Toxicology in Vitro xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Toxicology in Vitro

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

Round robin study to evaluate the reconstructed human epidermis (RhE) model as an in vitro skin irritation test for detection of irritant activity in medical device extracts

Wim H. De Jong[a,*], Sebastian Hoffmann[b], Michelle Lee[c], Helena Kandárová[d], Christian Pellevoisin[e], Yuji Haishima[f], Beau Rollins[g], Austin Zdawczyk[h], Jamin Willoughby[i], Michael Bachelor[j], Timothy Schatz[k], Shelby Skoog[l], Sherry Parker[m], Anita Sawyer[n], Paolo Pescio[o], Kristina Fant[p], Kwang-Mahn Kim[q], Jae Sung Kwon[q], Helge Gehrke[r], Hana Hofman-Hüther[r], Marisa Meloni[s], Conrad Julius[t], Damien Briotet[u], Silvia Letasiova[d], Reiko Kato[f], Atsuko Miyajima[f], Liset J.J. De La Fonteyne[a], Christelle Videau[e], Carine Tornier[e], Audrey P. Turley[c], Nicholas Christiano[v], Thor S. Rollins[c], Kelly P. Coleman[w]

[a] National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[b] Seh consulting + services, Paderborn, Germany
[c] Nelson Laboratories, Inc., Salt Lake City, UT, USA
[d] MatTek In Vitro Life Science Laboratories, Bratislava, Slovakia
[e] EPISKIN, Lyon, France
[f] National Institute of Health Sciences (NIHS), Division of Medical Devices, Tokyo, Japan
[g] Arthrex, Inc., Naples, FL, USA[1]
[h] NAMSA, Northwood, OH, USA
[i] Cyprotex US LCC, Kalamazoo, MI, USA[2]
[j] MatTek Corporation, Ashland, MA, USA
[k] American Preclinical Services LLC, Minneapolis, MN, USA
[l] US Food and Drug Administration, Center for Devices and Radiological Health, Silver Spring, MD, USA
[m] WuXi AppTec, St Paul, MN, USA
[n] Becton Dickinson, Research Triangle Park, NC, USA
[o] Eurofins Biolab Srl, Vimodrone, Milan, Italy
[p] SP Technical Research Institute of Sweden, Borås, Sweden
[q] Yonsei University, College of Dentistry, Seoul, South Korea
[r] Eurofins Biopharma, Planegg, Munich, Germany
[s] VitroScreen, Milan, Italy
[t] Envigo CRS GmbH, Rossdorf, Germany
[u] NAMSA, Chasse sur Rhône, France
[v] Arthrex, Inc., Naples, FL, USA
[w] Medtronic, plc, Minneapolis, MN, USA

ARTICLE INFO

Keywords:
Medical devices
Irritation
Alternative testing
In vitro
Reconstructed human epidermis

ABSTRACT

Assessment of skin irritation is an essential component of the safety evaluation of medical devices. OECD Test Guideline 439 describes the use of reconstructed human epidermis (RhE) as an in vitro test system for classification of skin irritation by neat chemicals. An international round robin study was conducted to evaluate the RhE method for determination of skin irritant potential of medical device extracts. Four irritant polymers and three non-irritant controls were obtained or developed that had demonstrated their suitability to act as positive or negative test samples. The RhE tissues (EpiDerm™ and SkinEthic™ RHE) were dosed with 100 µL aliquots of either saline or sesame oil extract. Incubation times were 18 h (EpiDerm™) and 24 h (SkinEthic™ RHE). Cell

Abbreviations: DPBS, Dulbecco's phosphate buffered saline; ET-50, Exposure time that induces 50% cell viability; I, Irritant; MTT, 3-[4, 5-dimethyl-thiazol-2-yl]-2,5-diphenyltetrazolium bromide; MDS, methods documentation sheet; NC, negative control; VC, vehicle control; NI, non-irritant; OD, optical density; PC, positive control; RhE, reconstructed human epidermis; SDS, Sodium dodecyl sulfate; PVC, Polyvinyl chloride

* Corresponding author.
E-mail address: Wim.de.jong@rivm.nl (W.H. De Jong).

[1] Currently Quality Tech Services, Inc., Bloomington, MN, USA.

[2] Currently IonTox, LCC, Kalamazoo, MI, USA.

https://doi.org/10.1016/j.tiv.2018.01.001
Received 23 October 2017; Received in revised form 29 December 2017; Accepted 3 January 2018

viability reduction > 50% was indicative of skin irritation. Both the EpiDerm™ and SkinEthic™ RHE tissues were able to correctly identify virtually all of the irritant polymer samples either in the saline, sesame oil or both solvent extracts. Our results indicate that RhE tissue models can detect the presence of strong skin irritants at low levels in dilute medical device polymer extracts. Therefore, these models may be suitable replacements for the rabbit skin irritation test to support the biological evaluation of medical devices.

1. Introduction

The various toxicological endpoints for the safety evaluation of medical devices are described in the EN/ISO 10993 series of standards. The tests to consider for the safety evaluation of medical devices are described in ISO 10993-1:2009 *Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process*. The endpoints considered depend upon the type of device, the contact of the device with the patient, and duration of the contact. Nevertheless, cytotoxicity, irritation and sensitization testing are recommended for almost all devices. For determining skin irritation potential of medical devices ISO 10993-10:2010 *Biological evaluation of medical devices – Part 10: Tests for irritation and skin sensitization*, provides information on the testing of medical devices and medical device extracts. For skin irritation testing the rabbit skin irritation assay developed by Draize et al. (1944) is used. For surface contacting devices, ISO 10993-10 describes using the rabbit irritation test with gauze patches and bandages (semi-occlusive or occlusive) to expose the skin for a minimum of 4 h. In addition, for implanted devices and devices with blood contact, an intracutaneous (intradermal) rabbit test is described using medical device extracts for which local reactions are evaluated at the injection site. Although the use of non-animal methods is mentioned, no guidance is provided due to a lack of medical device experience with these in vitro methods, which were developed and validated for testing pure chemical substances. In other areas of safety testing, especially for industrial chemicals and cosmetic ingredients, the rabbit test has been replaced by test methods using reconstructed human epidermis (RhE) models. After extensive validation activities (Spielmann et al., 2007), and a thorough review process, the OECD accepted these test methods as a full replacement by issuing Test Guideline (TG) 439 (OECD, 2015).

OECD TG 439 describes skin irritation testing using the RhE model. These models produced with human cells recapitulate the morphology and features of human epidermis. Fig. 1 shows two RhE models as they were used in the round robin study described in this paper. Clearly all the different layers of normal skin from the basal cell layer (stratum basale) onto the keratinizing outer skin (stratum corneum) layer can be identified. Survival of the epidermal tissues is used as a read-out system for irritant activity. Irritant chemicals are identified by their ability to decrease cell viability below a defined cut-off value (i.e. ≤50%) for United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS) Category 2, H315) (OECD, 2015).

Casas et al. (2013) demonstrated as a proof of principle that an RhE tissue model was able to identify the presence of irritant chemicals in complex dilute medical device extract mixtures. Well-known irritants were added to medical device extracts and tested in the EpiDerm™ model. The 50% cell viability threshold was used to determine the presence of an irritant compound in the test sample. It was concluded that the EpiDerm™ model using human epidermal tissues may be a suitable in vitro replacement for the rabbit skin irritation test for the assessment of the irritation potential of medical device extracts

