Appendix E

Background Review Document
Current Status of *In Vitro* Test Methods
for Identifying Mild/Moderate Ocular Irritants:
The Hen's Egg Test–Chorioallantoic Membrane (HET-CAM) Test Method

This page intentionally left blank

Background Review Document
Current Status of *In Vitro* Test Methods for Identifying Mild/Moderate Ocular Irritants:
The Hen's Egg Test–Chorioallantoic Membrane (HET-CAM) Test Method

Interagency Coordinating Committee on the Validation of Alternative Methods

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

National Institute of Environmental Health Sciences
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

2010

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

This page intentionally left blank

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

^{o}C	Degrees centigrade
BCOP	Bovine corneal opacity and permeability
BRD	Background review document
CAM	Chorioallantoic membrane
CASRN	Chemical Abstracts Service Registry Number
CEPI	Corneal epithelial cell line
CPSC	U.S. Consumer Product Safety Commission
EC	European Commission
EC/HO	European Commission/British Home Office
ECVAM	European Centre for the Validation of Alternative Methods
EEC	European Economic Council
EPA	U.S. Environmental Protection Agency
EU	European Union
FDA	U.S. Food and Drug Administration
FHSA	U.S. Federal Hazardous Substances Act
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
FR	*Federal Register*
GHS	United Nations Globally Harmonized System for Classification and Labelling of Chemicals
GLP	Good Laboratory Practice
HET-CAM	Hen's egg test–chorioallantoic membrane
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ICE	Isolated chicken eye
INVITOX	*In Vitro* Techniques in Toxicology (ERGATT FRAME ECVAM Data bank)
IOMA	Maximal ocular irritation index
IRE	Isolated rabbit eye
IS(A)	Irritation Score (A) Analysis Method
IS(B)	Irritation Score (B) Analysis Method
ITC	Irritation threshold concentration
JaCVAM	Japanese Center for the Evaluation of Alternative Toxicological Methods
MAS	Maximum average score
MCA	Mean chorioallantoic irritation index
MeSH	U.S. National Library of Medicine's Medical Subject Heading
MMTS	Maximum mean total score
mtc	Mean time of coagulation
NC	Not Classified (as irritant)

NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIH	National Institutes of Health
NL	Not Labeled (as irritant)
OD	Optical density
OECD	Organisation for Economic Co-operation and Development
OPPTS	EPA Office of Prevention, Pesticides and Toxic Substances
OSHA	U.S. Occupational Safety and Health Administration
OTWG	Ocular Toxicity Working Group
TNO	TNO Nutrition and Food
UN	United Nations
ZEBET	German Center for Documentation and Evaluation of Alternative Methods to Animal Experiments

Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives

Agency for Toxic Substances and Disease Registry
* Moiz Mumtaz, Ph.D.
Bruce Fowler, Ph.D.
Edward Murray, Ph.D.
Eric Sampson, Ph.D.

Consumer Product Safety Commission
* Marilyn L. Wind, Ph.D. (Chair)
+ Kristina Hatlelid, Ph.D.
Joanna Matheson, 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.
Peter J. Schultheiss, D.V.M., DACLAM

Department of Energy
* Michael Kuperberg, Ph.D.
+ Marvin Stodolsky, Ph.D.

Department of the Interior
* Barnett A. Rattner, Ph.D.
+ Sarah Gerould, Ph.D. (to Feb. 2009)

Department of Transportation
* George Cushmac, Ph.D.
+ Steve Hwang, Ph.D.

Environmental Protection Agency
Office of Pesticide Programs
* John R. "Jack" Fowle III, Ph.D., DABT
+ Vicki Dellarco, Ph.D.
+ Tina Levine, Ph.D.
Deborah McCall
Christine Augustyniak, Ph.D. (*U.S. Coordinator, OECD Test Guidelines Program*)
Office of Pollution Prevention and Toxics
Jerry Smrchek, Ph.D. (*U.S. Coordinator, OECD Test Guidelines Program*, to July 2009)
Office of Research and Development
Suzanne McMaster, Ph.D. (to Dec. 2008)
Julian Preston, Ph.D. (to July 2009)
Stephanie Padilla, Ph.D. (to July 2009)
Office of Science Coordination and Policy
Karen Hamernik, Ph.D. (to July 2009)

Food and Drug Administration
Office of the Commissioner
* Suzanne Fitzpatrick, Ph.D., DABT
Center for Biologics Evaluation and Research
Richard McFarland, Ph.D., M.D.
Ying Huang, Ph.D.
Center for Devices and Radiological Health
Melvin E. Stratmeyer, Ph.D.
Vasant G. Malshet, Ph.D., DABT
Center for Drug Evaluation and Research
+ Abigail C. Jacobs, Ph.D.
Paul C. Brown, Ph.D.
Center for Food Safety and Applied 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
Paul Howard, Ph.D.
Donna Mendrick, Ph.D.
William T. Allaben, Ph.D. (to Jan. 2009)
Office of Regulatory Affairs
Lawrence D'Hoostelaere, Ph.D.

National Cancer Institute
* T. Kevin Howcroft, Ph.D.
Chand Khanna, D.V.M., Ph.D.
Alan Poland, M.D. (to Oct. 2008)

National Institute of Environmental Health Sciences
* William S. Stokes, D.V.M., DACLAM
+ Raymond R. Tice, Ph.D.
Rajendra S. Chhabra, Ph.D., DABT
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
* Pertti (Bert) Hakkinen, Ph.D.
+ Jeanne Goshorn, M.S.

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

* Principal agency representative
+ Alternate principal agency representative

Acknowledgements

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Ocular Toxicity Working Group (OTWG)

U.S. Consumer Product Safety Commission
Marilyn L. Wind, Ph.D.
Adrienne Layton, Ph.D.

U.S. Department of Defense
Harry Salem, Ph.D.

U.S. Department of Transportation
Steve Hwang, Ph.D.

U.S. Environmental Protection Agency
Office of Pesticide Programs
Meta Bonner, Ph.D.
Jonathan Chen, Ph.D.
John R. "Jack" Fowle III, Ph.D., DABT
Masih Hashim, D.V.M., Ph.D.
Karen Hicks
Marianne Lewis
Debbie McCall
Timothy McMahon, Ph.D.
Mark Perry
John Redden
Jenny Tao, Ph.D.
Office of Research and Development
Andrew Geller, Ph.D.
Office of Science Coordination and Policy
Karen Hamernik, Ph.D.

U.S. Food and Drug Administration
Center for Drug Evaluation and Research
Paul Brown, Ph.D.
Wiley Chambers, M.D.
Abigail (Abby) Jacobs, Ph.D.
Jill Merrill, Ph.D., DABT (OTWG Chair)
Center for Food Safety and Applied Nutrition
Robert Bronaugh, Ph.D.
Donnie Lowther
Office of the Commissioner
Suzanne Fitzpatrick, Ph.D., DABT

National Institute of Environmental Health Sciences
Warren Casey, Ph.D., DABT
Mark F. Cesta, D.V.M, DACVP
Raymond (Buck) Grissom, Ph.D.
William Stokes, D.V.M., DACLAM

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

European Centre for the Validation of Alternative Methods – Liaison
João Barroso, Ph.D.
Thomas Cole, Ph.D.
Valerie Zuang, Ph.D.

Japanese Center for the Validation of Alternative Methods – Liaison
Hajime Kojima, Ph.D.

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences

William Stokes, D.V.M., DACLAM
Director; Project Officer

Warren Casey, Ph.D., DABT
Deputy Director

Deborah McCarley
Special Assistant; Assistant Project Officer

NICEATM Support Contract Staff (Integrated Laboratory Systems [ILS], Inc.)

David Allen, Ph.D.
Jonathan Hamm, Ph.D.
Nelson Johnson
Brett Jones, Ph.D.
Elizabeth Lipscomb, Ph.D.
Linda Litchfield
Steven Morefield, M.D.
Catherine Sprankle
James Truax, M.A.
Linda Wilson

Statistical Consultant for ILS, Inc.

Joseph Haseman, Ph.D.

Additional Reviewers for the *In Vitro* Ocular Corrosion and Irritation Test Methods Background Review Documents

Chantra Eskes, Eng., Ph.D.
ECVAM
Ispra, Italy

Robert L Guest Bsc, CBiol, MlBiol
SafePharm Laboratories, Ltd.
Derby, United Kingdom

John Harbell, Ph.D.
Institute for In Vitro Sciences
Gaithersburg, Maryland

Penny Jones
Unilever Research
Sharnbrook, United Kingdom

Menk Prinsen
TNO Nutrition & Food Research Institute
Zest, The Netherlands

Horst Spielmann, Dr.med.
ZEBET
Berlin, Germany

Companies and Individuals that Provided *In Vitro* and/or *In Vivo* Data for the HET-CAM Test Method Background Review Document

ICCVAM and NICEATM gratefully acknowledge the following individuals and institutions that submitted data to NICEATM for the evaluation of alternative ocular safety testing methods and approaches.

Cosmetics, Toiletry, and Fragrance Association
Carol Eisenmann, Ph.D.

ECVAM
Chantra Eskes, Ph.D.

Johnson & Johnson Pharmaceutical R&D
Freddy van Goethem, Ph.D.
Philippe Vanparys, Ph.D.

National Institute of Health Sciences
Yasuo Ohno, Ph.D.

U.S. Food and Drug Administration
Donnie Lowther

ZEBET
Manfred Liebsch, Ph.D.
Horst Spielmann, Dr. med.

This page intentionally left blank

Preface

Accidental contact with hazardous chemicals frequently causes eye injury and visual impairment. United States and international regulatory agencies currently use the Draize rabbit eye test (Draize et al. 1944) to identify potential ocular hazards associated with chemicals. The U.S. Consumer Product Safety Commission (CPSC), U.S. Environmental Protection Agency (EPA), U.S. Food and Drug Administration, and U.S. Occupational Safety and Health Administration have testing requirements and guidelines for assessing the ocular irritation potential of substances such as pesticides, household products, pharmaceuticals, cosmetics, and agricultural and industrial chemicals.

Although ocular safety assessment has clearly helped to protect consumers and workers, concerns have been raised about the humane aspects of the Draize rabbit eye test. Regulatory authorities have adopted various modifications that reduce the number of animals used and the potential pain and distress associated with the procedure. Significant progress has been made during the last decade. Now only one to three rabbits are required per test, compared to six rabbits in the original protocol. Provisions have been added that allow for animals with severe lesions or discomfort to be humanely euthanized.

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) previously evaluated the validation status of the bovine corneal opacity and permeability (BCOP), isolated chicken eye (ICE), isolated rabbit eye (IRE), and hen's egg test–chorioallantoic membrane (HET-CAM) assays for the identification of ocular corrosives or severe (irreversible) ocular irritants. ICCVAM's evaluation used the EPA (EPA 2003a), United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS) (UN 2007), and European Union (EU 2001) regulatory hazard classification systems. In ICCVAM's assessment, the performance of the BCOP and ICE test methods substantiated their use in testing some substances for regulatory hazard classification. The IRE and HET-CAM test methods lacked sufficient performance and/or sufficient data to substantiate their use for regulatory hazard classification.

ICCVAM recommended that the BCOP and ICE should be used in a tiered-testing strategy in which positive substances can be classified as ocular corrosives or severe irritants without animal testing. In accordance with the ICCVAM Authorization Act of 2000 (Public Law 106-545), these recommendations were made available to the public and provided to U.S. Federal agencies for consideration in the ICCVAM *Test Method Evaluation Report* – In Vitro *Ocular Toxicity Test Methods for Identifying Severe Irritants and Corrosives* (ICCVAM 2006b). The ICCVAM recommendations were accepted by U.S. Federal agencies, and *in vitro* test methods may now be used instead of the Draize rabbit eye test for certain regulatory testing purposes.

ICCVAM is now reviewing the validation status of these *in vitro* test methods for identification of nonsevere ocular irritants (that is, those that induce reversible ocular damage [EPA Category II, III; EU Category R36, GHS Category 2A, 2B]) and substances Not Classified as irritant (GHS NC or Not Labeled, EPA Category IV, FHSA Not Labeled, or EU Not Labeled) according to the GHS (UN 2007), EPA (EPA 2003a), FHSA (FHSA 2005), and EU (EU 2001) classification systems. The Federal Hazardous Substances Act (FHSA) classification system (FHSA 2005) as defined in the "Test for Eye Irritants" (i.e., "Irritant" or Not Labeled [as an irritant]) and published in 16 CFR 1500.42 (CPSC 2003) is also provided in the current background review documents. The FHSA classification system was not used in the previous analyses of test methods used for the identification of severe ocular irritants or corrosives because the FHSA classification is limited to irritants and is not intended to identify corrosive substances or to differentiate between severe and nonsevere irritants.

Accordingly, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the ICCVAM Ocular Toxicity Working Group (OTWG) prepared draft background review documents that summarize the current validation status of each test

method based on published studies and other data and information submitted in response to a June 7, 2007, *Federal Register* request (72 FR 31582, available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_10966.pdf). The background review documents form the basis for draft ICCVAM test method recommendations, which are provided in separate documents. Liaisons from the European Centre for the Validation of Alternative Methods and the Japanese Centre for the Validation of Alternative Methods will provide input and contribute to the OTWG throughout the evaluation process.

An international independent scientific peer review panel (Panel) met in public session on May 19-21, 2009, to develop conclusions and recommendations on the *in vitro* BCOP, ICE, IRE, and HET-CAM test methods. The Panel included expert scientists nominated by the European Centre for the Validation of Alternative Methods and the Japanese Center for the Validation of Alternative Methods. We anticipate that these organizations can use the subsequent independent Panel report to deliberate and develop their own test method recommendations (ICCVAM Peer Review Panel Report [ICCVAM 2009] available to the public for comment on July 12, 2009). The Panel considered these BRDs and evaluated the extent to which the available information supports the draft ICCVAM test method recommendations.

ICCVAM provided the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) with the draft BRD and draft Test Method Evaluation Report, the Panel's report, and all public comments. SACATM discussed these at their June 25-26, 2009, meeting, where public stakeholders were given another opportunity to comment. After SACATM's meeting, ICCVAM considered the SACATM comments, the Panel report, and all public comments before finalizing the Background Review Document and test method recommendations. These recommendations will be forwarded to Federal agencies for their consideration and acceptance decisions where appropriate.

We gratefully acknowledge the organizations and scientists who provided data and information for this document. We also acknowledge the efforts of those individuals who helped prepare this background review document, including the following staff from the NICEATM support contractor, Integrated Laboratory Systems, Inc.: David Allen, Jon Hamm, Nelson Johnson, Elizabeth Lipscomb, Linda Litchfield, Steven Morefield, Gregory Moyer, Catherine Sprankle, and Jim Truax. We also thank the members of the OTWG, chaired by Karen Hamernik, Ph.D. (U.S. EPA) and Jill Merrill, Ph.D. (U.S. Food and Drug Administration), and ICCVAM representatives who reviewed and commented on draft versions. We also thank Valerie Zuang, Ph.D., and Dr. Hajime Kojima, Ph.D., the liaisons to the OTWG from the European Centre for the Validation of Alternative Methods and the Japanese Center for the Validation of Alternative Methods, respectively, for their participation.

Marilyn Wind, Ph.D.
Deputy Associate Executive Director
Directorate for Health Sciences
U.S. Consumer Product Safety Commission
Chair, ICCVAM

William S. Stokes, D.V.M., DACLAM
Rear Admiral/Assistant Surgeon General, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Executive Summary

In October 2003, the U.S. Environmental Protection Agency (EPA) submitted to the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) a nomination requesting the evaluation of several activities related to reducing, refining, and replacing the use of rabbits in the current *in vivo* Draize rabbit eye test (69 FR 13859 [March 24, 2004]). In response to this nomination, ICCVAM evaluated the validation status of the bovine corneal opacity and permeability (BCOP), hen's egg test–chorioallantoic membrane (HET-CAM), isolated chicken eye (ICE), and isolated rabbit eye (IRE) test methods. To evaluate how well these test methods identify ocular corrosives and severe irritants, ICCVAM used the EPA (2003a), European Union (EU 2001), and United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS) (UN 2007) classification systems.

ICCVAM considered the performance of two of these *in vitro* test methods, the BCOP and the ICE, to be sufficient to support their use in testing certain types of substances for regulatory hazard classification. The IRE and HET-CAM test methods lacked sufficient performance and/or sufficient data to support their use for regulatory hazard classification. ICCVAM recommended that the BCOP and ICE test methods should be used in a tiered-testing strategy that would classify positive substances as ocular corrosives or severe irritants without animal testing. These recommendations were accepted by U.S. Federal agencies, and, as a result, *in vitro* test methods may now be used instead of conventional tests for certain regulatory testing purposes.

ICCVAM is now reviewing the validation status of these *in vitro* test methods to identify nonsevere ocular irritants (those that cause reversible ocular damage [EPA Category II and III; EU R36; GHS Category 2A and 2B]) and substances not labeled as irritants (EPA Category IV; EU Not Labeled; GHS Not Classified) according to the EPA (2003a), EU (2001), and GHS (UN 2007) classification systems. The FHSA classification system, which is based on the testing guidelines and associated criteria included in 16 CFR 1500.42 (CPSC 2003), is also included in these evaluations. The FHSA classification system was not used in the original analyses (ability of the test methods to identify ocular corrosives and severe irritants) because the FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test (Draize et al. 1944) does not distinguish between ocular corrosives and severe irritants and less severe irritants. For this reason, an evaluation to identify ocular corrosives and severe irritants using the FHSA classification system was not possible.

Because the FHSA classification system (FHSA 2005) is based on a sequential testing strategy that uses up to 18 animals, only a small percentage of the substances in the HET-CAM database would be classifiable if the FHSA criteria were strictly applied. To maximize the number of substances included in these analyses, "proportionality" criteria were applied for the purpose of assigning an FHSA classification to test results that would require additional testing according to the FHSA sequential testing strategy. These "proportionality" criteria (FHSA-20% and FHSA-67%) are as follows:

- FHSA-20% is based on the proportion of positive animals needed to identify a substance as an irritant using the FHSA sequential testing strategy, where 20% of the animals must demonstrate a positive response for a substance to be identified as an irritant. A substance tested using 3 to 6 animals would not be labeled if ≤1/6 animals were positive based on the FHSA criteria. The substance would be labeled as an irritant if there were ≥1 positive animal in a 3- to 5-animal test or ≥2 positive animals in a 6-animal test.
- FHSA-67% is based on the proportion of positive animals needed to identify a substance as an irritant using the "first test" of the FHSA sequential testing strategy, where 67% of the animals must demonstrate a positive response for a substance to be identified as an irritant. A substance tested using 3 to 6 animals would not be labeled as an irritant if ≤1/6

animals were positive based on the FHSA criteria. The substance would be labeled as an irritant if there were ≥2/3, 3/4, 4/5, or 4/6 positive animals. If 1/3, 1/4, 2/4, 1/5, 2/5, 3/5, 2/6, or 3/6 animals were positive, further testing would be required.

Together, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the ICCVAM Ocular Toxicity Working Group prepared draft background review documents (BRDs) that summarize the available data and information regarding the validity (usefulness and limitations) of each test method. This BRD summarizes all available information for the HET-CAM test method and its current validation status, including what is known about its reliability and accuracy, and the scope of the substances tested. Original data for the HET-CAM test method will be maintained for future use so that these performance statistics may be updated as additional information becomes available.

HET-CAM Test Method Protocol

The HET-CAM test method uses the vascular fetal membrane of chicken embryos. The HET-CAM test method is proposed to provide information on the effects that may occur in the conjunctiva of the eye following test substance administration. It is assumed that acute effects induced by a test substance on the small blood vessels and proteins of this soft tissue membrane are similar to effects induced by the same test substance in the eye of a treated rabbit. The membrane is evaluated for the development of irritant endpoints (hyperemia, hemorrhage, and coagulation) and qualitative assessments of the irritation potential of test substances are made.

Validation Database

No new HET-CAM data have been obtained since ICCVAM evaluated the HET-CAM test method for identifying ocular corrosives and severe irritants (ICCVAM 2006a). Therefore, the same database was used in the current evaluation. The HET-CAM validation database contains a total of 260 substances and formulations. The most commonly tested chemical classes are alcohols, carboxylic acids, and formulations. The most commonly tested product classes are solvents, shampoos, surfactants, and cosmetics. Analyses of each of the HET-CAM protocols indicate that the Irritation Score (A), or IS(A), analysis method performed best when evaluating substances not labeled as irritants. The available IS(A) database includes 63 test substances, 58 to 60 of which had sufficient *in vivo* data to be assigned an ocular irritancy hazard classification, depending on the classification system used. These 58 to 60 substances comprise 43 cosmetic and personal care product formulations (including 25 surfactant-based formulations and 18 oil/water emulsions) and 17 individual substances (including seven alcohols; no other classes were represented by more than three substances).

In order to calculate the appropriate EPA (2003a), EU (2001), FHSA (2005), and GHS (UN 2007) ocular irritancy hazard classifications, detailed *in vivo* data consisting of cornea, iris, and conjunctiva scores for each animal at 24, 48, and 72 hours following test substance administration and/or assessment of the presence or absence of lesions at 7, 14, and 21 days are needed. Some of the test substances had only limited *in vivo* data and could not be used to evaluate test method accuracy and reliability. To maximize the number of substances included in the FHSA analyses, “proportionality” criteria (FHSA-20% and FHSA-67%), as outlined above, were applied for the purpose of assigning a FHSA classification to test results that would require additional testing according to the FHSA sequential testing strategy.

HET-CAM Test Method Accuracy

Identification of All Ocular Hazard Categories

ICCVAM evaluated how well the HET-CAM test method identified all categories of ocular irritation potential as defined by the EPA (2003a), GHS (UN 2007), and EU (2001) classification systems. For

these evaluations, the IS(A) analysis method was used. Because the FHSA classification system does not distinguish between ocular corrosives and severe irritants and less severe irritants, an evaluation for all ocular hazard categories using the FHSA classification system was not possible. Analyses were also performed excluding specific chemical classes and/or physical properties that were previously identified as discordant in the HET-CAM test method (alcohols, surfactant formulations, and oil/water emulsions) relative to the *in vivo* hazard classification (ICCVAM 2006a).

