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

LM Schechtman[1]*, M Wind[2], JC Merrill[1], KL Hamernik[3], RR Tice[4], WS Stokes[4].

[1]U.S. FDA, Washington, DC, [2]U.S. CPSC, Bethesda, MD, [3]U.S. EPA, Washington, DC, [4]National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)/NIEHS/NIH/DHHS, Research Triangle Park, NC.

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 the ICCVAM's 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 Report 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.

Interagency Coordinating Committee on the Validation Of Alternative Methods (ICCVAM)

Designated Agency Representatives

Agency for Toxic Substances and Disease Registry
• Moiz Mumtaz, Ph.D.

Consumer Product Safety Commission
• Marilyn L. Wind, Ph.D. (Chair)◆
* Patricia Bittner, M.S.
* Kristina Hatlelid, Ph.D.

Department of Agriculture
• Jodie Kulpa-Eddy, D.V.M. (Vice-Chair)◆
◊ Elizabeth Goldentyer, D.V.M.

Department of Defense
• Robert E. Foster, Ph.D.
◊ Patty Decot
Harry Salem, Ph.D.

Department of Energy
• Marvin Stodolsky, Ph.D.

Department of the Interior
• Barnett A. Rattner, Ph.D.
◊ Sarah Gerould, Ph.D.

Department of Transportation
• George Cushmac, Ph.D.
◊ Steve Hwang, Ph.D.

Environmental Protection Agency

Office of Science Coordination and Policy
• Karen Hamernik, Ph.D.

Office of Research and Development
◊ Julian Preston, Ph.D.
* Suzanne McMaster, Ph.D.

OECD Test Guidelines Program
* Jerry Smrchek, Ph.D.

Office of Pesticides Programs
* Amy Rispin, Ph.D.
* Deborah McCall

◆ effective January 2007
Leonard Schechtman, Ph.D. [FDA] (former ICCVAM chair, retired, December 2006)

• Principal Agency Representative
◊ Alternate Principal Agency Representative
* Other Designated Agency Representatives

Food and Drug Administration

Office of Science and Health Coordination
• Suzanne Fitzpatrick, Ph.D., D.A.B.T.

Center for Drug Evaluation and Research
◊ Abigail C. Jacobs, Ph.D.

Center for Devices and Radiological Health
• vacant
* Melvin E. Stratmeyer, Ph.D.

Center for Biologics Evaluation and Research
* Richard McFarland, Ph.D., M.D.
* Ying Huang, Ph.D.

Center for Food Safety and Nutrition
* David G. Hattan, Ph.D.
* Robert L. Bronaugh, Ph.D.

Center for Veterinary Medicine
* Devaraya Jagannath, Ph.D.
* M. Cecilia Aguila, D.V.M.

National Center for Toxicological Research
* William T. Allaben, Ph.D.

Office of Regulatory Affairs
* Lawrence A. D'Hoostelaere, Ph.D.

National Cancer Institute
• Alan Poland, M.D.
◊ T. Kevin Howcroft, Ph.D.

National Institute of Environmental Health Sciences
• William S. Stokes, D.V.M., D.A.C.L.A.M.
◊ John R. Bucher, Ph.D., D.A.B.T.
* Rajendra S. Chhabra, Ph.D., D.A.B.T
* Jerrold J. Heindel, Ph.D.

National Institute for Occupational Safety and Health
• Paul Nicolaysen, V.M.D.
◊ K. Murali Rao, M.D., Ph.D.

National Institutes of Health
• Margaret D. Snyder, Ph.D.

National Library of Medicine
• Vera Hudson, M.S.
◊ Jeanne Goshorn, M.S.

Occupational Safety and Health Administration
• Surender Ahir, Ph.D.

ICCVAM Ocular Toxicity Working Group (OTWG)

Consumer Product Safety Commission
• Kailash Gupta, Ph.D. (Retired 2006)
• Marilyn Wind, Ph.D.
• Cassandra Prioleau, Ph.D.

Department of Defense
• Harry Salem, Ph.D.

Department of Transportation
• Steve Hwang, Ph.D.

Environmental Protection Agency

Office of Science Coordination and Policy
• Karen Hamernik, Ph.D. (Co-Chair)

Office of Research and Development
• Andrew Geller, Ph.D.

Office of Pesticides Programs
• Amy Rispin, Ph.D.
• Meta Bonner, Ph.D.
• Deborah McCall
• Mark Perry, Ph.D.
• John Redden, Ph.D.

Food and Drug Administration

Center for Drug Evaluation and Research
• Abigail C. Jacobs, Ph.D.
• Paul Brown, Ph.D.
• Wiley Chambers, M.D.
• Jill Merrill, Ph.D., D.A.B.T. (Co-Chair)

Center for Food Safety and Nutrition
• Robert L. Bronaugh, Ph.D.