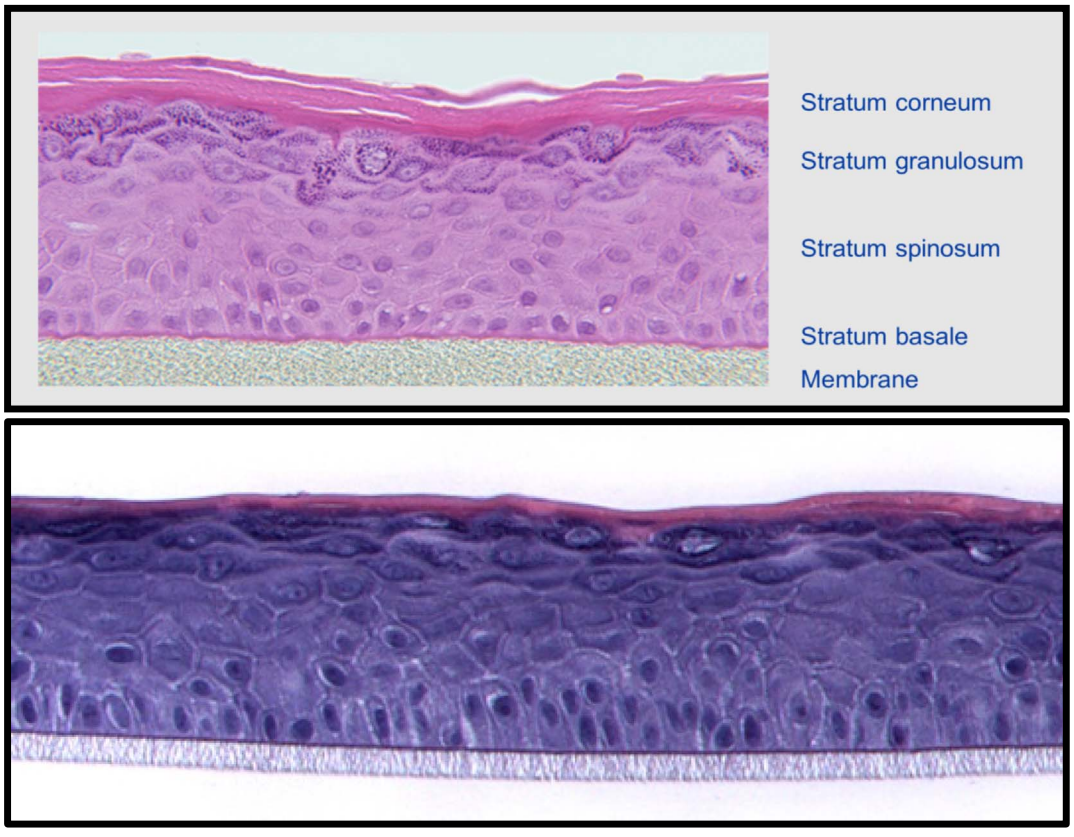

Fig. 1. Reconstructed human epidermis used in round robin study. Top: MatTek EpiDerm™, Bottom: EPISKIN SkinEthic™ RHE.

containing an irritant.

Based on the results of Casas et al. (2013), studies were initiated to evaluate the detection of irritants present in medical devices. Polymer samples were prepared to which known irritants were added during the polymer manufacturing process. Preliminary studies showed that irritant activity could be detected by using the in vitro RhE model in a modified protocol with an extension of the incubation time up to 18–24 h (Olsen et al., 2015; Pellevoisin et al., 2016, and Kandárová et al., 2015, 2016). The extension of the incubation time was considered necessary to optimize the detection of irritants in dilute mixtures. Chemical concentrations in medical device extracts may vary and maybe present in low concentrations as was demonstrated for ethylene oxide residues after sterilization (Lucas et al., 2003), so that they may be difficult to detect or to identify (Armstrong et al., 2013; Petrusevski et al., 2016). Based on these preliminary studies an international round robin study was initiated to assess the transferability and the laboratory reproducibility of RhE assays for measuring the irritant potential of medical device extracts. Two OECD TG 439 listed RhE models, EpiDerm™ (Kandárová et al., 2005) and SkinEthic™ RHE (Tornier et al., 2010), were used in the round robin study, the goal of which was to determine if RhE tissue models were suitable replacements for the rabbit skin irritation test for evaluating the irritant activity of medical device extracts.

2. Materials and methods

2.1. In vitro RhE models

Two RhE models were evaluated: EpiDerm™ (EPI-200) which was provided by MatTek In Vitro Life Science Laboratories (IVLSL, Bratislava, Slovakia) and MatTek Corporation (Ashland, MA, USA), has a surface area of 0.63 cm^2, and SkinEthic™ RHE, which was provided by EPISKIN (Lyon, France), has a surface area of 0.5 cm^2.

Both models have been validated for determining skin irritation of chemicals (Spielmann et al., 2007) and are included in OECD TG 439 and EU Guideline B.46. These models were validated with neat industrial chemicals for the purpose of classification and labeling of chemicals (e.g. H315, causes skin irritation).

For the EpiDerm™ model sixteen laboratories participated, while for the SkinEthic™ RHE model eight laboratories participated (Table 1). Six laboratories tested both RhE models. Participating laboratories had to demonstrate that they had successfully been trained in one or both test methods.

2.2. Polymer biomaterials (test materials) and sample preparation

Various polymer biomaterials were prepared as representative materials for medical devices. During manufacturing known irritants were added before the polymerization phase as reported by Coleman et al. (TIV this issue). In addition, the irritant chemicals (Genapol, Heptanoic acid (HA), and Sodium dodecyl sulfate (SDS)) were evaluated for their irritant activity in vivo in rabbits according to ISO 10993-10:2010 and by human patch test in volunteers (Kandárová et al. (TIV this issue). All chemicals used showed positive irritant activity in either the intracutaneous rabbit skin test or the 18 h human patch test. The resulting polymers containing the irritants are presented in Table 2. Seven test materials, four polymers containing an irritant, and three non-irritant containing polymers, were obtained or developed and in preliminary studies evaluated for their suitability prior to use.

In order to minimize bias in testing, the samples were individually coded for each participating laboratory and for each of three independent experiments. Blinded polymer samples were extracted with physiological saline (0.9% NaCl) and sesame oil (SO) per ISO 10993-12:2012 *Biological evaluation of medical devices – Part 12: Sample preparation and reference materials.* Each extract sample was tested in triplicate. Extraction and testing was performed three times in independent experiments (i.e. on different days). For each independent experiment two material samples were available (one for the saline extraction and one for the SO extraction).

2.3. Test sample preparation

Test samples were prepared according to ISO 10993-12:2012. Briefly, polymer samples with a total surface of 3 cm^2 (approximately 1 cm × 1.5 cm × 2 mm) or 0.2 g were incubated with either 1 mL of 0.9% NaCl or 1 mL of pharmaceutical grade SO (e.g. Sigma-Aldrich 85,067, CAS No. 8008-74-0, St, Louis, MO, USA) for 72 h (± 2 h) with continuous agitation/shaking at 37 °C (± 1 °C). After the extraction period extracts were collected and used within 24 h, usually the same day.

2.4. In vitro irritation test

For both the EpiDerm™ and the SkinEthic™ RHE model protocols were prepared (Supplementary Material 1, Supplementary Material 2, respectively).

After receiving the tissues, a quality control visual inspection of the tissues was performed. Excess agar was removed from the tissues. An overnight (18–24 h) preincubation step was included for EpiDerm™

Table 1
Participating laboratories.

Laboratory	RhE model tested
American Preclinical Services LLC, Minneapolis, MN, USA (APS)	EpiDerm™
Arthrex Inc., Naples, FL, USA	EpiDerm™, SkinEthic™ RHE
Becton Dickinson, Research Triangle Park, NC, USA (BD)	EpiDerm™
CDRH-FDA, Center for Devices and Radiological Health – Food and Drug Adminstration, Silver Spring, MD, USA (USFDA)	EpiDerm™
Cyprotex US LCC, Kalamazoo, MI, USA	EpiDerm™
Division of Medical Devices National Institutes of Health Services, Tokyo, Japan (NIHS)	EpiDerm™
Envigo CRS GmbH, Rossdorf, Germany	EpiDerm™
EPISKIN, Lyon, France	SkinEthic™ RHE
Eurofins BioPharma Product Testing GmbH, Planegg, Munich, Germany	EpiDerm™, SkinEthic™ RHE
Eurofins Biolab Srl, Vimodrone, Milan, Italy	EpiDerm™, SkinEthic™ RHE
MatTek In Vitro Life Science Laboratories, Bratislava, Slovakia	EpiDerm™
NAMSA, Northwood, OH, USA	EpiDerm™
Nelson Laboratories, Inc., Salt Lake City, UT, USA	EpiDerm™, SkinEthic™ RHE
RIVM, National Institute for Public Health and the Environment, Bilthoven, The Netherlands	EpiDerm™, SkinEthic™ RHE
SP Technical Research Institute of Sweden, Chemistry, Materials and Surfaces, Borås, Sweden (SP-TRI)	EpiDerm™
VitroScreen, Milan, Italy	SkinEthic™ RHE
WuXi Apptec Inc., St. Paul, MN, USA	EpiDerm™
Yonsei University College of Dentistry, Department & Research Institute for Dental Biomaterials & Bioengineering, Seoul, South Korea	EpiDerm™, SkinEthic™ RHE

Table 2
Polymer samples specifically prepared for the Round Robin study.