As shown in **Table 1**, overall correct classifications ranged from 38% (23/60) to 41% (24/59) when using the entire database, depending on the hazard classification system used. When discordant classes are excluded, overall correct classifications improved to a range of 62% (5/8) to 78% (7/9), depending on the classification system used. However, too few substances (0–2) are in the moderate category (EPA Category II, GHS Category 2A, EU R36) to adequately evaluate the performance of the HET-CAM test method for this irritant category. Similarly, while 18 substances are classified as mild (EPA Category III) for the EPA system, only five are classified as GHS Category 2B (the EU system does not distinguish mild irritants).

Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories

ICCVAM also evaluated how well the HET-CAM test method distinguished substances not labeled as irritants (EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other ocular hazard categories (EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B) as defined by the EPA (2003a), GHS (UN 2007), EU (2001), and FHSA (2005) classification systems. Analyses were also performed excluding specific chemical classes and/or physical properties that were previously identified as discordant in the HET-CAM test method (alcohols, surfactant formulations, and oil/water emulsions) relative to the *in vivo* hazard classification (ICCVAM 2006a).

As shown in **Table 2**, overall accuracy ranged from 62% (36/58) to 80% (44/55), depending on the hazard classification system used. The lowest false negative rate (0% [0/31 and 0/26]) was noted for the GHS and EU classification systems, followed by 3% (1/39) for FHSA-67% criteria, and 9% (4/45 and 4/47) for the EPA and FHSA-20% classification systems. All four false negatives for the EPA classification system were oil/water emulsions that were classified as EPA Category III substances based on Draize rabbit eye test data. The false negatives identified using the FHSA-20% and FHSA-67% criteria were the same oil/water emulsions identified by the EPA classification system. The lowest false positive rate (60% [9/15]) was noted for the EPA classification system, followed by 63% (10/16) for the FHSA-20% and FHSA-67% criteria, and 64% (18/28) and 69% (22/32) for the GHS and EU classification systems, respectively.

The exclusion of discordant classes improved accuracy (ranged from 75% [6/8] to 100% [9/9 and 10/10] when discordant classes were removed versus 62% [36/58] to 80% [44/55] for overall accuracy, depending on the hazard classification system used). However, the discordant substances comprised at least 84% of the substances in each classification system, so the performance of each classification system was based on ten or fewer substances.

Table 1 Performance of the HET-CAM Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA, GHS, and EU Classification Systems[1]

Hazard Classification System	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall (EPA)	38% (23/60)	48% (12/25)	52% (13/25)	50% (1/2)	50% (1/2)	0% (0/2)	56% (10/18)	22% (4/18)	22% (4/18)	60% (9/15)	40% (6/15)
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions[6]	78% (7/9)	100% (6/6)	0% (0/6)	50% (1/2)	50% (1/2)	0% (0/2)	100% (1/1)	0% (0/1)	0% (0/1)	- (0/0)	- (0/0)
Overall (GHS)	41% (24/59)	50% (13/26)	50% (13/26)	- (0/0)	- (0/0)	- (0/0)	80% (4/5)	20% (1/5)	0% (0/5)	64% (18/28)	36% (10/28)
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	67% (6/9)	86% (6/7)	14% (1/7)	- (0/0)	- (0/0)	- (0/0)	100% (1/1)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/1)
Overall (EU)	40% (23/58)	50% (12/24)	50% (12/24)	50% (1/2)	50% (1/2)	0% (0/2)	NA	NA	NA	69% (22/32)	31% (10/32)
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	62% (5/8)	100% (5/5)	0% (5/5)	100% (1/1)	0% (0/1)	0% (0/1)	NA	NA	NA	100% (2/2)	0% (0/2)

Abbreviations: EPA = U.S. Environmental Protection Agency; EU = European Union; GHS = Globally Harmonized System; HET-CAM = hen's egg test–chorioallantoic membrane; NA = not applicable.

[1] EPA classification system (EPA 2003a); GHS classification system (UN 2007); EU classification system (EU 2001). Because the FHSA classification system does not distinguish between ocular corrosives/severe irritants and less severe irritants, an evaluation for all ocular hazard categories using the FHSA classification system was not possible.

[2] Severe = EPA Category I; GHS Category 1; EU R41.

[3] Moderate = EPA Category II; GHS Category 2A; EU R36.

[4] Mild = EPA Category III; GHS Category 2B.

[5] Not Labeled = EPA Category IV; GHS Not Classified: EU Not Labeled.

[6] Alcohols, surfactant formulations, and oil/water emulsions were previously identified as discordant in the HET-CAM test method relative to the *in vivo* hazard classification (ICCVAM 2006a).

HET-CAM Test Method Reliability

Interlaboratory Reproducibility

Previous quantitative and qualitative evaluations of the reliability of the HET-CAM test method have been conducted (ICCVAM 2006a). Because the database used for the current evaluation of the HET-CAM test method has not changed, the quantitative evaluation of test method reliability remains unchanged. Additional qualitative analyses of interlaboratory reproducibility were conducted to evaluate how well the HET-CAM hazard classifications agreed among the five participating laboratories from the interlaboratory validation study (Hagino et al. 1999). These evaluations were based on the use of the HET-CAM test method (1) to identify all ocular hazard categories according to the EPA, EU, or GHS systems, and (2) to distinguish substances not labeled as irritants (EPA Category IV, GHS Not Classified, EU Not Labeled) from all other ocular hazard categories (EPA Categories I, II, III; GHS Categories 1, 2A, 2B; EU R41, R36). Because the performance of the HET-CAM test method was similar for the EPA and FHSA hazard classification systems, additional reliability analyses were not conducted for the FHSA hazard classification system.

Using the first approach (identifying all ocular hazard categories), there was 100% agreement among the five laboratories for a majority of the Draize ocular corrosives and severe irritants correctly classified by the HET-CAM test method based on all three classification systems. There was 100% agreement for 63% [5/8] of the correctly identified EPA Category I substances and 100% agreement for 71% [5/7] of the correctly identified GHS Category 1 or EU R41 substances. There was 100% agreement among the five laboratories for the one moderate irritant in the database (EPA Category II or EU R36; no GHS Category 2A substances were included), which was overpredicted by the HET-CAM test method. There was 100% agreement for the mild ocular irritants (EPA Category III, GHS Category 2B; the EU does not have a mild irritant category), which were uniformly overpredicted by the HET-CAM test method. For the Hagino et al. (1999) database, all of the substances not classified as irritants based on Draize data (EPA Category IV, EU Not Labeled, GHS Not Classified) were overpredicted by the HET-CAM test method. There was 100% agreement among the five laboratories for 86% (6/7) or 75% (3/4) of these substances for the EU and GHS classification systems, respectively. By comparison, for the two EPA Category IV substances tested, there was either 100% or 80% agreement among the five laboratories.

Using the second approach (distinguishing substances not labeled as irritants from all other ocular hazard categories), there was 100% agreement among the five laboratories for 76% (13/17) to 94% (16/17) of the substances tested by the HET-CAM test method, depending on the classification system used.

There was 100% agreement among the five laboratories for 100% (13/13) of the substances correctly identified as irritants according to the EPA classification system (Category I, II, or III). While neither of the EPA Category IV substances were correctly identified by the HET-CAM test method, there was 60% agreement among the five laboratories for 100% (2/2) of the EPA Category IV substances that were overpredicted by the HET-CAM test method.

There was 100% agreement among the five laboratories for 63% (5/8) of the substances correctly identified as an irritant according to the EU classification system (R36 or R41). There was at least 60% agreement among the five laboratories for the remaining three substances correctly classified as an irritant. While none of the EU Not Labeled substances were correctly identified by the HET-CAM test method, there was 100% agreement among the five laboratories for 86% (6/7) of these substances that were overpredicted by the HET-CAM test method.

There was 100% agreement among the five laboratories for 100% (11/11) of the substances correctly identified as irritants according to the GHS classification system (Category 1, 2A, or 2B). While none of the GHS Not Classified substances were correctly identified by the HET-CAM test method, there

was 100% agreement among the five laboratories for 75% (3/4) of these substances that were overpredicted by the HET-CAM test method.

Table 2 Accuracy of the HET-CAM IS(A) Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories, as Defined by the EPA, GHS, EU, and FHSA Classification Systems

Hazard Classification System	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall (EPA)[1]	60	78	47/60	91	41/45	40	6/15	60	9/15	9	4/45
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions[2]	9	100	9/9	100	9/9	-	0/0	0	0/9	-	0/0
Overall (GHS)[3]	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	89	8/9	100	8/8	0	0/1	100	1/1	0	0/8
Overall (EU)[4]	58	62	36/58	100	26/26	31	10/32	69	22/32	0	0/26
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	8	75	6/8	100	6/6	0	0/2	100	2/2	0	0/6
Overall (FHSA-20%)[5]	63	78	49/63	91	43/47	38	6/16	63	10/16	9	4/47
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	10	100	10/10	100	10/10	-[6]	-	-	-	0	0/10
Overall (FHSA-67%)[5]	55	80	44/55	97	38/39	38	6/16	63	10/16	3	1/39
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	100	9/9	100	9/9	-[6]	-	-	-	0	0/9

Abbreviations: EPA = U.S. Environmental Protection Agency; EU = European Union; FHSA = U.S. Federal Hazardous Substances Act; GHS = Globally Harmonized System; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] EPA classification system (EPA 2003a): Category IV vs. Category I/II/III.

[2] Alcohols, surfactant formulations, and oil/water emulsions were previously identified as discordant in the HET-CAM test method relative to the *in vivo* hazard classification (ICCVAM 2006a).

[3] GHS classification system (UN 2007): Not Classified vs. Category 1/2A/2B.

[4] EU classification system (EU 2001): Not Labeled vs. R41/R36.

[5] FHSA classification system (FHSA 2005): Not Labeled vs. Irritant. To maximize the number of substances included in the FHSA analyses, “proportionality” criteria (FHSA-20% and FHSA-67%) were applied for the purpose of assigning a FHSA classification to test results that would require additional testing according to the FHSA sequential testing strategy.

[6] No substances were classified as Not Labeled by FHSA or as nonirritants in HET-CAM, so specificity and the false positive rate could not be determined.

1.0 Introduction

1.1 Background

The current rabbit eye test method identifies both irreversible (e.g., corrosion) and reversible ocular effects. It also provides quantitative scoring with which to categorize the severity of reversible effects such as mild, moderate, or severe irritation. Current U.S. Environmental Protection Agency ocular testing guidelines and the United Nations (UN) Globally Harmonized System (GHS) of Classification and Labelling of Chemicals indicate that if serious ocular damage is anticipated (e.g., a lesion considered to be irreversible or persisting for 21 days), then a test on a single animal may be considered. If serious damage is observed, no further animal testing is necessary (EPA 1998; UN 2007). If no serious damage is observed, additional test animals (1 or 2 rabbits) may be evaluated sequentially until concordant irritant or nonirritant responses are observed based on the GHS (UN 2007) or until unequivocal results are obtained in a minimum of three animals according to the EPA test guideline (EPA 1998). In the FHSA classification system (FHSA 2005), which is based on the testing guidelines and associated criteria included in 16 CFR 1500.42 (CPSC 2003), corrosive substances are identified by other test methods (e.g., Draize skin test or human accidental exposure data) and excluded from further irritant testing.

In 2006, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) finished evaluating the hen's egg test–chorioallantoic membrane (HET-CAM) test method to identify ocular corrosives and severe irritants (ICCVAM 2006a). ICCVAM concluded that the HET-CAM test method was not suitable for identifying ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) (ICCVAM 2006b), but this recommendation could be revised as additional data become available.

ICCVAM is now evaluating the usefulness and limitations of the HET-CAM test method for identifying nonsevere irritants (i.e., those that induce reversible ocular damage [EPA Category II and III; EU R36; GHS Category 2A and 2B]) and substances not labeled as irritants (i.e., EPA Category IV; EU Not Labeled; FHSA Not Labeled; GHS Not Classified) according to the EPA, EU, FHSA, and GHS classification systems (EPA 2003a; EU 2001; FHSA 2005; UN 2007). However because the FHSA classification system (2005) is based on a sequential testing strategy, which uses up to 18 animals, only a small percentage of the substances in the ICE database would be classifiable if the FHSA criteria were strictly applied. In order to maximize the number of substances included in these analyses, "proportionality" criteria (i.e., FHSA-20% and FHSA-67%) were applied for the purpose of assigning an FHSA classification for test results that would require additional testing according to the FHSA sequential testing strategy (see **Section 4.1**).

As part of the evaluation process, this background review document (BRD) has been prepared to describe the current validation status of the HET-CAM test method, including what is known about its reliability and accuracy, its applicability domain, the numbers and types of substances tested, and the availability of a standardized protocol. An ICCVAM expert panel used this BRD when reviewing the HET-CAM as a method to identify all categories of ocular irritants and substances not labeled as irritants.

Parallel reviews of the bovine corneal opacity and permeability (BCOP), isolated chicken eye (ICE), and isolated rabbit eye (IRE), test methods are being conducted. The expert panel report and the analyses presented in the BRDs will be used to support ICCVAM recommendations on the proposed standardized test method protocols, proposed list of recommended reference substances, and additional optimization and/or validation studies that may be necessary to further develop and characterize the usefulness and limitations of these methods.

For a more detailed discussion of the background of the HET-CAM test method, including its scientific basis and regulatory rationale and applicability, see the *ICCVAM Background Review Document—Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane* (ICCVAM 2006a).

1.2 Use of the HET-CAM Test Method in Overall Strategy of Hazard or Safety Assessment

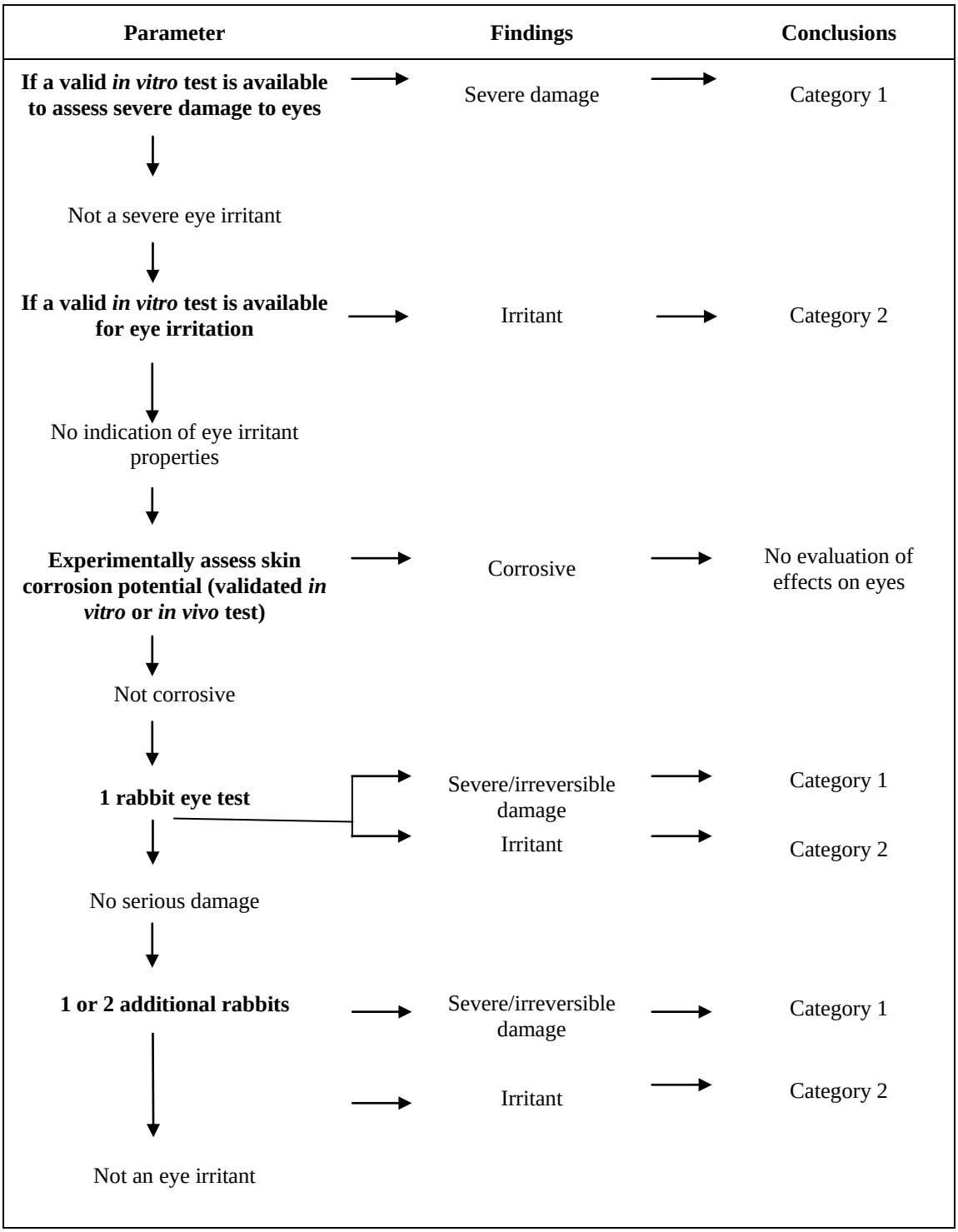
As shown in **Figure 1-1**, the GHS allows for use of validated and accepted *in vitro* methods to identify ocular corrosives/severe irritants and ocular irritants without further testing. The HET-CAM test method is currently not recommended for identification of ocular corrosives and severe irritants in a tiered-testing strategy for regulatory classification and labeling for use in the GHS testing scheme (UN 2007). ICCVAM is now further evaluating the usefulness and limitations of the HET-CAM test method for identifying nonsevere irritants and substances not labeled as irritants.

1.3 Validation of the HET-CAM Test Method

The ICCVAM Authorization Act of 2000 (Sec. 4([c]) mandates that "each Federal Agency … shall ensure that any new or revised … test method … is determined to be valid for its proposed use prior to requiring, recommending, or encouraging [its use]" [A16] (Public Law 106-545).

Validation is the process that establishes the reliability and relevance of a test method for a specific purpose (ICCVAM 2003). *Relevance* is defined as the extent to which a test method will correctly predict or measure the biological effect of interest (ICCVAM 2003). For the HET-CAM test method described in the ICCVAM 2006 BRD (ICCVAM 2006a), relevance is restricted to how well the test method identifies substances that are capable of producing corrosive or severe irritant effects to the eye. For the current BRD, relevance is based on how well the test method identifies substances that are capable of producing nonsevere ocular irritation or substances not labeled as irritants.

Figure 1-1 GHS Testing Strategy for Serious Eye Damage and Eye Irritation[1]



Abbreviations: GHS = Globally Harmonized System

[1] Adapted from UN (2007).

Reliability is defined as the reproducibility of a test method within and among laboratories. Reliability should be based on its performance with a diverse set of substances that (1) represent the types of

chemical and product classes likely to be tested and (2) cover the range of responses that need to be identified. The validation process will provide data and information to allow U.S. Federal agencies to develop guidance on the development and use of the HET-CAM test method as part of a tiered-testing approach to evaluating substances' eye irritation potential.

The first stage in this evaluation is the preparation of a BRD that presents and evaluates the relevant data and information about the test method, including its mechanistic basis, proposed uses, reliability, and performance characteristics (ICCVAM 2003). This BRD summarizes the available information on the HET-CAM test method. Where adequate data are available, the qualitative and quantitative performance of the test method are evaluated.

1.4 Search Strategies and Selection of Citations for the HET-CAM BRD

The HET-CAM test method data summarized in this BRD are based on information found in the peer-reviewed scientific literature as detailed in the ICCVAM *Background Review Document—Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane Test Method* (ICCVAM 2006a). A literature search for HET-CAM studies published between January 2005 and January 2009 used the same terminology and information databases used in the 2006 ICCVAM BRD (ICCVAM 2006a). The research revealed four studies that included information on HET-CAM protocols or contained data on test substances. While no *in vivo* reference data were included in any of the four citations, *in vivo* data for six of nine substances included in one study were available from the National Toxicology Program Interagency Center for the Validation of Alternative Toxicological Methods (NICEATM) database of Draize eye test results. However, because these substances were included in the original analyses (and the HET-CAM results from the new study agreed with the previous results), the database used in the HET-CAM performance analysis is the same as the database used in the ICCVAM *Background Review Document—Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane Test Method* (ICCVAM 2006a).

2.0 Hen's Egg Test–Chorioallantoic Membrane Test Method Protocol Components

The HET-CAM protocol first described by Luepke (1985) uses a vascular fetal membrane, the chorioallantoic membrane (CAM), which is composed of the fused chorion and allantois. The CAM has been proposed as a model for a living membrane (such as the conjunctiva) because it comprises a functional vasculature. Additionally, evaluation of coagulation (i.e., protein denaturation) may reflect corneal damage that may be produced by the test substance. The acute effects induced by a test substance on the small blood vessels and proteins of this soft tissue membrane are proposed to be similar to effects induced by the same test substance in the eye of a treated rabbit.

Since the initial description of the HET-CAM test method, several studies have been conducted to evaluate the feasibility of using HET-CAM as a complete replacement for the *in vivo* rabbit ocular test. Most of these reports describe a HET-CAM test method protocol that is similar but not identical to the original protocol. These differences include the breed of hen from which eggs are obtained, the endpoints evaluated, data collection procedures, and methods used to analyze the data.

To date, no single HET-CAM test method protocol has gained wide acceptance as a standardized protocol. However, for a general description of how the HET-CAM test method is conducted, see the ICCVAM *Background Review Document—Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane Test Method* (2006a). Briefly, during a HET-CAM study, the test substance is applied to the surface of the CAM. The CAM is subsequently evaluated for development of irritant endpoints: hemorrhage (bleeding), vascular lysis (blood vessel disintegration), and coagulation (intra- and extravascular protein

denaturation). Depending on the method used to collect data on the endpoints (e.g., time to development, severity of observed effect), qualitative assessments of the irritation potential of test substances are made. As detailed in **Section 6.0**, analyses of each of the HET-CAM analysis methods indicate that the irritation score (A) (IS[A]) analysis method achieved the best performance when evaluating substances not labeled as irritants. Therefore, the IS(A) method is described here. For a description of the other HET-CAM analysis methods (i.e., Q-score, mtc10, ITS, and S-score), see the 2006 ICCVAM BRD (ICCVAM 2006a).

