Draft 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

March 2009
NIH Publication No.

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

[This Page Intentionally Left Blank]

Table of Contents

List of Tables

193 # List of Figures

List of Abbreviations and Acronyms

BCOP	Bovine Corneal Opacity and Permeability
BRD	Background Review Document
CAM	Chorioallantoic membrane
CASRN	Chemical Abstracts Service Registry Number
CPSC	U.S. Consumer Product Safety Commission
^{o}C	Degrees centigrade
EC	European Commission
EC/HO	European Commission/British Home Office
ECVAM	European Center 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
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
INVITOXX	*In Vitro* Techniques in Toxicology (ERGATT FRAME ECVAM Data bank)
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

MeSH	U.S. National Library of Medicine's Medical Subject Heading
MAS	Maximum average score
mtc	Mean time of coagulation
NL	Not Labeled
NICEATM	National Toxicology Program Center for the Evaluation of Alternative Toxicological Methods
NIH	National Institutes of Health
OECD	Organisation for Economic Cooperation and Development
OPPTS	EPA Office of Prevention, Pesticides and Toxic Substances
OSHA	U.S. Occupational Safety & Hazards 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.

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
Peter J. Schultheiss, D.V.M., D.A.C.L.A.M.
Harry Salem, Ph.D.

Department of Energy
• Michael Kuperberg, Ph.D.
◊ Marvin Stodolsky, Ph.D.

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

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

Environmental Protection Agency
Office of Science Coordination and Policy
• Karen Hamernik, Ph.D.
Office of Research and Development
◊ Julian Preston, Ph.D.
Office of Pesticide Programs
Deborah McCall
OECD Test Guidelines Program
Jerry Smrchek, Ph.D.

• Principal agency representative
◊ Alternate principal agency representative

Food and Drug Administration
Office of Science
• Suzanne Fitzpatrick, Ph.D., D.A.B.T.
Center for Drug Evaluation and Research
◊ Abigail C. Jacobs, Ph.D.
Paul C. Brown, Ph.D.
Center for Devices and Radiological Health
Melvin E. Stratmeyer, Ph.D.
Vasant G. Malshet, Ph.D., D.A.B.T.
Center for Biologics Evaluation and Research
Richard McFarland, Ph.D., M.D.
Ying Huang, Ph.D.
Center for Food Safety and Nutrition
David G. Hattan, Ph.D.
Robert L. Bronaugh, Ph.D.
Center for Veterinary Medicine
Devaraya Jagannath, Ph.D.
M. Cecilia Aguila, D.V.M.
National Center for Toxicological Research
William T. Allaben, Ph.D.
Paul Howard, Ph.D.
Donna Mendrick, Ph.D.
Office of Regulatory Affairs
Lawrence D'Hoostelaere, Ph.D.

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

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

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

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

National Library of Medicine
• Pertti (Bert) Hakkinen, Ph.D.
◊ Jeanne Goshorn, M.S.

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

Acknowledgements

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

U.S. Consumer Product Safety Commission
Cassandra Prioleau, Ph.D.
Marilyn Wind, Ph.D., (ICCVAM Chair)

Department of Defense
Harry Salem, Ph.D.

Department of Transportation
Steve Hwang, Ph.D.

U.S. Environmental Protection Agency
Office of Pesticide Programs
Meta Bonner, Ph.D.
Jonathan Chen, Ph.D.
Masih Hashim, D.V.M., Ph.D.
Karen Hicks
Marianne Lewis
Deborah McCall
Timothy McMahon, Ph.D.
Mark Perry, Ph.D.
John Redden, Ph.D.
Amy Rispin, Ph.D.
Jenny Tao, Ph.D.

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

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

U.S. Food and Drug Administration
Center for Drug Evaluation and Research
Paul C. Brown, Ph.D.
Abigail Jacobs, Ph.D.
Jill Merrill, Ph.D. (OTWG Co-Chair)

Center for Food Science and Nutrition
Robert Bronaugh, Ph.D.
Donnie Lowther

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

National Institute of Environmental Health Sciences
Mark Cesta, D.V.M, D.A.C.V.P.
Raymond (Buck) Grissom, Ph.D.
William S. Stokes, D.V.M., D.A.C.L.A.M. (Director, NICEATM)
Raymond R. Tice, Ph.D.

Occupational Safety and Health Administration (OSHA)
Surrender Ahir, Ph.D.

