

Comparative Performance of Four *In Vitro* Test Methods For the Classification of Ocular Corrosives and Severe Irritants.

WS Stokes¹, NY Choksi^{1,2}, DG Allen^{1,2}, JF Truax^{1,2}, RR Tice¹

¹National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), NIEHS/NIH/DHHS, Research Triangle Park, North Carolina; ²Integrated Laboratory Systems, Inc., Research Triangle Park, North Carolina

Abstract

The Draize rabbit eye test is accepted by U.S. and international regulatory authorities for the assessment of ocular hazard potential. However, concerns about animal welfare and the reliability of the in vivo rabbit eve test have led researchers to develop in vitro test methods as alternatives to the currently used Draize rabbit eye test method. The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) evaluated the Isolated Rabbit Eye, Isolated Chicken Eye, Bovine Corneal Opacity and Permeability, and the Hens Egg Test-Chorioallantoic Membrane test methods for their ability to detect ocular corrosives or severe irritants. In vivo results were classified based on U. S. and international ocular hazard classification systems (i.e., U.S. Environmental Protection Agency [EPA], the European Union [EU], and the United Nations Globally Harmonized System [GHS]). Ocular corrosives and severe irritants were defined as Category 1 according to the GHS, as Category I according to the EPA, or as R41 according to the EU. In vitro results were classified as severe irritants based on decision criteria obtained from a literature review and/or personal communications with developers of the individual test methods. Accuracy of the four evaluated test methods ranged from 53% to 81% for the GHS classification system and was similar across all three regulatory classification systems for each in vitro test method. Accuracy analyses based on the physicochemical characteristics of a test substance suggested limitations

for each test method. Intralaboratory repeatability and reproducibility and interlaboratory reproducibility were evaluated, when adequate data were available, using both quantitative and qualitative approaches. The analyses indicated that the test methods were generally reproducible within and between testing laboratories. Supported by NIEHS contract N01-ES-35504.

Introduction

Accidental eye injury is the leading cause of visual impairment in the United States (BLS 2004). In 2003, eye injuries from chemicals and their products (6,080) accounted for 16% of all eye injuries (36,940) reported as the cause of Days Away From Work for employees. The ocular irritation or corrosion potential of substances to which humans may be exposed has been evaluated since 1944 by the Draize rabbit eye test (Draize et al. 1944). Efforts to develop in vitro alternatives to this in vivo test have yet to result in adequately validated and accepted nonanimal test methods for ocular irritancy. Since a test method must be demonstrated to be adequately validated before it can be considered for regulatory acceptance, the U.S. Environmental Protection Agency (EPA) formally nominated to ICCVAM four in vitro test methods, the Isolated Rabbit Eye (IRE), Isolated Chicken Eye (ICE), Bovine Corneal Opacity and Permeability (BCOP), and the Hen's Egg Test-Chorioallantoid Membrane (HET-CAM), for evaluation of their ability to identify ocular corrosives and severe irritants in a tiered testing strategy.

For each of the evaluated test methods (i.e., ICE, IRE, BCOP, and HET-CAM), the National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods (NICEATM), which provides scientific support to ICCVAM, prepared a comprehensive Background Review Document (BRD) reviewing the available data and information. NICEATM released the draft BRD for public comment on November 1. 2004². On January 11-12, 2005, ICCVAM convened an Expert Panel to independently evaluate the validation status of these in vitro test methods for identifying ocular corrosives or severe irritants³. Since public comments indicated that additional data could be made available, the Expert Panel recommended that the additional data be used in a reanalysis of the accuracy and reliability of each test method.

In response, an FR notice was published on February 28, 2005 (FR Vol. 70, No. 38, pp. 9661-9662) requesting the submission of all available in vitro test data and corresponding in vivo rabbit eye test data to NICEATM. In addition to considering any data received in response to the FR notice, the reanalysis of the accuracy and reliability of this test method took into account (1) changes that occurred in the ocular irritancy classification of a few substances in response to clarification of the European Union (EU) (2001) and United Nations (UN) Globally Harmonized System for Classification and Labelling (GHS) (UN 2003) ocular irritation classification rules; (2) a decision to use classifications based on *in vivo* rabbit eye test data only, and not on physico-chemical properties such as pH extremes or other test methods (e.g., dermal corrosivity); and (3) revised chemical class assignments for some substances. The accuracy and reliability reanalyses and a revised reference substances list for

validation of *in vitro* tests to detect ocular corrosives and severe irritants were presented in a BRD Addendum that was released on July 26, 2005². Additional information on the reanalysis can be obtained at http://iccvam.niehs.nih.gov/methods/ocudocs/reanalysis.htm.