National Center for Toxicological Research
• Leonard Schechtman, Ph.D. (Retired 2006)

National Institute of Environmental Health Sciences
• William S. Stokes, D.V.M., D.A.C.L.A.M.
• Raymond Grissom, Ph.D.

Occupational Safety and Health Administration
• Surender Ahir, Ph.D.

In Vitro Ocular Test Method Expert Panel Roster

Robert Scala, Ph.D. (Panel Chair),
Consultant, Tucson, Arizona

Sally Atherton, Ph.D.,
Professor, Medical College of Georgia, Augusta, Georgia

Roger Beuerman, Ph.D.,
Professor, Louisiana State University, New Orleans, Louisiana

June Bradlaw, Ph.D.,
International Foundation for Ethical Research, Rockville, Maryland

Ih Chu, Ph.D.,
Health Canada, Ottawa, Canada

Henry Edelhauser, Ph.D.,
Professor, Emory University, Atlanta, Georgia

Nancy Flournoy, Ph.D.,
Professor, University of Missouri, Columbia, Missouri

Donald Fox, Ph.D.,
Professor, University of Houston, Houston, Texas

James Freeman, Ph.D.,
Section Head, ExxonMobil Biomedical Sciences, Inc., Annandale, New Jersey

Shayne Gad, Ph.D., D.A.B.T., A.T.S.,
Consultant, Gad Consulting Services, Cary, North Carolina

Sidney Green, Ph.D., A.T.S.,
Graduate Professor, Howard University, Washington, DC

Frederick Guerriero, M.S.,
Senior Occupational Toxicologist, GlaxoSmithKline, King of Prussia, Pennsylvania

A. Wallace Hayes, Ph.D., D.A.B.T., F.A.T.S., F.I.Biol., F.A.C.F.E., E.R.T.,
Scientist, Harvard School of Public Health, Andover, Massachusetts

Hiroshi Itagaki, Ph.D.,
Principal Scientist, Shiseido Co., Ltd., Japan

David Lovell, Ph.D.,
Reader in Medical Statistics, University of Surrey, United Kingdom

Yasuo Ohno, Ph.D., D.J.S.T.S.,
Director of Japanese Society of Alternatives to Animal Experiments and Director of Division of Pharmacology, National Institute of Health Sciences, Japan

Robert Peiffer, D.V.M., D.A.C.V.O.,
Senior Investigator, Merck Research Laboratories, West Point, Pennsylvania

Lionel Rubin, V.M.D., D.A.C.V.O.,
Emeritus Professor of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania

Horst Spielmann, Dr. Med.,
Director and Professor, ZEBET at the BfR, Germany

Martin Stephens, Ph.D.,
Vice President for Animal Research, Humane Society of the United States, Washington DC

Katherine Stitzel, D.V.M.,
Consultant, West Chester, Ohio

Peter Theran, V.M.D., D.A.C.V.I.M.,
Vice President Animal Science, Massachusetts Society for the Prevention of Cruelty to Animals, Novato, California

Scheffer Tseng, M.D., Ph.D.,
Director, Ocular Surface Research and Education Foundation, Miami, Florida

Philippe Vanparys, Ph.D.,
Senior Research Fellow, Johnson and Johnson, Belgium

Timeline for Development of the ICCVAM Test Method Evaluation Report (TMER)

2003 EPA nominated four *in vitro* methods proposed for identifying potential ocular corrosives and severe irritants
- Bovine Corneal Opacity and Permeability (BCOP) assay
- Hen's Egg Test Chorioallantoic Membrane (HET-CAM) assay
- Isolated Chicken Eye (ICE) assay
- Isolated Rabbit Eye (IRE) assay

2004 NICEATM and ICCVAM prepared draft Comprehensive Background Review Documents (BRDs)
- Each BRD described the information available to assess the current validation status of each of the nominated methods