Polymer	Chemical added	Irritant activity of sample	Identification	Supplier
Polyurethane E80A	–	Negative sample	Polyurethane E80A	Medtronic[a]
One-part silicone	–	Negative sample	100% silicone	Medtronic
Polyvinyl chloride	–	Negative sample	Y-1	NIHS[b]
Polyvinyl chloride	Genapol X-80 (5.8%)	Positive sample	Y-4[c]	NIHS
Polyvinyl chloride	Genapol X-100 (4%)	Positive sample	PVC + 4% Genapol	Arthrex[d]
One part silicone	Heptanoic acid (25%)	Positive sample	Silicone + 25% Heptanoic acid	Medtronic
Two part silicone	Sodium dodecyl sulfate (SDS) (15%)	Positive sample	Silicone + 15% SDS	Arthrex

[a] Medtronic plc, Minneapolis, MN, USA.
[b] National Institutes of Health Services, NIHS, Division of Medical Devices, Tokyo, Japan.
[c] Y-4 pellet contains 55 parts of Di (2-ethylhexyl) phthalate (DEHP), 8 parts of Epoxidized soybean oil (ESBO), and 10 parts of Genapol X-80 against 100 parts of PVC by weight. Genapol X-80 is present at a percentage of 5.8% (Nomura et al., 2017).
[d] Arthrex, Inc. Naples, FL, USA.

tissues, and a minimum of 2–24 h preincubation step was included for the SkinEthic™ RHE tissues. After the preincubation the tissues ($n = 3$) were dosed with 100 µL extract aliquots. Positive chemical controls (PC, 1% *w*/w sodium dodecyl sulfate (SDS) in 0.9% NaCl and SO), negative controls (NC, Dulbecco's phosphate buffered saline (DPBS) for the EpiDerm™ method, and phosphate buffered saline (PBS) for the SkinEthic™ RHE method), and solvent vehicle controls (VC, 0.9% NaCl and SO) were included (n = 3). Incubation times were 18 h ± 30 min (EpiDerm™) and 24 h ± 1 h (SkinEthic™ RHE) at 37 °C, 5% CO_2, 95% air humidified atmosphere, according to the protocol of the supplier. After the incubation the tissues were rinsed with PBS and cell viability was determined by the MTT method.

The following acceptance criteria were set.

NC and VC acceptance criteria: the optical density (OD) of the NC and VC reflects the viability of the tissues used in the test conditions. For the EpiDerm™ tissue model the NC and VC were considered to meet the acceptance criteria, if the mean OD value of the three tissues was ≥0.8 and ≤2.8. For the SkinEthic™ RHE tissue model the NC and VC data met the acceptance criteria if the mean OD value of the three tissues was ≥0.8 and ≤3.0. For all assays the data were acceptable if the standard deviation (SD) value of the percent viability was ≤20%.

PC acceptance criteria: OD of the PC (1% SDS-treated) reflects the sensitivity of the tissues used in the test conditions. The PC was considered to meet the acceptance criteria if the mean viability expressed as percent of the NC, was < 50% and the SD value was ≤20%.

Extraction test substance data acceptance criteria: In each test the mean viability was calculated from the triplicate incubations, and the SD of each mean result had to be ≤20%. The overall mean viability was calculated from the three independent experiments on different days for each test sample. For a given extract, if only one batch (among the three batches used) gave an SD > 20%, then the extract was retested once (to address a possible technical problem or error). If two or three batches gave SDs > 20% the assay was not repeated, concluding that the variability was linked to the extract itself.

2.5. Viability determination

Viability of the tissues was based on cellular reduction of MTT (3-[4, 5-dimethyl-thiazol-2-yl]-2,5-diphenyltetrazolium bromide), a yellow tetrazole, and subsequent conversion to a purple formazan salt that is quantitatively measured after extraction from tissues (Faller et al., 2002; Mosmann, 1983). The insoluble formed formazan crystals are solubilized by adding isopropanol and the intensity was measured at 570 nm in a spectrophotometer.

After the incubation period the tissues were rinsed thoroughly (EpiDerm™ 15 times, SkinEthic™ RHE 25 times) with DPBS to remove any residual test material. The rinsed tissues were placed in 24-well plates prefilled with MTT (0.3 mL of 1 mg/mL), and incubated for 3 h (± 5 min) for the EpiDerm™ method, and 3 h (up to 3 h and 15 min) for the SkinEthic method, both at 37 °C, 5% CO_2, 95% humidified atmosphere.

After 3 h the residual MTT media was removed and tissues were transferred to a 24-well plate prefilled with isopropanol for extraction of the formazan. EpiDerm™ tissues were incubated for 2 h (± 5 min) at room temperature in prefilled (2 mL isopropanol) wells with gentle shaking. As an alternative, overnight extraction (18–24 h) was also possible by incubating at room temperature with gentle shaking or in the refrigerator in the dark, without shaking.

SkinEthic™ RHE tissues were transferred into wells and incubated in 1.5 mL isopropanol for 2 h (± 5 min) at room temperature with gentle shaking was done. For SkinEthic™ RHE tissues it was also possible to perform the formazan extraction overnight for about 16–18 h at room temperature with gentle agitation shaking.

After the extraction the OD was determined at 570 nm in a spectrophotometer. To determine viability 200 µL (2 × 200 µL for EpiDerm™ per tissue, 3 × 200 µL for SkinEthic™ RHE per tissue) was measured for determination of the viability. Relative tissue viability was determined for each tissue by using the following formula:

$$\%\text{viability} = [\text{OD treated tissue}/\text{OD negative control tissue}] \times 100$$

2.6. Data processing

All data obtained by the various laboratories were centrally collected and analyzed. Results of the various participating laboratories are presented anonymously and randomly. Variability of measurements (i.e. the percent cell viability per sample and experiment), was investigated on two levels. Reproducibility of results within each laboratory was evaluated for each test method by concordance of classifications, where a score of 100% concordance was given if all seven samples were classified consistently (either non-irritant (NI) or irritant (I)) in all three experiments. Concordance of classifications across laboratories was assessed using the combined outcomes of the three experiments. The results were compared with the expected outcomes for the various samples as presented in Table 3. Expectation of the outcomes for the various samples investigated was based on the preliminary studies as reported in Kandárová et al. (TIV this issue).

3. Results

The original results of the participating laboratories from which the Tables and Figures are derived are presented in the supplementary material (Supplementary Material 3). During the preliminary preparation and testing period five laboratories ultimately were lost and did not participate in the final round robin study. Reasons for not participating were: change of staff, building activities, time limitations, insufficient training and restrictions regarding tissue deliveries.

Table 3
Polymer samples used and their predicted[a] outcome.

Polymer	Predicted outcome for classification		
	Saline extract	Sesame oil extract	Overall
Polyurethane E80A	NI[b]	NI	NI
Polyvinyl chloride (Y-1)	NI	NI	NI
Polyvinyl chloride 5.8% Genapol X-80 (Y-4)	I[b]	I	I
Polyvinyl chloride 4% Genapol X-100	I	I	I
100% One part silicone	NI	NI	NI
One part silicone 25% heptanoic acid	NI	I	I
Two part silicone + 15% SDS	I	NI	I

[a] The predicted outcome was based on preliminary testing with the prepared polymer samples containing the experimentally added irritant.
[b] NI, non-irritant, I irritant

3.1. EpiDerm™ test method

3.1.1. Positive and vehicle control outcomes for EpiDerm™ tissues

In all sixteen laboratories the results of the positive and vehicle controls (0.9% NaCl and SO controls with or without 1% SDS added) were in accordance with the expected outcome (Fig. 2). Positive and vehicle controls were correctly identified by all laboratories using the EpiDerm™ tissues. Vehicle controls of both saline and SO were always clearly non-irritating with a cell viability in the individual experiments ranging from 80 to 145%, and the vast majority of the vehicle controls falling in the range of 100 ± 10% (Fig. 2). The results for the vehicle control SO showed slightly more variation compared to the saline controls. For the positive controls survival of the tissues was approximately 3% with a maximum viability of 8%, when compared to the negative control which was set at 100% and used for comparison.