¹Validation is the process by which the reliability and relevance of a test method are established for a specific purpose (ICCVAM 1997, 2003).

²The draft BRDs and BRD addendum can be obtained at http://iccvam.niehs.nih.gov/methods/ocudocs/ocu brd.htm ³The January 2005 Expert Panel Report can be obtained at http://iccvam.niehs.nih.gov/methods/ocudocs/EPreport/ocureport.htm

In Vitro Test Method Overviews and Databases

BCOP Test Method Overview and Database

The basic procedure for the BCOP test method is provided in Figure 1. Historically, negative control corneas have been used to correct opacity and permeability values measured on treated corneas. Mean corrected opacity and mean corrected permeability values are calculated for each treatment group. An In Vitro Irritancy Score (IVIS) is calculated using the following empirically-derived formula (Sina et al. 1995): IVIS = Opacity value + (15 x OD₄₉₀ value).

An *in Vitro* Irritancy Score ≥ 55.1 is considered a severe eye irritant. Some substances, such as anionic and nonionic surfactants, increase permeability without significant opacity; thus, only permeability values are used for certain chemical classes. In such situations, a test substance that increases permeability $(OD_{490}) > 0.600$ is considered a severe irritant. In addition, histopathological evaluation of the treated cornea (conducted after permeability is assessed) is used on a case-by-case basis (Curren

- The following studies were used for the various analyses:
- Gautheron et al. (1994) Southee (1998)

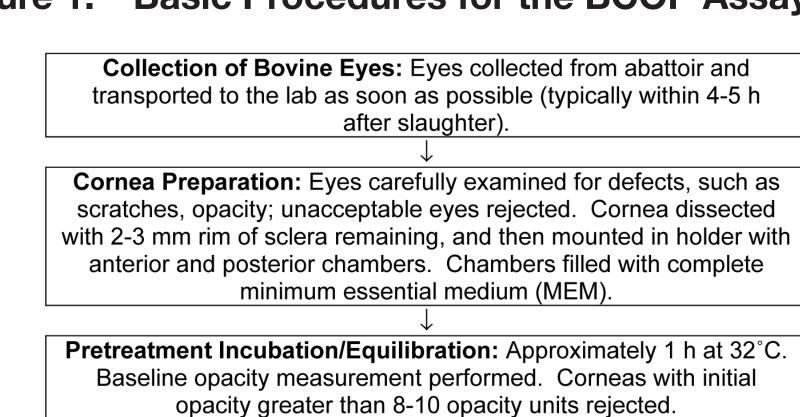
transport, different negative controls).

- Balls et al. (1995)
- Swanson and Harbell (2000)
- Swanson et al. (1995)
- - Bailey et al. (2004) Submission from Dr. Joseph Sina
- Gettings et al. (1996) Casterton et al. (1996)

A total of 158 substances were used to evaluate the BCOP test method accuracy. Data for 127 substances were appropriate for evaluation of interlaboratory reproducibility, while data for 96 and 41 substances were appropriate for analysis of intralaboratory repeatability and reproducibility respectively. The BCOP test method protocols used in these studies were similar to each other, but not identical (differences included number

Figure 1. Basic Procedures for the BCOP Assay

of corneas used [n=3-5], storage conditions of bovine eyes during



Treatment Groups: 2-3 corneas selected as negative controls. 3-5 corneas used per test substance and positive control.

Treatment of Corneas: MEM removed from both chambers. Fresh MEM is added to posterior chamber. Test substance added to

anterior chamber.

→	
LIQUIDS: 750 µL tested at 100%.	SOLIDS: 750 µL tested at 20%.
SURFACTANTS: 750 µL tested at	(4-h exposure at 32°C)
10%. <i>(10-min exposure at 32°C)</i>	
\	\downarrow
Corneas rinsed at least 3X with MEM;	Corneas rinsed at least 3X with MEM;
Fresh MEM added to anterior chamber	Fresh MEM added to both chambers
↓	
Post-treatment opacity taken.	Final opacity measurement taken.