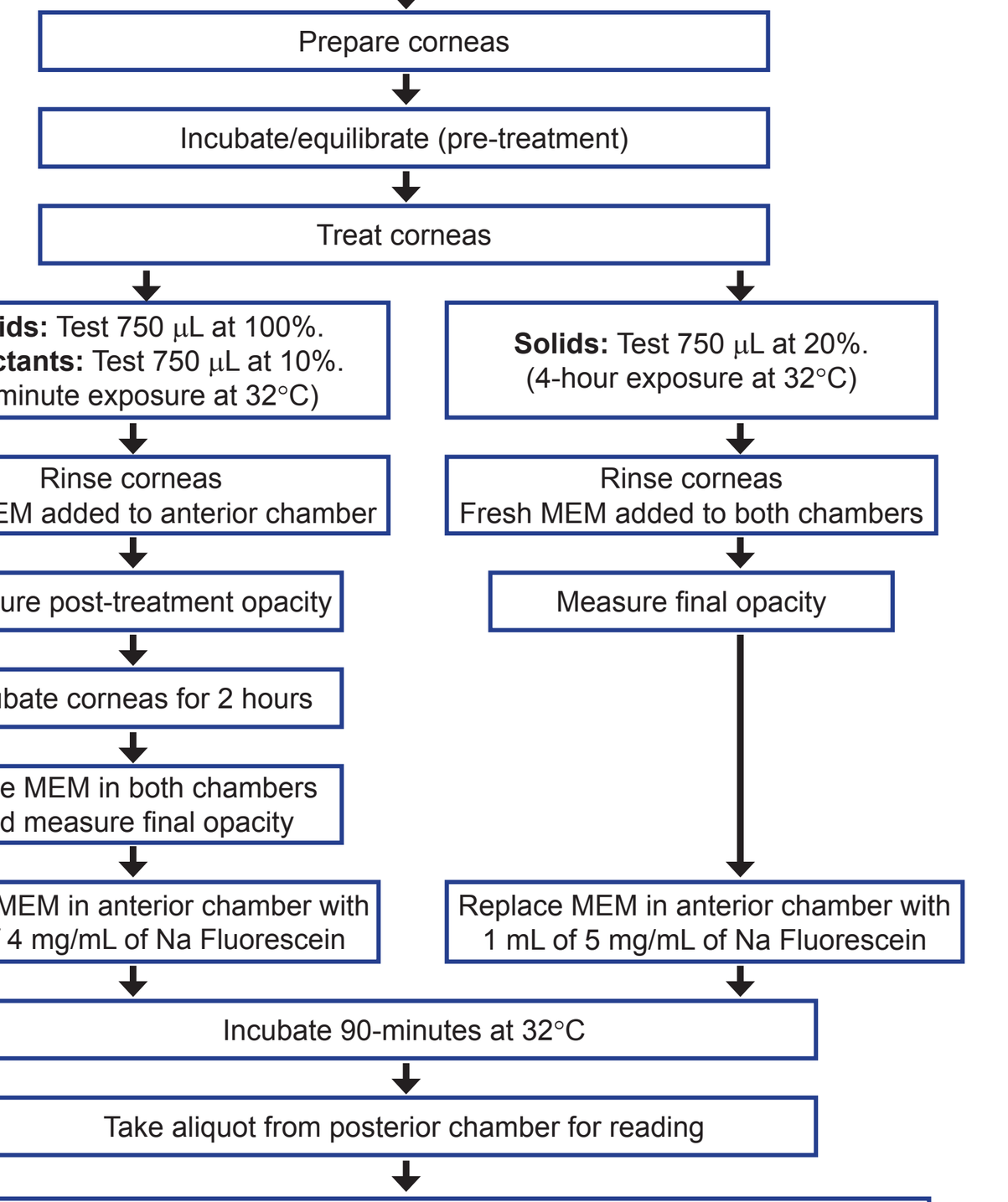
2005 Independent Expert Peer Review Panel was convened to assess the validation status of the four methods
- Expert Panel Report published in March
- Expert Panel Report Addendum published in November

2006 Expert Panel Reports, SACATM comments, and public comments reviewed by ICCVAM
- ICCVAM published final BRDs on each of the four *in vitro* ocular toxicity methods in March
- ICCVAM published a Test Method Evaluation Report on the four *in vitro* ocular toxicity methods in November

In Vitro Test Method Overviews, Databases and Analysis Results

A complete description of all the databases used for each test method can be obtained from each test method BRD (See http://iccvam.niehs.nih.gov/methods/ocudocs/).
Tables 1 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).

Bovine Corneal Opacity and Permeability (BCOP) Test Method Protocol



BCOP Database

A total of 158 substances were available to evaluate the BCOP test method accuracy.[1] The BCOP test method protocols used in these studies were similar to each other, but not identical (e.g., number of corneas used [n=3-5], storage conditions of bovine eyes during transport, different negative controls).

In Vitro Test Method Accuracy

For the BCOP test method, the highest overpredicted classes are alcohols and ketones, while the highest underpredicted class is solids (**Table 1**).

Table 1. 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 assay; GHS = Globally Harmonized System.
[1]N = number of substances, the numbers in parentheses in each row indicates the data on which the percentage calculation is based.
[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 substances from the classes that gave the most discordant results were excluded (i.e., alcohols, ketones, solids). When all three discordant 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 2**).

Table 2. 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.

