ICCVAM Recommendations on the Use of Four *In Vitro* Test Methods for the Identification and Classification of Ocular Corrosives and Severe Irritants

K Hamernik[1], J Merrill[2], A Jacobs[2], D Lowther[3], M Wind[4], A Rispin[1], R Tice[5], W Stokes[5]

[1]U.S. Environmental Protection Agency, Washington, DC; [2]U.S. Food and Drug Administration, Silver Spring, MD, USA; [3]U.S. Food and Drug Administration, College Park, MD, USA; [4]U.S. Consumer Products Safety Commission, Bethesda, MD, USA; [5]National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences/National Institutes of Health/Department of Health and Human Services, Research Triangle Park, NC, USA.

Introduction

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is charged by the ICCVAM Authorization Act of 2000[1] with evaluating the scientific validity of new, revised, and alternative toxicological test methods with potential applicability to U.S. Federal agency safety testing. ICCVAM is also required to provide recommendations to U.S. Federal agencies regarding the usefulness and limitations of such test methods. The ICCVAM test method evaluation report (TMER), *In Vitro Ocular Toxicity Test Methods for Identifying Severe Irritants and Corrosives* provides ICCVAM recommendations for using four *in vitro* test methods to identify severe ocular irritants and corrosives in a tiered-testing strategy.

These recommendations are based on a comprehensive evaluation of the scientific validation status of the test methods by ICCVAM, and take into consideration the comments and recommendations received from an independent expert peer review panel, ICCVAM's Scientific Advisory Committee on Alternative Toxicological Methods (SACATM), and the general public.

The TMER contains ICCVAM recommendations for:

- Test method uses
- Standardized test method protocols
- Future studies
- Proposed reference substances

[1]42 U.S.C. § 2851-2, 2851-5 (2000) http://iccvam.niehs.nih.gov/about/PL106545.pdf

In Vitro Test Method Performance

A complete description of all databases and the resulting accuracy and reliability analyses conducted for each of these test methods can be obtained at http://iccvam.niehs.nih.gov/methods/ocudocs/.

Test Method Accuracy

Accuracy of the four *in vitro* test methods when compared to *in vivo* rabbit eye test classifications using the United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS; UN 2003) classification system are provided in **Table 1**. Similar results were obtained for the EPA (1996) and European Union (2001) classification systems.

Table 1. 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)
Accuracy	65%[4] (70/107)	83% (120/144)	68% (69/101)	54% (75/138)	81% (119/147)
Sensitivity	70% (33/47)	50% (15/30)	70% (28/40)	87% (34/39)	84% (36/43)
Specificity	62% (37/60)	92% (105/114)	67% (41/61)	41% (41/99)	80% (83/104)
False Positive Rate	38% (23/60)	8% (9/114)	33% (20/61)	59% (58/99)	20% (21/104)
False Negative Rate	30% (14/47)	50% (15/30)	30% (12/40)	13% (5/39)	16% (7/43)

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability test method; GHS = Globally Harmonized System; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane test method; ICE = Isolated Chicken Eye test method; IRE = Isolated Rabbit Eye test method.
[1]N = number of substances tested; the numbers in parentheses in each row indicates the data on which the percentage calculation is based.
[2]These data are for the IS(B) method (described by Kalweit et al. 1987) when testing substances as a 10% solution *in vitro*.
[3]These data are for the IS(B) method (described by Kalweit et al. 1987) when testing substances at a 100% concentration *in vitro*.
[4]These results are for the Pooled Data Set (see http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_ire.htm for additional information).

Tables 2 to **7** 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.

BCOP TEST METHOD

For the BCOP test method, the highest overpredicted classes are alcohols and ketones, while the highest underpredicted class is solids (**Table 2**).

Table 2. False Negative and False Positive Rates of the BCOP 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	147	20% (21/104)	16% (7/43)
Chemical Class[3]			
Alcohols	18	53% (8/15)	67% (2/3)
Amine/Amidine	8	0% (0/4)	0% (0/4)
Carboxylic acids	15	38% (3/8)	14% (1/7)
Esters	12	12% (1/8)	0% (0/4)
Ether/Polyether	6	0% (0/5)	0% (0/1)
Heterocyclic compounds	12	33% (2/6)	17% (1/6)
Hydrocarbons	12	8% (1/12)	- (0/0)
Inorganic Salt	5	0% (0/3)	0% (0/2)
Ketones	10	40% (4/10)	- (0/0)
Onium compounds	11	0 % (0/3)	0% (0/8)
Properties of Interest			
Liquids	92	26% (18/68)	4% (1/24)
Solids	32	10% (2/20)	42% (5/12)
Pesticide	8	33% (1/3)	40% (2/5)
Surfactants[4]	35	5% (1/21)	7% (1/14)

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability test method; GHS = Globally Harmonized System.
[1]N = number of substances tested.
[2]False 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.
[3]Chemical classes included in this table are represented by at least five substances tested by the method and assignments are made based on the Medical Subject Heading categories (http://www.nlm.nih.gov/mesh/).
[4]Combines single chemicals labeled as surfactants along with surfactant-containing formulations.