Corneas incubated for 2 h MEM replaced in both chambers and final opacity measurement taken

MEM in anterior chamber replaced MEM in anterior chamber replaced with 1 mL of 4 mg/mL NaF with 1 mL of 5 mg/mL NaF

90 min incubation at 32° Aliquot taken from posterior chamber for OD₄₉₀ reading If histology is to be performed, corneas are fixed after permeability measurements are completed

IRE Test Method Overview And Database

During an IRE study, a test substance is applied to the cornea of eyes isolated from rabbits. Test substances are applied as a single dose (100 μL or 100 mg) for 10 sec followed by rinsing with 20 mL. Corneas are then evaluated for opacity and swelling (measured as a change in thickness), fluorescein penetration, and epithelial damage at 0.5, 1, 2, 3, and 4 hours.

Substances that induce a response that exceeds a cutoff score in any one of four ocular endpoints (corneal opacity score [opacity x area] > 3, corneal swelling \geq 25%, fluorescein penetration score [intensity x area] \geq 4, or any sign of epithelial damage [stippling, mottling, ulceration, etc.]) are identified as corrosive or severe ocular irritants.

The database for the IRE accuracy analysis consisted of a total of 149 test substances obtained from four studies (CEC 1991; Balls et al. 1995; Gettings et al. 1996; Guerriero et al. 2004). However, only Guerriero et al. (2004) used all four ocular endpoints to identify corrosives or severe irritants. An analysis based on a "Pooled Data Set" was conducted, which included all available data from the four studies. Only one study (number of substances = 59) could be used to assess the interlaboratory reproducibility of the IRE test method (Balls et al. 1995).

ICE Test Method Overview and Database

During an ICE study, a test substance is applied to the cornea of eyes isolated from chickens processed for human consumption. Test substances are applied as a single dose (30 µL or 30 mg) for 10 sec followed by rinsing with isotonic saline. A single negative control eye (treated with saline) is used to verify assay conditions. Corneal reactions (swelling and opacity) are measured at 0, 30, 75, 120, 180, and 240 min post-treatment, and mean values (at each time point for all eyes) for each endpoint are determined. Fluorescein retention is evaluated at 0 and 30 min. The maximum mean value for each endpoint is used to categorize the response and then the categories for all the endpoints are used to assign an in vitro irritancy classification (See Table 1). Morphological (e.g., loosening of the epithelium; roughening of the corneal surface) and histopathological assessments can also be included on a case-by-case basis to discriminate borderline cases, although decision criteria to assign an irritancy classification have not been established for histopathological endpoints.

A total of 175 substances from five different studies (Prinsen and Koëter 1993: Balls et al. 1995: Prinsen 1996: 2000: 2005) were used to evaluate the accuracy of the ICE test method; data for 59 substances were appropriate for evaluation of interlaboratory reproducibility, while data for four substances were appropriate for analysis of intralaboratory reproducibility. The primary difference among various ICE studies was the number of treated eyes per test substance (3 to 5).

Table 1. ICE Decision Criteria for Classifying Ocular **Corrosives and Severe Irritants**

Corneal Swelling		Corneal Opacity	
Max. Mean Swelling* (%)	Category	Max. Mean Score*	Category
0 - 5	ı	0 - 0.5	I
>5 – 12	II	0.6 - 1.5	II
>12 – 18 (>75 min post-treatment)	II	1.6 - 2.5	III
>12 – 18 (≤ 75 min post-treatment)	III	2.6 – 4.0	IV
>18 – 26	III	Fluorescein Retention	
>26 – 32 (>75 min post-treatment)	III	Mean Score** Category	
>26 - 32 (≤ 75 min post-treatment)	IV	0 – 0.5	ı
>32	IV	0.6 - 1.5	II
	•	1.6 - 2.5	III
		2.6 - 3.0	IV

any time point (maximum mean value) is used for categorization. **Recorded at 30 min post-treatment.