In Vitro Test Method Reliability

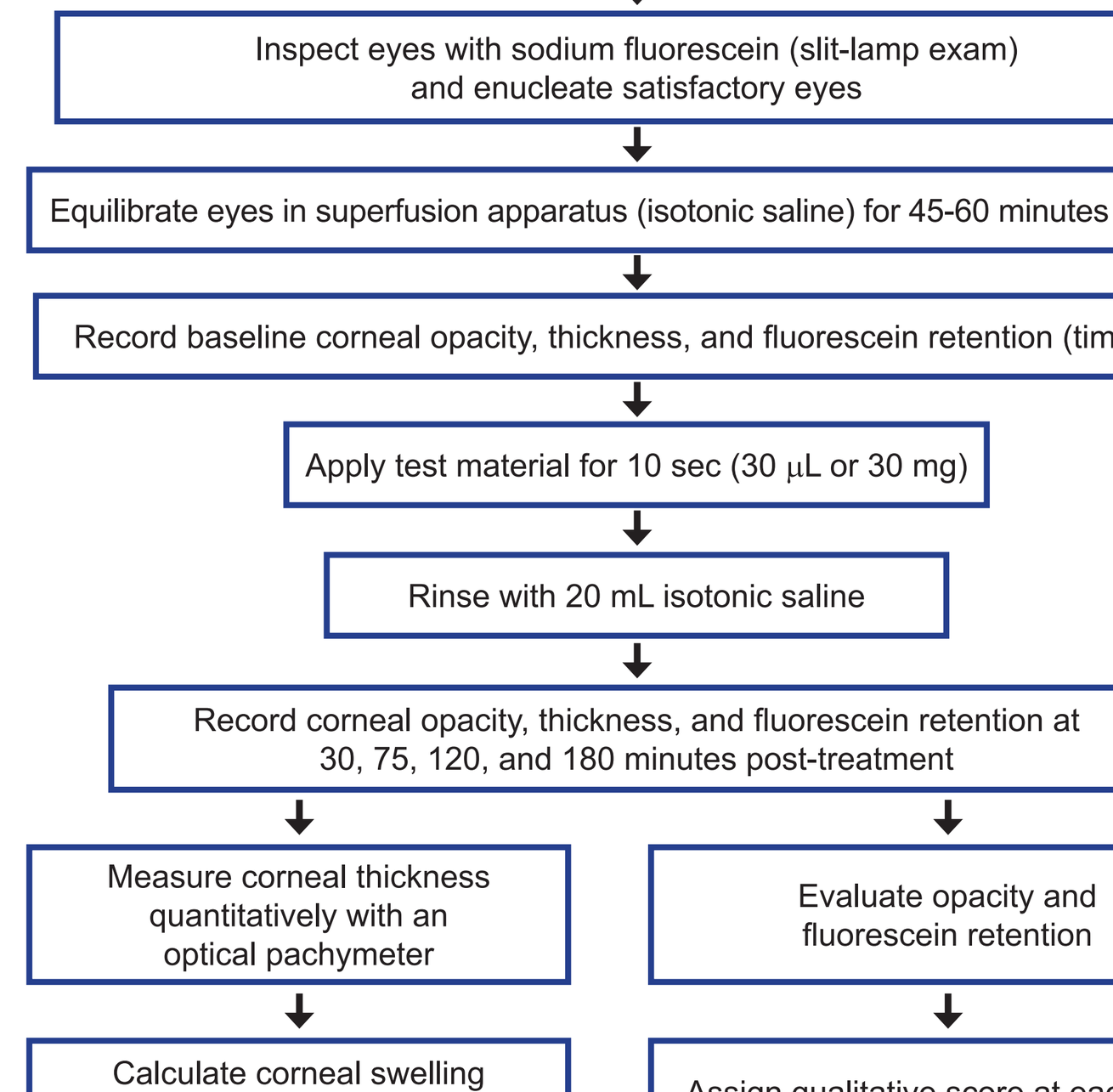
Two studies assessed intralaboratory repeatability. The median coefficient of variation (CV) values for *In Vitro* Irritation Score (IVIS; Opacity + (15*OD_{490})) for replicate corneas ranged from 11.8% to 14.2% (evaluated in three laboratories) for the first study and was 35% for the second study.

Intralaboratory reproducibility evaluations indicated mean and median CV values for permeability values were 33.4% and 29.0%, respectively, for 25 surfactant-based personal care cleaning formulations in one study. Mean CV values of IVIS for 16 substances tested two or more times in three laboratories ranged from 12.6% to 14.8%, while the median CV values ranged from 6.7% to 12.4%.

In a qualitative assessment of interlaboratory reproducibility of hazard classification category, 67% to 94% of the substances were classified the same by the participating laboratories. A quantitative evaluation of interlaboratory reproducibility was conducted for three studies by performing a CV analysis of IVIS obtained for substances tested in multiple laboratories. In these studies, the mean and median CV values were (a) 36% and 17%, respectively, for results obtained in either 11 or 12 laboratories, (b) 25% and 22%, respectively, for results obtained in five laboratories, and (c) 32.4% and 22.8%, respectively, for results obtained in three laboratories.

[1]The total number of substances used in the accuracy and reliability evaluations for each of the methods evaluated may differ from the total number of substances in the database since appropriate in vivo data for each substance, for each classification system, may not have been available.

Isolated Chicken Eye (ICE) Test Method Protocol



ICE Test Method Database

A total of 154 substances from five different studies were available to evaluate the accuracy of the ICE test method. The primary difference among various ICE studies was the number of treated eyes per test substance (3 to 5).