BCOP test method performance statistics also were evaluated when discordant chemical classes were excluded (i.e., alcohols, ketones, solids). When all three classes were excluded from the data set, accuracy increased to 92% (78/85), the false positive and false negative rates decreased to 12% (7/58) and 0% (0/27), respectively (**Table 3**).

Table 3. Effect of Exclusion of Discordant Classes (Alcohols, Ketones, and Solids) on False Negative and False Positive Rates of the BCOP Test Method, for the GHS Classification System

Data Set	Accuracy		False Positive Rate[1]		False Negative Rate[1]	
	%	No.[2]	%	No.	%	No.
Overall	81	119/147	20	21/104	16	7/43
Excluding Alcohols	86	109/126	14	12/86	13	5/40
Excluding Ketones	81	113/138	19	18/95	16	7/43
Excluding Solids	82	93/113	23	19/84	4	1/29
Excluding Alcohols, Ketones, and Solids	92	78/85	12	7/58	0	0/27

Abbreviation: GHS = Globally Harmonized System.
[1]False 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*.
[2]Data used to calculate the percentage.

References

Draize et al. 1944. J Pharmacol Exper Therapeut 82:377-390.
EPA. 1996. EPA737-B-96-001.
EU. 2001. Official Journal of the European Communities L255:1-333.
ICCVAM. 2003. NIH Publication No: 03-4508. http://iccvam.niehs.nih.gov/.
ICCVAM. 1997. NIH Publication No.: 97-3981. http://iccvam.niehs.nih.gov/.
Kalweit et al. 1987. Mol Toxic 1:597-603.
Luepke N. 1985. Food Chem Toxic 23:287-291.
National Research Council. 1996. Guide for the Care and Use of Laboratory Animals.
PHS. 2002. Public Health Service (PHS) Policy on Humane Care and Use of Laboratory Animals.
Sina et al. 1995. Fundam Appl Toxicol 26:20-31.
Ubels et al. 2002. Toxicol In Vitro 16:621-628.
UN. 2003. http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/00files_e.html.

Acknowledgments

This poster was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. ILS staff were supported by NIEHS contract N01-ES 35504. The views expressed above do not necessarily represent the official positions of any federal agency.

In Vitro Test Method Performance

ICE TEST METHOD

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[3]			
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 test method.
[1]N = number of substances tested.
[2]False 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.
[3]Chemical classes included in this table are represented by at least five substances tested by the method and assignments are made based on the Medical Subject Heading categories (http://www.nlm.nih.gov/mesh/).

ICE test method performance statistics also were evaluated when discordant chemical classes were excluded (i.e., alcohols, surfactants, solids). When all three classes were excluded from the data set, accuracy increased to 92% (69/75), the false negative and false positive rates decreased to 29% (2/7) and 6% (4/68), respectively (**Table 5**).

Table 5. Effect of Exclusion of Discordant Classes on False Negative and False Positive Rates of the ICE Test Method, for the GHS Classification System

Data Set	Accuracy		False Positive Rate[1]		False Negative Rate[1]	
	%	No.[2]	%	No.	%	No.
Overall	83	120/144	8	9/114	50	15/30
Excluding Alcohols	86	114/132	4	4/104	50	14/28
Excluding Surfactants	85	104/123	9	9/102	48	8/18
Excluding Solids	84	91/108	10	9/90	44	8/18
Excluding Alcohols, Surfactants, and Solids	92	69/75	6	4/68	29	2/7

Abbreviation: GHS = Globally Harmonized System.
[1]False 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*.
[2]Data used to calculate the percentage.

IRE TEST METHOD

For the IRE test method, alcohols, amines, ketones, and liquids were the most overpredicted chemical classes, while carboxylic acids and organic compounds were the most underpredicted chemical classes (**Table 6**). Due to the high false positive and false negative rates, additional chemical class assessments were not conducted.

