize – Cat. 2A	No triggers for Cat. 1 prediction (1)/ **Cat. 1** [c] **- July 2015**	No triggers for Cat. 1 prediction [a]/**Cat. 1 (1 out of 3 labs)** [c]
Cetylpyridinium bromide (1%) – study 1	140-72-7	Draize – Cat. 2A	No triggers for Cat. 1 prediction - July 2015	No triggers for Cat. 1 prediction
Cetylpyridinium bromide (1%) – study 2 (unwarmed)	140-72-7	Draize – Cat. 2A	No triggers for Cat. 1 prediction - Sept 2015	No triggers for Cat. 1 prediction
Polysorbate 20 (Tween 20) (100%)	9005-64-5	Draize No Cat. and LVET Non-Cat. 1 [b]	No triggers for Cat. 1 prediction - Sept. 2014	No triggers for Cat. 1 prediction (2 out of 3 labs) [a]/**Cat. 1 (1 out of 3 labs)** [c]

Cat.: Category; L: liquid; LVET; Low Volume Eye Test; NPCM: No Prediction Can be Made; S: solid.
[a] Based on histopathological criteria as described in the revised OECD GD 160 (2017) if not indicated otherwise.
[b] Representing a No Category if the Draize criteria for classification was used.
[c] Predicted as Cat. 1 if necrosis threshold of 1 is used instead of 2 (see section 3.3).

further decreased to 18% (4/22) but with an increase in the false positive rate to 25% (2/8), as shown in Table 11. As shown in Table 12, similar findings were obtained with the entire dataset that included test materials having LVET *in vivo* non-Cat. 1 data. Table 12 also shows that the predictive capacity of ICE histopathology for non-extreme pH detergent formulations is higher than the one used with the additivity approach.

It is interesting to note that when evaluating the histopathology results obtained with surfactants, epithelial necrosis was observed in 8 out of the 20 tested surfactants and in 6 of the 13 EU CLP/UN GHS Cat. 1 tested surfactants. Such findings are in contrast to the observations obtained with the 48 tested mixtures of the use domain of detergents, where epithelial necrosis was found only for 2 out of the 48 tested mixtures (Cazelle et al., 2014, 2015) and did not affect the predicted classification based on the ICE histopathology criteria developed by A.I.S.E. (Table 3). As a consequence, at the time when the histopathology criteria were originally developed for identification of EU CLP/UN GHS Cat. 1 non-extreme pH detergents, only a limited dataset was available on epithelial necrosis to make conclusive decisions. This meant that only preliminary decision criteria could be established for the epithelial necrosis effects, in contrast to epithelial erosion and vacuolation for which a considerably larger amount of data was available.

4. Conclusions

With the enlarged dataset including surfactants the results previously reported for non-extreme pH (2 < pH < 11.5) detergents in Cazelle et al. (2014) were confirmed: use of ICE histopathology in addition to

Table 10
Overview of combined new and existing dataset of ICE histopathology results obtained with the surfactants and detergents[e] having Draize *in vivo* data (Cat. 1 and non-Cat. 1), and LVET *in vivo* Cat. 1 data.