Possible combinations of the three ICE endpoint categories yielding a severe irritant/corrosive classification:

- 3 x IV
- 2 x IV, 1 x III or II or 1
- CO ≥ 3 at 30 min
- CO = 4 at any time Severe loosening of the epithelium

HET-CAM Test Method Overview and Database

During a HET-CAM study, a test substance is applied to the chorioallantoic membrane (CAM) as a single dose. Adverse effects on the CAM are measured up to 300 sec after application of the test substance and damage to the CAM is assessed by visual inspection. Each endpoint (e.g., hyperemia, hemorrhage, and coagulation) evaluated is used to develop an overall irritancy score that is used to assign an *in vitro* irritancy

The following studies were used for this reanalysis:

- CEC (1991)
- Kojima et al. (1995)

Gettings et al. (1996)

- Gettings et al. (1991)
- Gilleron et al. (1996) • Bagley et al. (1992)

and 99 substances were evaluated, respectively.

- Gettings et al. (1994) Spielmann et al. (1996)
- Gilleron et al. (1997) Vinardell and Macián (1994)
- Balls et al. (1995) Hagino et al. (1999)

These studies included a number of variations in test method protocol (e.g., relative humidity of eggs during incubation, endpoints evaluated) and methods of data analysis (i.e., IS(A), IS(B), Q-Score, S-Score, mtc10, and IS & ITC⁴). Due to these variations, not all studies were suitable for the accuracy and reliability analyses reported here. For the IS(B)-10 and IS(B)-100 analysis methods, which are presented here, 101 and 138 substances were evaluated for the accuracy analyses, respectively. For the reliability analyses of the IS(B)-10 and IS(B)-100 methods, 107

⁴Analysis methods: **IS(A):** Irritation responses are evaluated at 0.5, 2, and 5 minutes and time-dependent scores are assigned to each endpoint. The total score is calculated by adding assigned scores. IS(B): Time of first appearance of endpoint is noted after application of test substance. Total score is calculated by using empirically derived formula. Q-Score: Calculated as ratio of test substance irritation score to investigator determined reference standard irritation score. **S-Score:** Calculated as the highest total score for any endpoint evaluated. **mtc10:** Mean detection time for appearance of coagulation endpoint when using a 10% solution. **IS &ITC:** Two different analysis methods used. IS value calculated as IS(A) or IS(B) (described above). ITC defined as lowest concentration required to produce a slight response after application of test substance.).

Comparative Test Method Accuracy

The accuracy of the four *in vitro* test methods for the various data analysis methods described, when compared to in vivo rabbit eye test classifications using the UN GHS (UN 2003) classification system are provided in Table 2. Similar results were obtained for the EPA and EU classification systems.

Table 2. Comparison of Performance Characteristics of Four In Vitro Test Method for Identification of GHS Severe Ocular Irritants or Corrosives

Statistic	IRE (n = 107) ¹	ICE (n = 144)	HET-CAM (n = 101) ²	HET-CAM (n = 138) ³	BCOP (n = 147)
A	65%	83%	68%	54%	81%
Accuracy	$(70/107)^4$	(120/144)	(69/101)	(75/138)	(119/147)
Concidinate	70%	50%	70%	87%	84%
Sensitivity	(33/47)	(15/30)	(28/40)	(34/39)	(36/43)
Chaoifiaitre	62%	92%	67%	41%	80%
Specificity	(37/60)	(105/114)	(41/61)	(41/99)	(83/104)
Positive	59%	63%	58%	37%	63%
Predictivity	(33/56)	(15/24)	(28/48)	(34/92)	(36/57)
Negative	73%	88%	77%	89%	92%
Predictivity	(37/51)	(105/120)	(41/53)	(41/46)	(83/90)
False Positive	38%	8%	33%	59%	20%
Rate	(23/60)	(9/114)	(20/61)	(58/99)	(21/104)
False Negative	30%	50%	30%	13%	16%
Rate	(14/47)	(15/30)	(12/40)	(5/39)	(7/43)

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability assay; GHS = Globally Harmonized System; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane assay; ICE = Isolated Chicken Eye assay: IRE = Isolated Rabbit Eve assay. ¹n = number of substances tested; the numbers in parentheses in each row indicates the data on which

²These data are for the IS(B) method (described by Kalweit et al. 1987) when testing substances as a 10% ³These data are for the IS(B) method (described by Kalweit et al. 1987) when testing substances at a These data are for the Pooled Data Set.