In Vitro Test Method Accuracy

For the ICE test method, alcohols tend to be overpredicted, while surfactants and solids tend to be underpredicted (**Table 3**).

Table 3. 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 assay.
[1]N = number of substances, the numbers in parentheses in each row indicates the data on which the percentage calculation is based.
[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 substances from the classes that gave the most discordant results were excluded (i.e., alcohols, surfactants, solids). When all three discordant 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 4**).

Table 4. 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.

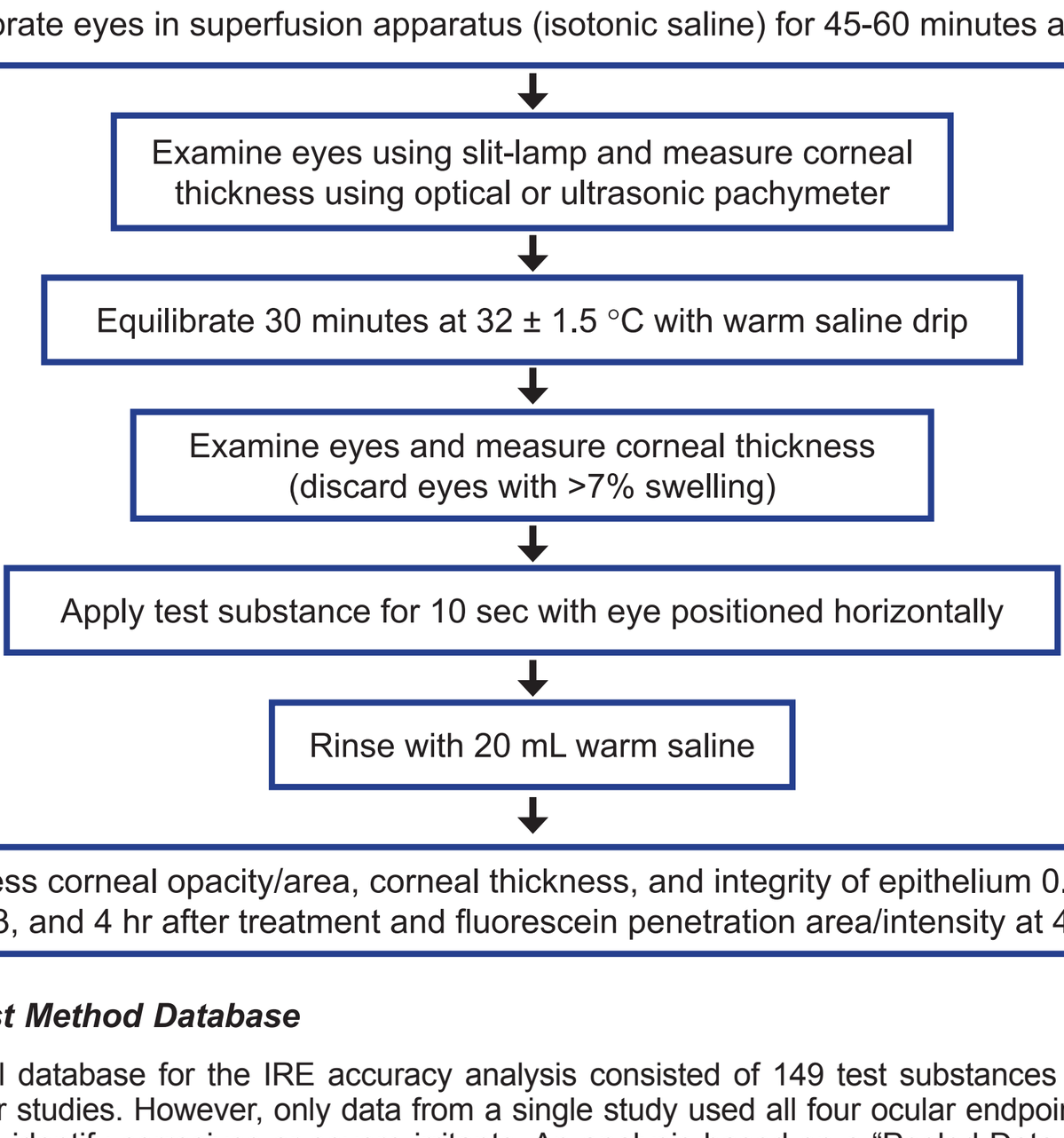
In Vitro Test Method Reliability

The range of CV values for the corneal thickness measurement, when results were compared within experiments (intralaboratory repeatability), was from 0.9% to 6.1%. The other endpoints evaluated produced ranges of CV values that were larger, with variability most prominent with the nonirritating substance tested.

The range of CV values for the corneal thickness measurement, when results were compared across experiments (intralaboratory reproducibility), was from 1.8% to 6.3%. The CV values for the remaining endpoints had a larger range (e.g., corneal swelling CV = 13.9% to 138.7%). However, if the nonirritating substance was removed, the range of CV values was reduced (e.g., corneal swelling CV = 13.9% to 22.4%).