Chemical name	CAS	Physical state	Surfactant content	Type of surfactant	*In vivo* EU CLP/UN GHS Cat.	Tissue driver for Cat. 1 classification	ICE prediction [a]	ICE histopathology [b]
Benzensulphonylchloride	98-09-9	L	100%	Anionic	**Draize – Cat. 1**	Severity + persistence	Cat. 1	**Cat. 1**
Cetyl pyridinium bromide 100%	140-72-7	S	100%	Cationic	**Cat. 1**	By extrapolation	Cat. 1	**Cat. 1**
Domiphen bromide 10%	538-71-6	L	10%	Cationic	**Draize – Cat. 1**	Severity + persistence	Cat. 1	**Cat. 1**
Stearyltrimethylammonium chloride - 10% unwarmed	112-03-8	L	10%	Cationic	**Draize – Cat. 1**	Persistence (3/3 animals)	Cat. 1	**Cat. 1**
Benzethonium chloride (10%)	121-54-0	L	10%	Cationic	**Draize – Cat. 1**	Severity + persistence	Cat. 1	**Cat. 1**
Distearyldimethylammonium chloride (neat)	107-64-2	S	100%	Cationic	**Draize – Cat. 1**	Severity of effects	Cat. 1	**Cat. 1**
Benzalkonium chloride (5%)	63449-41-2	L	5%	Cationic	**Draize – Cat. 1**	Severity + persistence	Cat. 1	**Cat. 1**
Alkaline extreme pH # 5	n.a.	S	~20%	n.a.	**LVET – Cat. 1**	Severity of effects	Cat. 1	**Cat. 1**
HDWL 1	n.a.	L	~35%	n.a.	**LVET – Cat. 1**	Severity of effects	NPCM	**Cat. 1**
Laundry liquid 1	n.a.	L	~50%	n.a.	**LVET – Cat. 1**	Persistence (2/6 animals)	NPCM	**Cat. 1**
HDWL 3	n.a.	L	~45%	n.a.	**LVET – Cat. 1**	Persistence (2/3 animals)	NPCM	**Cat. 1**
HDWL 4	n.a.	L	~35%	n.a.	**LVET – Cat. 1**	Persistence (2/3 animals)	NPCM	**Cat. 1**
HDWL 2	n.a.	L	~40%	n.a.	**LVET – Cat. 1**	Persistence (1/3 animals)	NPCM	**Cat. 1**
HDWL 5	n.a.	L	~30%	n.a.	**LVET – Cat. 1**	Persistence (1/3 animals)	NPCM	**Cat. 1**
Cetylpyridinium bromide (10%)	140-72-7	L	10%	Cationic	**Draize – Cat. 1**	Severity + persistence	NPCM	**Cat.1**
Cetylpyridinium bromide (6%)	140-72-7	L	6%	Cationic	**Draize – Cat. 1**	Severity (iritis in 3/4 animals)	NPCM	**Cat.1**
Amines, C12-14 (even numbered)-alkyldimethyl, N-oxides (~28%)	308062-28-4 (1643-20-5)	L	~28%	Non-ionic	**Draize – Cat. 1**	Persistence (3/3 animals)	NPCM [b]/**Cat. 1** [d]	
Coco alkyl dimethyl betaine (~30%)	68424-94-2	L	~30%	Amphoteric	**Draize – Cat. 1**	Persistence (4/4 animals)	NPCM [b]/**Cat. 1** [d]	
Laundry powder 1	n.a.	S	~25%	n.a.	**LVET – Cat. 1**	Severity of effects	NPCM	NPCM
Laundry liquid 2	n.a.	L	~55%	n.a.	**LVET – Cat. 1**	Persistence (2/3 animals)	NPCM	NPCM
Benzalkonium chloride (1%)	63449-41-2	L	1%	Cationic	**Draize – Cat. 1**	Persistence (4/10 animals)	NPCM	NPCM
50% active N,N-Dimethyl-N-coco-N-(3-sulfopropyl) ammonium betaine	68201-55-8	L	50%	Amphoteric	**LVET – Cat. 1**	Persistence (2/6 animals)	NPCM	NPCM
N-Lauroyl sarcosine Na salt (10%)	137-16-6	L	10%	Anionic	Draize – Cat. 2A	n.a.	NPCM	NPCM [b]/**Cat.1** [d]
Cetylpyridinium bromide (1%)	140-72-7	L	1%	Cationic	Draize – Cat. 2A	n.a.	NPCM	NPCM
HDWL 13	n.a.	L	~20%	n.a.	Draize – Cat. 2A	n.a.	NPCM	**Cat.1**
HDWL 14	n.a.	L	~20%	n.a.	Draize – Cat. 2A	n.a.	No Cat.	NPCM
1-Ethyl-3-methylimidazolium ethyl sulphate	342573-75-5	L	100%	Cationic	Draize - No Cat.	n.a.	NPCM	NPCM
HDWL 17	n.a.	L	~30%	n.a.	Draize - No Cat.	n.a.	No Cat.	NPCM
Cetyl pyridinium bromide −0.1%	140-72-7	L	0.1%	Cationic	Draize - No Cat.	n.a.	No Cat.	NPCM
Polysorbate 20 (Tween 20) (100%)	9005-64-5	L	100%	Non-ionic	Draize No Cat. and LVET Non-Cat. 1 [c]	n.a.	No Cat.	NPCM

Cat.: Category; HDWL: Hand Dish Wash Liquid; L: liquid; LVET; Low Volume Eye Test; n.a.: not applicable; NPCM: No Prediction Can be Made; S: solid.

[a] Based on the Decision Criteria as described in OECD TG 438 (2018a)".

[b] Based on histopathological criteria as described in the revised OECD GD 160 (2018b) if not indicated otherwise.

[c] Representing a No Category if the Draize criteria for classification was used.

[d] Predicted as Cat. 1 if necrosis threshold of 1 is used (instead of 2).

[e] According to revised TG 438 (2018a), "Detergents: a mixture (excluding dilutions of single surfactant) containing one or more surfactants at a final concentration of >3%, intended for washing and cleaning processes. Detergents may be in any form (liquid, powder, paste, bar, cake, moulded piece, shape, etc.) and marketed for or used in household, or institutional or industrial purposes."

Table 11
Performances of the ICE with or without histopathology (histo.) for surfactants and detergents having Draize *in vivo* data (Cat. 1 and non-Cat. 1) and LVET *in vivo* Cat. 1 data as shown in Table 10.