2.1 The Irritation Score (IS) Analysis Method

For those test method protocols that assigned a score to each of the endpoints evaluated at preset time intervals, the values assigned to each endpoint were added to give an irritation score (IS) value for the test substance (i.e., IS[A] analysis method). The possible IS values range from 0 (for test substances that do not induce development of any of the toxic endpoints of interest over the range of time intervals) to 21 (for test substances that induced development of all three toxic endpoints within 30 seconds of application of the test substance) (Luepke 1985).

For those test method protocols that noted the time that a specific endpoint was first observed, the IS value was calculated (i.e., IS[B] analysis method) using the following formula (Kalweit et al. 1987, 1990):

$$\left(\left(\frac{(301-\textit{Hemorrhage time})}{300}\right)\times 5\right)+\left(\left(\frac{(301-\textit{Lysis time})}{300}\right)\times 7\right)+\left(\left(\frac{(301-\textit{Coagulation time})}{300}\right)\times 9\right)$$

where:
Hemorrhage time = time (in seconds) of the first appearance of blood hemorrhages
Lysis time = time (in seconds) of the first appearance of vessel lysis
Coagulation time = time (in seconds) of the first appearance of protein coagulation

The IS value, when calculated using this formula, has a maximal value of 21.

When the development of hyperemia, injection, or another toxic endpoint was evaluated instead of vessel lysis, the time to first appearance for the alternative endpoint replaced the lysis time point.

2.1.1 IS Classification Scheme

For studies that used the analysis methods developed by Luepke (1985) or Kalweit et al. (1987, 1990), the accuracy analysis presented in this BRD (see **Section 6.0**) used the ocular irritancy classification scheme described in **Table 2-1**. Therefore, substances with IS(A) or IS(B) values of 9 or greater were classified as severe irritants for the purposes of this analysis. The rationale for the decision criteria used in this classification scheme were not provided, and the correlation of these categories to irritancy categories described by the EPA (2003), GHS (UN 2007), and EU (2001) classification systems is unknown.

Table 2-1 IS Classification Scheme Used to Classify Substances for Accuracy Analysis[1]

HET-CAM Score Range	Irritation Category
0 to 0.9	Not Labeled
1 to 4.9	Slight Irritation
5 to 8.9	Moderate Irritation
9 to 21	Severe Irritation

[1] According to Luepke (1985) and Kalweit et al. (1987, 1990).

3.0 Substances Used for Validation of the HET-CAM Test Method

3.1 Rationale for the Substances or Products Selected for Use

Validation studies for *in vitro* ocular test methods should ideally evaluate an adequate sample of test substances and products from chemical and product classes that would be evaluated using the *in vivo* rabbit eye test method. Test substances with a wide range of *in vivo* ocular responses (e.g., corrosive/severe irritant to not labeled) also should be assessed to determine any limit to the range of responses that can be evaluated by the *in vitro* test method.

Although new HET-CAM data were identified among four studies published since the ICCVAM evaluation of HET-CAM for identification of ocular corrosives and severe irritants (ICCVAM 2006a), the only substances for which *in vivo* reference data were available were already included in the original HET-CAM database. Therefore, the same database was used in the current evaluation (i.e., Bagley et al. 1992; Balls et al. 1995; CEC 1991; Gettings et al. 1991, 1994, 1996; Gilleron et al. 1996, 1997; Hagino et al. 1999; Kojima et al. 1995; Spielmann et al. 1996;Vinardell and Macián, 1994). As detailed in **Section 6.0**, analyses of each of the multiple HET-CAM protocols indicates that the IS(A) analysis method achieved the best performance when evaluating substances not labeled as irritants. The available database for the IS(A) includes a total of 63 test substances, of which *in vivo* reference data sufficient to assign an ocular irritancy classification are available for 58 - 60 substances depending upon the classification system.

Table 3-1 and **Table 3-2** show the chemical classes and product classes for the test substances included in the original assessment. Information, including substance name, Chemical Abstracts Service Registry Number (CASRN), chemical and/or product class, concentration(s) tested, purity, supplier or source, and literature reference for the test substance are provided in **Annex I**. If not assigned in the study report, the product class was sought from other sources, including the National Library of Medicine's ChemIDplus® database. Chemical classes were assigned to each substance using a standard classification scheme based on the National Library of Medicine Medical Subject Headings (MeSH®) classification system (available at: http://www.nlm.nih.gov/mesh), which ensures consistency in classifying substances among all *in vitro* ocular test methods under consideration. Importantly, a substance could be assigned to more than one chemical or product class.

As shown in **Table 3-1**, the chemical classes with the greatest amount of HET-CAM data are alcohols (n=75), carboxylic acids (n= 51), and formulations (n=53). Of the 504 substances included in **Annex II**, 28 substances, including formulations and mixtures of unknown composition, could not be assigned a specific chemical class.

Table 3-1 Chemical Classes Tested in the HET-CAM Test Method

Chemical Class	# of Substances
Acyl halide	2
Alcohol	75
Aldehyde	9
Alkali	4
Amide	2
Amidine	6
Amine	34
Amino acid	7
Carbohydrate	1
Carboxylic acid	51
Ester	34
Ether	38
Formulation	53
Heterocyclic compound	37
Hydrocarbon, acyclic	5
Hydrocarbon, cyclic	5
Inorganic boron compound	2

Chemical Class	# of Substances
Inorganic salt	14
Imide	4
Ketone	15
Lactone	5
Nitrile	3
Nitro compound	3
Onium compound	22
Organic salt	50
Organometallic compound	2
Organophosphorous compound	1
Organosilicon compound	6
Phenol	4
Polycyclic compound	11
Organic sulfur compound	18
Unknown	28
Urea	3

As shown in **Table 3-2**, the most common product classes tested in the HET-CAM test method are solvents (n=13), hair shampoos (n=13), surfactants (n=17), and cosmetics (n=14). Of the 504 substances included in **Annex II**, 167 were unable to be classified within a product class.

As described in **Section 6.0**, analyses of each of the multiple HET-CAM protocols indicates that the IS(A) analysis method achieved the best performance when evaluating substances not labeled as irritants. The total available database for the IS(A) analysis method includes 63 substances, for which 58–60 substances have available *in vivo* reference data sufficient to assign an ocular irritancy classification depending upon the classification system. Among these substances are 43 cosmetic and personal care product formulations (including 25 surfactant-based formulations and 18 oil/water emulsions) and 17 individual substances (including seven alcohols; no other classes represented by more than three substances).

Table 3-2 Product Classes Tested in the HET-CAM Test Method

Product Class	# of Substances
Aerosol formulation ingredient	1
Antifreezing agent	1
Anti-infective agent, Anti-bacterial agent	2
Antiperspirant	1
Bactericide, Biocide, Fungicide, Germicide	4
Beverage	1
Cationic surface active agent	1
Chemical intermediate	6
Cleaner	1
Conditioner, Hair	2
Cosmetics	14
Cream	1
Disinfectant	1
Drug vehicle	1
Emollient	2
Fertilizer	1
Flavor ingredient	5
Fragrances	4
Industrial explosive	1

Product Class	# of Substances
Laboratory reagent	7
Lotion	3
Lubricant	1
Mouthwash	1
Neurotransmitter	2
Pesticide	5
Pharmaceutical agent, Pharmaceutical intermediate, Pharmaceutical metabolite	4
Plasticizer	2
Polymer	1
Preservative	1
Raw material	1
Shampoo, Hair	13
Solvent	13
Sunscreen	3
Surfactant	17
Synthetic flavor ingredient, Flavor ingredient	4
Synthetic intermediate	1
Unknown	167

4.0 *In Vivo* Reference Data Used for an Assessment of HET-CAM Test Method Accuracy

A detailed description of the test method protocol predominantly used to generate the *in vivo* reference data (i.e., the Draize rabbit eye test) is provided in the *ICCVAM Background Review Document—Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane Test Method* (ICCVAM 2006a). There also are a number of national and international test guidelines that describe this procedure (EPA 1998; OECD 2002; CPSC 2003; EU 2004). The scoring system used for assigning an ocular hazard classification is subjective and based on a discrete scale for grading the severity of ocular lesions on the cornea, iris, and conjunctiva.

Most of the HET-CAM studies evaluated in this BRD include *in vivo* reference data generated using the basic procedures for the *in vivo* rabbit eye test method described above. These data were used by NICEATM to assign an ocular hazard classification according to the EPA (2003a), EU (2001), FHSA (2005), and the GHS (UN 2007) ocular irritancy classification systems (**Annex III**). Exceptions included the *in vivo* data used by Gilleron et al. (1996), which were obtained from the studies of Gautheron et al. (1994). According to the report by Gilleron et al., the studies were performed according to the French and European directives (European Economic Council [EEC] 1984, 1991). Substances were classified by the authors according to the EU (1993) classification system and were used to assess the *in vitro* test method accuracy.

4.1 *In Vivo* Classification Criteria Used for BRD Analysis

As described in the ICCVAM 2006 BRD (2006a), the *in vivo* rabbit eye test database that was used to analyze the accuracy of the HET-CAM test method includes studies that were conducted using from one to six rabbits. However, some of the *in vivo* classification systems considered for the accuracy analyses are designed for application to studies using no more than three rabbits. Thus, to maximize the amount of data used to evaluate the HET-CAM test method, the decision criteria for each classification system were expanded to include studies that used more than three rabbits in their evaluation. The criteria used for classification according to the EPA (2003a), GHS (UN 2007), and EU (2001) classification systems were detailed in the 2006 ICCVAM BRD. Each of these classification systems requires that the Draize scoring system be used. For these classification systems, scoring continues until the effect is cleared, but usually not beyond 21 days after the substance is applied to the eye of the rabbit. In order for a substance to be included in the accuracy evaluations in the 2006 ICCVAM BRD (2006a), the following four criteria must have been met.

- At least three rabbits were tested in the study unless a severe effect (e.g., corrosion of the cornea) was noted in a single rabbit. In such cases, substance classification could proceed based on the effects observed in fewer than three rabbits.
- A volume of 0.1 mL or 0.1 g was tested in each rabbit. A study in which a lower volume was applied to the eye could be accepted for substance classification provided that a severe effect (e.g., corrosion of the cornea, lesion persistence) was observed in a rabbit.
- Observations of the eye were made at least 24, 48, and 72 hours after test substance application if no severe effect was observed.
- Observations of the eye were made until reversibility was assessed, typically meaning that all endpoint scores were cleared. Results from a study terminated early were not used unless the reason for the early termination was documented.

If any of the above criteria were not fulfilled, then the data for that substance were not used for the accuracy analyses. The rules used for classification according to the EPA, EU, or GHS classification systems are detailed in the ICCVAM 2006 BRD (2006a).

For the FHSA classification system (FHSA 2005), the testing guidelines and associated criteria are included in 16 CFR 1500.42 (CPSC 2003). The FHSA classification system is based on using up to three sequential tests for each test substance with six animals used per test (**Table 4-1**). Decisions on further sequential testing are based on the number of positive responses in each test. The severity of effects for each endpoint (i.e., corneal ulceration and opacity, conjunctival redness and/or swelling, and iritis) is measured at 24, 48, and 72 hours after test substance administration. Positive responses include corneal ulceration (other than a fine stippling), corneal opacity or iritis ≥1, and conjunctival swelling and/or redness ≥2. In the first test, six animals are tested. If ≥4 animals are positive, the test is positive. If ≤1 animal tests positive, the test is negative. If 2/6 or 3/6 animals are positive, then a second test is performed with six additional animals. A third test is needed if 1/6 or 2/6 animals are positive with the second test.

The FHSA classification system (FHSA 2005) is a binary system, which classifies substances that test positive (according to the criteria provided in **Table 4-1)** as an irritant and substances that test negative as not requiring labeling (i.e. FHSA Not Labeled). Based on the FHSA sequential testing strategy, a substance can be classified as an eye irritant hazard with a few as 22% of the animals having a positive response (i.e., 2/6 [first test] +1/6 [second test] +1/6 [third test] = 4/18 or 22%).

Because the FHSA classification system is based on a sequential testing strategy, which uses up to 18 animals, only a small percentage of the substances in HET-CAM database would be classifiable if the FHSA criteria were strictly applied. In order to maximize the number of substances include in these analyses, "proportionality" criteria were developed by NICEATM for the purpose of assigning an FHSA classification for test results that would require additional testing according to the FHSA sequential testing strategy (**Table 4-2**).

These "proportionality" criteria (i.e., FHSA-20% and FHSA-67%) are as follows:

- (FHSA-20%) – FHSA-20% is based on the proportion of positive animals needed to identify a substance as an irritant using the FHSA sequential testing strategy, where 20% of the animals need to demonstrate a positive response for a substance to be identified as an irritant. A substance tested using 3 to 6 animals would not be labeled if ≤1/6 animals were positive based on the FHSA criteria. The substance would be labeled as an irritant if there were ≥1 positive animal in a 3- to 5-animal test or ≥2 positive animals in a 6-animal test.
- (FHSA-67%) – FHSA-67% is based on the proportion of positive animals needed to identify a substance as an irritant using the "first test" of the FHSA sequential testing strategy, where 67% of the animals need to demonstrate a positive response for a substance to be identified as an irritant. A substance tested using 3 to 6 animals would not be labeled if ≤1/6 animals were positive based on the FHSA criteria. The substance would be labeled as an irritant if there were ≥2/3, 3/4, 4/5, or 4/6 positive animals. If 1/3, 1/4, 2/4, 1/5, 2/5, 3/5, 2/6, or 3/6 animals were positive, further testing would be required.

Table 4-1 FHSA Classification System (16 CFR 1500.42)[1,2]

Positive Response for a Single Rabbit[3] ≥1 of the following at 24, 48, and/or 72 hours	*In Vivo* Effect
• Corneal ulceration (other than a fine stippling) • Corneal opacity (CO) ≥1 • Iritis (IR) ≥1 • Conjuctival redness (CR) and/or chemosis (CC) ≥2	First Test – If ≥4/6 animals are positive, the test is positive. If ≤1 animal is positive, the test is negative. If 2/6 or 3/6 animals are positive, the test is repeated using a different group of six animals. Second Test – If ≥3/6 animals are positive, the test is positive. If 0/6 animals are positive, the test is negative. If 1/6 or 2/6 animals are positive, the test is repeated using a different group of six animals. Third Test – Should a third test be needed, the test is positive if ≥1/6 animals are positive. If 0/6 animals are positive, the test is negative.

Abbreviations: CC = conjunctival chemosis; CFR = Code of Federal Regulations; CO = corneal opacity; CR = conjunctival redness; FHSA = Federal Hazardous Substances Act; IR = iritis.

[1] For the FHSA Classification System (2005), the testing guidelines and associated criteria are included in 16 CFR 1500.42 (CPSC 2003).

[2] At least three animals per test (one animal screen for corrosive/severe irritants permitted). Maximum score in any animal used for classification.

[3] The following scores are considered positive: CO or IR ≥1 or CR or CC ≥2. Therefore, CO and IR scores of 0 or CR and CC scores ≤1 are considered negative.

Table 4-2 Proposed FHSA "Proportionality" Criteria

No. of Animals in Test	FHSA-20%[1]		FHSA-67%[1]		
	NL	Irritant	NL	Irritant	Further Testing Required[2]
3	0/3	≥1 (≥33%)	0/3	≥2 (≥67%)	1/3
4	0/4	≥1 (≥25%)	0/4	≥3 (≥75%)	1/4, 2/4
5	0/5	≥1 (≥20%)	0/5	≥4 (≥80%)	1/5, 2/5, 3/5
6	0/6, 1/6	≥2 (≥33%)	0/6, 1/6	≥4 (≥67%)	2/6, 3/6

Abbreviations: CPSC = U.S. Consumer Product Safety Commission; FHSA = Federal Hazardous Substances Act; NL = Not Labeled (as an irritant); No. = number.

[1] FHSA-20% and FHSA-67% analysis methods are based on the proportionality of positive animals needed to identify a substance as an irritant.

[2] For FHSA-67%, Further Testing Required refers to substances that do not meet adequate positive or negative criteria to be classified.

4.2 *In Vivo* Data Quality

Ideally, all data supporting the validity of a test method should be obtained and reported in accordance with Good Laboratory Practice (GLP) guidelines. GLP guidelines are nationally and internationally recognized rules designed to produce high-quality laboratory records (OECD 1998; EPA 2003b, 2003c; FDA 2003). These guidelines provide an internationally standardized approach

for the conduct of studies, reporting requirements, archival of study data and records, and information about the test protocol, thereby ensuring the integrity, reliability, and accountability of a study.

The extent to which the *in vivo* rabbit eye studies that were used to provide the comparative data in the published HET-CAM validation studies complied with GLP guidelines is based on the information provided in the published reports. Based on the available information, the reports that were identified as following GLP guidelines or used data obtained according to GLP guidelines were Gettings et al. (1991, 1994, 1996), Balls et al. (1995), Spielmann et al. (1996), and Hagino et al. (1999).

5.0 Hen's Egg Test–Chorioallantoic Membrane Test Method Data and Results

The following twelve published reports contained sufficient data for an accuracy analysis of the HET-CAM test method for the identification of all categories of ocular irritation: CEC (1991), Gettings et al. (1991, 1994, 1996), Bagley et al. (1992), Vinardell and Macián (1994), Balls et al. (1995), Kojima et al. (1995), Gilleron et al. (1996, 1997), Spielmann et al. (1996), and Hagino et al. (1999).

5.1 Availability of Copies of Original Data Used to Evaluate the Accuracy and Reliability

On March 24, 2004, NICEATM published a *Federal Register* notice requesting original HET-CAM data for substances that also had been tested *in vivo* using the standard rabbit eye test (69 FR 13589; available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_04_6487.pdf). A second request was published on February 28, 2005 (70 FR 9661; available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_05_3831.pdf). In addition, NICEATM contacted authors of selected published HET-CAM studies and requested the original HET-CAM data. In response to these efforts, the following *in vitro* data were obtained:

- Summaries of HET-CAM results (e.g., Q-scores) for the 60 substances evaluated by Balls et al. (1995) from the European Centre for the Validation of Alternative Methods (ECVAM). The summary data included the substance name and the average HET-CAM score for the substance.
- *In vitro* data for the substances evaluated in Spielmann et al. (1996) from Drs. H. Spielmann and M. Liebsch. The data included the overall HET-CAM scores obtained by each laboratory for each substance evaluated. *In vitro* data for two control substances also were provided.
- Individual endpoint scores for each egg evaluated for substances described in Gilleron et al. (1996, 1997) from Drs. Philippe Vanparys and Freddy Van Goethem. *In vitro* data for four control substances also were provided.

5.2 Description of the Statistical Approaches Used to Evaluate the Resulting Data

The approach used to analyze HET-CAM study data varied and depended on the method used to collect the data. For test method protocols that evaluated the time to development of endpoints (i.e., hemorrhage, lysis, coagulation) that are correlated with ocular corrosivity or irritation, an IS, Q-score, or mean time of coagulation (mtc) value was calculated. For test method protocols that evaluated the severity of the toxic response, an S-score was calculated. For test method protocols that evaluated the lowest test substance concentration needed to produce a minimal response on the CAM, the irritation threshold concentration was determined. The irritation threshold concentration was typically combined with the IS for the test substance to evaluate ocular irritation or corrosivity potential of a substance.

The accuracy analysis in this BRD focuses on the ability of the HET-CAM test method to identify all irritant hazard categories (i.e., moderate and mild irritants) and/or substances not labeled as irritants as defined by the EPA, GHS, and EU classification systems (EPA 2003a; UN 2007; EU 2001). However, multiple irritancy schemes have been developed for HET-CAM, and different scoring methods and decision criteria were used. No single uniform irritancy classification scheme was developed for HET-CAM. Furthermore, the *in vitro* hazard classifications were not always consistent with or applicable to those based on Draize rabbit eye test data used by the U.S. (EPA 2003a), the GHS (UN 2007), or the EU (EU 2001). However, some investigators have tried to correlate HET-CAM scores with the ocular irritation classification scheme described by the Federal Hazardous Substances Act classification system (CPSC 1988) and the EU classification system (EU 1992) (Gettings et al. 1991, 1994, 1996; Spielmann et al. 1996, respectively).

To evaluate the ability of HET-CAM to identify all ocular hazard categories or substances not labeled as irritants, NICEATM assigned HET-CAM results obtained using each of the different analysis methods an ocular irritancy classification based on the *in vitro* classification system most commonly used for that particular data analysis method. Thus, substances were classified in categories based on the *in vitro* score. Categories ranged from substances not labeled as irritants to ocular corrosives or severe irritants (see **Section 2.0**). Some investigators (e.g., Gettings et al. 1996) classified the ocular irritancy potential of test substances using two or more different analysis methods. In such cases, these data were reclassified according to the approach used most commonly for each *in vitro* classification scheme, and an accuracy assessment was conducted for each analysis method.

NICEATM's preliminary evaluation using the various analysis methods (see **Section 6.1** and **Annex III**) indicated that only the IS(A) analysis method had adequate accuracy to conduct a study of mild/moderate ocular irritation based on rabbit eye test data. Therefore, the data was limited to 63 test substances obtained from Bagley et al. (1992), Gettings et al. (1994, 1996), Kojima et al. (1995), and Hagino et al. (1999).

5.3 Summary of Results

A total of 260 test substances were evaluated in 383 HET-CAM studies for which comparative *in vivo* data were available (ICCVAM 2006a). A summary of results used to evaluate test method accuracy appears in **Annex III**. This table, sorted by reference, provides the following specifics, if provided:

- Name
- CASRN (if available)
- Chemical class
- Product class
- Concentration tested
- Form tested
- Calculated *in vitro* score
- *In vitro* irritation classification of the test substance (based on the irritation classification schemes in **Section 5.3**)
- *In vivo* reference classifications (i.e., EPA, GHS, EU)
- Literature source

Other supporting information, such as purity of the test substance, was included in the table to the extent that this information was available. If not provided, the CASRN was obtained from various sources, including the National Library of Medicine's ChemIDplus® database (available at http://chem2.sis.nlm.nih.gov/chemidplus). All substances with the same CASRN were listed under the same name, regardless of the synonym used in the original report. Chemical and product classes were assigned to each test substance based on the National Library of Medicine's Medical Subject Heading classification system (MeSH®; available at http://www.nlm.nih.gov/mesh). **Annex I**

provides information on the names, synonyms, CASRN, and chemical/product class, where available, for each substance. **Annex II** provides the *in vitro* HET-CAM test method data sorted by reference and alphabetically by substance name.