Tables 3 to 6 provide results for each *in vitro* test method when accuracy was evaluated for a variety of physical and chemical classes. The small number of substances representing most chemical classes allows for only limited conclusions with respect to the accuracy of test methods by chemical class or property of interest (e.g., solids vs. liquids, basic vs. acidic pH, surfactants). Highlighted chemical classes are those that might be considered problematic based on: (a) their associated false positive and/or false negative rates are greater than the overall false positive and/or false negative rates and (b) there were greater than five substances used to determine the rate.

For the BCOP test method, the highest overpredicted classes are alcohols and ketones, while the highest underpredicted class is solids

Table 3. False Negative and False Positive Rates of the BCOP Test Method, by Chemical Class and Properties of Interest, for the GHS Classification System

Overall	147	20% (21/104)	16% (7/43)
		Chemical Class	
hols	18	53% (8/15)	67% (2/3)
ne/Amidine	8	0% (0/4)	0% (0/4)
oxylic acids	15	38% (3/8)	14% (1/7)
rs	12	12% (1/8)	0% (0/4)
r/Polyether	6	0% (0/5)	0% (0/1)
rocyclic pounds	12	33% (2/6)	17% (1/6)
rocarbons	12	8% (1/12)	- (0/0)
ganic Salt	5	0% (0/3)	0% (0/2)
nes	10	40% (4/10)	- (0/0)
ım compounds	11	0 % (0/3)	0% (0/8)
		Properties of Interest	
ids	92	26% (18/68)	4% (1/24)
ds	32	10% (2/20)	42% (5/12)
icide	8	33% (1/3)	40% (2/5)
actants ³	35	5% (1/21)	7% (1/14)
viations: BCOP = Bo	vine Corneal	Opacity and Permeability assa	ay; GHS = Globally Harmonized
m			

False Positive Rate² False Negative Rate²

²False Positive Rate = the proportion of all negative substances that are falsely identified as positive *in* vitro; False Negative Rate = the proportion of all positive substances that are falsely identified as negative in vitro. The data used to calculate the percentage are provided in parenthesis 3Combines single chemicals labeled as surfactants along with surfactant-containing formulations.

More information on ICCVAM and NICEATM can be accessed at: http://iccvam.niehs.nih.gov/





For the ICE test method, alcohols tend to be overpredicted, while surfactants and solids tend to be underpredicted (Table 4)

Table 4. False Negative and False Positive Rates of the ICE Test Method, by Chemical Class and **Properties of Interest, for the GHS** Classification System

Category	$ N^1 $	False Positive Rate ²	False Negative Rate ²
Overall	144	8% (9/114)	50% (15/30)
		Chemical Class	
Alcohols	12	50% (5/10)	50% (1/2)
Amine/Amidine	5	0% (0/2)	33% (1/3)
Carboxylic acids	10	0% (0/3)	43% (3/7)
Esters	9	13% (1/8)	0% (0/1)
Heterocyclic compounds	9	0% (0/3)	33% (2/6)
Onium compounds	8	0% (0/2)	33% (2/6)
-		Properties of Interest	
Liquids	108	10% (9/90)	44% (8/18)
Solids	36	0% (0/24)	58% (7/12)
Pesticides	11	0% (0/6)	60% (3/5)
Surfactants	21	0% (0/12)	56% (5/9)

Abbreviations: GHS = Globally Harmonized System; ICE = Isolated Chicken Eye assay. ²False Negative Rate = the proportion of all positive substances that are falsely identified as negative in vitro: False Positive Rate = the proportion of all negative substances that are falsely identified as positive

in vitro. The data used to calculate the percentage are provided in parenthesis.