One interlaboratory comparative study involving four laboratories contained test data on 59 substances for a qualitative and quantitative assessment of interlaboratory reproducibility. Based on a qualitative analysis, 60% to 70% of the substances classified as ocular corrosives or severe irritants, depending on the regulatory classification system employed (i.e., EPA, EU, GHS), were correctly identified by all four participating laboratories. A CV analysis of these same data indicated that the mean and median CV for severe substances tested was less than 35% for all test method endpoints, with the exception of corneal swelling.

Isolated Rabbit Eye (IRE) Test Method Protocol



IRE Test Method Database

The total database for the IRE accuracy analysis consisted of 149 test substances obtained from four studies. However, only data from a single study used all four ocular endpoints noted above 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.

In Vitro Test Method Accuracy

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 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 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)

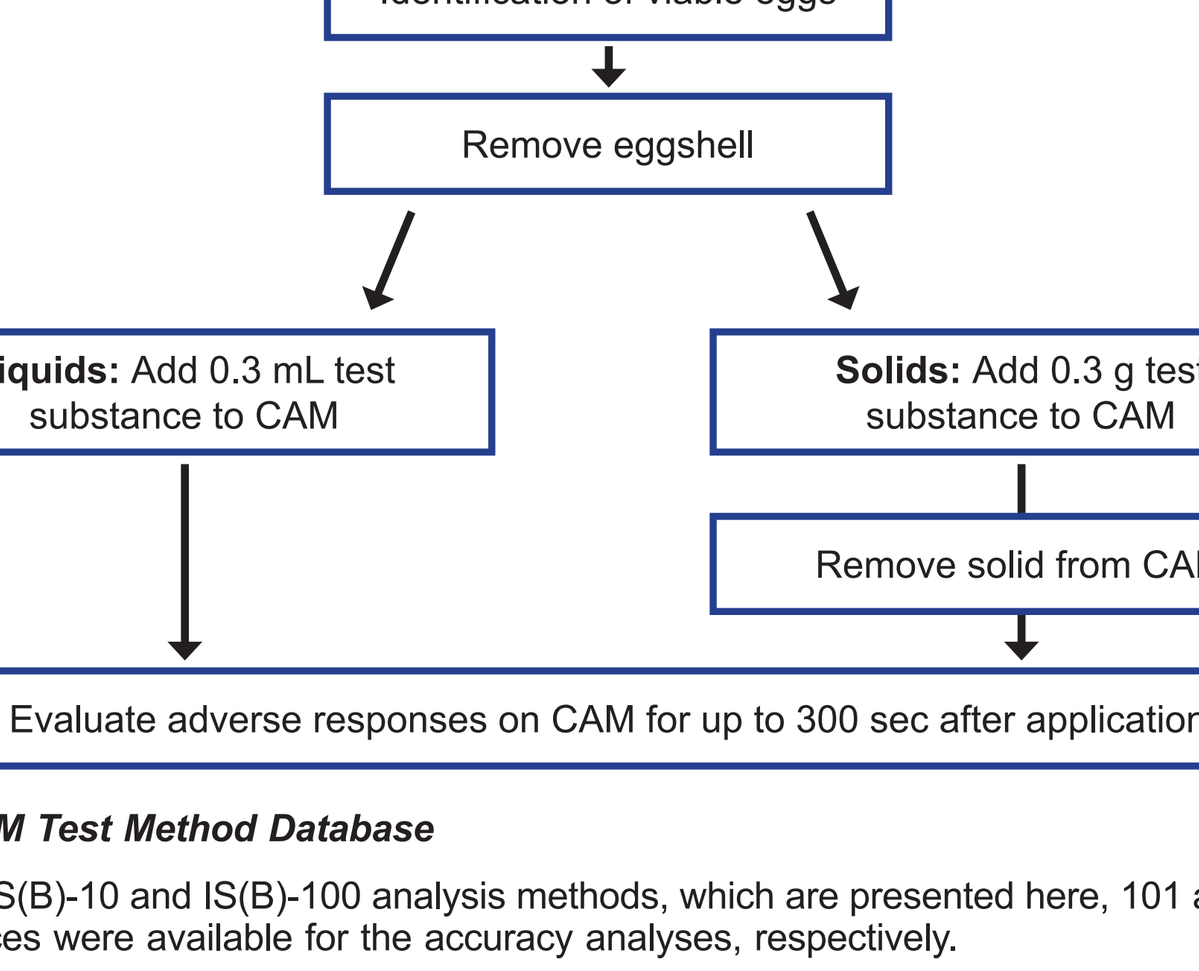
Abbreviations: GHS = Globally Harmonized System; IRE = Isolated Rabbit Eye assay.
[1]N = number of substances, the numbers in parentheses in each row indicates the data on which the percentage calculation is based.
[2]False Negative Rate = the proportion of all positive substances that are falsely identified 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/).