5.4 Use of Coded Chemicals and Compliance with GLP Guidelines

Ideally, all data supporting the validity of a test method should be obtained and reported in accordance with GLP guidelines and with the use of coded chemicals (OECD 1998; EPA 2003b, 2003c; FDA 2003). The data quality was evaluated by reviewing the methods section in literature references and the submitted reports. Thus, data quality presented in the reviewed literature references can be evaluated only to the extent such information was provided in the published reports. Based on the available information, the following reports were identified as following GLP guidelines or using data obtained according to GLP guidelines: Gettings et al. (1991, 1994, 1996), Balls et al. (1995), Spielmann et al. (1996), and Hagino et al. (1999).

Detailed information on coding procedures used in different studies is provided in Section 3.4 of the ICCVAM 2006 BRD (2006a).

6.0 Hen's Egg Test–Chorioallantoic Membrane Test Method Accuracy

6.1 Accuracy of the HET-CAM Test Method

A critical component of an ICCVAM evaluation of a test method's validation status is an assessment of the proposed test method's accuracy compared to that of the current reference test method (ICCVAM 2003). This aspect of assay performance is typically evaluated by calculating:

- *Accuracy* (concordance): the proportion of correct outcomes (positive and negative) of a test method
- *Sensitivity*: the proportion of all positive substances that are classified as positive
- *Specificity*: the proportion of all negative substances that are classified as negative
- *Positive predictivity*: the proportion of correct positive responses among substances testing positive
- *Negative predictivity*: the proportion of correct negative responses among substances testing negative
- *False positive rate*: the proportion of all negative substances that are falsely identified as positive
- *False negative rate*: the proportion of all positive substances that are falsely identified as negative

ICCVAM evaluated the ability of the HET-CAM test method to identify all categories of ocular irritation potential as defined by the EPA, GHS, and EU classification systems (EPA 2003a; UN 2007; EU 2001). Given that the "Test for Eye Irritants" (16 CFR 1500.42) used for FHSA classification does not discriminate severe or corrosive effects from eye irritation in the rabbit, an evaluation for all ocular hazard categories using the FHSA classification system was not performed. This same analysis was also performed with specific chemical classes and/or physical properties excluded based on their previous identification as discordant in the HET-CAM test method (ICCVAM 2006a).

These evaluations were conducted on the overall data set created by combining results from the reports discussed in **Section 5.0**, then assigning an overall ocular irritancy classification for each substance (see **Annex II** and **III**). When the same substance was evaluated in multiple laboratories, an overall HET-CAM classification was based on the majority classification among all of the studies. When there were an equal number of differing irritancy classifications for substances (e.g., two tests

classified a substance as not labeled and two tests classified the same substance as a mild irritant), the more severe irritancy classification was used for the overall classification for the substance (mild irritant, in this case).

ICCVAM analyzed HET-CAM performance compared to the Draize rabbit eye test for each classification system (i.e., EPA, GHS, EU) using each of the six HET-CAM protocols (i.e., IS[A], IS[B], Q-score, S-score, IS, and irritation threshold concentration protocols, see **Annex III**). With the exception of the IS(A) and IS(B) protocols, all analysis methods had at least one *in vivo* moderate or severe irritant substance classified *in vitro* as not labeled as an irritant (i.e., EPA Category IV, GHS Not Labeled as Irritant, EU Not Labeled). However, the IS(B) overclassified most of the Not Classified Substances (e.g., HET-CAM IS[B] overclassified 93% [39/42] of the GHS Not Labeled as Irritant substances). Therefore, more extensive analyses of the HET-CAM test method described in the following sections were restricted to the IS(A) protocol.

6.1.1 GHS Classification System: HET-CAM Test Method Accuracy

Five studies (Bagley et al. 1992; Gettings et al. 1994; Gettings et al. 1996; Hagino et al. 1999; Kojima et al. 1995) contained HET-CAM data for 63 substances, 59 of which had sufficient *in vivo* data to be assigned GHS ocular irritant classifications (UN 2007) (see **Annex III**). For three of these studies (Gettings et al.1994, 1996; Hagino et al. 1999), ICCVAM evaluated each individual study separately. Individual analyses were not conducted on the other two studies (Bagley et al. 1992; Kojima et al. 1995) because they contained data for only one and two substances, respectively. Based on *in vivo* rabbit eye test data, 44% (26/59) of substances were classified as Category 1; none was classified as Category 2A; 8% (5/59) were classified as Category 2B, and 47% (28/59) were not classified as irritants. Four substances could not be classified due to lack of adequate animal data and are so noted in **Annex III**.

Identification of Category 1 Substances (Ocular Corrosives/Severe Irritants)

The HET-CAM test method correctly identified 50% (13/26) of the Category 1 substances (**Table 6-1**). Among the remaining 50% (13/26) of Category 1 substances underpredicted by HET-CAM, 42% (11/26) were classified as Category 2A and 8% (2/26) were classified as Category 2B.

Identification of Category 2A Substances (Moderate Ocular Irritants)

No substances were identified as GHS Category 2A irritants *in vivo*, and the HET-CAM test method did not mislabel any other substances as moderate ocular irritants (**Table 6-1**).

Identification of Category 2B Substances (Mild Ocular Irritants)

For the five substances that could be evaluated, the HET-CAM test method correctly identified 20% (1/5) as Category 2B, while 80% (4/5) were overpredicted and 0% (0/5) were underpredicted (**Table 6-1**).

Identification of Not Classified Substances

For the 28 substances that could be evaluated, the HET-CAM test method correctly identified 36% (10/28) as substances not classified as irritants, while 64% (18/28) were overpredicted (**Table 6-1**).

Ability to Distinguish Substances Not Classified as Irritants from All Other Classes

In addition to evaluating the ability of the HET-CAM test method to identify each individual ocular hazard category according to the GHS classification system, ICCVAM also evaluated the ability of the HET-CAM test method to distinguish ocular substances not classified as irritants from all irritant

classes.[1] For the 59 substances considered, the HET-CAM test method had an overall accuracy of 69% (41/59), a sensitivity of 100% (31/31), a specificity of 36% (10/28), a false positive rate of 64% (18/28), and a false negative rate of 0% (0/31) (**Table 6-2**).

As detailed below, the results from each individual study were also evaluated separately.

Gettings et al. (1994): Based upon the *in vivo* rabbit data, 18 substances were assigned a GHS classification. The HET-CAM test method, by comparison, has an accuracy of 50% (9/18), sensitivity of 100% (1/1), specificity of 47% (8/17), false positive rate of 53% (9/17), and a false negative rate of 0% (0/1) (**Table 6-2**).

Gettings et al. (1996): Based on the *in vivo* rabbit data, 24 substances could be assigned a GHS classification. Among these 24 substances, the HET-CAM test method has an accuracy of 83% (20/24), sensitivity of 100% (18/18), specificity of 33% (2/6), false positive rate of 67% (4/6), and a false negative rate of 0% (0/18) (**Table 6-2**).

[1] The ICCVAM 2006 BRD provides an evaluation of the HET-CAM test method for distinguishing ocular corrosives and severe irritants from all other classes (ICCVAM 2006a). Because the database of HET-CAM test method results has not changed, this analysis is not repeated here.

Table 6-1 Performance of the HET-CAM Test Method (IS[A]) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the GHS Classification System,[1] by Study and Overall

Data Source	Overall Correct Classification	Severe (Category 1)		Moderate (Category 2A)			Mild (Category 2B)			Not Classified as Irritant	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Gettings et al. (1994)	50% (9/18)	100% (1/1)	0% (0/1)	0% (0/0)	0% (0/0)	0% (0/0)	0% (0/0)	0% (0/0)	0% (0/0)	53% (9/17)	47% (8/17)
Gettings et al. (1996)	29% (7/24)	25% (4/16)	75% (12/16)	0% (0/0)	0% (0/0)	0% (0/0)	50% (1/2)	50% (1/2)	0% (0/2)	67% (4/6)	33% (2/6)
Hagino et al. (1999)	53% (8/15)	100% (8/8)	0% (0/8)	0% (0/0)	0% (0/0)	0% (0/0)	100% (3/3)	0% (0/3)	0% (0/3)	100% (4/4)	0% (0/4)
Overall[2]	41% (24/59)	50% (13/26)	50% (13/26)	0% (0/0)	0% (0/0))	0% (0/0)	80% (4/5)	20% (1/5)	0% (0/5)	64% (18/28)	36% (10/28)

Abbreviations: GHS = Globally Harmonized System; HET-CAM = hen's egg test–chorioallantoic membrane.

[1] GHS classification system (UN 2007).

[2] Overall data set contains 59 test substances that were assigned a GHS classification and includes one additional test substance from Bagley et al. (1992) and one from Kojima et al. (1995) that were not included as individual data sources. One additional substance from Kojima et al. (1995) was not included because it was classified *in vitro* as Category1/Category 2A in the rabbit eye test.

Table 6-2 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Substances Not Classified as Irritants from All Other Irritant Classes, as Defined by the GHS Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gettings et al. (1994)	18	50	9/18	100	1/1	47	8/17	53	9/17	0	0/1
Gettings et al. (1996)	24	83	20/24	100	18/18	33	2/6	67	4/6	0	0/18
Hagino et al. (1999)	15	73	11/15	100	11/11	0	0/4	100	4/4	0	0/11
Overall[2]	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31

Abbreviations: GHS = Globally Harmonized System; HET–CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Not Classified as Irritant vs Category 1/2A/2B.

[2] Overall data set contains 59 test substances that were assigned a GHS hazard classification. Data from one additional test substance from Bagley et al. (1992) and one from Kojima et al. (1995) were not included as individual data sources. One additional substance from Kojima et al. (1995) was not included because it was classified *in vitro* as Category1/Category 2A in the rabbit eye test.

Hagino et al. (1999): Based upon the *in vivo* rabbit data, 15 substances could be assigned a GHS classification. Among these 15 substances, the HET-CAM test method has an accuracy of 73% (11/15), sensitivity of 100% (11/11), specificity of 0% (0/4), false positive rate of 100% (4/4), and a false negative rate of 0% (0/11) (**Table 6-2**).

Performance of the HET-CAM Test Method with Discordant Classes Excluded

Because the IS(A) analysis method is the focus of the evaluation of HET-CAM for identifying all hazard categories, separate analyses were also conducted for all chemical classes and specific physical properties of interest represented in this database of 59 substances by at least five substances (i.e., surfactant-based formulations, oil/water emulsions, and alcohols). The results indicate that alcohols tend to be overpredicted by HET-CAM: 75% (4/6) of alcohols classified as Category 2B or Not Classified as Irritant based on Draize test results, and depending on the classification system used, were overpredicted by HET-CAM by at least one hazard category. Similarly, the HET-CAM test method overpredicted 53% (9/17) of the oil/water emulsions identified as Not Classified as Irritant by at least one hazard category. By comparison, surfactant formulations classified as Category 1 based on Draize results tended to be underpredicted by HET-CAM: 75% (12/16) were underpredicted by HET-CAM as Category 2A or 2B. However, none of these substances was underpredicted as Not Classified as Irritant.

Given the proportion of substances in the HET-CAM IS(A) database represented by these chemical and product classes (i.e., 85% [50/59] of the substances are included in one of these three categories), separate analyses without these discordant substances are not particularly informative. However, because of the associated discordance with each type, overall performance, particularly for Category 1 substances, can be improved by excluding surfactant-based formulations (see **Table 6-3**).

When the ability of the HET-CAM test method to distinguish Not Classified as Irritant substances from all other irritant classes was evaluated with the specific chemical and product classes removed, the greatest improvement in false positive rate occurred when alcohols and surfactant formulations were excluded. The false positive rate decreased from 64% (18/28) to 56% (10/18). However, because the false negative rate for the overall database is 0% (0/31), this rate remained constant regardless of which chemical or product class(es) were excluded (**Table 6-4**).

Further analysis of substances for which hazard classification was underpredicted by HET-CAM according to chemical class indicated that carboxylic acids had the highest proportion of underpredicted substances (25% [1/4]). Because 98% of the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of underpredicted substances was liquids. Among the 16 Category 1 surfactants, HET-CAM underpredicted 75% (12/16) (**Table 6-5**).

According to the GHS classification system, the most overpredicted substances (false positives) were alcohols, of which HET-CAM overpredicted 75% (6/8). Because 98% of the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of underpredicted substances was liquids. Only one of the surfactants tested in HET-CAM was overpredicted (**Table 6-5**).

Table 6-3 Performance of the HET-CAM Test Method (IS[A]) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the GHS Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	Overall Correct Classification	Severe (Category 1)		Moderate (Category 2A)			Mild (Category 2B)			Not Classified as Irritant	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall	41% (24/59)	50% (13/26)	50% (13/26)	- (0/0)	- (0/0)	- (0/0)	80% (4/5)	20% (1/5)	0% (0/5)	64% (18/28)	36% (10/28)
Without Alcohols	43% (22/51)	46% (11/24)	54% (13/24)	- (0/0)	- (0/0)	- (0/0)	67% (2/3)	33% (1/3)	0% (0/3)	58% (14/24)	42% (10/24)
Without Surfactant Formulations	49% (17/35)	90% (9/10)	10% (1/10)	- (0/0)	- (0/0)	- (0/0)	100% (3/3)	0% (0/3)	0% (0/3)	64% (14/22)	36% (8/22)
Without Oil/Water Emulsions	41% (15/41)	48% (12/25)	52% (13/25)	- (0/0)	- (0/0)	- (0/0)	80% (4/5)	20% (1/5)	0% (0/5)	82% (9/11)	18% (2/11)
Without Alcohols and Surfactant Formulations	56% (15/27)	87% (7/8)	12% (1/8)	- (0/0)	- (0/0)	- (0/0)	100% (1/1)	0% (0/1)	0% (0/1)	56% (10/18)	44% (8/18)
Without Alcohols and Oil/Water Emulsions	39% (13/33)	44% (10/23)	56% (13/23)	- (0/0)	- (0/0)	- (0/0)	67% (2/3)	33% (1/3)	0% (0/3)	71% (5/7)	29% (2/7)
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	67% (6/9)	86% (6/7)	14% (1/7)	- (0/0)	- (0/0)	- (0/0)	100% (1/1)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/1)

Abbreviations: GHS = Globally Harmonized System; HET-CAM = hen's egg test–chorioallantoic membrane.

[1] GHS classification system (UN 2007).

Table 6-4 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Substances Not Classified as Irritants from All Other Irritant Classes, as Defined by the GHS Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31
Without Alcohols	51	73	37/51	100	27/27	42	10/24	58	14/24	0	0/27
Without Surfactant Formulations	35	60	21/35	100	13/13	36	8/22	64	14/22	0	0/13
Without Oil/Water Emulsions	41	78	32/41	100	30/30	18	2/11	82	9/11	0	0/30
Without Alcohols and Surfactant Formulations	27	63	17/27	100	9/9	44	8/18/	56	10/18	0	0/9
Without Alcohols and Oil/Water Emulsions	33	85	28/33	100	26/26	29	2/7	71	5/7	0	0/26
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	89	8/9	100	8/8	0	0/1	100	1/1	0	0/8

Abbreviations: GHS = Globally Harmonized System; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007).

Table 6-5 Performance of the HET-CAM Test Method (IS[A]) Using the GHS Classification System[1] in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property

Category	N	Underprediction (*In Vivo/In Vitro)*						Overprediction (*In Vivo/In Vitro)*					
		Severe (Category 1)			Moderate (Category 2A)		Mild (Category 2B)	Moderate (Category 2A)	Mild (Category 2B)		NC (Not Classified)		
		NC	2B	2A	NC	2B	NC	1	2A	1	2B	2A	1
Overall	59	0% (0/26)	8% (2/26)	42% (11/26)	-	-	0% (0/5)	-	20% (1/5)	60% (3/5)	32% (9/28)	14% (4/28)	18% (5/28)
Chemical Class[2]													
Alcohol	8	0% (0/2)	0% (0/2)	0% (0/2)	-	-	0% (0/2)	-	0% (0/2)	100% (2/2)	0% (0/4)	50% (2/4)	50% (2/4)
Carboxylic acid	5	0% (0/4)	0% (0/4)	25% (1/4)	-	-	0% (0/1)	-	0% (0/1)	100% (1/1)	-	-	-
Organic salt	6	0% (0/6)	0% (0/6)	17% (1/6)	-	-	-	-	-	-	-	-	-
Properties of Interest													
Liquids	58	0% (0/25)	8% (2/25)	40% (10/25)	0% (0/5)	-	0% (0/2)	-	20% (1/5)	60% (3/5)	32% (9/28)	14% (4/28)	18% (5/28)
Solids	0	-	-	-	-	-	-	-	-	-	-	-	-
Pesticide	0	-	-	-	-	-	-	-	-	-	-	-	-
Surfactant—Total	24	0% (0/16)	12% (2/16)	62% (10/16)	-	-	0% (0/2)	-	50% (1/2)	0% (0/2)	0% (0/6)	0% (0/6)	0% (0/6)
-nonionic	-	-	-	-	-	-	-	-	-	-	-	-	-
-anionic	-	-	-	-	-	-	-	-	-	-	-	-	-
-cationic	-	-	-	-	-	-	-	-	-	-	-	-	-

Category	N	Underprediction (*In Vivo/In Vitro*)						Overprediction (*In Vivo/In Vitro*)					
		Severe (Category 1)			Moderate (Category 2A)		Mild (Category 2B)	Moderate (Category 2A)	Mild (Category 2B)		NC (Not Classified)		
		NC	2B	2A	NC	2B	NC	1	2A	1	2B	2A	1
Overall	59	0% (0/26)	8% (2/26)	42% (11/26)	-	-	0% (0/5)	-	20% (1/5)	60% (3/5)	32% (9/28)	14% (4/28)	18% (5/28)
Oil/Water Emulsion	18	0% (0/1)	0% (0/1)	0% (0/1)	-	-	-	-	-	-	24% (4/17)	12% (2/17)	18% (3/17)
pH—Total	0	-	-	-	-	-	-	-	-	-	-	-	-
-acidic (pH <7.0)	-	-	-	-	-	-	-	-	-	-	-	-	-
-basic (pH >7.0)	-	-	-	-	-	-	-	-	-	-	-	-	-

Abbreviations: GHS = Globally Harmonized System; HET-CAM = hen's egg test – chorioallantoic membrane; NC = Not Classified as Irritant.

[1] GHS classification system (UN 2007).

[2] Chemical classes included in this table are represented by at least five substances tested in the HET-CAM test method, and assignments are based upon National Library of Medicine medical subject heading (MeSH) categories (www.nlm.nih.gov/mesh) as defined in **Annex I**.

Table 6-6 Performance of the HET-CAM Test Method (IS[A]) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System[1], by Study and Overall

Data Source	Overall Correct Classification	Severe (Category I)		Moderate (Category II)			Mild (Category III)			Not Labeled (Category IV)	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Gettings et al. (1994)	33% (6/18)	100% (1/1)	0% (0/1)	0% (0/0)	0% (0/0)	0% (0/0)	38% (3/8)	12% (1/8)	50% (4/8)	56% (5/9)	44% (4/9)
Gettings et al. (1996)	36% (9/25)	24% (4/17)	76% (13/17)	0% (0/0)	0% (0/0)	0% (0/0)	25% (1/4)	75% (3/4)	0% (0/4)	50% (2/4)`	50% (2/4)
Hagino et al. (1999)	47% (7/15)	100% (7/7)	0% (0/7)	100% (1/1)	0% (0/1)	0% (0/1)	100% (5/5)	0% (0/5)	0% (0/5)	100% (2/2)	0% (0/2)
Overall[2]	38% (23/60)	48% (12/25)	52% (13/25)	50% (1/2)	50% (1/2)	0% (0/2)	56% (10/18)	22% (4/18)	22% (4/18)	60% (9/15)	40% (6/15)

Abbreviations: EPA = U.S. Environmental Protection Agency; HET-CAM = hen's egg test–chorioallantoic membrane.

[1] EPA classification system (EPA 2003a).

[2] Overall data set includes 60 test substances that were assigned an EPA hazard classification based on rabbit eye test data. Data from one test substance from Bagley et al. (1992) and one from Kojima et al. (1995) were not included as individual data sources. One substance from Kojima et al. (1995) was classified as a GHS Category 1/2A and could not be used in the analysis.

6.1.2 EPA Classification System: HET-CAM Test Method Accuracy

Five studies (Bagley et al. 1992; Gettings et al. 1994; Kojima et al. 1995; Gettings et al. 1996; Hagino et al. 1999) contained HET-CAM test method data on 63 substances, 60 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the EPA classification system (EPA 2003a) (see **Annex III**). Based on results from *in vivo* rabbit eye experiments, 42% (25/60) were classified as severe irritants (i.e., Category I), 3% (2/60) were classified as moderate irritants (Category II), 30% (18/60) were classified as mild irritants (Category III), and 25% (15/60) were classified as not labeled as irritant (Category IV). Three substances could not be classified according to the EPA classification system due to the lack of adequate animal data and are so noted in **Annex III**.

Identification of Category I Substances (Ocular Corrosives/Severe Irritants)

The HET-CAM test method correctly identified 48% (12/25) of the Category I substances (**Table 6-6**). Among the remaining 52% (13/25) Category I substances that were underpredicted by HET-CAM, 40% (10/25) were classified as Category II, and 12% (3/25) were classified as Category III.

Identification of Category II Substances (Moderate Ocular Irritants)

For the two substances that could be evaluated, the HET-CAM test method correctly identified 50% (1/2) as Category II while 50% (1/2) were overpredicted and 0% (0/2) were underpredicted (**Table 6-6**).

Identification of Category III (Mild Ocular Irritants)

For the 18 substances that could be evaluated, the HET-CAM test method correctly identified 22% (4/18) as Category III while 56% (10/18) were overpredicted and 22% (4/18) were underpredicted (**Table 6-6**).

Identification of Category IV Substances

For the 15 substances that could be evaluated, the HET-CAM test method correctly identified 40% (6/15) as substances not labeled as irritants while 60% (9/15) were overpredicted (**Table 6-6**).