For the IRE test method, alcohols, amines, ketones, and liquids were the most overpredicted classes, while carboxylic acids and organic compounds were the most underpredicted chemical classes (Table 5)

Table 5. False Negative and False Positive Rates of the IRE Test Method, by Chemical Class and Properties of Interest, for the GHS Classification System (Analysis Based on the Pooled

Category	N ¹	False Positive Rate ²	False Negative Rate ²
rerall	107	38% (23/60)	30% (14/47)
	Ch	emical Class	
cohol	13	55% (6/11)	50% (1/2)
nide	5	0% (0/3)	0% (0/2)
nine	11	50% (3/6)	20% (1/5)
rboxylic acid	12	33% (2/6)	67% (4/6)
ter	10	30% (3/10)	- (0/0)
her	9	33% (2/6)	0% (0/3)
rmulation	24	25% (2/8)	38% (6/16)
terocyclic compound	18	44% (4/9)	11% (1/9)
tone	6	67% (4/6)	- (0/0)
nium compound	10	33% (1/3)	0% (0/7)
ganic compound	12	17% (1/6)	50% (3/6)
Ifur compound	8	20% (1/5)	33% (1/3)
	Prope	erties of Interest	
quid/Solution	65	49% (18/37)	29% (8/28)
lids	42	22% (5/23)	32% (6/19)
rfactant-based mulation	24	25% (2/8)	38% (6/16)
rfactant	13	40% (2/5)	12% (1/8)

²False Positive Rate = the proportion of all negative substances that are falsely identified as positive in vitro: False Negative Rate = the proportion of all positive substances that are falsely identified

Alcohols, ethers, heterocyclcic compounds, and organic salts were the highest overpredicted chemical classes in HET-CAM IS(B)-10. while alcohols, aldehydes, amines, esters, heterocyclic compounds, and organic salts were the highest overpredicted chemical classes in HET CAM IS(B)-100 (**Table 6**).

Table 6. False Negative and False Positive Rates of the HET-CAM Test Method, by Chemical Class and Properties of Interest, for the GHS Classification System

Category	N ¹	False Positive Rate ²	False Negative Rate ²
	Chemic	cal Class-IS(B)-10	
Entire database	101	33% (20/61)	30% (12/40)
Alcohols	16	89% (8/9)	25% (2/7)
Aldehyde	5	0% (0/4)	100% (1/1)
Amines	7	60% (3/5)	50% (1/2)
Ethers	14	50% (5/10)	50% (2/4)
Heterocyclic compound	7	86% (6/7)	- (0/0)
Organic salts	7	57% (4/7)	- (0/0)
_	Chemic	al Class-IS(B)-100	
Entire database	138	59% (58/99)	13% (5/39)
Alcohols	24	88% (14/16)	13% (1/8)
Aldehydes	6	80% (4/5)	0% (0/1)
Amines	9	83% (5/6)	33% (1/3)
Carboxylic acid salt	11	60% (3/5)	17% (1/6)
Esters	12	90% (9/10)	0% (0/2)
Ethers	16	50% (6/12)	25% (1/4)
Formulations	27	26% (6/23)	0% (0/4)
Heterocyclic compound	12	78% (7/9)	33% (1/3)
Inorganic salt	5	100% (2/2)	0% (0/3)
Ketones	6	67% (4/6)	- (0/0)
Organic salts	9	86% (6/7)	0% (0/2)
	Prope	erties of Interest	
IS(B)-10 Physical Form:			
Liquid/Soultion	35	19% (3/16)	37% (7/19)
Solid	27	58% (11/19)	13% (1/8)
Unknown	39	23% (6/26)	31% (4/13)
IS(B)-100 Physical Form:			
Liquid	60	65% (33/51)	0% (0/9)
Solid	41	67% (16/24)	24% (4/17)
Unknown	37	38% (9/24)	8% (1/13)
Abbreviations: GHS = Globally Ha	armonized S	ystem; HET-CAM = Hen's Egg ⁻	Test – Chorioallantoic

²False Positive Rate = the proportion of all negative substances that are falsely identified as positive *in* vitro; False Negative Rate = the proportion of all positive substances that are falsely identified as negative in vitro. The data used to calculate the percentage are provided in parenthesis.