Table 6. 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 Data Set)

Category	N[1]	False Positive Rate[2]	False Negative Rate[2]
Overall	107	38% (23/60)	30% (14/47)
Chemical Class[3]			
Alcohol	13	55% (6/11)	50% (1/2)
Amide	5	0% (0/3)	0% (0/2)
Amine	11	50% (3/6)	20% (1/5)
Carboxylic acid	12	33% (2/6)	67% (4/6)
Ester	10	30% (3/10)	- (0/0)
Ether	9	33% (2/6)	0% (0/3)
Formulation	24	25% (2/8)	38% (6/16)
Heterocyclic compound	18	44% (4/9)	11% (1/9)
Ketone	6	67% (4/6)	- (0/0)
Onium compound	10	33% (1/3)	0% (0/7)
Organic compound	12	17% (1/6)	50% (3/6)
Sulfur compound	8	20% (1/5)	33% (1/3)
Properties of Interest			
Liquid/Solution	65	49% (18/37)	29% (8/28)
Solids	42	22% (5/23)	32% (6/19)
Surfactant-based formulation	24	25% (2/8)	38% (6/16)
Surfactant	13	40% (2/5)	12% (1/8)

Abbreviations: GHS = Globally Harmonized System; IRE = Isolated Rabbit Eye test method.
[1]N = number of substances tested.
[2]False 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.
[3]Chemical classes included in this table are represented by at least five substances tested by the method and assignments are made based on the Medical Subject Heading categories (http://www.nlm.nih.gov/mesh/).

HET-CAM TEST METHOD

Alcohols, heterocyclic compounds, and organic salts were the most overpredicted chemical classes by the IS(B)-10 and IS(B)-100 analysis methods, which are based on the method of Kalweit et al. (1987) where substances were tested at 10% and 100% concentration, respectively. Ethers also were overpredicted by the IS(B)-10 method, while aldehydes, amines and esters were overpredicted by the IS(B)-100 analysis method (**Table 7**). Due to the high false positive and false negative rates for the evaluated analysis methods, additional chemical class assessments were not conducted.

Table 7. 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]
Chemical Class-IS(B)-10[3]			
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)
Formulation	24	0% (0/8)	44% (7/16)
Heterocyclic compound	7	86% (6/7)	- (0/0)
Organic salts	7	57% (4/7)	- (0/0)
Chemical Class-IS(B)-100[3]			
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/ 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)
Properties of Interest			
IS(B)-10 Physical Form:			
Liquid/Solution	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 Harmonized System; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane test method.
[1]N = number of substances tested.
[2]False 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.
[3]Chemical classes included in this table are represented by at least five substances tested by the method and

Test Method Reliability

BCOP TEST METHOD

- Intralaboratory repeatability evaluated for two studies; coefficient of variation (CV) values ranged from 12% to 35%
- Intralaboratory reproducibility evaluated for two studies; CV values ranged from 13% to 33%
- Interlaboratory reproducibility evaluated quantitatively and qualitatively
  - Qualitative: ≥ 67% of the substances were classified the same by the participating laboratories
  - Quantitative: mean and median CV values ≤ 36% and ≤ 23%, respectively

ICE TEST METHOD

- Intralaboratory repeatability CV values ranged from 0.9% to 6.1% for corneal thickness endpoint; all other endpoints produced larger CV ranges
- Intralaboratory reproducibility CV values ranged from 1.8% to 6.3% for corneal thickness endpoint; all other endpoints produced larger CV ranges
  - Exclusion of nonirritating substance reduced the CV ranges for other endpoints evaluated
- Interlaboratory reproducibility evaluated quantitatively and qualitatively
  - Qualitative: ≥ 60% of the substances were classified the same by the participating laboratories
  - Quantitative: mean and median endpoint CV values ≤ 35% (except for corneal swelling)

IRE TEST METHOD

- Intralaboratory repeatability and reproducibility were not evaluated
- Interlaboratory reproducibility evaluated quantitatively and qualitatively
  - Qualitative: 100% of the substances were classified the same by the participating laboratories
  - Quantitative: mean and median endpoint CV values ≤ 50%

HET-CAM TEST METHOD

- Intralaboratory repeatability and reproducibility studies indicated the highest CV values were for the hemorrhage endpoint
- Interlaboratory reproducibility, for both analysis methods, were evaluated quantitatively and qualitatively
  - Qualitative: Approximately 80% of the substances were classified the same by the participating laboratories for either analysis method
  - Quantitative: IS(B)-10 mean and median CV values ≤ 66% and ≤ 61%, IS(B)-100 mean and median CV values ≤ 35% and ≤ 33%

ICCVAM Test Method Recommendations

Current uses

- **None of the four *in vitro* test methods evaluated can be considered to be complete replacements for the *in vivo* rabbit eye test.** However, based on available data, BCOP and ICE can be used, in appropriate circumstances and with certain limitations, as screening tests for the detection of ocular corrosives and severe irritants in a tiered-testing[4] strategy, as part of a weight-of-evidence approach.
- Although IRE and HET-CAM test methods cannot currently be recommended for meeting regulatory requirements, there may be non-regulatory uses for these two test methods. Therefore, all four *in vitro* test methods should be considered prior to conducting *in vivo* ocular testing and used where determined appropriate for the specific testing situation.
- Users should be aware that the performance characteristics for each of these test methods could be revised as additional data become available. Consult the ICCVAM/NICEATM website (http://iccvam.niehs.nih.gov) to review the most current validation database.
- All raw data generated using any of the recommended standardized *in vitro* ocular testing protocols and the *in vivo* rabbit eye test on the same substance should be submitted to NICEATM to expand the available validation database for these four test methods.