3.1.2. Qualification of experiments and test materials within experiments for EpiDerm™ tissues

For the laboratories using EpiDerm™ tissues four non-qualified experiments were observed. These were repeated with a new set of coded test materials. Excessive test material variability (i.e. SD > 20%) was observed in five cases (data not shown), two of which were not repeated and excluded from analysis. One was considered acceptable as the SD was 20.03% which was only slightly above 20%.

Several protocol deviations were reported by eight of the sixteen participating laboratories, such as data reporting with blank already subtracted, reporting positive control as negative control and vice versa, tissues held up for one day in transit, missing one test material because of leakage, defects and irregular coloring of very few individual tissues, attachment of test material to bottom of container for shipment, loss of data due to error in Information and Communication Technology (ICT). These issues were considered to not have impacted the respective experiments. Some experiments were repeated with an additional set of test materials specifically prepared in reserve for unexplained issues (e.g. like the data loss due to ICT error). In summary, data were complete, because three valid independent experiments were available for each test material with both extracts, in thirteen laboratories. The remaining laboratories provided only two valid experiments for one or two test materials (see legend of Fig. 3). These data were excluded from further evaluation according to the respective performance standards recommended by the OECD (2015). All results of the individual experiments with the EpiDerm™ test method are presented in the Supplementary Table S1.

3.1.3. Within laboratory reproducibility of the test method using EpiDerm™ tissues

The combined results of the three independent experiments are presented in Fig. 3. In general, test materials extracted with SO showed slightly higher replicate variability (Fig. 3).

When saline was used as the vehicle for extraction the Within-Laboratory Reproducibility (WLR) was very high (Table 4). The only case of discordant repeat experiments was observed with the test material PVC + 4% Genapol X-100. However, the respective cell viability measurements were well reproducible (Experiment 1: 48,8%; Experiment 2: 41,3%; Experiment 3: 51,2%).

WLR was also high for the extract SO. Across all laboratories, classifications were not reproducible in 16 cases, half of these were with the test material 'Silicone + 25% Heptanoic Acid'. In the other cases, cell viabilities were close to the classification threshold of 50%.

This high mean WLR of 99.1% (across the sixteen included laboratories) for the saline extracts and 85.3% for the SO extracts, respectively (Table 4) was confirmed using the SD of the cell viabilities of experiments per laboratory and test material as a quantitative measure. In analogy to variability threshold for replicates, a SD > 20% was defined as a threshold. Saline extracted test materials exceeded this threshold in one case (110/111, 99.1%), whereas SO extracted samples

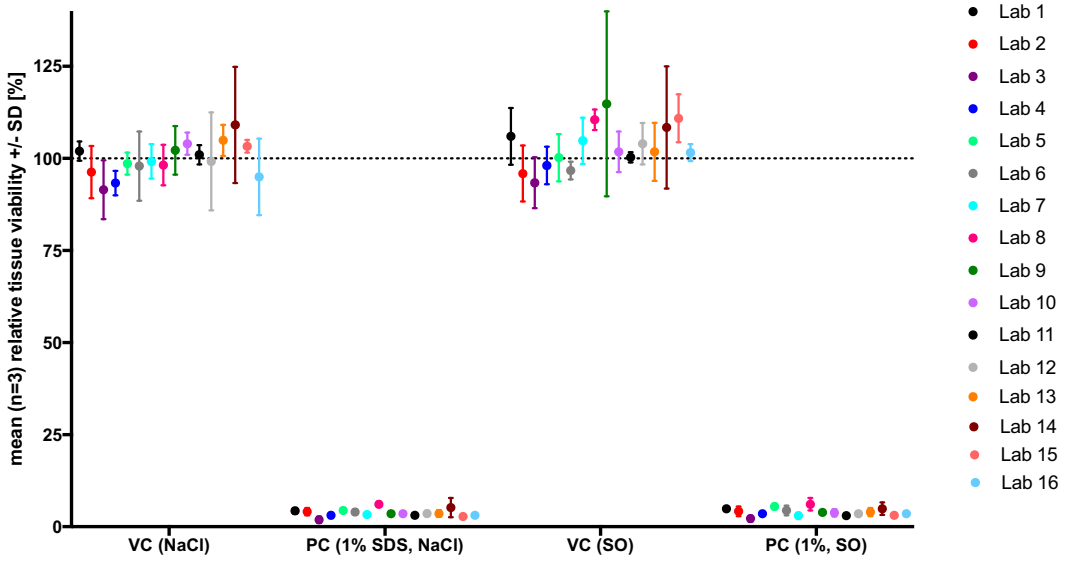

Fig. 2. EpiDerm™ test method. Cell viability results for vehicle (VC) and positive controls (PC) for the participating 16 laboratories (Lab 1-16). Mean cell viability and standard deviation (SD) of three individual experiments per laboratory are displayed. The dotted line at 100% represents the respective negative (DPBS) control, which is used for normalisation.