Comparative Test Method Reliability

was evaluated based on the available information and databases for each of the test methods. Two types of interlaboratory reproducibility analyses were conducted:

- Qualitative analysis: Extent of agreement among testing laboratories for classification of substances
- Quantitative analysis: Evaluated using a coefficient of variation

A through description of the reliability analyses conducted for each of these test methods can be obtained at http://iccvam.niehs.nih.gov/ methods/ocudocs/. For comparative purposes here, Tables 7-10 summarize the results of the qualitative evaluation for each of the

The qualitative analysis for BCOP indicated that 65% to 94% of the substances were classified the same by the participating laboratories (**Table 7**). When only severe irritants (based on *in vivo* rabbit eye test results) were considered, the participating laboratories were in 100% agreement for 67% to 100% of the substances tested, and were in at least 80% agreement for at least 83% of the substances tested.

BCOP Qualitative Interlaboratory Reproducibility Evaluation – GHS Classification System

% Interlaboratory Agreement	Gautheron et al. (1994) (11 or 12 labs)	Balls et al. (1995) (5 labs)	Southee (1998) (3 labs)
100% (all substances)	65% (34/52)	68% (41/60)	94% (15/16)
≥80% (all substances)	87% (45/52)	85% (51/60)	94% (15/16)
100% (GHS Category 1 Substances)	67% (4/6)	76% (13/17)	100% (4/4)
≥80% (GHS Category 1 Substances)	83% (5/6)	94% (16/17)	100% (4/4)

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability assay; GHS = Globally Harmonized

The numbers in parenthesis indicate the data on which the % value is based.

For the ICE test method, the four participating laboratories were in 100% agreement in regard to the ocular irritancy classification of ~75% of the substances tested, and were in at least 75% agreement for 90% of the substances tested (**Table 8**). When only severe irritants (based on in vivo rabbit eye test results) were considered, the participating laboratories were in 100% agreement for ~70% of the substances tested, and were in at least 75% agreement for at least 95% of the substances tested.

Table 8. ICE Qualitative Interlaboratory Reproducibility **Evaluation – GHS Classification System**

% Interlaboratory Agreement	Balls et al. (1995) (59 substances)
100% (all substances)	75% (44/59)
≥75% (all substances)	90% (53/59)
100% (severe <i>in vivo</i> and <i>in vitro</i> substances)*	72% (16/22)
≥75% (severe <i>in vivo</i> and <i>in vitr</i> o substances)*	95% (21/22)
Abbreviations: GHS = Globally Harmonized	System; ICE = Isolated Chicken Eye

The numbers in parenthesis indicate the data on which the % value is based. *Scores for fluorescein retention and corneal swelling were not provided for one severe irritant/corrosive (30% trichloroacetic acid). Classification based on results

For the IRE test method, the four testing laboratories in the Balls et al (1995) study were in 100% agreement with respect to the in vivo/in vitro outcomes (severe/nonsevere) 59% (35/59) of the time (Table 9) and were in 75% agreement for 85% of the substances tested. For the CEC (1991) study, the three testing laboratories agreed 81% (17/21) of the time when the EU classification system was used and were in 67%

Table 9. IRE Qualitative Interlaboratory Reproducibility **Evaluation – GHS Classification System**

% Interlaboratory Agreement ¹	Balls et al. (1995) (59 substances) ²	
100%	59% (35/59)	
75%	85% (50/59)	
50%	100% (59/59)	
Abbreviations: GHS = Globally Harmonized System; IRE = Isolated Rabbit Eye		

The numbers in parenthesis indicate the data on which the % value is based. ¹Percent of agreement with all outcomes combined. ²Corneal opacity and corneal swelling were measured. Four laboratories

agreement for 95% (20/21) of the substances tested.

from only 3 laboratories.

used in evaluation.

82% (81/99) 6% (6/107)

Reproducibility Evaluation – GHS Classification

For the HET-CAM test method, the two to three testing laboratories in

GHS classification system) outcomes for 79% to 82% of the tested

Table 10. HET-CAM Qualitative Interlaboratory

substances, for both analysis methods (Table 10).

the Spielmann et al. (1996) study were in 100% agreement with respect

15% (16/107)

 Chorioallantoic Membrane assav. The numbers in parenthesis indicate the data on which the % value is based.

¹Percent of agreement with all outcomes combined.

References

Baglev D. et al. 1992. Toxic In Vitro 6:275-284. Bailey PT, et al. 2004. Poster presentation at the Society of Toxicology 2004 meeting.