Test Method Protocols

- The recommended test method protocols are provided in the ICCVAM TMER.
  - Exceptions and/or changes to the proposed standardized test method protocol should be accompanied by a scientific rationale.
  - Users should be aware that the test method protocols could be revised based on future optimization and/or validation studies. Therefore, test method users should consult the ICCVAM/NICEATM website to obtain the most current recommended protocol.

Future Studies

- Interested stakeholders are encouraged to support research and development of alternative test methods and technologies that may provide for a more accurate assessment of ocular toxicity and/or advantage in terms of time and cost.
- Additional research and development, optimization, and/or validation efforts should use reference substances with existing rabbit data. Additional *in vivo* studies should be conducted only if important data gaps are identified, and such studies should be designed to minimize the number of rabbits tested, to minimize or avoid pain and distress, and to maximize the information collected.
- The potential usefulness of a battery approach that combines multiple *in vitro* test methods to identify ocular corrosives and severe irritants should be evaluated.
- For BCOP, ICE, and IRE, a histopathological evaluation using a standardized scoring scheme should be conducted. Such data will allow for the development of standardized decision criteria and a more comprehensive evaluation of the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.
- For BCOP, studies should be conducted to evaluate the impact of using a corneal holder that maintains normal corneal curvature (e.g., the corneal mounting system designed by Ubels et al. 2002) on accuracy and/or reliability of the BCOP test method.
  - An evaluation should also be conducted on the effect of modifying various test method protocol components (e.g., duration of test substance exposure) on the accuracy and/or reliability of the BCOP test method.
- For ICE and IRE, centering lights should be installed on the optical pachymeter, which is used to measure corneal thickness, to ensure consistent central corneal thickness measurements across laboratories.
- For ICE, additional optimization studies/evaluations should be conducted in an attempt to decrease the 29% to 50% false negative rate of the ICE test method. After optimization, additional studies to further assess the reliability and accuracy of the test method are recommended.
- For IRE, the decision criteria should be optimized. Once optimized, additional validation studies should be conducted to further evaluate the relevance and reliability of the IRE test method.
- For HET-CAM, additional studies should be conducted to further optimize the prediction models and the decision criteria that would be used to identify ocular corrosives and severe irritants.

Substances For Validation Of *In Vitro* Ocular Toxicity Test Method For The Evaluation Of Ocular Corrosives And Severe Irritants

- ICCVAM developed a list of 122 reference substances for the optimization and/or validation of *in vitro* tests to identify ocular corrosives and severe irritants (see Appendix H of the ICCVAM Test Method Evaluation Report).
- The list includes:
  - 79 GHS Category 1 substances (UN 2003)
  - 28 GHS Category 2 substances (UN 2003)
  - 15 GHS nonirritant substances (UN 2003)
  - 34 chemical classes
  - 79 liquids and 43 solids
- ICCVAM further endorses the use of the reference substance list as a source for generating a subset of substances to be used for evaluating *in vitro* ocular toxicity test methods on a case-by-case basis. The subset of substances that are developed from the reference substance list should comprise a scientifically sound distribution of substances among various properties.
- In situations where a listed substance is not available, other substances of the same class (e.g., chemical) for which there is high quality *in vivo* reference data may be substituted.
- Following completion of optimization and/or validation studies, substances from this list can be selected for inclusion in performance standards and proficiency testing (ICCVAM 2003).

[4]A tiered-testing strategy for ocular irritation/corrosion (e.g., as described in the Globally Harmonized System of Classification and Labelling of Chemicals; UN 2003) allows for the use of validated and accepted *in vitro* methods prior to the use of animals for ocular safety testing. In a tiered-testing strategy, when a positive result is obtained in an appropriately validated *in vitro* test, a test substance may be classified as an ocular hazard without testing in rabbits. A substance that tests negative in the *in vitro* ocular toxicity test would need to be tested in the *in vivo* ocular test to identify possible *in vitro* false negatives and to identify moderate and mild ocular irritants. As is appropriate for any test system, there is the opportunity for confirmatory testing if false positive results are suggested based on a weight-of-evidence evaluation of supplemental information (e.g., pH, structure-activity relationships, other testing data). Using *in vitro* data in a tiered-testing strategy with a weight-of-evidence decision process to classify substances as ocular corrosives or severe irritants will avoid the potential pain and distress that might be experienced by rabbits who otherwise would have been administered these test substances. A tiered-testing strategy may not be applicable to purposes other than regulatory classification and labeling.

NIEHS National Institute of Environmental Health Sciences
ICCVAM The Interagency Coordinating Committee on the Validation of Alternative Methods
NICEATM The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

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