In Vitro Test Method Reliability

Based on a qualitative analysis of interlaboratory reproducibility in one study, 100% of the 12 to 18 tested substances were correctly identified as ocular corrosives or severe irritants by the IRE test method by all participating laboratories, for each of the regulatory classification systems evaluated (i.e., EPA, EU, GHS).

A quantitative evaluation of interlaboratory reproducibility was conducted for two studies by performing a CV analysis. The CV analysis of the first study indicated that the median CV for all substances tested was 43.4% for a 4-hour corneal opacity endpoint and 49.7% for a 4-hour swelling endpoint. When only ocular corrosives or severe irritants were considered, the CV values were 33.6% for the 4-hour corneal opacity endpoint and 35.5% for the 4-hour corneal swelling endpoint. In the second study, the median CV values for the endpoints evaluated (corneal opacity, corneal swelling, and fluorescein penetration) ranged from 24.0% to 40.0% when all substances were considered and from 15.4% to 35.5% when only severe ocular corrosives or severe irritants were considered.

Hen's Egg Test-Chorioallantoic Membrane (HET-CAM) Test Method Protocol



HET-CAM Test Method Database

For the IS(B)-10 and IS(B)-100 analysis methods, which are presented here, 101 and 138 substances were available for the accuracy analyses, respectively.

In Vitro Test Method Accuracy

Alcohols, ethers, heterocyclic compounds, and organic salts were the most overpredicted chemical classes for HET-CAM IS(B)-10[2] analysis method, while alcohols, aldehydes, amines, esters, heterocyclic compounds, and organic salts were the most overpredicted chemical classes in HET-CAM IS(B)-100[3] analysis method (**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]
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 assay.
[1]N = number of substances, the numbers in parentheses in each row indicates the data on which the percentage calculation is based.
[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/).

In Vitro Test Method Reliability

The analysis of intralaboratory repeatability was evaluated using data from two different studies for the IS(B) analysis method. In both studies, the hemorrhage endpoint had the highest CV value (109.10%-117.56%). Similar results were obtained for an analysis of intralaboratory reproducibility for the same two studies.

A qualitative analysis of interlaboratory reliability for the IS(B)-10 analysis method, showed that the participating laboratories were in 100% agreement for 84 to 85 (79% to 81%) of 104 to 107 substances evaluated, depending on the hazard classification system evaluated. For the IS(B)-100 analysis method, the participating laboratories were in 100% agreement for 80 to 81 (82% to 84%) of the 95 to 99 substances evaluated, depending on the hazard classification system evaluated.

A quantitative evaluation of interlaboratory reproducibility for 14 substances in one study, evaluated at 100% concentration (IS(B)-100), indicated that the mean and median CV values were 31.86% and 33.04%, respectively. For 12 substances evaluated at 10% concentration (IS(B)-10) in the same study, the mean and median CV values were 66.29% and 60.75%, respectively. For the substances evaluated in another study, which used the IS(B) analysis method, the mean and median CV values for substances tested at 100% concentration were 35.21% and 26.22%, respectively. Comparatively, the mean and median CV values for substances tested at 10% concentration in the same study were 60.17% and 42.65%, respectively.

[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*.

Comparative Accuracy Performance

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 7**. Different numbers of chemicals were evaluated for the EPA (1996) and European Union (2001) classification systems. However, results similar to those observed for the GHS classification system were obtained.

Table 7. 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 assay; GHS = Globally Harmonized System; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane assay; ICE = Isolated Chicken Eye assay; IRE = Isolated Rabbit Eye assay.
[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.

NIEHS National Institute of Environmental Health Sciences

U.S. Department of Health and Human Services National Institutes of Health

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/

ICCVAM Test Method Recommendations

OVERALL RECOMMENDATIONS

Current uses
- The 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.
- 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 strategy[4], as part of a weight-of-evidence approach. Positive results can be used to classify a substance as an ocular hazard without testing in rabbits.
- IRE and HET-CAM test methods cannot currently be recommended for meeting regulatory hazard classification requirements; however, there may be non-regulatory uses for these two test methods.
- None of the four *in vitro* test methods evaluated can be considered to be complete replacements for the *in vivo* eye test.
- Users should be aware that performance characteristics for each of the four 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, overall performance characteristics, and chemical and physical class performance characteristics
- 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 Test Method Evaluation Report.
  - 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 ensure use of the most current recommended test method 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 rabbit studies should be conducted only if important data gaps are identified. If such studies are conducted, they 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 combining two or more *in vitro* test methods in a battery to identify ocular corrosives and severe irritants should be evaluated.

BCOP TEST METHOD

Current Use
- There are sufficient data to support the use of the BCOP test method, in appropriate circumstances and with certain limitations, as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach.