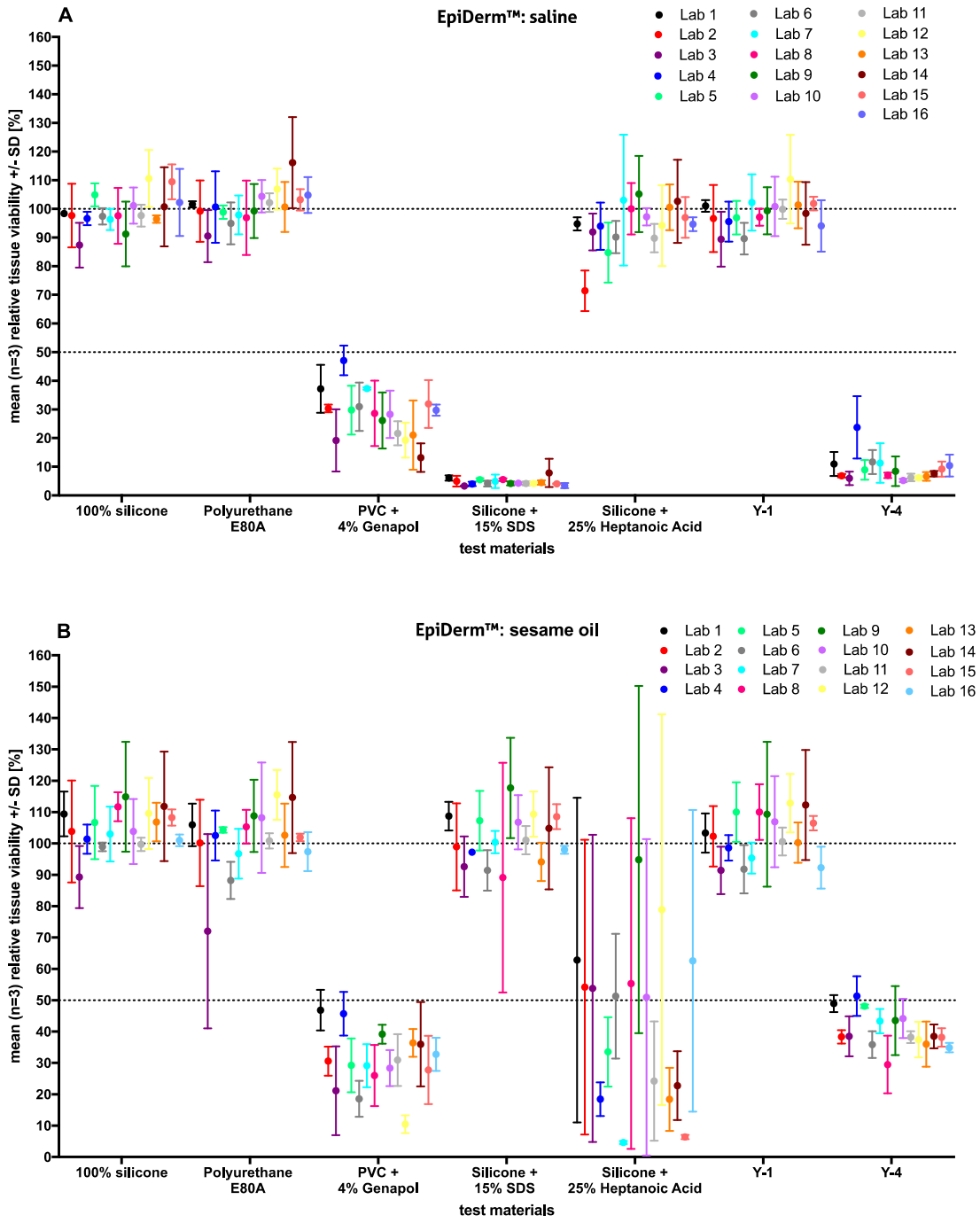

Fig. 3. EpiDerm™ test method. Cell viability results for sixteen laboratories, test materials and both extracts, saline (A) and sesame oil (B). The data show the mean ± SD of three independent experiments (n = 3), except for Lab7: PVC + 4% Genapol (in A) and Lab3: 100% silicone, and Lab9; Silicone + 25% heptanoic acid., for which n = 2.

were below this threshold in 90.9% (100/110) of the cases (data not shown).

3.1.4. Between laboratory reproducibility of the test method using EpiDerm™ tissues

Between Laboratory Reproducibility (BLR) in terms of concordance of classifications was assessed by two approaches per test material (Table 5). The first was based on the mode approach (i.e. classification of a sample judged irritant and/or a non-irritant in three independent experiments and two out of three determining the final score), while the second one was based on the mean approach (i.e. using the mean percentage of tissue viability of the three independent experiments). For saline extracts, the BLR was 100% regardless of the calculation approach. For SO extracts, the BLR was lower (96.2% for the mode approach and 92.4% for the mean approach, respectively). The quantitative measurement of the SD of the mean cell viabilities of all laboratories supported this analysis (Table 5). For example, the SD of 'Silicone + 25% Heptanoic Acid' extracted with SO was highest among all SO extracted test materials, while the concordance was lowest.

3.1.5. Predictivity of the test using EpiDerm™ tissues

The predictivity of the test method was assessed by comparing the obtained classification with the expected results (see Table 3). Three different types of comparisons were applied: a) using the mode of the classifications from the three experiments per test material and laboratory (mode approach), b) using the classification corresponding to

Table 4
Within laboratory reproducibility of the EpiDerm™ test method based on concordance of predictions for the seven test materials.

Laboratory	Concordance of outcomes	
	Saline extracted samples	Sesame oil extracted samples
1	7/7	4/7
2	7/7	6/7
3	7/7	4/6[a]
4	6/7	5/7
5	7/7	7/7
6	7/7	6/7
7	6/6[a]	7/7
8	7/7	5/7
9	7/7	5/6[a]
10	7/7	5/7
11	7/7	7/7
12	7/7	6/7
13	7/7	7/7
14	7/7	7/7
15	7/7	7/7
16	7/7	6/7
Mean	99.1%	85.3%

[a] test materials excluded as only two valid experiments were available.

Table 5
Between laboratory reproducibility of the EpiDerm™ test methods (mode and mean approach) for 16 laboratories per test material, including the SD as a quantitative measurement.

Extract vehicle	Test material	Concordance of outcomes		SD
		Mode approach	Mean approach	
Saline	100% silicone	100% = 16/16	100% = 16/16	5.9
	Polyurethane E80A	100% = 16/16	100% = 16/16	5.7
	PVC + 4% Genapol X-100	100% = 15/15	100% = 15/15	8.2
	Silicone + 15% SDS	100% = 16/16	100% = 16/16	1.1
	Silicone + 25% heptanoic acid	100% = 16/16	100% = 16/16	8.2
	Y-1	100% = 16/16	100% = 16/16	5.1
	Y-4	100% = 16/16	100% = 16/16	4.4
	BLR_{saline} (mean)	100%	100%	–
Sesame oil	100% silicone	100% = 15/15	100% = 15/15	4.9
	Polyurethane E80A	100% = 16/16	100% = 16/16	10.3
	PVC + 4% Genapol X-100	100% = 16/16	100% = 16/16	9.3
	Silicone + 15% SDS	100% = 16/16	100% = 16/16	7.8
	Silicone + 25% heptanoic acid	73.3% = 11/15	53.3% = 8/15	22.9
	Y-1	100% = 16/16	100% = 16/16	7.4
	Y-4	100% = 16/16	94% = 15/16	5.8
	$BLR_{sesame\ oil}$ (mean)	96.2%	92.4%	–

the mean of the viabilities of the three experiments (mean approach) and c) using the classification of each individual experiment. The results, both per extract and across extracts (total), are presented in Table 6.

When extracting with saline, the sixteen laboratories correctly predicted all test materials, regardless the comparison approach. In terms of individual experiments, one laboratory classified the irritant test material 'PVC + 4% Genapol X-100' in one experiment as non-irritant. In summary, 99.7% of the predictions with saline were correct, regardless the comparison approach. When extracting with SO, the highest predictivity (96.4%) was obtained with the mode comparison approach, with four laboratories identifying the test material 'Silicone + 25% Heptanoic Acid' incorrectly as non-irritating. Applying the mean approach resulted in a predictivity of 91.8%, with eight laboratories misidentifying 'Silicone + 25% Heptanoic Acid' and one laboratory misidentifying 'Y-4'. When considering the results of individual samples, twenty-one samples were not correctly identified, with fourteen misidentifications of 'Silicone + 25% Heptanoic Acid', four misidentifications of 'Y-4', and three misidentifications of 'PVC- + 4% Genapol X-100'. Combining the data of both extracts from individual samples resulted in a predictivity of 95.5%.

3.2. SkinEthic™ RHE test method

3.2.1. Positive and negative control outcomes for SkinEthic™ RHE tissues

In all eight laboratories the results of the positive and vehicle controls (saline and SO controls with or without 1% SDS added) concurred with the expected outcome (Fig. 4). Positive and vehicle controls were correctly identified by all laboratories using the SkinEthic™ RHE tissues. For the saline and SO vehicle controls the cell viability in the individual experiments ranged, with one exception, from 85 to 125%, with the vast majority of the vehicle controls falling in the range of 100 ± 10% (Fig. 4). The results for the vehicle control SO showed slightly more variation compared to the saline controls. For the positive controls maximum survival of the tissues was approximately 2% when compared to the negative control which was set at 100%.

3.2.2. Qualification of experiments and test materials within experiments for SkinEthic™ RHE tissues

There were no non-qualified experiments. Excessive test material variability (i.e. SD > 20%) was observed in two cases, which were therefore retested. The combined results of the three independent experiments are presented in Fig. 5. In general, test materials extracted with SO showed slightly higher replicate variability. For the experiments performed with the SkinEthic™ RHE tissues some minor protocol deviations were reported (e.g. an OD of 1.199 for the saline vehicle control in one experiment, which was considered valid and did not to fail the 1.20 OD acceptance criterion). None of these protocol deviations resulted in exclusion of data. Therefore, the data of all eight laboratories were complete (i.e. three valid independent experiments available for each test material with both extracts). All results of the individual experiments with the SkinEthic™ RHE test method are presented in the Supplementary Table S2.

3.2.3. Within-laboratory reproducibility of the test method using SkinEthic™ RHE tissues

When saline was used for extraction the WLR was high (89.3%), with six laboratories discordantly classified test material 'Silicone + 25% Heptanoic Acid' (Table 7). In the SO extract all laboratories predicted the test material 'Silicone + 25% Heptanoic Acid' correctly according to the prediction presented in Table 3 (i.e. NI in saline and I in SO extractant). In addition, the high WLR was confirmed quantitatively using the standard deviation of the cell viabilities of experiments per laboratory and test material. Once again, a mean SD > 20% was defined as a cut-off. When extracting samples with saline, 89% (50/56) of the SD did not exceed this threshold and 86% (48/56) when using SO (data not shown).