		ICE	ICE + histo[c]	ICE + histo[d]
Non-extreme pH (2< pH < 11.5) detergents[a] and surfactants	False negative rate	63.6% (14/22)	**27.3% (6/22)**	**18.2% (4/22)**
	False positive rate	0.0% (0/8)	**12.5% (1/8)[b]**	**25.0% (2/8)**
	Accuracy	53.3% (16/30)	**76.7% (23/30)**	**80.0% (24/30)**

[a] Includes 8 non-extreme pH detergents from the training set distributed as 5 LVET Cat. 1 (5/5 and 0/5 false negatives with ICE and ICE + histo respectively) and 3 Draize No Cat. 1 non-extreme pH detergents (0/3 and 1/3 false positives with ICE and ICE + histo respectively).
[b] The false positive observed with ICE histopathology is a Draize Cat. 2A non-extreme pH detergent.
[c] Histopathological criteria as described in the revised OECD GD 160 (2018b).
[d] Considering epithelial necrosis scores of 1 (instead of 2) as threshold based on the original score system, i.e., only necrosis of attached cells/tissues being scored.

Table 12
Performances of the ICE with or without histopathology (histo.) for the overall available dataset of detergents and surfactants (including LVET non-Cat. 1 detergents and surfactants).

		Accuracy	Sensitivity	False negatives	Specificity	False positives
Surfactants	ICE	70.0% (14/20)	53.8% (7/13)	46.1% (6/13)	100% (7/7)	0% (0/7)
	ICE + ICE histo.[b]	80.0% (16/20)	69.2% (9/13)	30.8% (4/13)	100% (7/7)	0% (0/7)
	ICE + ICE histo.[c]	85.0% (17/20)	84.6% (11/13)	15.4% (2/13)	85.7% (6/7)	14.3% (1/7)
Non-extreme pH (2< pH < 11.5) detergents[a]	ICE	74.2% (23/31)	11.1% (1/9)	88.9% (8/9)	100% (22/22)	0% (0/22)
	ICE + ICE histo.[d]	74.2% (23/31)	77.8% (7/9)	22.2% (2/9)	72.7% (16/22)	27.3% (6/22)
	Additivity approach	32.3% (10/31)	100% (9/9)	0% (0/9)	4.5% (1/22)	95.5% (21/22)
Combined dataset	**ICE**	**72.5% (37/51)**	**36.4% (8/22)**	**63.6% (14/22)**	**100% (29/29)**	**0% (0/29)**
	ICE + ICE histo.[b]	**76.5% (39/51)**	**72.7% (16/22)**	**27.3% (6/22)**	**79.3% (23/29)**	**20.7% (6/29)**
	ICE + ICE histo.[c]	**78.4% (40/51)**	**81.8% (18/22)**	**18.2% (4/22)**	**75.9% (22/29)**	**24.1% (7/29)**

[a] Including the training set of 11 non-extreme pH detergents distributed as 5 Cat. 1 (0/5 and 5/5 sensitivity with ICE and ICE + histo respectively) and 6 No Cat. 1 non-extreme pH detergents (6/6 and 5/6 specificity with ICE and ICE + histo respectively).
[b] Histopathological criteria as described in the revised OECD GD 160 (2018b).
[c] Considering epithelial necrosis scores of 1 (instead of 2) as threshold.
[d] No difference in outcome observed for formulations if using the necrosis score of 1 as a threshold, as very few necrosis observed with formulations.

the standard ICE test method enabled a reduction in false negative rates observed with the standard ICE test method and increase the sensitivity for the identification of EU CLP/UN GHS Cat. 1 non-extreme pH (2 < pH < 11.5) detergents and surfactants.

Good reproducibility was found between pathologists and their peer-reviewers from three independent laboratories (10/12 or 83%) and over time (17/18 for non-extreme pH detergents and 6/6 for surfactants) for the ICE histopathology derived predictions.

These findings confirm that ICE histopathology improves the predictive capacity when compared to the standard ICE alone and can be used effectively to identify non-extreme pH (2 < pH < 11.5) detergents requiring EU CLP/UN GHS Cat. 1 classification. The present work was used as a basis for revising the OECD TG 438, GD 160 and GD 188, to include the possibility of using ICE histopathology for the identification of UN GHS Cat. 1 non-extreme pH (2 < pH < 11.5) detergents and surfactants (OECD, 2018a). In the revised adopted version of OECD TG 438 (2018a), the ICE histopathology criteria developed by A.I.S.E. is applied as shown in Table 3, using a necrosis score of 2.