- Balls M, et al. 1995. Toxicol In Vitro 9(6):871-929.
- Burton ABG, et al. 1981. Food Cosmet Toxicol 19:417-480.
- Casterton PL. et al. 1996. J Toxicol Cut and Ocular Toxicol 15(2):147-163. CEC. 1991. Doc. XI/632/91/V/E/1/131/91 Part I and II.
- Curren R, et al. 2000. Veterinary Pathology 37(5):557. Draize J, et al. 1944. J Pharmacol Exper Therapeut 82:377-390.
- EPA. 1996. EPA737-B-96-001. Washington, DC: U.S. Environmental Protection Agency.
- EU. 2001. Commission Directive 2001/59/EC of 6 August 2001. Official Journal of the
- Gautheron P. et al. 1994. Toxicol In Vitro 8(3):381-392.
- Gettings S, et al. 1991. In Vitro Toxic 4:247-288.
- Gettings S, et al. 1994. Food Chem Toxic 32:943-976. Gettings S, et al. 1996. Food Chem Toxic 34:79-117.
- Guerriero FJ, et al. 2004. [Abstract No. 1282]. Toxicol Sci (The Toxicologist Supplement)
- ICCVAM. 1997. NIH Publication No: 97-3981. Research Triangle Park, NC: National Toxicology
- 2003. NIH Publication No: 03-4508. Research Triangle Park, NC: National h Institute. INVITTOX 1992. Available: https://ecvam-sis.irc.it/invittox/published/indexed 47.html.
- Kalweit S, et al. Mol Toxic 1:597-603.
- Kalweit S. et al. Toxic In Vitro 4:702-706 Kojima H, et al. Toxic In Vitro 9:333-340.
- Luepke N and Kemper F. 1986. Food Chem Toxic 24:495-496. Luepke N. 1985, Food Chem Toxic 23:287-291.
- McDonald TO, et al. 1977. Eye irritation. In: Advances in Modern Toxicolog
- Dermatotoxicology. 1st ed. (Marzulli FN, Maibach HI. eds.). Washington: Hemisphere
- Prinsen MK. Koëter BWM. 1993. Food Chem Toxicol 31:69-76. Prinsen MK. 1996. Food Chem Toxicol 34:291-296.
- by M Prinsen. TNO Nutrition and Food Research Institute. Sina JF, et al. 1995, Fundam Appl Toxicol 26:20-31.

Southee JA. 1998. European Community contract no. 11279-95-10F 1ED ISP GB.

- Spielmann H, et al. 1996. ATLA 24:741-858. Swanson JE, et al. 1995. J Toxicol - Cut and Ocular Toxicol 14(3):179-195.
- Swanson JE, Harbell JW. 2000. The Toxicologist 54(1):188-189.
- UN. 2003. New York & Geneva: United Nations Publications. Available
- http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/00files_e.html. Vinardell M and Macián M. 1994. Toxic In Vitro 8:467-470.

Acknowledgments

We would like to thank the generous contributions of the companies and individuals who provided data for the review of these test methods

 Access Business Group (Luann Potts) Cosmetics, Toiletry, and Fragrance Association (Dr. Carol Eisenmann)

 ECVAM (Dr. Chantra Eskes) ExxonMobil Biomedical Sciences, Inc. (Dr. James Freeman)

GlaxoSmithKline (Mr. Frederick Guerriero)

 Johnson & Johnson (Drs. Philippe Vanparys and Freddy Van Goethem) Merck (Dr. Joseph Sina)

National Institute of Health Sciences (Japan) (Dr. Yasuo Ohno)

Proctor and Gamble (Dr. Daniel Marsman)

S.C. Johnson & Son, Inc. (Dr. Nicole Cuellar and Dr. Judith Swanson)

 SafePharm Laboratories (Mr. Robert Guest, Dr. Andrew Wittingham) TNO Nutrition and Food Research (Mr. Menk Prinsen)

U.S. Food and Drug Administration (Ms. Donnie Lowther)

 Unilever (Ms. Penny Jones) ZEBET (Dr. med Horst Spielmann and Dr. Manfred Liebsch)

This poster was supported by the Intramural Research Program of the NIH. National Institute of Environmental Health Sciences. ILS staff supported by NIEHS contract N01-ES 35504. The views expressed

above do not necessarily represent the official positions of any federal