Ability to Distinguish Category IV Substances from All Other Classes

In addition to evaluating the ability of the HET-CAM test method to identify each individual ocular hazard category according to the EPA classification system, ICCVAM also evaluated the ability of the HET-CAM test method to distinguish ocular substances not labeled as irritants from all irritant classes.[2] Among the 60 substances considered, the HET-CAM test method had an overall accuracy of 78% (47/60), a sensitivity of 91% (41/45), a specificity of 40% (6/15), a false positive rate of 60% (9/15), and a false negative rate of 9% (4/45) (**Table 6-7**).

As detailed below, the results from each individual study were also evaluated separately.

Gettings et al. (1994): Based upon the *in vivo* rabbit data, 18 substances were assigned an EPA classification. The HET-CAM test method, by comparison, has an accuracy of 50% (9/18), sensitivity of 56% (5/9), specificity of 44% (4/9), false positive rate of 56% (5/9), and a false negative rate of 44% (4/9) (**Table 6-7**).

[2] The ICCVAM 2006 BRD (2006a) provides an evaluation of the HET-CAM test method for distinguishing ocular corrosives and severe irritants from all other classes (ICCVAM 2006a). Because the database of HET-CAM test method results has not changed, this analysis is not repeated here.

Table 6-7 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Category IV Substances from All Other Irritant Classes as Defined by the EPA Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gettings et al. (1994)	18	50	9/18	56	5/9	44	4/9	56	5/9	44	4/9
Gettings et al. (1996)	25	92	23/25	100	21/21	50	2/4	50	2/4	0	0/21
Hagino et al. (1999)	15	87	13/15	100	13/13	0	0/2	100	2/2	0	0/13
Overall[2]	60	78	47/60	91	41/45	40	6/15	60	9/15	9	4/45

Abbreviations: EPA = U.S. Environmental Protection Agency; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] EPA classification system (EPA 2003a): Category IV vs. Categories I/II/III.

[2] Overall database includes 60 test substances that were assigned an EPA hazard classification based on rabbit eye test data. Data on one test substance from Bagley et al. (1992) and another substance from Kojima et al. (1995) were not included as individual data sources. One substance from Kojima et al. (1995) was classified as a GHS Category 1/2A and, therefore, was not used in the analysis either.

Gettings et al. (1996): Based upon the *in vivo* rabbit data, 25 substances were assigned an EPA classification. The HET-CAM test method, by comparison, has an accuracy of 92% (23/25), sensitivity of 100% (21/21), specificity of 50% (2/4), false positive rate of 50% (2/4), and a false negative rate of 0% (0/21).

Hagino et al. (1999): Based upon the *in vivo* rabbit data, 15 substances were assigned an EPA classification. The HET-CAM test method, by comparison, has an accuracy of 87% (13/15), sensitivity of 100% (13/13), specificity of 0% (0/2), false positive rate of 100% (2/2), and a false negative rate of 0% (0/13).

Performance of the HET-CAM Test Method with Discordant Classes Excluded

Because the IS(A) analysis method is the focus of the evaluation of HET-CAM for identifying all hazard categories, separate analyses were also conducted for all chemical classes and specific physical properties of interest represented in this database of 60 substances by at least five substances (i.e., surfactant-based formulations, oil/water emulsions, and alcohols).

Given the proportion of substances in the HET-CAM IS(A) database represented by these chemical and product classes (i.e., 85% [51/60] of the substances are included in one of these three categories), separate analyses without these discordant substances are not particularly informative. However, because of the associated discordance with each type, overall performance, particularly for the ocular corrosive and severe irritant category, can be improved by excluding certain product types (see **Table 6-8**). The results indicate that HET-CAM tends to overpredict alcohols. All seven alcohols (100%) classified as Category III or IV based on Draize test results were overpredicted by HET-CAM by at least one hazard category. Similarly, 47% (8/17) of the oil/water emulsions classified as Category III or IV based on Draize test results were overpredicted by HET-CAM by at least one hazard category. By comparison, surfactant formulations classified as Category I based on Draize results tended to be underpredicted by HET-CAM (73% [13/17] were underpredicted by HET-CAM as Category II or III). However, none of these substances was underpredicted as Category IV.

When the ability of the HET-CAM test method to distinguish Category IV substances from all other irritant classes was evaluated with the specific chemical and product classes removed, the greatest improvement in false positive rate occurred when alcohols and surfactant-based formulations were excluded. The false positive rate decreased from 60% (9/15) to 56% (5/9). The false negative rate for the overall database, 9% (4/45), could be reduced to 0% (0/30) by excluding oil/water emulsions from the database (**Table 6-9**).

Among the four false negatives for the EPA system, 100% (4/4) were EPA Category III substances based on Draize data. For 100% (4/4) of these substances, the categorization was based on conjunctival redness (**Table 6-10**). All of the false negative substances were oil/water emulsions.

Table 6-8 Performance of the HET-CAM Test Method (IS[A]) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	Overall Correct Classification	Severe (Category I)		Moderate (Category II)			Mild (Category III)			Not Labeled (Category IV)	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall	41% (24/59)	50% (13/26)	50% (13/26)	0% (0/0)	0% (0/0)	0% (0/0)	80% (4/5)	20% (1/5)	0% (0/5)	64% (18/28)	36% (10/28)
Without Alcohols	42% (22/52)	46% (11/24)	54% (13/24)	50% (1/2)	50% (1/2)	0% (0/2)	38% (5/13)	31% (4/13)	31% (4/13)	54% (7/13)	46% (6/13)
Without Surfactant Formulations	40% (14/35)	100% (8/8)	0% (0/8)	50% (1/2)	50% (1/2)	0% (0/2)	64% (9/14)	7% (1/14)	29% (4/14)	64% (7/11)	36% (4/11)
Without Oil/Water Emulsions	37% (15/41)	48% (12/25)	52% (13/25)	0% (0/0)	0% (0/0)	0% (0/0)	80% (4/5)	10% (1/5)	0% (0/5)	82% (9/11)	18% (2/11)
Without Alcohols and Surfactant Formulations	48% (13/27)	100% (7/7)	0% (0/7)	50% (1/2)	50% (1/2)	0% (0/2)	44% (4/9)	11% (1/9)	44% (4/9)	56% (5/9)	44% (4/9)
Without Alcohols and Oil/Water Emulsions	47% (16/34)	43% (10/23)	57% (13/23)	50% (1/2)	50% (1/2)	0% (0/2)	40% (2/5)	60% (3/5)	0% (0/5)	50% (2/4)	50% (2/4)
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	78% (7/9)	100% (6/6)	0% (0/6)	50% (1/2)	50% (1/2)	0% (0/2)	100% (1/1)	0% (0/1)	0% (0/1)	-	-

Abbreviations: EPA = U.S. Environmental Protection Agency; HET-CAM = hen's egg test – chorioallantoic membrane

[1] EPA classification system (EPA 2003a).

Table 6-9 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing EPA Category IV from All Other Irritant Classes as Defined by the EPA Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.[3]	%	No.	%	No.	%	No.	%	No.
Overall	60	78	47/60	91	41/45	40	6/15	60	9/15	9	4/45
Without Alcohols	52	87	45/52	100	39/39	46	6/13	54	7/13	10	4/39
Without Surfactant Formulations	35	80	28/35	100	24/24	29	4/14	82	9/11	17	4/24
Without Oil/Water Emulsions	41	78	32/41	100	30/30	18	2/11	82	9/11	0	0/30
Without Alcohols and Surfactant Formulations	27	81	22/27	100	18/18	44	4/9	56	5/9	44	4/18
Without Alcohols and Oil/Water Emulsions	34	94	32/34	100	30/30	50	2/4	50	2/4	0	0/30
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	100	9/9	100	9/9	-	0/0	0	0/9	-	0/0

Abbreviations: EPA = U.S. Environmental Protection Agency; HET-CAM = hen's egg test – chorioallantoic membrane; N = number of substances included in this analysis/the total number of substances in the study; No. = data used to calculate the percentage.

[1] EPA classification system (EPA 2003a): Category IV vs. Categories I/II/III.

Table 6-10 HET-CAM False Negative Substances[1] Using the EPA Classification System[2]

Substance	*In Vivo* Scores		
	N	Corneal Opacity: Score (Day Cleared)[3]	Conjunctival Redness: Score (Day Cleared)[3]
HZA	6	-	N=1 2(2) N=1 2(3)
HZC	6	-	N=1 2(2)
HZV	6	-	N=2 2(2)
HZW	6	-	N=4 2(2) N=1 2(3)

Abbreviations: EPA = U.S. Environmental Protection Agency; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of animals

[1] False negative compounds are those that test as nonirritants *in vitro* but are mild, moderate, or severe ocular irritants/corrosive *in vivo*, i.e., EPA Category I, II, or III.

[2] EPA classification system (EPA 2003a).

[3] For the purposes of this evaluation, *clearing* is defined in the EPA hazard classification system as opacity or iritis scores = 0 or redness or chemosis scores = 1.

Further analysis of substances for which hazard classification was underpredicted by HET-CAM according to chemical class indicated that carboxylic acids had the highest proportion of underpredicted substances (25% [1/4]). Because the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of underpredicted substances was liquids. Among the 17 Category I surfactants, 73% (13/17) were underpredicted (**Table 6-11**).

According to the EPA classification system, the most overpredicted substances (false positives) were alcohols, of which 100% (7/7) were overpredicted. Because 98% (59/60) of the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of overpredicted substances was liquids. Three of the surfactants tested in HET-CAM were overpredicted (**Table 6-11**).

6.1.3 EU Classification System: HET-CAM Test Method Accuracy

Five studies (Bagley et al. 1992; Gettings et al. 1994; Kojima et al. 1995; Gettings et al. 1996; Hagino et al. 1999) contained HET-CAM test method data on 63 substances, 58 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the EU classification system (EU 2001) (see **Annex III**). Based on results from *in vivo* rabbit eye tests, 41% (24/58) were classified as R41 (severe irritants), 3% (2/58) were classified as R36 (moderate irritants), and 55% (32/58) were classified as Not Labeled. Five substances could not be classified according to the EU classification system due to the lack of adequate animal data and are so noted in **Annex III**.

Identification of Category R41 Substances (Ocular Corrosives/Severe Irritants)

The HET-CAM test method correctly identified 50% (12/24) of the R41 substances (**Table 6-12**). Among the remaining 50% (12/24) of R41 substances that were underpredicted by HET-CAM, 42% (10/24) were classified as R36, and 8% (2/24) were classified as Not Labeled.

Identification of Category R36 Substances (Moderate Ocular Irritants)

For the two substances that could be evaluated, the HET-CAM test method correctly identified 50% (1/2) as R36, while 50% (1/2) were underpredicted and 0% (0/2) were overpredicted (**Table 6-12**).

Table 6-11 Under- and Overprediction of the HET-CAM Test Method Using the EPA Classification System[1] in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property

Category	N	Underprediction (*In Vivo/In Vitro*)						Overprediction (*In Vivo/In Vitro*)					
		Severe (Category I)			Moderate (Category II)		Mild (Category III)	Moderate (Category II)	Mild (Category III)		Not Labeled (Category IV)		
		IV	III	II	IV	III	IV	I	II	I	III	II	I
Overall	60	0% (0/25)	12% (3/25)	40% (10/25)	0% (0/2)	0% (0/2)	40% (4/10)	50% (1/2)	50% (5/10)	50% (5/10)	40% (6/15)	0% (0/15)	20% (3/15)
Chemical Class[2]													
Alcohol	8	0% (0/1)	0% (0/1)	0% (1/1)	-	-	0% (0/5)	-	40% (2/5)	60% (3/5)	50% (1/2)	0% (0/2)	50% (1/2)
Carboxylic acid	6	0% (0/4)	0% (0/4)	25% (1/4)	-	-	0% (0/2)	-	0% (0/2)	100% (2/2)	-	-	-
Organic salt	6	0% (0/6)	0% (0/6)	17% (1/6)	-	-	-	-	-	-	-	-	-
Properties of Interest													
Liquids	59	0% (0/25)	12% (3/25)	40% (10/25)	-	-	22% (4/18)	-	28% (5/18)	28% (5/18)	40% (6/15)	0% (0/15)	20% (3/15)
Solids	0	-	-	-	-	-	-	-	-	-	-	-	-
Pesticide	0	-	-	-	-	-	-	-	-	-	-	-	-
Surfactant—Total	25	0% (0/17)	18% (3/17)	59% (10/17)	-	-	0% (0/4)	-	25% (1/4)	0% (0/4)	50% (2/4)	0% (0/4)	0% (0/4)
-nonionic	-	-	-	-	-	-	-	-	-	-	-	-	-
-anionic	-	-	-	-	-	-	-	-	-	-	-	-	-
-cationic	-	-	-	-	-	-	-	-	-	-	-	-	-

Continued

Table 6-11 Under- and Overprediction of the HET-CAM Test Method Using the EPA Classification System[1] in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property (continued)

Category	N	Underprediction (*In Vivo/In Vitro)*						Overprediction (*In Vivo/In Vitro)*					
		Severe (Category I)			Moderate (Category II)		Mild (Category III)	Moderate (Category II)	Mild (Category III)		Not Labeled (Category IV)		
		IV	III	II	IV	III	IV	I	II	I	III	II	I
Overall	60	0% (0/25)	12% (3/25)	40% (10/25)	0% (0/2)	0% (0/2)	40% (4/10)	50% (1/2)	50% (5/10)	50% (5/10)	40% (6/15)	0% (0/15)	20% (3/15)
Properties of Interest (continued)													
Oil/Water Emulsion	18	0% (0/1)	0% (0/1)	0% (0/1)	-	-	50% (4/8)	-	25% (2/8)	13% (1/8)	33% (3/9)	0% (0/9)	22% (2/9)
pH—Total	0	-	-	-	-	-	-	-	-	-	-	-	-
-acidic (pH <7.0)	-	-	-	-	-	-	-	-	-	-	-	-	-
-basic (pH >7.0)	-	-	-	-	-	-	-	-	-	-	-	-	-

Abbreviations: EPA = Environmental Protection Agency; HET-CAM = hen's egg test – chorioallantoic membrane; N = number of animals.

[1] EPA classification system (EPA 2003a).

[2] Chemical classes included in this table are represented by at least five substances tested in the HET-CAM test method, and assignments are based on the National Library of Medicine's medical substance headings (MeSH) classifications (www.nlm.nih.gov/mesh) as defined in **Annex I**.

Table 6-12 Performance of the HET-CAM Test Method (IS[A]) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System,[1] by Study and Overall

Data Source	Overall Correct Classification	Severe (R41)		Moderate (R36)			Mild			Not Labeled	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Gettings et al. (1994)	50% (9/18)	100% (1/1)	0% (0/1)	0% (0/0)	0% (0/0)	0% (0/0)	NA	NA	NA	53% (9/17)	47% (8/17)
Gettings et al. (1996)	29% (7/24)	25% (4/16)	75% (10/16)	0% (0/1)	100% (1/1)	0% (0/1)	NA	NA	NA	71% (5/7)	29% (2/7)
Hagino et al. (1999)	47% (7/15)	100% (7/7)	0% (0/7)	100% (1/1)	0% (0/1)	0% (0/1)	NA	NA	NA	100% (7/7)	0% (0/7)
Overall[2]	40% (23/58)	50% (12/24)	50% (12/24)	50% (1/2)	50% (1/2)	0% (0/2)	NA	NA	NA	69% (22/32)	31% (10/32)

Abbreviations: EU = European Union; HET-CAM = hen's egg test–chorioallantoic membrane; NA = not applicable.

[1] EU classification system (EU 2001).

[2] Overall data set includes one additional test substance from Bagley et al. (1992).

Identification of Not Labeled Substances

For the 32 substances that could be evaluated, the HET-CAM test method correctly identified 31% (10/32) as substances not labeled as irritants, while 69% (22/32) were overpredicted (**Table 6-12**).

Ability to Distinguish Not Labeled Substances from All Other Classes

In addition to evaluating the ability of the HET-CAM test method to identify each individual ocular hazard category according to the EU classification system, ICCVAM also evaluated the ability of the HET-CAM test method to distinguish ocular substances not labeled as irritants from all other irritant classes.[3] Among the 58 substances considered, the HET-CAM test method has an overall accuracy of 62% (36/58), a sensitivity of 100% (26/26), a specificity of 31% (10/32), a false positive rate of 69% (22/32), and a false negative rate of 0% (0/26) (**Table 6-13**).

As detailed below, the results from each individual study were also evaluated separately.

Gettings et al. (1994): Based upon the *in vivo* rabbit data, 18 substances were assigned an EU classification. The HET-CAM test method, by comparison, has an accuracy of 50% (9/18), sensitivity of 100% (1/1), specificity of 47% (8/17), false positive rate of 53% (9/17), and a false negative rate of 0% (0/1) (**Table 6-13**).

Gettings et al. (1996): Based upon the *in vivo* rabbit data, 24 substances were assigned a EU classification. The HET-CAM test method, by comparison, has an accuracy of 79% (19/24), sensitivity of 100% (17/17), specificity of 29% (2/7), false positive rate of 61% (5/7), and a false negative rate of 0% (0/17) (**Table 6-13**).

Hagino et al. (1999): Based upon the *in vivo* rabbit data, 15 substances were assigned a EU classification. The HET-CAM test method, by comparison, has an accuracy of 53% (8/15), sensitivity of 100% (8/8), specificity of 0% (0/7), false positive rate of 100% (7/7), and a false negative rate of 0% (0/26) (**Table 6-13**).

Performance of the HET-CAM Test Method with Discordant Classes Excluded

Because the IS(A) analysis method is the focus of the evaluation of HET-CAM for identifying all hazard categories, separate analyses were also conducted for all chemical classes and specific physical properties of interest represented in this database of 58 substances by at least five substances (i.e., surfactant-based formulations, oil/water emulsions, and alcohols).

Given the proportion of substances in the HET-CAM IS(A) database represented by these chemical and product classes (i.e., 88% [51/58] of the substances are included in one of these three categories), separate analyses without these discordant substances are not particularly informative. However, because of the associated discordance with each type, overall performance, particularly for the ocular corrosive and severe irritant category, can be improved by excluded certain product types (see **Table 6-14**). The results indicate that HET-CAM tends to overpredict alcohols (i.e., 83% [5/6] of alcohols classified as Not Labeled based on Draize test results were overpredicted by HET-CAM by at least one hazard category). Similarly, 53% (9/17) of the oil/water emulsions were overpredicted by HET-CAM by at least one hazard category. By comparison, surfactant formulations classified as R41 based on Draize results tended to be underpredicted by HET-CAM (75% [12/16] were underpredicted by HET-CAM as R36). However, none of these substances was underpredicted as Not Labeled.

When the ability of the HET-CAM test method to distinguish Not Labeled substances from all other irritant classes was evaluated with the specific chemical and product classes removed, the greatest

[3] The ICCVAM 2006 BRD provides an evaluation of the HET-CAM test method for distinguishing ocular corrosives and severe irritants from all other classes (ICCVAM 2006a). Because the database of HET-CAM test method results has not changed, this analysis is not repeated here.

improvement in false positive rate occurred when alcohols and surfactant formulations were excluded. The false positive rate decreased from 69% (22/32) to 58% (11/19). However, because the false negative rate for the overall database is 0% (0/31), this rate remained constant regardless of which chemical or product class(es) were excluded (**Table 6-15**).

Further analysis of substances for which hazard classification was underpredicted by HET-CAM according to chemical class indicated that carboxylic acids had the highest proportion of underpredicted substances (25% [1/4]). Because the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of underpredicted substances was liquids. Among the 16 R41 surfactant formulations, 75% (12/16) were underpredicted (**Table 6-16**).

According to the EU classification system, the most overpredicted substances (false positives) were alcohols, of which 83% (5/6) were overpredicted. Because the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of underpredicted substances was liquids. One of the Not Labeled surfactant formulations tested in HET-CAM was overpredicted (**Table 6-16**).

6.1.4 FHSA Classification System: HET-CAM Test Method Accuracy

The three studies (Gettings et al. 1994; Gettings et al. 1996; Hagino et al. 1999) contained HET-CAM test method data on 64 substances, 63 and 55 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the FHSA-20% and FHSA-67% criteria, respectively. Based on results from *in vivo* rabbit eye experiments using the FHSA-20% criteria, 68% (43/63) were classified as Irritants and 10% (6/63) were classified as Not Labeled. The remaining 24% (15/64) could not be classified using the FHSA-20% criteria and are so noted in **Annex III**. Using the FHSA-67% criteria, 69% (38/55) were classified as Irritants and 11% (6/55) were classified as Not Labeled. The remaining 17% (11/64) could not be classified using the FHSA-20% criteria and are so noted in **Annex III**.

Ability to Distinguish Not Labeled Substances From Irritants

ICCVAM evaluated the ability of the HET-CAM test method to distinguish substances not labeled as irritants from irritants. Using this approach for the 63 substances classified according to the FHSA-20% criteria, the HET-CAM test method has an overall accuracy of 78% (49/63), a sensitivity of 91% (43/47), a specificity of 38% (6/16), a false positive rate of 63% (10/16), and a false negative rate of 9% (4/47) (**Table 6-17**).

Using this approach for the 55 substances classified according to the FHSA-67% criteria, the HET-CAM test method has an overall accuracy of 80% (44/55), a sensitivity of 97% (38/39), a specificity of 38% (6/16), a false positive rate of 63% (10/16), and a false negative rate of 3% (1/39) (**Table 6-18**).

As detailed below, the results from each individual study were evaluated separately.

Gettings et al. (1994): Based upon *in vivo* rabbit data using the FHSA-20% criteria (**Table 6-17**), 18 substances could be assigned a classification. Among these 18 substances, the HET-CAM test method has an accuracy of 44% (8/18), sensitivity of 50% (4/8), specificity of 40% (4/10), a false positive rate of 60% (6/10), and a false negative rate of 50% (4/8).

Based upon *in vivo* rabbit data using the FHSA-67% analysis method (**Table 6-18**), 15 substances could be assigned a classification. Among these 15 substances, the HET-CAM test method has an accuracy of 53% (8/15), sensitivity of 80% (4/5), specificity of 40% (4/10), a false positive rate of 60% (6/10), and a false negative rate of 20% (1/5).