When using the ICE *in vitro* ocular toxicity test method as described within the OECD TG 438 (2018a), users are encouraged to preserve eyes and prepare histopathology specimens that can be used to develop an internal database that may improve the performance of this test method and/or help to further investigate the usefulness of ICE histopathology for chemistries other than non-extreme pH detergents and surfactants.

In order to ensure ICE histopathology reproducibility, the following recommendations are made to test facilities using willing to implement and use it on a regular basis:

- Have an internal histopathology peer-review system in place as recommended by e.g. Morton et al. (2010) and by the OECD Advisory document n. 16 on GLP requirements for peer review of histopathology (OECD, 2014);
- Assess the original slides in order to enable the evaluation of three dimensional effects;
- Have appropriate training & proficiency appraisal.

To help in this process, end users are invited to consult the OECD Guidance Document 160 when using the ICE *in vitro* ocular toxicity test method, which includes detailed procedures on the collection and processing of histopathology specimens for evaluation and collection of data (OECD, 2018b). In particular, Annex II of this GD provides an Atlas with typical photomicrographs of untreated as well as treated Isolated Chicken Eyes illustrating the various possible histopathological effects as described in Table 2 for the semi-quantitative histopathology scoring system used for ICE that were fixed, trimmed, embedded in paraffin wax, sectioned and stained.

Funding body information:

This study was funded by A.I.S.E. - The International Association for Soaps, Detergents and Maintenance Products.

In turn A.I.S.E. receives its operation budget from the member companies and national associations that are active in the domain of detergents and maintenance products.

For more information about A.I.S.E., please see: https://www.aise.eu/

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to warmly acknowledge Pauline McNamee (The Procter & Gamble Company, Egham, United Kingdom), Wendy Cameron (A.I.S.E., Brussels, Belgium), Penny Jones (Unilever, Colworth Science Park, Sharnbrook, Bedford, UK) and Olaf Holtkoetter (Henkel AG & Co. KGaA, Duesseldorf, Germany) for their contributions to the present work. Furthermore, the authors would like to acknowledge ERASM and Cosmetics Europe for the fruitful collaboration and sharing of information that made the present work possible.

References

Cazelle, E., Eskes, C., Hermann, M., Jones, P., McNamee, P., Prinsen, M., Taylor, H., Wijnands, M.V.W., 2014. Suitability of histopathology as an additional endpoint to the isolated chicken eye test for classification of non-extreme pH detergent and cleaning products. Toxicol. Vitro 28, 657–666.

Cazelle, E., Eskes, C., Hermann, M., Jones, P., McNamee, P., Prinsen, M., Taylor, H., Wijnands, M.V.W., 2015. Suitability of the isolated chicken eye test for classification of extreme pH detergents and cleaning products. Toxicol. Vitro 29, 609–616.

EC, 2008. Regulation (EC) No 1272/2008 of the European parliament and of the council of 16 december 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing directives 67/548/EEC and 1999/45/EC, and amending regulation (EC) No 1907/2006. Official Journal of the European Union L353, 1–1355.

Morton, D., Sellers, R.S., Barale-Thomas, E., Bolon, B., George, C., Hardisty, J.F., Irizarry, A., McKay, J.S., Odin, M., Teranishi, M., 2010. Recommendations for pathology peer review. Toxicol. Pathol. 38, 1118–1127.

OECD, 2014. Advisory Document N. 16. Guidance on the Good Laboratory Practice (GLP) Requirements for Peer Review of Histopathology. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring. Advisory Document of the Working Group on GLP. ENV/JM/MONO(2014)30. Organisation for Economic Cooperation and Development, Paris.

OECD, 2018a. Guideline for Testing of Chemicals No. 438. Isolated Chicken Eye Test Method for Identifying I) Chemicals Inducing Serious Eye Damage and Ii) Chemicals Not Requiring Classification. Organisation for Economic Cooperation and Development, Paris. Available at: https://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788. (Accessed 24 May 2021).

OECD, 2018b. Revised guidance document on the bovine corneal opacity and permeability (BCOP) and isolated chicken eye (ICE) test methods: collection of tissues for histological evaluation and collection of data on non-severe irritants. Series on testing and assessment no. 160. Available at: http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm. (Accessed 24 May 2021).

OECD, 2018c. Series on testing and assessment No 188. Revised streamlined summary document supporting OECD test guideline 438 on the isolated chicken eye for eye irritation/corrosion. Available at: http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm. (Accessed 24 May 2021).

Prophet, E.B., Mills, B., Arrington, J.B., Sobin, L.H., 1992. Laboratory Methods in Histotechnology. Armed Forces Institute of Pathology, Washington, USA, p. 279.

Prinsen, M.K., Schipper, M.E.I., Wijnands, M.V.W., 2011. Histopathology in the isolated chicken eye test and comparison of different stainings of the cornea. Toxicol. Vitro 25, 1475–1479.