3.2.4. Between-laboratory reproducibility of the test method using SkinEthic™ RHE tissues

The BLR in terms of concordance of classifications, was assessed by two approaches per test material (Table 8). The first was based on the mode approach (i.e. classification of a sample a being an irritant and/or a non-irritant based on the viability of the samples being < 50% of the negative control), while the second one was based on the mean approach (i.e. using the mean percentage of tissue viability of the three independent experiments). When extracting with saline, the BLR was generally very high, 89.3% and 92.9% for the mode and mean approach respectively (Table 8). For SO, the BLR was 100% regardless of the approach. The SD of the mean cell viabilities of all laboratories supported this analysis. For example, the SD of 'Silicone + 25% Heptanoic Acid' extracted with saline was highest among all saline extracted test materials, while the concordance was lowest.

Table 6
Predictivity of the EpiDerm™ test method.

Laboratory	Concordance of outcomes with predicted outcome						
	Mode approach		Mean approach		Classifications of individual samples[a]		
	Saline extract	Sesame oil extract	Saline extract	Sesame oil extract	Saline extract	Sesame oil extract	Overall results
1	7/7	6/7	7/7	6/7	21/21	17/21	19/21
2	7/7	7/7	7/7	6/7	21/21	20/21	20/21
3	7/7	6/6	7/7	5/6	21/21	18/20	19/20
4	7/7	7/7	7/7	6/7	20/21	19/21	20/21
5	7/7	7/7	7/7	7/7	21/21	21/21	21/21
6	7/7	6/7	7/7	6/7	21/21	19/21	19/21
7	6/6	7/7	6/6	7/7	20/20	21/21	21/21
8	7/7	7/7	7/7	6/7	21/21	19/21	20/21
9	7/7	6/6	7/7	6/6	21/21	19/20	18/20
10	7/7	7/7	7/7	6/7	21/21	19/21	20/21
11	7/7	7/7	7/7	7/7	21/21	21/21	21/21
12	7/7	6/7	7/7	6/7	21/21	19/21	19/21
13	7/7	7/7	7/7	7/7	21/21	21/21	21/21
14	7/7	7/7	7/7	7/7	21/21	21/21	21/21
15	7/7	7/7	7/7	7/7	21/21	21/21	21/21
16	7/7	6/7	7/7	6/7	21/21	19/21	19/21
Mean predictivity	100%	96.4%	100%	91.8%	99.7%	94.0%	95.5%

[a] Three individual samples for each of the seven test materials for which the assays were performed on three independent days.

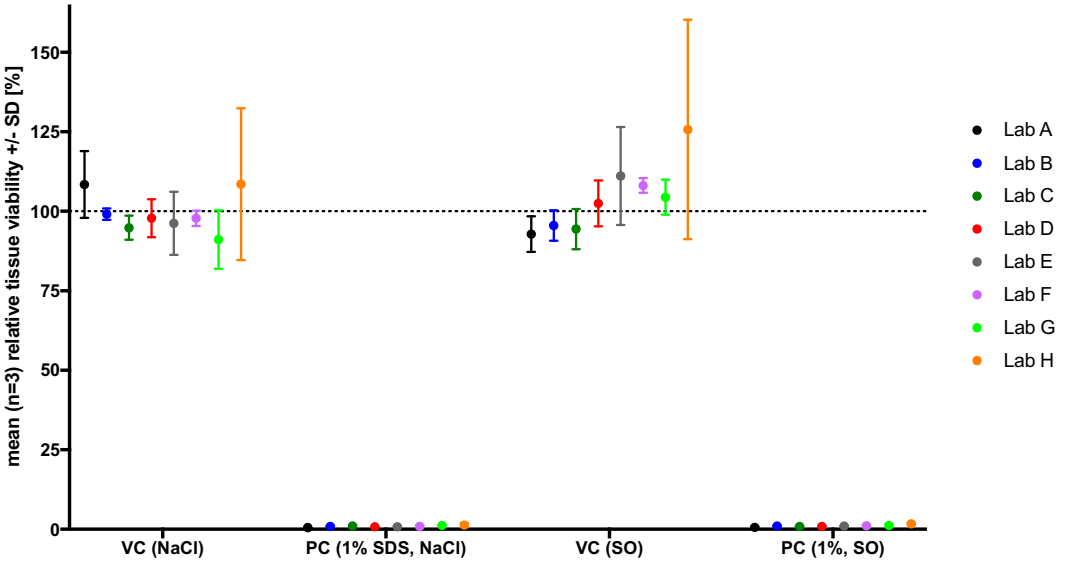

Fig. 4. SkinEthic™ RHE test method. Results for vehicle (VC) and positive controls (PC) for the participating 8 laboratories (Lab A -H). Mean cell viability and standard deviation (SD) of three individual experiments per laboratory are displayed. The dotted line at 100% represents the respective negative control, which is used for normalisation.

3.2.5. Predictivity of the test using SkinEthic™ RHE tissues

The predictivity of the test method was assessed by comparing the obtained classification with the expected results (see Table 3). Three different types of comparisons were applied: a) using the mode of the classifications from the three experiments per test material and laboratory (mode approach), b) using the classification corresponding to the mean of the viabilities of the three experiments (mean approach) and c) using the classification of each individual experiment. The results, both per extract and across extracts (total), are presented in Table 9.

When the test materials were extracted with saline, the lowest predictivity of 89.3% was observed for the mode approach, as six laboratories classified the non-irritant test material 'Silicone + 25% Heptanoic Acid' as an irritant which it is not in polar extracts like saline. The predictivity was similar when considering the individual experiments, but slightly higher with the mean approach (92.9%). When extracting with SO, all laboratories correctly predicted all samples for both the mode and the mean approach. Considering the results of individual samples, the predictivity was still 98.2%. Combining the data of both extracts from individual samples resulted in a predictivity of 100%.

4. Conclusions and discussion

In total, twenty organizations participated in the round robin study evaluating RhE tissues for the determination of skin irritation potential of medical device extracts. These comprised three governmental laboratories, nine contract research organizations, three medical device companies, two university laboratories, and the two RhE model providers, and one consulting company for the statistical evaluation. All stakeholders in the area of safety evaluation of medical devices such as manufacturers, test houses and governmental organizations were included. The round robin tests were performed by eighteen laboratories. All participating laboratories either already had extensive experience with RhE models for the irritant testing of chemicals, or received training at the facilities of the RhE tissue providers prior to performing