Gettings et al. (1996): Based upon *in vivo* rabbit data using the FHSA-20% criteria (**Table 6-17**), 25 substances could be assigned a classification. Among these 25 substances, the HET-CAM test

method has an accuracy of 92% (23/25), sensitivity of 100% (21/21), specificity of 50% (2/4), a false positive rate of 50% (2/4), and a false negative rate of 0% (0/21).

Based upon *in vivo* rabbit data using the FHSA-67% criteria (**Table 6-18**), 23 substances could be assigned a classification. Among these 23 substances, the HET-CAM test method has an accuracy of 91% (21/23), sensitivity of 100% (19/19), specificity of 50% (2/4), a false positive rate of 50% (2/4), and a false negative rate of 0% (0/19).

Hagino et al. (1999): Based upon *in vivo* rabbit data using the FHSA-20% criteria (**Table 6-17**), 17 substances could be assigned a classification. Among these 17 substances, the HET-CAM test method has an accuracy of 88% (15/17), sensitivity of 100% (15/15), specificity of 0% (0/2), a false positive rate of 100% (2/2), and a false negative rate of 0% (0/15).

Based upon *in vivo* rabbit data using the FHSA-67% criteria (**Table 6-18**), 15 substances could be assigned a classification. Among these 15 substances, the HET-CAM test method has an accuracy of 87% (13/15), sensitivity of 100% (13/13), specificity of 0% (0/2), a false positive rate of 100% (2/2), and a false negative rate of 0% (0/13).

Performance of the HET-CAM Test Method with Discordant Classes Excluded

The previous ICCVAM BRD identified limitations of the HET-CAM test method based upon the false positive rate for alcohols and the false negative rates for surfactant-based formulations, many of which were oil/water emulsions when the HET-CAM is used to identify ocular corrosives and severe irritants (ICCVAM 2006a). For this reason, the performance of the HET-CAM test method in identifying FHSA irritants using the FHSA-20% criteria was evaluated with these substances excluded from the database. The overall accuracy and sensitivity improve with exclusion of all substances belonging to these discordant classes (**Table 6-19**). However, the number of available substances was reduced to ten with none classified as Not Labeled that precluded determination of specificity and the false positive rate when all of the discordant substances were removed.

Exclusion of oil/water emulsions improved performance with an increase in accuracy from 78% (49/63) to 91% (41/45) and decreased the false negative rate from 9% (4/47) to 0% (0/39) with only a 4% increase in the false positive rate (**Table 6-19**). Removal of alcohols did not affect performance significantly, but the false positive rate was reduced 21% when alcohols and oil/water emulsions were excluded while the false negative rate remained the same and accuracy increased 17%. Removal of surfactant formulations reduced accuracy to 68% (26/38) and marginally decreased sensitivity and specificity at the expense of an increase in the false negative rate from 9% (4/47) to 15% (4/26). The false negative rate increased further to 22% (4/18) if alcohols and surfactant formulations were excluded.

The four false negative substances identified using the FHSA-20% criteria overall (i.e., HZA, HZC, HZV, and HZW) are the same four substances identified as false negative substances using the EPA classification system (EPA 2003a) shown in **Table 6-10**.

The performance of the HET-CAM test method in identifying FHSA irritants using the FHSA-67% criteria also was evaluated with these substances excluded from the database. The overall accuracy and sensitivity improve with exclusion of all substances belonging to these discordant classes (**Table 6-20**). However, the number of available substances was reduced to nine with none classified as Not Labeled that precluded determination of specificity and the false positive rate when all of the discordant substances were removed.

Using the FHSA-67% criteria, the exclusion of oil/water emulsions improved performance with an increase in accuracy from 80% (44/55) to 90% (36/40) and decreased the false negative rate from 3% (1/39) to 0% (0/34) with only a 4% increase in the false positive rate (**Table 6-20**). Removal of alcohols did not affect performance significantly, but the false positive rate was reduced 21% when

alcohols and oil/water emulsions were excluded while the false negative rate remained the same and accuracy increased 15%. Removal of surfactant formulations reduced accuracy to 72% (23/32) and marginally decreased sensitivity and increased the false negative rate. The false negative rate increased further to 7% (1/14) if alcohols and surfactant formulations were excluded.

The false negative substance using the FHSA-67% criteria overall was HZW, one of the four false negative substances identified using the EPA classification system shown in **Table 6-10**.

Table 6-13 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the EU Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gettings et al. (1994)	18	50	9/18	100	1/1	47	8/17	53	9/17	0	0/1
Gettings et al. (1996)	24	79	19/24	100	17/17	29	2/7	61	5/7	0	0/17
Hagino et al. (1999)	15	53	8/15	100	8/8	0	0/7	100	7/7	0	0/8
Overall[2]	58	62	36/58	100	26/26	31	10/32	69	22/32	0	0/26

Abbreviations: EU = European Union; HET-CAM = hen's egg test – chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] EU classification system (EU 2001): Not Labeled vs. R41/R36.

[2] Overall data set includes one additional test substance from Bagley et al. (1992).

Table 6-14 Performance of the HET-CAM Test Method (IS[A]) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	Overall Correct Classification	Severe (R41)		Moderate (R36)			Mild			Not Labeled	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall	40% (23/58)	50% (12/24)	50% (12/24)	50% (1/2)	50% (1/2)	0% (0/2)	NA	NA	NA	69% (22/32)	31% (10/32)
Without Alcohols	42% (21/50)	45% (10/22)	55% (12/22)	50% (1/2)	50% (1/2)	0% (0/2)	NA	NA	NA	62% (16/26)	38% (10/26)
Without Surfactant Formulations	47% (16/34)	100% (8/8)	0% (0/8)	100% (1/1)	0% (0/1)	0% (0/1)	NA	NA	NA	68% (17/25)	32% (8/25)
Without Oil/Water Emulsions	35% (14/40)	48% (11/23)	52 (12/23)	50% (0/2)	50% (1/2)	0% (0/2)	NA	NA	NA	87% (13/15)	13% (2/15)
Without Alcohols and Surfactant Formulations	54% (14/26)	100% (6/6)	0% (0/6)	100% (0/1)	0% (0/1)	0% (0/1)	NA	NA	NA	58% (11/19)	42% (8/19)
Without Alcohols and Oil/Water Emulsions	37% (12/32)	43% (9/21)	57% (12/21)	50% (1/2)	50% (1/2)	0% (0/2)	NA	NA	NA	78% (7/9)	22% (2/9)
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	62% (5/8)	100% (5/5)	0% (0/5)	100% (1/1)	0% (0/1)	0% (0/1)	NA	NA	NA	100% (2/2)	0% (0/2)

Abbreviations: EU = European Union; HET-CAM = hen's egg test–chorioallanotic membrane; NA = not applicable.

[1] EU classification system (EU 2001).

Table 6-15 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the EU Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	58	62	36/58	100	26/26	31	10/32	69	22/32	0	0/26
Without Alcohols	50	42	21/50	100	24/24	38	10/26	62	16/26	0	0/24
Without Surfactant Formulations	34	50	17/34	100	9/9	32	8/25	68	17/25	0	0/9
Without Oil/Water Emulsions	40	67	26/39	100	25/25	13	2/15	87	13/15	0	0/25
Without Alcohols and Surfactant Formulations	26	58	15/26	100	7/7	42	8/19	58	11/19	0	0/7
Without Alcohols and Oil/Water Emulsions	32	78	25/32	100	23/23	22	2/9	78	7/9	0	0/23
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	8	75	6/8	100	6/6	0	0/2	100	2/2	0	0/6

Abbreviations: EU = European Union; HET-CAM = hen's egg test – chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] EU classification system (EU 2001): Not Labeled vs. R41/R36.

Table 6-16 Performance of the HET-CAM Test Method Using the EU Classification System[1] in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property

Category	N	Underprediction (*In Vivo/In Vitro*)			Overprediction (*In Vivo/In Vitro*)		
		Severe (R41)		Moderate (R36)	Moderate (R36)	Not Labeled (NL)	
		NL	R36	NL	R41	R36	R41
Overall	58	8% (2/24)	42% (10/24)	50% (1/2)	0% (0/2)	68% (15/22)	32% (7/22)
Chemical Class[2]							
Alcohol	8	0% (0/2)	0% (0/2)	50% (1/2)	0% (0/2)	33% (2/6)	50% (3/6)
Carboxylic Acid	5	0% (0/4)	25% (1/4)	-	-	0% (0/1)	100% (1/1)
Organic salt	2	0% (0/5)	20% (1/5)	100% (1/1)	0% (0/1)	-	-
Properties of Interest							
Liquids	58	8% (2/24)	42% (10/24)	50% (1/2)	50% (1/2)	16% (5/32)	25% (8/32)
Solids	0	-	-	-	-	-	-
Pesticide	0	-	-	-	-	-	-
Surfactant-Total	24	0% (0/16)	62% (12/16)	100% (1/1)	0% (0/1)	14% (1/7)	0% (0/7)
-nonionic	-	-	-	-	-	-	-
anionic	-	-	-	-	-	-	-
cationic	-	-	-	-	-	-	-

continued

Table 6-16 Performance of the HET-CAM Test Method Using the EU Classification System[1] in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property (continued)

Category	N	Underprediction (*In Vivo/In Vitro)*			Overprediction (*In Vivo/In Vitro)*		
		Severe (R41)		Moderate (R36)	Moderate (R36)	Not Labeled (NL)	
		NL	R36	NL	R41	R36	R41
Overall	58	8% (2/24)	42% (10/24)	50% (1/2)	0% (0/2)	68% (15/22)	32% (7/22)
Properties of Interest (continued)							
Oil/Water Emulsion	18	0% (0/1)	0% (0/1)	-	-	35% (6/17)	18% (3/17)
pH-Total	0	-	-	-	-	-	-
-acidic (pH <7.0)	-	-	-	-	-	-	-
-basic (pH >7.0)	-	-	-	-	-	-	-

Abbreviations: EU = European Union; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of animals; NL = Not Labeled (as irritant).

[1] EU classification system (EU 2001).

[2] Chemical classes included in this table are represented by at least five substances tested in the HET-CAM test method, and assignments are based upon National Library of Medicine medical subject heading (MeSH) categories (www.nlm.nih.gov/mesh) as defined in **Annex I**.

Table 6-17 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Substances Not Labeled as Irritants from Irritants, as Defined by the FHSA-20% Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gettings et al. (1994)	18	44	8/18	50	4/8	40	4/10	60	6/10	50	4/8
Gettings et al. (1996)	25	92	23/25	100	21/21	50	2/4	50	2/4	0	0/21
Hagino et al. (1999)	17	88	15/17	100	15/15	0	0/2	100	2/2	0	0/15
Overall[2]	63	78	49/63	91	43/47	38	6/16	63	10/16	9	4/47

Abbreviations: FHSA = U.S. Federal Hazardous Substances Act; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of classifiable substances; No. = number on which the percentage is calculated.

[1] FHSA classification system (16 CFR 1500.42): Irritant or Not Labeled as an Irritant. FHSA-20% analysis method is based on use of proportionality in which an irritant is identified by a positive response (i.e., CO or IR >1 and/or CR or CC≥2) in ≥1/3, 1/4, 1/5 or ≥2/6 animals (20 to 33% positive). Substances that do not produce a positive response in 3, 4, 5, or 6 animals or that produce a positive response in 1/6 animals are not classified as irritants, and therefore do not require labeling.

[2] Because Bagley et al. (1992) and Kojima et al. (1995) contain only one and two classifiable substances, respectively, data from these studies were included only in the overall analysis and were not evaluated separately.

Table 6-18 Accuracy of the HET-CAM Test Method in Distinguishing Substances Not Labeled as Irritants from Irritants, as Defined by the FHSA-67% Classification System,[1] by Study and Overall

Data Source	N[2]	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.[3]	%	No.	%	No.	%	No.	%	No.
Gettings et al. (1994)	15	53	8/15	80	4/5	40	4/10	60	6/10	20	1/5
Gettings et al. (1996)	23	91	21/23	100	19/19	50	2/4	50	2/4	0	0/19
Hagino et al. (1999)	15	87	13/15	100	13/13	0	0/2	100	2/2	0	0/13
Overall[2]	55	80	44/55	97	38/39	38	6/16	63	10/16	3	1/39

Abbreviations: FHSA = Federal Hazardous Substances Act; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = number on which the percentage is calculated.

[1] FHSA classification system (16 CFR 1500.42): Irritant or not labeled. FHSA-67% analysis method is based on use of proportionality in which an irritant is identified by a positive response (i.e., CO or IR >1 and/or CR or CC≥2) in ≥2/3, 3/4, 4/5 or 4/6 animals (67% to 80% positive). Substances that do not produce a positive response in 3, 4, 5, or 6 animals or that produce a positive response in 1/6 animals are not classified as irritants, and therefore do not require labeling.

[2] Because Bagley et al. (1992) and Kojima et al. (1995) contain only one and two classifiable substances, respectively, data from these studies were included only in the overall analysis and were not evaluated separately. The FHSA-67% Inconclusive substances were not included in the calculations. One of these was from the Bagley et al. (1992) study; therefore, the overall correct classification values increase by two rather than by three substances.

Table 6-19 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-20% Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	63	78	49/63	91	43/47	38	6/16	63	10/16	9	4/47
Without Alcohols	53	77	41/53	90	35/39	43	6/14	57	8/14	10	4/39
Without Surfactant Formulations	38	68	26/38	85	22/26	33	4/12	67	8/12	15	4/26
Without Oil/Water Emulsions	45	91	41/45	100	39/39	33	2/6	67	4/6	0	0/39
Without Alcohols and Surfactant Formulations	28	64	18/28	78	14/18	40	4/10	60	6/10	22	4/18
Without Alcohols and Oil/Water Emulsions	35	94	33/35	100	31/31	50	2/4	50	2/4	0	0/31
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	10	100	10/10	100	10/10	-[2]	-	-	-	0	0/10

Abbreviations: FHSA = U.S. Federal Hazardous Substances Act; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] FHSA classification system (16 CFR 1500.42): Irritant or Not Labeled as an Irritant. FHSA-20% analysis method is based on use of proportionality in which an irritant is identified by a positive response (i.e., CO or IR >1 and/or CR or CC≥2) in ≥1/3, 1/4, 1/5 or ≥2/6 animals (20% to 33% positive). Substances that do not produce a positive response in 3, 4, 5, or 6 animals or that produce a positive response in 1/6 animals are not classified as irritants, and are therefore do not require labeling.

[2] No substances were classified as Not Labeled by FHSA or as nonirritants in HET-CAM, therefore specificity and the false positive rate could not be determined.

Table 6-20 Accuracy of the HET-CAM Test Method (IS[A]) in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-67% Classification System,[1] with Discordant Chemical and Physical Classes Excluded

HET-CAM Database	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	55	80	44/55	97	38/39	38	6/16	63	10/16	3	1/39
Without Alcohols	47	81	38/47	97	32/33	43	6/14	57	8/14	3	1/33
Without Surfactant Formulations	32	72	23/32	95	19/20	33	4/12	67	8/12	5	1/20
Without Oil/Water Emulsions	40	90	36/40	100	34/34	33	2/6	67	4/6	0	0/34
Without Alcohols and Surfactant Formulations	24	71	17/24	93	13/14	40	4/10	60	6/10	7	1/14
Without Alcohols and Oil/Water Emulsions	32	94	30/32	100	28/28	50	2/4	50	2/4	0	0/28
Without Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	100	9/9	100	9/9	-[2]	-	-	-	0	0/9

Abbreviations: FHSA = U.S. Federal Hazardous Substances Act; HET-CAM = hen's egg test–chorioallantoic membrane; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] FHSA classification system (16 CFR 1500.42): Irritant or not labeled. FHSA-67% analysis method is based on use of proportionality in which an irritant is identified by a positive response (i.e., CO or IR >1 and/or CR or CC≥2) in ≥2/3, 3/4, 4/5 or 4/6 animals (67% to 80% positive). Substances that do not produce a positive response in 3, 4, 5, or 6 animals or that produce a positive response in 1/6 animals are not classified as irritants, and are therefore do not require labeling.

[2] No substances were classified as Not Labeled by FHSA or as Nonirritants in HET-CAM; therefore, specificity and the false positive rate could not be determined.

7.0 HET-CAM Test Method Reliability

An assessment of test method reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) is essential to any evaluation of the performance of an alternative test method (ICCVAM 2003). Quantitative and qualitative evaluations of HET-CAM test method reliability have been conducted previously (ICCVAM 2006a). Because the database used for the current evaluation of the HET-CAM test method has not changed, the quantitative evaluation of test method reliability remains unchanged. However, additional qualitative analyses of test method reproducibility were conducted to evaluate the extent of agreement in HET-CAM hazard classifications among the laboratories. Given that the performance of the BCOP test method was similar for the EPA and FHSA hazard classification systems, additional reliability analyses were not conducted for the FHSA hazard classification system.

7.1 Interlaboratory Reproducibility of Hazard Classification Category Using the GHS Classification System

Fifteen of 17 substances tested had sufficient data to be classified using the GHS system (UN 2007). Of four Not Classified and three Category 2B substances, none was correctly identified by HET-CAM. None of the 15 GHS-classified substances tested was classified Category 2A by HET-CAM. However, eight substances classified as GHS Category 1 were correctly identified by the HET-CAM test method.

To evaluate the extent of agreement in irritant classifications among laboratories (i.e., Category 1, 2A, and 2B = + and Not Classified = -), regardless of the individual hazard classification, NICEATM compared *in vivo* and *in vitro* data (**Table 7-1**).

For 11 substances, there was 100% agreement between the *in vivo* and *in vitro* classifications (i.e., +/+). For four substances that were overpredicted *in vitro* (i.e., -/+), there was 100% agreement for 75% (3/4) of the substances and 80% agreement for 25% (1/4) of the substances. For two substances that could not be assigned GHS classifications, there was 100% agreement on the *in vitro* classifications (i.e., ?/+).

NICEATM could not assess the agreement between laboratories for substances not labeled as irritants compared to all other classes, because the HET-CAM test method did not produce any Not Classified classifications. Overall, however, there was 100% agreement for 94% (16/17) of the substances and 80% agreement for 6% (1/17) of the substances.[4]

The extent of agreement for a test substance was also evaluated among the five laboratories based on prediction of the individual GHS hazard category (**Table 7-2**). Of four Not Classified substances, all were overpredicted with 100% agreement by 75% (3/4) of the laboratories and 80% agreement by 25% (1/4) of the laboratories. All three Category 2B substances were overpredicted with 100% (3/3) agreement among the five laboratories. No Category 2A substances were identified.

All eight substances were correctly predicted as Category 1 with 100% agreement for 63% (5/8) of the substances, 80% agreement for 13% (1/8) of the substances, and 60% agreement for 25% (2/8) of the substances.

[4] Because the database of HET-CAM test method results has not changed since the 2006 ICCVAM BRD, the qualitative evaluation of reproducibility is not repeated here.

Table 7-1 Interlaboratory Variability of Hagino et al. (1999) Compared to the *In Vivo* Rabbit Eye Test in Predicting Ocular Hazard Categories for Severe Irritants or Corrosives (1) from Nonsevere Irritants (2A, 2B) and Substances Not Classified, as Defined by the GHS Classification System[1]

Report	Analysis Method[2]	Classification (*In Vivo*/*In Vitro*)[3]	# of Labs	N	Substances with 100% Agreement among Labs[4]	Substances with 80% Agreement among Labs[4]
Hagino et al. (1999)	IS(A)	+/+	5	11	11 (100%)	0
		+/-	5	0	0	0
		-/+	5	4	3 (75%)	1 (25%)
		-/-	5	0	0	0
		?/-	5	0	0	0
		?/+	5	2	2 (100%)	0
		Total	5	17	16 (94%)	1 (6%)

Abbreviations: GHS = Globally Harmonized System; N = number of substances.

[1] GHS classification system (UN 2007).

[2] Analysis method used to transform the sample data into HET-CAM scores. IS(A) = method described in Luepke (1985); IS(B)-10 and IS(B)-100 = method described in Kalweit et al. (1987).

[3] A "+" indicates that the substance was assigned an overall classification of corrosive or a severe irritant (Category 1). A "-" indicates that the substance was assigned an overall classification of nonsevere irritant (Category 2A or 2B) or Not Classified. A "?" indicates that, due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects; insufficient dose volume), a GHS classification could not be made. See **Section 6.1** for a description of the rules followed to classify the ocular irritancy of test substances tested multiple times *in vitro.*

[4] Number in parentheses indicates percentage of tested chemicals.

None of the eight Category 1 substances was incorrectly identified. However, all four Not Classified substances and the three Category 2B substances, 4/4 (100%) and 3/3 (100%), respectively, were incorrectly identified (**Table 7-2**).

There was no agreement among the five participating laboratories in incorrect classification of 0/8 (0%) of the GHS Category 1 substances. All were correctly classified. There was 100% agreement in overclassifying 100% (3/3) of the GHS Category 2B substances, 100% agreement in overclassifying 75% (3/4) of the substances, and 80% agreement in overclassifying 25% (1/4) of the Not Classified substances (**Table 7-2**).

7.2 Interlaboratory Reproducibility of Hazard Classification Category Using the EPA Classification System

Fifteen of 17 substances tested had sufficient data to be classified using the EPA system (EPA 2003a). Of two Category IV, five Category III, and one Category II substances, none (0% [0/2], 0% [0/5], and 0% [0/1], respectively) was correctly identified by the HET-CAM test method. However, seven substances classified as EPA Category I were correctly identified by HET-CAM (100% [7/7]).

To evaluate the extent of agreement in irritant classifications among laboratories (i.e., Category 1, 2A, and 2B = + and Not Labeled = -), regardless of the individual hazard classification, NICEATM compared *in vivo* and *in vitro* data (**Table 7-3**).

For 13 substances, there was 100% agreement among the *in vivo* and *in vitro* classifications (i.e., +/+). There was 60% agreement for both (100% [2/2]) of the substances that were overpredicted *in vitro* (i.e., -/+). For two substances that could not be assigned an EPA classification, there was 100% agreement on the *in vitro* classifications (i.e., ?/+) for 50% (1/2) of the substances and 60% agreement for 50% (1/2) of the substances.