Future Studies
- To further characterize and improve the usefulness of this test method and to evaluate its potential future use for identifying mild and moderate ocular irritants, ICCVAM recommends the following studies:
  - A histopathological evaluation of the corneal tissue, 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.
  - 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 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.

ICE TEST METHOD

Current Use
- ICCVAM concludes that there are sufficient data to support the use of the ICE test method, in appropriate circumstances and with certain limitations, as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach.

Future Studies
- To further characterize and improve the usefulness of this test method and to evaluate its potential future use for identifying mild and moderate ocular irritants, ICCVAM recommends the following studies:
  - A histopathological evaluation of the corneal tissue, 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.
  - 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.
- ICCVAM also recommends that centering lights be installed on the optical pachymeter, which is used to measure corneal thickness, to ensure consistent central corneal thickness measurements across laboratories.

IRE TEST METHOD

Current Use
- Based on the current accuracy, false negative, and false positive rates across the EU, EPA, and GHS classification systems, the IRE test method is not recommended for regulatory hazard classification purposes; but may have applicability for other uses.
- There also are insufficient data using all four recommended IRE endpoints (corneal opacity, fluorescein penetration, corneal swelling, and observations of significant effect on corneal epithelium) to assess test method accuracy and reliability when all these endpoints are evaluated in a single study.

Future Studies
- To potentially improve the usefulness of the IRE test method for identifying severe ocular irritants and corrosives and its possible future use for the identification of mild and moderate ocular irritants, the following evaluations should be conducted:
  - A histopathological evaluation of the corneal tissue, 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.
  - The IRE test method 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.
- ICCVAM also recommends that centering lights be installed when an optical pachymeter is used to measure corneal thickness, to ensure consistent central corneal thickness measurements across laboratories.

HET-CAM TEST METHOD

Current Use
- Based on the accuracy, false negative, and false positive rates when the decision criteria of Luepke (1985) are used, the IS(B)-10 and IS(B)-100 analysis methods are not recommended for screening and identifying ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) for regulatory hazard classification purposes.

Future Studies
- Additional studies should be conducted to further optimize the HET-CAM prediction models and the decision criteria that would be used to identify ocular corrosives and severe irritants for the EPA, GHS, or EU classification systems.

SUBSTANCES FOR VALIDATION OF IN VITRO OCULAR TOXICITY TEST METHODS FOR IDENTIFYING OCULAR CORROSIVES AND SEVERE IRRITANTS
- ICCVAM developed a list of reference substances for the optimization and/or validation of *in vitro* tests to identify ocular corrosives and severe irritants. The list of substances (see Appendix H of the ICCVAM Test Method Evaluation Report) includes:
  - 79 GHS Category 1 substances (UN 2003)
    - Classification of 10 substances is based solely on human data
  - 28 GHS Category 2 substances (UN 2003)
    - 15 GHS Category 2A substances
    - 13 GHS Category 2B substances
  - 15 GHS nonirritant substances (UN 2003)
- These 122 substances include:
  - 34 chemical classes
  - 24 product classes
  - 79 liquids
  - 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 proficiency testing and inclusion in performance standards (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.

References and Acknowledgements

Draize J, Woodard G, Calvery H. 1944. Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J Pharmacol Exper Therapeut 82:377-390.

EPA. 1996. Label Review Manual: 2nd Ed. EPA737-B-96-001. Washington, DC: U.S. Environmental Protection Agency.

EU. 2001. Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Communities L255:1-333.

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No: 03-4508. Research Triangle Park: National Toxicology Program.

ICCVAM. 1997. Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No.: 97-3981. Research Triangle Park: National Toxicology Program.

Kalweit S, Gerner I, Spielmann H. 1987. Validation project of alternatives for the Draize eye test. Mol Toxic 1:597-603.

Luepke N. 1985. Hen's egg chorioallantoic membrane test for irritation potential. Food Chem Toxic 23:287-291.

National Research Council. 1996. Guide for the Care and Use of Laboratory Animals. Washington, DC: National Academy Press.

PHS. 2002. Public Health Service (PHS) Policy on Humane Care and Use of Laboratory Animals.

Sina JF, Galer DM, Sussman RG, Gautheron PD, Sargent, EV, Leong B, Shah PV, Curren RD, Miller K. 1995. A collaborative evaluation of seven alternatives to the Draize eye irritation test using pharmaceutical intermediates. Fundam Appl Toxicol 26:20-31.

Ubels JL, Paauw JD, Casterton PL, Kool DJ. 2002. A redesigned corneal holder for the bovine cornea opacity and permeability assay that maintains normal corneal morphology. Toxicol In Vitro 16:621-628.

UN. 2003. Globally Harmonized System of Classification and Labelling of Chemicals (GHS). New York & Geneva: United Nations Publications. Available: http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/00files_e.html. [accessed 16 October 2006].

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 agency.

This poster has not been reviewed by the US Consumer Product Safety Commission.

*As of January 2007, contact information has changed to Innovative Toxicology Consulting, LLC, Lake Worth, FL