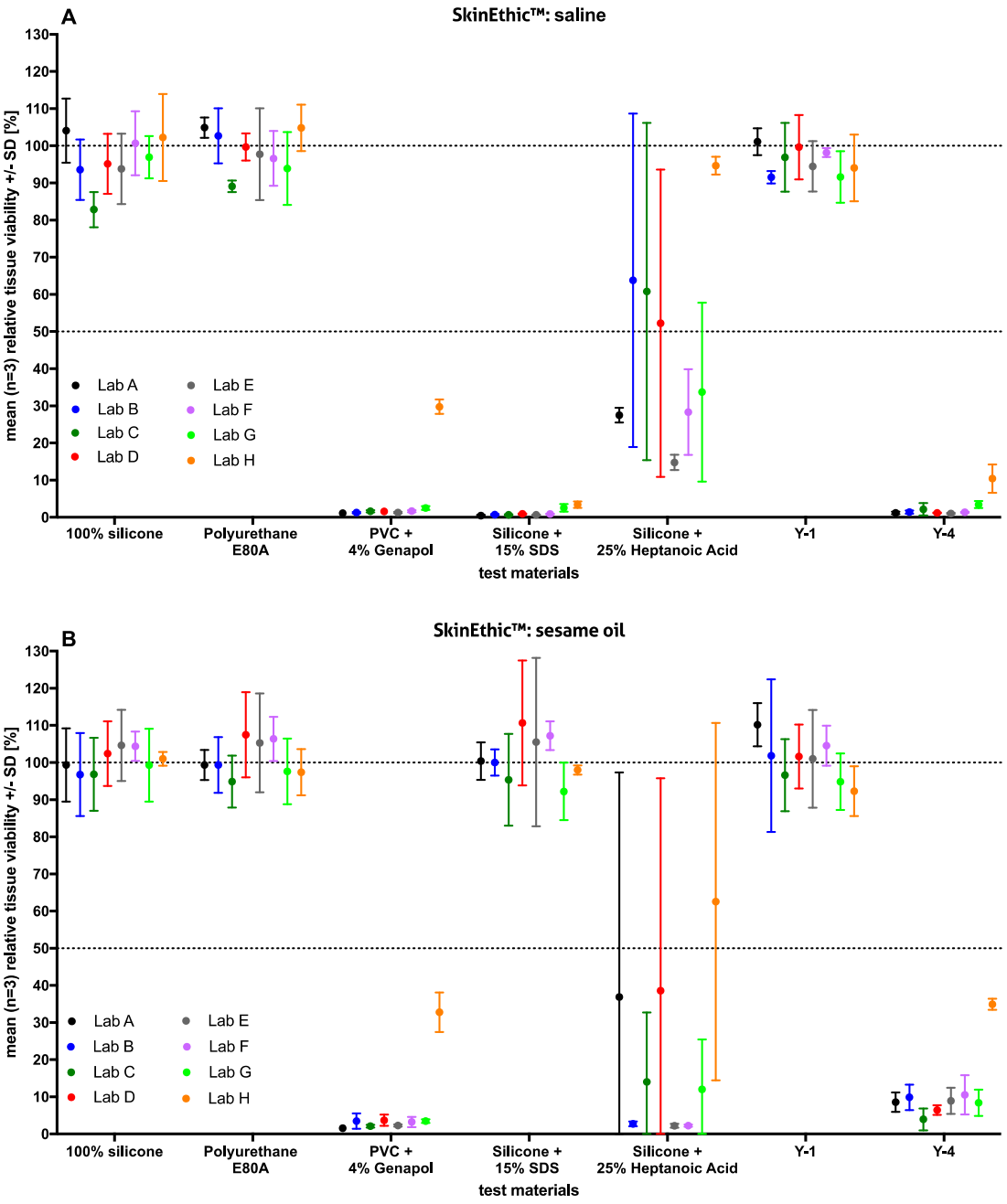

Fig. 5. SkinEthic™ RHE test method. Cell viability results for eight laboratories, test materials and both extracts (A: saline; B: sesame oil). The data show the mean ± SD of three independent experiments.

Table 7
Within laboratory reproducibility of the SkinEthic™ RHE test method based on concordance of predictions for the 7 test materials.

Laboratory	Concordance of outcomes	
	Saline extract	Sesame oil extract
A	6/7	7/7
B	7/7	7/7
C	7/7	7/7
D	6/7	7/7
E	6/7	7/7
F	6/7	7/7
G	6/7	7/7
H	6/7	7/7
Mean	89.3%	100%

the study.

With the RhE tissues a reduction of cell viability below 50% compared to the control was indicative of irritant activity by the tested sample. In total seven materials of which four positive and three negative samples for irritant activity were included in the round robin study. The number of positive samples was relatively low, because it was difficult to prepare acceptable irritant polymers. The primary challenge encountered in preparing these samples was identifying irritant loading levels that would yield extracts with high enough irritant concentrations to produce positive tissue responses. For example, if the loading level was too high (e.g. > 35% by weight), then the polymer's matrix was often compromised with the result being a soupy or paste-like material that was unsuitable for extraction. As a result, the polymerization of various polymers with the added irritants was difficult to establish. Another difficulty was dissolving crystalline irritants if their

Table 8
Between laboratory reproducibility of SkinEthic™ RHE test method (mode and mean approach) for eight laboratories per test material, including also SD as a quantitative measurement.

Extract vehicle	Test material	Concordance of outcomes		SD
		Mode approach	Mean approach	
Saline	100% silicone	100% = 8/8	100% = 8/8	13.2
	Polyurethane E80A	100% = 8/8	100% = 8/8	10.2
	PVC + 4% Genapol X-100	100% = 8/8	100% = 8/8	0.4
	Silicone + 15% SDS	100% = 8/8	100% = 8/8	0.7
	Silicone + 25% heptanoic acid	25% = 2/8	50% = 4/8	18.1
	Y-1	100% = 8/8	100% = 8/8	7.6
	Y-4	100% = 8/8	100% = 8/8	0.8
	BLR_{saline} (mean)	89.3%	92.9%	–
Sesame oil	100% silicone	100% = 8/8	100% = 8/8	5.9
	Polyurethane E80A	100% = 8/8	100% = 8/8	8.6
	PVC + 4% Genapol X-100	100% = 8/8	100% = 8/8	0.8
	Silicone + 15% SDS	100% = 8/8	100% = 8/8	9.0
	Silicone + 25% heptanoic acid	100% = 8/8	100% = 8/8	16.8
	Y-1	100% = 8/8	100% = 8/8	8.1
	Y-4	100% = 8/8	100% = 8/8	2.4
	$BLR_{sesame\ oil}$ (mean)	100%	100%	–

polarity was different to that of the polymer. For lactic and heptanoic acid, a number of failed attempts were made, also several irritant oils were tried as well with little or no success. In preliminary evaluations liquid and powdered lactic acid in one part silicone did not yield reproducible results. So, it was discarded. Overall, it took three years of biomaterial preparation and exploratory preliminary testing before the round robin was conducted in 2016. During the preparation of test samples several biomaterials including PVC containing lactic acid, were found not to be suitable to serve as positive test materials. In addition, the results obtained with heptanoic acid incorporated into one-part silicone showed considerable variation. When extracted in saline these samples produced more consistent results than those obtained from SO extractions. This variation might be due to an uneven distribution of the heptanoic acid in the polymer matrix, solubility issues, or due to other yet unknown reasons.

With the exception of heptanoic acid in silicone, the other irritant containing polymer biomaterials (silicone and PVC containing SDS and Genapol X-80 and X-100, respectively) were accurately detected (Figs. 3 and 5 in which cell viability below 50% indicates irritant activity). The irritant activity was detected either in the saline extract (SDS in silicone), in the SO extract (heptanoic acid in silicone, although with some variation in the outcome), or in both the saline and SO extracts (PVC + 4% Genapol X-100, Y-4). Sesame oil extraction did not detect SDS as an irritant in the silicone polymer, however the saline extraction did. This difference may possibly be explained by the water solubility of SDS. Also, the negative samples not containing an irritant were all correctly identified. The overall predictability for the assay was above 90% (95.5% with EpiDerm and 100% with SkinEthic RHE). The laboratories were able to identify irritants and non-irritants with an overall accuracy rate of 97.4% (data not shown).

Although only two RhE models were evaluated in this round robin study, the outcome of the study showed that the positive and negative coded samples were identified with a high degree of accuracy (De Jong et al., 2017, this study). Thus, it seems likely that other available RhE models as listed in OECD TG 439 and/or validated for neat chemicals, may also be suitable for the testing of medical device extracts. However, as indicated by our preliminary studies, exposure times may need to be adapted (i.e. prolonged) for the low concentrations of irritants that are likely to be present in the medical device extracts. Reference materials used in this study will be helpful to benchmark performances of such new protocols.

In conclusion, the results demonstrate that RhE tissues are a robust model for the detection of irritant activity, and can be used for the identification of low levels of strong irritants in medical device extracts. The round robin study was an initiative of Working Group 8 for Irritation and Sensitization of ISO Technical Committee 194 on Biological and Clinical Evaluation of Medical Devices. Based on the results as presented here, a new international standard, ISO 10993-23, for in vitro irritation testing of medical devices, will be drafted as a replacement for the animal irritation studies now indicated in ISO 10993-10:2010 *Biological evaluation of medical devices – Part 10: Tests for irritation and skin sensitization.*

Conflict of interest

Wim H. De Jong, Yuji Haishima, Beau Rolllins, Shelby Skoog, Anita Sawyer, Kristina Fant, Kwang-Mahn Kim, Jea sung Kwon, Reiko Kato, Atsuko Miyajima, Liset De La Fonteyne, Nicholas Christiano, and Kelly P. Coleman, employees of governmental organizations, universities or medical device companies, declare no interest. Michelle Lee, Austin Zdawzcyk, Jamin Willoughby, Timothy Schatz, Sherry Parker, Paolo Pescio, Helge Gehrke, Hana Hofman-Hüther, Marisa Meloni, Conrad Julius, Damien Briotet, Audrey Turley, and Thor Rollins are employees of Contract Research Organizations (CRO's). Helena Kandárová, Christian Pellevoisin, Michael Bachelor, Sylvia Letasiova, Christelle Videau, and Carine Tornier are employees of the manufacturers of the reconstructed human epidermis (RhE) model. Sebastian Hoffmann as independent statistician received a fee from one of the participating CRO companies.