NICEATM could not assess the agreement between laboratories for substances not labeled as irritants compared to all other classes, because the HET-CAM test method did not produce any Not Labeled classifications. Overall, however, there was 100% agreement for 82% (14/17) of the substances and 60% agreement for 18% (3/17) of the substances.[5]

The extent of agreement for a test substance was also evaluated among the five laboratories based on prediction of the individual EPA hazard category (**Table 7-4**). Both Category IV substances were overpredicted with 100% agreement by 50% (1/2) of the laboratories and with 80% agreement by 50% (1/2) of the laboratories. All five Category III substances were overpredicted with 100% agreement among the five laboratories. One Category II substance was overpredicted with 100% agreement among the five laboratories. All seven substances were correctly predicted as Category I substances with 100% agreement for 71% (5/7) of the substances and 80% agreement for 29% (2/7) of the substances.

None of the seven Category 1 substances was incorrectly identified. However, both Category IV, all five Category III, and the one Category II substance (i.e., 100% [2/2], 100% [5/5], and 100%, respectively) were incorrectly identified by the HET-CAM test method (**Table 7-4**).

There was no agreement among the five participating laboratories in incorrectly classifying any (0% [0/7]) of the EPA Category I substances. All were correctly classified. There was 100% agreement in overclassifying 50% (1/2) and 80% agreement in overclassifying 50% (1/2) of the EPA Category IV substances. For Category III substances, there was 100% agreement in overclassifying 5/5 substances. There was 100% agreement in overclassifying the Category II substance.

---

[5] Because the database of HET-CAM test method results has not changed since the 2006 ICCVAM BRD (2006a), the qualitative evaluation of reproducibility is not repeated here.

Table 7-2 Interlaboratory Variability of Hagino et al. (1999) Compared to the *In Vivo* Rabbit Eye Test in Predicting Each Ocular Hazard Category (1, 2A, 2B) and Substances Not Classified, as Defined by the GHS Classification System[1]

***In Vivo* Classification (No.)[2]**	***In Vitro* Classification**	**N**	**# of Labs**	**Substances with 100% Agreement Among Labs**	**Substances with 80% Agreement Among Labs**	**Substances with 60% Agreement Among Labs**
Not Classified (4)	Actual	0	5	0	0	0
	Over	4	5	3 (75%)	1 (25%)	0
Category 2B (3)	Under	0	5	0	0	0
	Actual	0	5	0	0	0
	Over	3	5	3 (100%)	0	0
Category 2A (0)	Under	0	5	0	0	0
	Actual	0	5	0	0	0
	Over	0	5	0	0	0
Category 1 (8)	Under	0	5	0	0	0
	Actual	8	5	5 (63%)	1 (13%)	2 (25%)

Abbreviations: GHS = Globally Harmonized System; N = number of substances; No. = number of substances classified.

[1] GHS classification system (UN 2007).

[2] Due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects), a GHS classification could not be made for two substances. See **Section 6.1** for a description of the rules followed to classify the ocular irritancy of test substances tested multiple times *in vitro.*

Table 7-3 Interlaboratory Variability of Hagino et al. (1999) Compared to the *In Vivo* Rabbit Eye Test in Predicting Ocular Hazard Category I (Severe Irritants or Corrosives) from Nonsevere Irritants (Category II, III) and Substances Not Labeled (Category IV), as Defined by the EPA Classification System[1]

Report	Analysis Method[2]	Classification (*In Vivo*/*In Vitro*)[3]	# of Labs	N	Substances with 100% Agreement Among Labs[4]	Substances with 60% Agreement Among Labs[4]
Hagino et al. (1999)	IS(A)	+/+	5	13	13 (100%)	0
		+/-	5	0	0	0
		-/+	5	2	0	2 (100%)
		-/-	5	0	0	0
		?/-	5	0	0	0
		?/+	5	2	1 (50%)	1 (50%)
		Total	5	17	14 (82%)	3 (18%)

Abbreviations: EPA = U.S. Environmental Protection Agency; N = number of substances.

[1] EPA classification system (EPA 2003a).

[2] Analysis method used to transform the sample data into HET-CAM scores. IS(A) = method described in Luepke (1985); IS(B)-10 and IS(B)-100 = method described in Kalweit et al. (1987).

[3] A "+" indicates that the substance was assigned an overall classification of corrosive or a severe irritant (Category 1). A "-" indicates that the substance was assigned an overall classification of nonsevere irritant (Category 2A or 2B) or Not Labeled. A "?" indicates that, due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects; insufficient dose volume), a GHS classification could not be made. See **Section 6.1** for a description of the rules followed to classify the ocular irritancy of test substances tested multiple times *in vitro.*

[4] Number in parentheses indicates percentage of tested chemicals.

Table 7-4 Interlaboratory Variability of Hagino et al. (1999) Compared to the *In Vivo* Rabbit Eye Test in Predicting Each Ocular Hazard Category for Severe Irritants or Corrosives (I), Irritants (II, III), and Substances Not Labeled (Category IV), as Defined by the EPA Classification System[1]

In Vivo Classification (No.)[2]	*In Vitro* Classification	# of Labs	N	Substances with 100% Agreement Among Labs	Substances with 80% Agreement Among Labs
Category IV (2)	Actual	5	0	0	0
	Over	5	2	1 (50%)	1 (50%)
Category III (5)	Under	5	0	0	0
	Actual	5	0	0	0
	Over	5	5	5 (100%)	0
Category II (1)	Under	5	0	0	0
	Actual	5	0	0	0
	Over	5	1	1 (100%)	0
Category 1 (7)	Under	5	0	0	0
	Actual	5	7	5 (71%)	2 (29%)

Abbreviations: EPA = U.S. Environmental Protection Agency; N = number of substances; No. = number of substances classified.

[1] EPA classification system (EPA 2003a).

[2] Due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects), an EPA classification could not be made for two substances. See **Section 6.1** for a description of the rules followed to classify the ocular irritancy of test substances tested multiple times *in vitro.*

7.3 Interlaboratory Reproducibility of Hazard Classification Category Using the EU Classification System

Fifteen of 17 substances tested had sufficient data to be classified using the EU system (EU 2001). Of seven Not Labeled and one R36 substances, none (0% [0/7] and 0% [0/1], respectively) were correctly identified by HET-CAM. However, all seven substances classified as EU R41 were correctly identified by the HET-CAM test method (100%).

To evaluate the extent of agreement in irritant classifications among laboratories (i.e., Category 1, 2A, and 2B = + and Not Labeled = -), regardless of the individual hazard classification, NICEATM compared *in vivo* and *in vitro* data (**Table 7-5**).

For eight substances, there was 100% agreement among the *in vivo* and *in vitro* classifications for 63% (5/8), 80% agreement for 25% (2/8), and 60% agreement for 13% (1/8). For seven substances that were overpredicted *in vitro* (i.e., -/+), there was 100% agreement for 86% (6/7) and 80% agreement for 14% (1/7) of the substances. There was 100% agreement on the *in vitro* classification (i.e.,?/+) of both substances that could not be assigned an EU classification.

NICEATM could not assess the agreement between laboratories for substances not labeled as irritants compared to all other classes, because the HET-CAM test method did not produce any Not Labeled classifications.

The extent of agreement for a test substance was also evaluated among the five laboratories based on prediction of the individual EU hazard category (**Table 7-6**).

All seven Not Labeled substances were overpredicted with 100% agreement by 86% (6/7) of the laboratories and with 80% agreement by 14% (1/7) of the laboratories.

The one R36 substance was overpredicted with 100% agreement among the five laboratories.

Seven R41 substances were overpredicted with 100% agreement among the five laboratories for 71% (5/7), 80% agreement for 14% (1/7), and 60% agreement for 14% (1/7) of the substances.

None of the seven R41 substances was incorrectly identified. However, all seven Not Labeled, one Category R36, and seven R41 substances (i.e., 100% [7/7], 100% [1/1], and 100% [7/7], respectively) were incorrectly identified by HET-CAM (**Table 7-6**).

There was no agreement among the five participating laboratories in incorrectly classifying any (0/7) of the EU R41 substances; all were correctly classified. There was 100% agreement in overclassifying 86% (6/7) and 80% agreement in overclassifying 14% (1/7) of the EPA substances not labeled as irritants. For R36 substances, there was 100% agreement in overclassifying 1/1 substance.

7.4 Common Chemical or Product Classes Among Test Substances with Discordant Interlaboratory Results Using the GHS Classification System

There were insufficient data with which to determine the effect of discordant chemicals on the interlaboratory analyses.

Table 7-5 Interlaboratory Variability of Hagino et al. (1999) Compared to the *In Vivo* Rabbit Eye Test in Predicting Ocular Hazard Categories for Severe Irritants or Corrosives (R41) from Irritants (R36) and Substances Not Labeled, as Defined by the EU Classification System[1]

Report	Analysis Method[2]	Classification (*In Vivo*/*In Vitro*)[3]	# of Labs	N	Substances with 100% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 60% Agreement Among Labs
Hagino et al. (1999)	IS(A)	+/+	5	8	5 (63%)	2 (25%)	1 (13%)
		+/-	5	0	0	0	0
		-/+	5	7	6 (86%)	1 (14%)	0
		-/-	5	0	0	0	0
		?/-	5	0	0	0	0
		?/+	5	2	2 (100%)	0	0
		Total	5	17	13 (76%)	3 (18%)	1 (6%)

Abbreviations: EU = European Union; N = number of substances.

[1] EU classification system (2001).

[2] Analysis method used to transform the sample data into HET-CAM scores. IS(A) = method described in Luepke (1985); IS(B)-10 and IS(B)-100 = method described in Kalweit et al. (1987).

[3] A "+" indicates that the substance was assigned an overall classification of severe irritant or corrosive (R41). A "-" indicates that the substance was assigned an overall classification of nonsevere irritant (R36) or Not Labeled. A "?" indicates that, due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects; insufficient dose volume), an EU classification could not be made. See **Section 6.1** for a description of the rules followed to classify the ocular irritancy of test substances tested multiple times *in vitro.*

Table 7-6 Interlaboratory Variability of Hagino et al. (1999) Compared to the *In Vivo* Rabbit Eye Test in Predicting Each Ocular Hazard Category for Severe Irritants or Corrosives (R41), Irritants (R36), and Substances Not Labeled, as Defined by the EU Classification System[1]

In Vivo Classification (No.)[2]	Classification *(In Vitro)*	# of Labs	N	Substances with 100% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 60% Agreement Among Labs
NL (7)	Actual	5	0	0	0	0
	Over	5	7	6 (86%)	1 (14%)	0
R36 (1)	Under	5	0	0	0	0
	Actual	5	0	0	0	0
	Over	5	1	1 (100%)	0	0
R41 (7)	Under	5	0	0	0	0
	Actual	5[2]	7	5 (71%)	1 (14%)	1 (14%)

Abbreviations: EU = European Union; N = number of substances; NL = Not Labeled (as irritant); No. = number of substances classified.

[1] EU classification system (2001).

[2] Due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects), an EU classification could not be made for two substances. See **Section 6.1** for a description of the rules followed to classify the ocular irritancy of test substances tested multiple times *in vitro.*

8.0 Test Method Data Quality

The same database was used in this assessment and the 2006 ICCVAM *Background Review Document: Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane*, in which test method data quality is evaluated (ICCVAM 2006a).

9.0 Other Scientific Reports and Reviews

NICEATM obtained two studies that were not discussed in the 2006 BRD (ICCVAM 2006a) but that contain HET-CAM data: de Silva et al. (1992) and Boue-Grabot et al. (1995).

De Silva et al. (1992) presented the results of a HET-CAM study of 60 chemicals and 41 cosmetic formulations. The chemicals were tested at 10% of their *in vivo* test concentration, whereas the cosmetic formulations were tested neat. The researchers used the test method of Luepke (1985) with a fixed time point IS(A) analysis method (i.e., 0.5, 2, and 5 minutes). Intralaboratory reproducibility was evaluated using a double-blind study of 20 surfactants tested at concentrations of 1% and 10%. Spearman's coefficient rho was greater than 0.9 ($p < 10^{-8}$) for the two concentrations. For the 60 chemicals, HET-CAM scores (i.e., maximum score of 21) were correlated with three EEC ocular irritation classes (i.e., Class I = Not Labeled, Class II = R36, and Class III = R41). Class I substances were clearly distinguished from Class II substances. Sensitivity, specificity, and concordance were 91%, 88%, and 90%, respectively, when an IS(A) score of 9 was used to differentiate Class I from Class II substances.

In de Silva et al., the false positive rate was 5% (3/60), and the false negative rate was 5% (3/60). The false negative substances were one Class II or severe irritant (acetaldehyde) and two nonsevere irritants (n-butanol and a nonionic surfactant). The HET-CAM scores for 21 formulations (i.e., make-up removers, shower gels, and shampoos) studied without rinsing, and 20 formulations (i.e., creams and body milks) washed off after a 20-second contact were compared to Draize MAS values, resulting in Spearman rank correlation coefficients of rho = 0.77 ($p < 10^{-2}$) and rho = 0.76 ($p < 10^{-2}$), respectively. The authors suggest that the HET-CAM test method, with optimization, is potentially useful in a battery of *in vitro* test methods for the screening of new ingredients and formulations. These data were not used in the HET-CAM performance analyses in this BRD because original Draize data were not available to derive regulatory classifications based on the current EPA, GHS, and EU classification systems (EPA 2003a; UN 2007; EU 2001).

In Boue-Grabot (1995), 103 cosmetics and toiletries were tested in the HET-CAM test method using the fixed time point method (i.e., 0, 0.5, 2 and 5 minutes) of Luepke (1985). In this method, the CAM is observed for the appearance of vasodilation, hemorrhage, or coagulation at each time point, and numerical scores are assigned. The IS was converted to a mean chorioallantoic irritation index (MCA), and the HET-CAM results (i.e., nonirritant, slightly irritant, moderately irritant, or very irritant) were compared to the Draize test using the maximal ocular irritation index (IOMA) with an identical irritation classification scheme. Results were expressed in terms of correlation ($r = 0.657$, $p < 0.001$) between the MCA and IOMA values. Accuracy was 92%, sensitivity was 80%, specificity was 94%, the false negative rate was 2%, and the false positive rate was 6%. A cytotoxicity test was used to further reduce the false positive and false negative rates. No individual HET-CAM or Draize data were provided in this study, so the data could not be used in the performance analysis.

NICEATM found five additional studies containing HET-CAM data in the peer-reviewed literature from 2005 to 2009 (Dahl 2007; Debbasch et al. 2005; Mancebo et al. 2008; Mehling et al. 2007; Vinardell and Mitjans 2006). From these studies, seven test substances were identified with *in vitro* scores and *in vivo* data using the Draize rabbit eye test. However, the Draize rabbit eye test data and

HET-CAM results for all seven test substances were included in the accuracy analyses reported in the ICCVAM BRD (2006a). Consequently, they have already been considered in the current evaluation.

Getttings et al. (1996b) used the original Draize data and new low volume eye test (LVET) data to evaluate new *in vitro* test method data, including HET-CAM using the IS(A) and IS(B) analysis methods, on 10 hydroalcoholic formulations that were originally published in Gettings et al. (1991). The authors suggest that the performance of the *in vitro* test methods, including HET-CAM, conformed no better (or worse) with the LVET than with the Draize test method. No individual animal data were provided to enable regulatory classification. Therefore, these data were not used in the current HET-CAM performance analyses.

In Debbasch et al. (2005), 12 coded make-up removers were applied to the external eyelid and tested in the HET-CAM, BCOP, and the corneal epithelial cell line (CEPI) test methods, as well as a clinical in-use test under ophthalmological control. Three hundred microliters of undiluted test product was applied to the CAM of 9-day-old fertilized eggs (White Leghorn chicken, four per product). Corneal opacity was determined using an adapted spectrophotometer and barrier disruption by fluorescein uptake using OD_{490} nm. *In vitro* scores were classified according to Gautheron et al. (1994) and Harbell and Curren (1998). However, no *in vivo* rabbit eye data were reported, and these data have not been obtained. For this reason, the results from this study were not included in the HET-CAM performance analyses detailed in this BRD.

In Vinardell and Mitjans (2006), several industrial and laboratory solvents were tested for potential eye irritation using the HET-CAM test method. The test substances were applied on the membrane of fertile eggs (Leghorn SA31, six per solvent) in a constant volume of 0.3 mL at 37ºC. The membrane, blood vessels, and albumen were examined for 5 minutes. The time of appearance, in seconds, of each irritant effect was recorded. No *in vivo* rabbit reference data were reported, but the Draize rabbit eye test data and HET-CAM results for 7/9 of these substances were included in the accuracy analyses reported in the ICCVAM BRD (2006a). Consequently, they have in turn already been considered in the current evaluation.

In Dahl (2007), 27 dental adhesive products in a total of 36 solutions based on four adhesive concepts (i.e., self-etch 1 step, self-etch 2 step, etch and rinse 2 steps, or etch and rinse 3 steps) were evaluated in the HET-CAM test method. The potential of dental adhesives to evoke irritation relevant to the biocompatibility of dental adhesives with regard to pulpal and mucous membrane exposure was assessed. An IS was obtained over a 5-minute observation period based on the time of first appearance of hemorrhage, vascular lysis, or coagulation in the chorioallantoic membrane. Substances were applied in a volume of 0.3 mL (n=3 eggs in two experiments). Products were classified based on conversion of the HET-CAM IS to a mean irritation score (i.e., nonirritant, slight irritant, moderate irritant, or strong irritant). Sixteen solutions were identified as strong irritants and found among all adhesive concept groups except the newest, self-etch 1 step. However, all substances in the self-etch 1 step group were classified as moderate irritants with IS scores close to those of a strong irritant. The results suggested that dental adhesives have the potential to cause an irritant reaction if exposed to oral mucosa. This HET-CAM data could not be used in the BRD performance analysis because no corresponding Draize data were provided.

Mehling et al. (2007) tested 18 proprietary surfactants using the red blood cell test, HET-CAM, and the SkinEthic™ ocular tissue model. Following the standard operating procedure of the Colipa project (INVITTOX Protocol No. 96), 300 microliters of test solution diluted in water were applied to the exposed CAM. The intensity of the subsequent reactions (i.e., hemorrhage, lysis, and coagulation) was semiquantitatively assessed on a scale of 0 to 3. No *in vivo* rabbit reference data were reported in this study; therefore, it was not included in the HET-CAM performance analysis detailed in this BRD.

In Mancebo et al. (2008), 14 proprietary formulations generally used in agriculture were tested in acute dermal toxicity and in eye irritation/corrosion tests. Three substances were tested using the

HET-CAM method and the acute eye irritation/corrosion test. Three hundred microliters of each test substance was applied to the CAM of fertile eggs (Lohman, six per substance) and observed for 5 minutes. The three endpoints for this study were hemorrhage, vessel lyses, and coagulation. Although mean *in vivo* rabbit eye data and corresponding irritation levels and HET-CAM IS values were reported in the study, the original animal data were not provided. Thus the study was not included in the HET-CAM performance analyses detailed in this BRD.

Several other studies on HET-CAM were reported. For example, Budai et al. (2004) tested three pesticide formulations in the HET-CAM test method using the IS(B) analysis method, but only qualitative results and no corresponding Draize data were provided. Tavaszi and Budai (2006) provided IS(B) scores for HET-CAM data but no corresponding Draize data on six agrochemical pesticides. Tavaszi and Budai (2007) reported HET-CAM data on six additional agrochemical formulations using the IS(B) analysis method and converted the scores to qualitative irritation indices that were compared to qualitative Draize results based on the maximum mean total score (MMTS). This data could not be used for regulatory classification and was not included in the performance analyses. Tavaszi et al. (2008) performed similar analyses on six additional agrochemical formulations.

10.0 How the HET-CAM Test Method Will Refine, Reduce, or Replace Animal Use

ICCVAM promotes the scientific validation and regulatory acceptance of new methods that refine, reduce, or replace animal use where scientifically feasible. Refinement, reduction, and replacement are known as the "three Rs" of animal protection. These principles of humane treatment of laboratory animals are described as:

- Refining experimental procedures such that animal suffering is minimized
- Reducing animal use through improved science and experimental design
- Replacing animal models with non-animal procedures (e.g., *in vitro* technologies), where possible (Russell and Burch 1992)

The HET-CAM test method has the potential to refine and reduce animal use in eye irritation testing. The HET-CAM test method would refine animal use by the *in vitro* identification of ocular corrosives and severe irritants, nonsevere irritants, or substances not labeled as irritants when used in a tiered-testing scheme. Substances identified as corrosives or severe irritants would be excluded from *in vivo* testing. Furthermore, the ability to identify mild and moderate ocular irritants would eliminate the need for *in vivo* testing, thus sparing rabbits from the pain associated with these types of substances. The HET-CAM test method can also reduce animal use because the test method does not use live animals. Use of the HET-CAM test method in lieu of one that uses live animals or animals used as a food source (e.g., BCOP, ICE, IRE) would further reduce the number of animals in a tiered-testing strategy.

10.1 Requirement for the Use of Animals

The HET-CAM test method has been designed so as not to require the use of animals. International regulations provide for the protection of animals used for experimental or other scientific purposes. For test methods using an animal embryo or fetus, some provisions indicate when an animal embryo or fetus is considered an animal and is therefore protected by the regulations. According to some of these regulations, a bird is considered a protected animal (thus the test is considered an *in vivo* and not *in vitro* test) when more than half of the gestation or incubation period has elapsed (Day 10.5 of the 21-day incubation period for a chicken embryo) (Animals [Scientific Procedures] Act 1986; EU 1986). The Public Health Service Policy, with which all National Institutes of Health (NIH)-funded research projects must comply, applies to all live vertebrate species. The NIH Office of Laboratory

Animal Welfare has provided written guidance in this area, interpreting “live vertebrate animal” to apply to avians (e.g., chick embryos) only after hatching (Kulpa-Eddy J, personal communication; NIH 2000).

It has been proposed that at incubation Day 9, the embryonic differentiation of the chicken central nervous system is sufficiently incomplete that suffering from pain perception is unlikely to occur (MSPCA 2005; Liebsch M, personal communication). Evaluations suggest that there are few sensory fibers present at Day 9 in the avian embryo and that significant development of the sensory nerve ending occurs between incubation Days 11 and 14 (Romanoff 1960). Studies also have suggested that the extraembryonal vascular systems (e.g., yolk sac, CAM) are not sensitive to pain (Rosenbruch 1997; Spielmann H, personal communication). Combined, these studies suggest that at incubation Day 9 the developing embryo perceives little or no pain during the conduct of the HET-CAM test method.

11.0 Practical Considerations

Practical considerations for the HET-CAM test method are detailed in the *Background Review Document: Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen’s Egg Test – Chorioallantoic Membrane Test Method* (ICCVAM 2006a).

12.0 References

Animals (Scientific Procedures) Act 1986. 1986. Chapter 14.

Bagley D, Bruner L, DeSilva O, Cottin M, O'Brien K, Uttley M, et al. 1992. An evaluation of five potential alternatives *in vitro* to the rabbit eye irritation test *in vivo*. Toxic In Vitro 6:275-284.

Balls M, Botham P, Bruner L, Spielmann H. 1995. The EC/HO international validation study on alternatives to the Draize eye irritation test. Toxic In Vitro 9:871-929.

Boue-Grabot M, Bernardin G, Chaumond S, Pinon JF. 1995. Alternative methods: hen's egg chorioallantoic membrane and in vitro cytotoxicity - a complementary approach. Intl J Cosmetic Science 17:207-215.

Budai P, Varnagy L, Fejes S, Somlyay IM, Linczmayer K, Pongracz A. 2004. Irritative effects of some pesticides and a technical component on tissue structure of the chorioallantoic membrane. Comm App Biol Sci 69:807-809.

CEC. 1991. Collaborative Study on the Evaluation of Alternative Methods to the Eye Irritation Test. Doc. XI/632/91/V/E/1/131/91 Parts I and II.

CPSC. 2003. Commercial practices. Hazardous Substances and Articles; Administration and Enforcement Regulations. 16 CFR 1500.42.

Dahl JE. 2007. Potential of dental adhesives to induce mucosal irritation evaluated by the HET-CAM method. Acta Odontologica Scandinavica 65:275-283.

De Silva O, Rougier A, Dossou KG. 1992. The HET-CAM test: a study of the irritation potential of chemicals and formulations. Altern Lab Anim 20:432-437.

Debbasch C, Ebenhahn C, Dami N, Pericoi M, Van den Berghe C, Cottin M, et al. 2005. Eye irritation of low-irritant cosmetic formulations: correlation of *in vitro* results with clinical data and product comparison. Food Chem Toxicol 43:155-165.

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 Pharm Exp Ther 82:377-390.

EEC. 1984. Commission Directive 84/449/EEC of 25 April 1984 adapting to technical progress for the sixth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (Acute Toxicity–Eye Irritation. Annex V, Part B.5). Official Journal of the European Communities L251:109.

EEC. 1991. Commission Directive 91/325/EEC of 1 March 1991 adapting to technical progress for the twelfth 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 L180:52–56.

EPA. 1998. Health Effects Test Guideline, OPPTS 870.2400 Acute Eye Irritation. EPA 712-C-98-195. Washington, DC:U.S. Environmental Protection Agency.

EPA. 2003a. Label Review Manual: 3rd ed. EPA735-B-03-001. Office of Prevention, Pesticides & Toxic Substances (OPPTS). Washington, DC:U.S. Environmental Protection Agency.

EPA. 2003b. Good Laboratory Practice Standards. Toxic Substances Control Act. 40 CFR 792.

EPA. 2003c. Good Laboratory Practice Standards. Federal Insecticide, Fungicide, and Rodenticide Act. 40 CFR 160.

EU. 1986. Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. Official Journal of the European Communities L358:1-28.

EU. 1992. Commission Directive 92/69/EEC of 31 July 1992 adapting to technical progress for the seventeenth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Communities L383:113-115.

EU. 1993. Commission Directive 93/21/EEC of 27 April 1993 adapting to technical progress for the 18th time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Communities L110A:20–21.

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 L225:1-333.

EU. 2004. Manual of Decisions for Implementation of the Sixth and Seventh Amendments to Directive 67/548/EEC on Dangerous Substances (Directives 79/831/EEC and 92/32/EEC). European Commission, Joint Research Centre. Report n. EUR 20519 EN, updated version of July 2004. Available: http://ecb.jrc.it/new-chemicals [accessed 14 October 2004].

FDA. 2003. Good Laboratory Practice for Nonclinical Laboratory Studies. 21 CFR 58.

Federal Hazardous Substances Act (FHSA). 2005. Public Law 86-613.

Gautheron P, Giroux J, Cottin M, Audegond L, Morilla A, Mayordomo-Blanco L, et al. 1994. Interlaboratory assessment of the bovine corneal opacity and permeability (BCOP) assay. Toxicol In Vitro 8:381-392.

Gettings S, Dipasquale L, Bagley D, Casterton P, Chudkowski M, Curren R, et al. 1994. The CTFA evaluation of alternatives program: an evaluation of *in vitro* alternatives to the Draize primary eye irritation test. (Phase II) oil/water emulsions. Food Chem Toxicol 32:943-976.

Gettings S, Lordo R, Hintze K, Bagley D, Casterton P, Chudkowski M, et al. 1996. The CTFA evaluation of alternatives program: an evaluation of *in vitro* alternatives to the Draize primary eye irritation test. (Phase III) surfactant-based formulations. Food Chem Toxicol 34:79-117.

Gettings S, Teal J, Bagley D, Demetrulias J, DiPasquale L, Hintze K, et al. 1991. The CTFA evaluation of alternatives program: an evaluation of *in vitro* alternatives to the Draize primary eye irritation test. (Phase 1) hydro-alcoholic formulations; (Part 2) data analysis and biological significance. In Vitro Toxicol 4:247-288.

Gilleron L, Coecke S, Sysmans M, Hansen E, van Oproy S, Marzin D, et al. 1996. Evaluation of a modified HET-CAM assay as a screening test for eye irritancy. Toxicol In Vitro 10:431-446.

Gilleron L, Coecke S, Sysmans M, Hansen E, van Oproy S, Marzin D, et al. 1997. Evaluation of the HET-CAM-TSA method as an alternative to the Draize eye irritation test. Toxicol In Vitro 11:641-644.

Hagino S, Kinoshita S, Tani N, Nakamura T, Ono N, Konishi K, et al. 1999. Interlaboratory validation of *in vitro* eye irritation tests for cosmetic ingredients. (2) chorioallantoic membrane (CAM) test. Toxicol In Vitro 13:99-113.

Harbell J, Curren RD. 1998. The bovine corneal opacity and permeability assay: observations on assay performance. In Vitro Mol Toxicol 11:337–341.

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 Institute of Environmental Health Sciences.

ICCVAM. 2006a. Background Review Document: Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test – Chorioallantoic Membrane Test Method. NIH Publication No. 06-4515. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

ICCVAM. 2006b. ICCVAM Test Method Evaluation Report: *In Vitro* Ocular Toxicity Methods for Identifying Severe Irritants and Corrosives. NIH Publication No. 07-4517. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

ICCVAM Authorization Act of 2000. 2000. Public Law 106-545. Available: http://iccvam.niehs.nih.gov/about/overview.htm. [accessed 21 October 2004].

Kalweit S, Besoke R, Gerner I, Spielmann H. 1990. A national validation project of alternative methods to the Draize rabbit eye test. Toxicol In Vitro 4:702-706.

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

Kojima H, Hanamura A, Miyamoto S, Sato A, Konishi H, Yoshimura I. 1995. Evaluation of seven alternative assays on the main ingredients in cosmetics as predictors of Draize eye irritation scores. Toxicol In Vitro 9:333-340.

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

Mancebo A, Hernandez O, Gonzalez Y, Aldana L, Carballo O. 2008. Assessment of skin and eye irritation of 14 products under the stepwise approach of the OECD. Cutan Ocul Toxicol 27:173-185.

Mehling A, Kleber M, Hensen H. 2007. Comparative studies on the ocular and dermal irritation potential of surfactants. Food Chem Toxicol 45:747-758.

Massachusetts Society for the Prevention of Cruelty to Animals (MSPCA). 2005. Animals in Education: Laws and Regulations. Available: http//www.mspca.org.

NIH. 2000. The Public Health Service Responds to Commonly Asked Questions. Available: http://grants.nih.gov.

OECD. 1998. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 1: OECD Principles on Good Laboratory Practice (as revised in 1997). ENV/MC/CHEM(98)17. Paris:Organisation for Economic Co-operation and Development.

OECD. 2002. Test Guideline 405. Acute eye irritation/corrosion, adopted April 24, 2002. In: OECD Guidelines for Testing of Chemicals. Paris:Organisation for Economic Co-operation and Development.

Romanoff AL. 1960. The Avian Embryo: Structural and Functional Development. New York:The Macmillan Company.

Rosenbruch M. 1997. The sensitivity of chicken embryos in incubated eggs. ALTEX 14:111-113.

Russell WMS, Burch RL. 1992. The Principles of Humane Experimental Technique. 14th ed. South Mimms, Potters Bar, Herts, England:Universities Federation for Animal Welfare.

Spielmann H, Liebsch M, Kalweit S, Moldenhauer F, Wirnsberger T, Holzhütter H, et al. 1996. Results of a validation study in Germany on two *in vitro* alternatives to the Draize eye irritation test, HET-CAM test and the 3T3 NRU cytotoxicity test. Altern Lab Anim 24:741-858.

Tavaszi J, Budai P. 2006. Toxicity study of agrochemicals on chorioallantoic membrane of the egg. Comm App Biol Sci 71:101-105.

Tavaszi J, Budai P. 2007. The use of HET-CAM test in detecting the ocular irritation. Comm App Biol Sci 72:137-141.

Tavaszi J, Budai P, Palovics A, Kismanyoki A. 2008. An alternative test battery in detecting ocular irritancy of agrochemicals. Comm App Biol Sci 73:891-895.

UN. 2007. Globally Harmonized System of Classification and Labelling of Chemicals (GHS). New York & Geneva:United Nations Publications.

Vinardell M, Macián M. 1994. Comparative study of the HET-CAM test and the Draize eye test for assessment of irritancy potential. Toxicol In Vitro 8:467-470.

Vinardell MP, Mitjans M. 2006. The chorioallantoic membrane test as a model to predict the potential human eye irritation induced by commonly used laboratory solvents. Toxicol In Vitro 20:1066-1070.

13.0 Glossary[6]

Accuracy:[*] (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of "relevance." The term is often used interchangeably with *concordance* (see also *two-by-two* table). Accuracy is highly dependent on the prevalence of positives in the population being examined.

Assay:[*] The experimental system used. Often used interchangeably with *test* and *test method*.

---

[6] The definitions in this Glossary are restricted to their uses with respect to the Draize rabbit eye test method and the HET-CAM test method.

[*] Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Benchmark control: A sample containing all components of a test system and treated with a known substance (i.e., the benchmark substance) to induce a known response. The sample is processed with test substance-treated and other control samples to compare the response produced by the test substance to the benchmark substance to allow for an assessment of the sensitivity of the test method to assess a specific chemical class or product class.

Benchmark substance: A substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties:

- a consistent and reliable source(s)
- structural and functional similarity to the class of substances being tested
- known physical/chemical characteristics
- supporting data on known effects
- known potency in the range of the desired response

Blepharitis: Inflammation of the eyelids.

Bulbar conjunctiva: The portion of the conjunctiva that covers the outer surface of the eye.

Chorioallantoic membrane (CAM): A vascularized respiratory fetal membrane that is composed of the chorion and allantois.

Classification system: An arrangement of quantified results or data into groups or categories according to previously established criteria.

Coagulation: The process of a liquid becoming viscous, jellylike, or solid by chemical reaction.

Coded substances: Substances labeled by code rather than name so that they can be tested and evaluated without knowledge of their identity or anticipation of test results. Coded substances are used to avoid intentional or unintentional bias when evaluating laboratory or test method performance.

Coefficient of variation: A statistical representation of the precision of a test. It is expressed as a percentage and is calculated as follows:

$$\left(\frac{standard\ deviation}{mean}\right) \times 100\%$$

Concordance:[*] The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance and one aspect of "relevance." The term is often used interchangeably with *accuracy* (see also *two-by-two table*). Concordance is highly dependent on the prevalence of positives in the population being examined.

Conjunctiva: The mucous membrane that lines the inner surfaces of the eyelids and folds back to cover the front surface of the eyeball, except for the central clear portion of the outer eye (the cornea). The conjunctiva is composed of three sections: palpebral conjunctiva, bulbar conjunctiva, and fornix.

Conjunctival sac: The space located between the eyelid and the conjunctiva-covered eyeball. Substances are instilled into the sac to conduct an *in vivo* eye test.

Cornea: The transparent part of the coat of the eyeball that covers the iris and pupil and admits light to the interior.

Corneal opacity: Measurement of the extent of opaqueness of the cornea following exposure to a test substance. Increased corneal opacity is indicative of damage to the cornea. Opacity can be evaluated subjectively, as done in the Draize rabbit eye test, or objectively with an instrument such as an opacitometer.

Corrosion: Destruction of tissue at the site of contact with a substance.

Corrosive: A substance that causes irreversible tissue damage at the site of contact.

Endpoint:* The biological process, response, or effect assessed by a test method.

False negative:* A substance incorrectly identified as negative by a test method.

False negative rate:* The proportion of all positive substances falsely identified by a test method as negative (see *two-by-two table*). It is one indicator of test method accuracy.

False positive:* A substance incorrectly identified as positive by a test method.

False positive rate:* The proportion of all negative substances that are falsely identified by a test method as positive (see *two-by-two table*). It is one indicator of test method accuracy.

Fibrous tunic: The outer of the three membranes of the eye, comprising the cornea and the sclera; called also *tunica fibrosa oculi*.

Globally Harmonised System (GHS): A classification system presented by the United Nations that provides (a) harmonized criteria for classifying substances and mixtures according to their health, environmental, and physical hazards; and (b) harmonized hazard communication elements, including requirements for labeling and safety data sheets.

Good Laboratory Practices (GLP):* Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the Organization for Economic Cooperation and Development and Japanese authorities that describe record-keeping and quality assurance procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.

Hazard:* The potential for an adverse health or ecological effect. A hazard potential results only if an exposure occurs that leads to the possibility of an adverse effect being manifested.

Hemorrhage: Discharge of blood from a vessel.

Hyperemia: Excess of blood in a body part.

Interlaboratory reproducibility:* A measure of whether different qualified laboratories using the same protocol and test substances can produce qualitatively and quantitatively similar results. Interlaboratory reproducibility is determined during the prevalidation and validation processes and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory repeatability:* The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Intralaboratory reproducibility:* The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.

In vitro: In glass. Refers to assays that are carried out in an artificial system (e.g., in a test tube or petri dish) and typically use single-cell organisms, cultured cells, cell-free extracts, or purified cellular components.

In vivo: In the living organism. Refers to assays performed in multicellular organisms.

Iris: The contractile diaphragm perforated by the pupil and forming the colored portion of the eye.

---

* Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Irritation score: Value calculated by different analysis methods, which is used to classify the irritancy potential of a test substance. Also referred to as *IS*.

Irritation Threshold Concentration: The lowest concentration of a test substance required to produce a weak or slight irritant response on the CAM. Also referred to as *ITC*.

IS(A) analysis method: HET-CAM analysis method where endpoints are observed at specified time points after application of the test substance (typically 0.5, 2, and 5 minutes post exposure). At the time points, presence of an endpoint is determined and a score assigned, if it is present. The scores are totaled to yield an overall irritation score.

IS(B) analysis method: HET-CAM analysis method where endpoints are observed over the entire observation period after application of the test substance (typically 5 minutes). The time (in seconds) when an endpoint develops is noted, and the times are used to yield an overall irritation score using a mathematical formula.

Lysis: The disintegration of blood vessels.

Mean Time to Coagulation (mtc): Mean detection time for appearance of coagulation endpoint.

Negative control: An untreated sample containing all components of a test system, except the test substance solvent, which is replaced with a known nonreactive material, such as water. This sample is processed with test substance-treated samples and other control samples to determine whether the solvent interacts with the test system.

Negative predictivity:[*] The proportion of correct negative responses among substances testing negative by a test method (see *two-by-two table*). It is one indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

Neuroectodermal tunic: The innermost of three membranes of the eye, comprising the retina.

Nictating membrane: The membrane that moves horizontally across the eye in some animal species (e.g., rabbit, cat) to provide additional protection in particular circumstances. It may be referred to as the *third eyelid*.

Not Labeled: (a) A substance that produces no changes in the eye following application to the anterior surface of the eye. (b) Substances that are not classified as GHS Category 1, 2A, or 2B; or EU R41 or R36 ocular irritants.

Nonsevere irritant: (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye; the tissue damage is reversible within 21 days of application and the observed adverse effects in the eye are less severe than observed for a severe irritant. (b) Substances that are classified as GHS Category 2A or 2B; EPA Category II, III, or IV; or EU R36 ocular irritants.

Ocular: Of or relating to the eye.

Ocular corrosive: A substance that causes irreversible tissue damage in the eye following application to the anterior surface of the eye.

Ocular irritant: A substance that produces a reversible change in the eye following application to the anterior surface of the eye.

Palpebral conjunctiva: The part of the conjunctiva that covers the inner surface of the eyelids.

---

[*] Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Pannus: A specific type of corneal inflammation that begins within the conjunctiva, and with time spreads to the cornea. Also referred to as *chronic superficial keratitis*.

Performance:* The accuracy and reliability characteristics of a test method (see *accuracy*, *reliability*).

pH: A measure of the acidity or alkalinity of a solution; pH 7.0 is neutral, higher pHs are alkaline, lower pHs are acidic.

Positive control: A sample containing all components of a test system and treated with a substance known to induce a positive response, which is processed with the test substance-treated and other control samples to demonstrate the sensitivity of each experiment and to allow for an assessment of variability in the conduct of the assay over time.

Positive predictivity:* The proportion of correct positive responses among substances testing positive by a test method (see *two-by-two table*). It is one indicator of test method accuracy. Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.

Prevalence:* The proportion of positives in the population of substances tested (see *two-by-two table*).

Protocol:* The precise, step-by-step description of a test, including the listing of all necessary reagents, criteria and procedures for the evaluation of the test data.

Q-score: HET-CAM analysis method that calculates the ratio from the irritation score of a test substance compared to the irritation score of a reference substance. This HET-CAM analysis method is typically used with transparent test substances.

Quality assurance:* A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures is assessed independently by individuals other than those performing the testing.

Reduction alternative:* A new or modified test method that reduces the number of animals required.

Reference test method:* The accepted *in vivo* test method used for regulatory purposes to evaluate the potential of a test substance to be hazardous to the species of interest.

Refinement alternative:* A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being.

Relevance:* The extent to which a test method correctly predicts or measures the biological effect of interest in humans or another species of interest. Relevance incorporates consideration of the accuracy or concordance of a test method.

Reliability:* A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. It is assessed by calculating intra- and interlaboratory reproducibility and intralaboratory repeatability.

Replacement alternative:* A new or modified test method that replaces animals with nonanimal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate).

Reproducibility:* The consistency of individual test results obtained in a single laboratory (intralaboratory reproducibility) or in different laboratories (interlaboratory reproducibility) using the same protocol and test substances (see intra- and interlaboratory reproducibility).

---

* Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

Sclera: The tough, fibrous tissue that extends from the cornea to the optic nerve at the back of the eye.

Sensitivity:* The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy (see *two-by-two table*).

Secondary bacterial keratitis: Inflammation of the cornea that occurs secondary to another insult that compromised the integrity of the eye.

Severe irritant: (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye that is not reversible within 21 days of application or causes serious physical decay of vision. (b) Substances that are classified as GHS Category 1, EPA Category I, or EU R41 ocular irritants.

Solvent control: An untreated sample containing all components of a test system, including the solvent that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the same solvent. When tested with a concurrent negative control, this sample also demonstrates whether the solvent interacts with the test system.

Specificity:* The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy (see *two-by-two table*).

S-score: HET-CAM analysis method that totals the severity scores for each endpoint evaluated. The highest total score is used as the S-score. This HET-CAM analysis method is typically used with nontransparent test substances.

Test:* The experimental system used; used interchangeably with *test method* and *assay*.

Test method:* A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with *test* and *assay*. See also *validated test method* and *reference test*.

Test method components: Structural, functional, and procedural elements of a test method that are used to develop the test method protocol. These components include unique characteristics of the test method, critical procedural details, and quality control measures.

Tiered testing: A testing strategy where all existing information on a test substance is reviewed, in a specified order, prior to *in vivo* testing. If the irritancy potential of a test substance can be assigned, based on the existing information, no additional testing is required. If the irritancy potential of a test substance cannot be assigned, based on the existing information, a step-wise animal testing procedure is performed until an unequivocal classification can be made.

Toxic keratoconjunctivitis: Inflammation of the cornea and conjunctiva due to contact with an exogenous agent. Used interchangeably with *contact keratoconjunctivitis*, *irritative keratoconjunctivitis*, and *chemical keratoconjunctivitis*.

Transferability:* The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Two-by-two table:* The two-by-two table can be used for calculating accuracy (concordance) ([a+d]/[a+b+c+d]), negative predictivity (d/[c+d]), positive predictivity (a/[a+b]), prevalence

---

* Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).

([a+c]/[a+b+c+d]), sensitivity (a/[a+c]), specificity (d/[b+d]), false positive rate (b/[b+d]), and false negative rate (c/[a+c]).

		New Test Outcome		
		Positive	Negative	Total
Reference Test Outcome	Positive	a	c	a + c
	Negative	b	d	b + d
	Total	a + b	c + d	a + b + c + d

Uvea tract: The middle of three membranes of the eye, comprising the iris, ciliary body, and choroid. Also referred to as the *vascular tunic*.

Validated test method:[*] An accepted test method for which validation studies have been completed to determine the relevance and reliability of this method for a specific proposed use.

Validation:[*] The process by which the reliability and relevance of a procedure are established for a specific purpose.

Vascular tunic: The middle of three membranes of the eye, comprising the iris, ciliary body, and choroid. Also referred to as the *uvea*.

Weight of evidence (process): The strengths and weaknesses of a collection of information are used as the basis for a conclusion that may not be evident from the individual data.

[*] Definition used by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM 2003).