Table 9
Predictivity of the SkinEthic™ RHE test method.

Laboratory	Concordance of outcomes with predicted outcome						
	Mode approach		Mean approach		Classifications of individual[a] samples		
	Saline extract	Sesame oil extract	Saline extract	Sesame oil extract	Saline extract	Sesame oil extract	Overall
A	6/7	7/7	6/7	7/7	18/21	20/21	21/21
B	7/7	7/7	7/7	7/7	20/21	21/21	21/21
C	7/7	7/7	7/7	7/7	20/21	21/21	21/21
D	6/7	7/7	7/7	7/7	19/21	20/21	21/21
E	6/7	7/7	6/7	7/7	18/21	21/21	21/21
F	6/7	7/7	6/7	7/7	18/21	21/21	21/21
G	6/7	7/7	6/7	7/7	19/21	21/21	21/21
H	6/7	7/7	7/7	7/7	20/21	20/21	21/21
Mean predictivity	89.3%	100%	92.9%	100%	90.5%	98.2%	100%

[a] three individual samples for each of the seven test materials for which the assays were performed on three independent days.

Transparency document

The http://dx.doi.org/10.1016/j.tiv.2018.01.001 associated with this article can be found, in online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tiv.2018.01.001.

References

Armstrong, B.L., Senyurt, A.F., Narayan, V., Wang, X., Alquier, L., Vas, G., 2013. Stir bar sorptive extraction combined with GC-MS/MS for determination of low level leachable components from implantable medical devices. J. Pharm. Biomed. Anal. 74, 162–170.

Casas, J.W., Lewerenz, G.M., Rankin, E.A., Willoughby Sr., J.A., Blakeman, L.C., JM Jr., McKim, Coleman, K.P., 2013. In vitro human skin irritation test for evaluation of medical device extracts. Toxicol. in Vitro 27, 2175–2183.

De Jong, W.H., Hoffmann, S., Lee, M., Kandárová, H., Letasiova, S., De La Fonteyne, L.J., Pellevoisin, C., Tornier, C., Bremond, C., Haishima, Y., Zdawczyk, A.M., Turley, A., Willoughby, J.A., Bachelor, M.A., Rollins, B., Coleman, K.P., 2017. Round Robin study to evaluate the reconstructed human epidermis (RhE) model as in vitro skin irritation test for detection of irritant activity in medical device extracts. In: The Toxicologist. Abstract Vol. 2153b. pp. 272.

Draize, J.H., Woodard, G., Clavery, H.O., 1944. Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J. Pharmacol. Exp. Ther. 82, 377–390.

Faller, C., Bracher, M., Dami, N., Roguet, R., 2002. Predictive ability of reconstructed human epidermis equivalents for the assessment of skin irritation of cosmetics. Toxicol. in Vitro 16, 557–572.

ISO 10993-1, 2009. Biological Evaluation of Medical Devices – Part 1: Evaluation and Testing Within a Risk Management Process. International Organization for Standardization, Geneva, Switzerland.

ISO 10993-10, 2010. Biological Evaluation of Medical Devices – Part 10: Tests for Irritation and Skin Sensitization. International Organization for Standardization, Geneva, Switzerland.

ISO 10993-12, 2012. Biological Evaluation of Medical Devices – Part 12: Sample Preparation and Reference Materials. International Organization for Standardization, Geneva, Switzerland.

Kandárová, H., Liebsch, M., Gerner, I., Schmidt, E., Genschow, E., Traue, D., Spielmann, H., 2005. The EpiDerm test protocol for the upcoming ECVAM validation study on in vitro skin irritation tests—an assessment of the performance of the optimised test. Altern. Lab. Anim 33, 351–367.

Kandárová, H., Willoughby, J.A., De Jong, W.H., Bachelor, M.A., Letasiova, S., Milasova, T., Breyfogle, B., De La Fonteyne, L., Coleman, K.P., 2015. Development, optimization, and standardization of an in vitro skin irritation test for medical devices using the reconstructed human tissue model EpiDerm. In: The Toxicologist. Abstract 2017. pp. 433.

Kandárová, H., Willoughby, J.A., De Jong, W.H., Bachelor, M.A., Letasiova, S., Milasova, T., Breyfogle, B., De La Fonteyne, L., Haishima, Y., Coleman, K.P., 2016. Development, optimization and validation of an in vitro skin irritation test for medical devices using the reconstructed human tissue model Epiderm. In: The Toxicologist, pp. 172 (Late breaking abstract Submissions).

Lucas, A.D., Merritt, K., Hitchins, V.M., Woods TO, McNamee, S.G., Lyle, D.B., Brown, S.A., 2003. Residual ethylene oxide in medical devices and device material. J Biomed Mater Res B Appl Biomater 66, 548–552.

Mosmann, T., 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J. Immunol. Methods 65, 55–62.

Nomura, Y., Lee, M., Fukui, C., Watanabe, K., Olsen, D., Turley, A., Morishita, Y., Kawakami, T., Yuba, T., Fujimaki, H., Inoue, K., Yoshida, M., Ogawa, K., Haishima, Y., 2017. Proof of concept testing of a positive reference material for *in vivo* and *in vitro* skin irritation testing. J. Biomed. Mater. Res., Part B 2017 (00B), 000-000. http://dx.doi.org/10.1002/jbm.b.34061/epdf.

OECD, 2015. OECD guidelines for the testing of chemicals: *in vitro* skin irritation: reconstructed human *epidermis* test method. In: Test Guideline 439. Organisation for Economic Co-operation and Development, Paris, France (21 pp).

Olsen, D.S., Lee, M., Turley, A., Sasaki, S., Yamasaki, K., Fukui, C., Nomura, Y., Kato, R., Niimi, S., Yuba, T., Sakaguchi, K., Haishima, Y., 2015. Extractable positive control for in vitro skin irritation testing of medical devices. In: The Toxicologist. Abstract 2016. pp. 433.

Pellevoisin, C., Tornier, C., Gregoire, C., Rollins, B., Briotet, D., Turley, A., Lee, M., Haishima, Y., Seyler, N., 2016. Skin irritation of medical devices: in vitro assay with episkin reconstructed human epidermis (RHE). In: 2016 Toxicologist. Abstract 3547. pp. 598.

Petrusevski, P., Jolevsk, S.T., Ribarsk, J.T., Chachorovsk, M., Petkovska, A., Ugarkovica, S., 2016. Development of complementary HPLC-DAD/APCI MS methods for chemical characterization of pharmaceutical packaging materials. J. Pharm. Biomed. Anal. 124, 228–235.

Spielmann, H., Hoffmann, S., Liebsch, M., Botham, P., Fentem, J.H., Eskes, C., Roguet, R., Cotovio, J., Cole, T., Worth, A., Heylings, J., Jones, P., Robles, C., Kandárová, H., Gamer, A., Remmele, M., Curren, R., Raabe, H., Cockshott, A., Gerner, I., Zuang, V., 2007. The ECVAM international validation study on in vitro tests for acute skin irritation: report on the validity of the EPISKIN and EpiDerm assays and on the skin integrity function test. ATLA Altern. Lab. Anim. 35, 559–601.

Tornier, C., Amsellem, C., Fraissinette Ade, B., Alépée, N., 2010. Assessment of the optimized SkinEthic Reconstructed Human Epidermis (RHE) skin irritation protocol over 39 test substances. Toxicol. in Vitro 24, 245–256.