European Centre for the Validation of Alternative Methods – Liaison
João Barroso, Ph.D.
Thomas Cole, Ph.D.
Chantra Eskes, 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., D.A.C.L.A.M.
Director; Project Officer

Deborah McCarley
Special Assistant; Assistant Project Officer

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

David Allen, Ph.D.
Senior Toxicologist/Principal Investigator

Jonathan Hamm, Ph.D.
Senior Toxicologist

Nelson Johnson
Senior Project Coordinator/Technical Writer

Elizabeth Lipscomb, Ph.D.
Staff Toxicologist

Linda Litchfield
Meeting Coordinator/Admin. Asst.

Greg Moyer, M.B.A.
Project Manager

Catherine Sprankle
Senior Communications Specialist

James Truax
Senior Project Coordinator/Technical Writer

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
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

NICEATM gratefully acknowledges the generous contributions of the companies and individuals who provided data for this review. Their time and efforts to sort through study archives and compile data for this document are greatly appreciated.

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

ECVAM
Chantra Eskes

Johnson & Johnson Pharmaceutical R&D
Philippe Vanparys
Freddy van Goethem

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

U.S. Food and Drug Administration
Donnie Lowther

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

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, U.S. Environmental Protection Agency (EPA), U.S. Food and Drug Administration, and U.S. Occupational Health and Safety 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 severe (irreversible) ocular irritants/corrosives using the EPA, United Nations Globally Harmonized System of Classification and Labeling of Chemicals (GHS), and European Union regulatory hazard classification systems. In ICCVAM's assessment, the performance of the BCOP and ICE assays substantiated their use in testing some substances for regulatory hazard classification. The IRE and HET-CAM assays 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* (NIH Publication No: 07-4517, available at http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmer.htm). 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.

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) and nonirritants. 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 (BRDs) 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 BRDs 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 (ECVAM) and the Japanese Centre for the Validation of Alternative Methods (JaCVAM) will provide input and contribute to the OTWG throughout the evaluation process.

An international independent scientific peer review panel (Panel) will convene in public forum on May 19–21, 2009, to develop conclusions and recommendations on the *in vitro* BCOP, ICE, IRE, and HET-CAM test methods. The Panel includes expert scientists nominated by ECVAM and JaCVAM. We anticipate that these organizations can use the subsequent independent Panel report to deliberate and develop their own test method recommendations. The Panel will consider these BRDs and evaluate the extent to which the available information supports the draft ICCVAM test method recommendations. ICCVAM will consider the conclusions and recommendations of the Panel, along with comments from the public and the Scientific Advisory Committee on Alternative Toxicological Methods, and then finalize the BRD and test method recommendations. These 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 contributing to the preparation of this summary 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, Gregory Moyer, Catherine Sprankle, and Jim Truax. We also thank the members of the OTWG, chaired by Karen Hamernik, Ph.D. (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 ECVAM and the JaCVAM, 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., D.A.C.L.A.M.
Rear Admiral, U.S. Public Heath Service
Director, NICEATM
Executive Director, ICCVAM
March 2009

Executive Summary

Background

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, replacing, and refining the use of rabbits in the current *in vivo* eye irritation test method (69 FR 13859 [March 24, 2004]). In response to this nomination, ICCVAM 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. ICCVAM evaluated the test methods' ability to identify severe (irreversible) ocular irritants/corrosives using the EPA, United Nations Globally Harmonized System of Classification and Labeling of Chemicals (GHS), and European Union (EU) regulatory classification systems. ICCVAM considered two of the alternative test methods, the BCOP assay and ICE assay, to have sufficient performance to substantiate their use for regulatory hazard classification testing of some types of substances. The IRE and HET-CAM assays lacked sufficient performance and/or sufficient data to substantiate their use for regulatory hazard classification. ICCVAM subsequently recommended that the BCOP and ICE methods should be used in a tiered-testing strategy, where positive substances can be classified as ocular corrosives or severe irritants without the need for animal testing. These recommendations were forwarded to U.S. Federal agencies for consideration, 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 for identifying nonsevere ocular irritants (i.e., those that induce reversible ocular damage) and substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified). Accordingly, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), in conjunction with an ICCVAM Ocular Toxicity Working Group (OTWG) 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 the available information for the HET-CAM test method.

HET-CAM Test Method Protocol

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) since 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. The CAM is evaluated for the development of irritant endpoints (hyperemia, hemorrhage, and coagulation). 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.

Validation Database

A total of 260 substances and formulations were evaluated among all of the studies under consideration. The chemical classes with the greatest amount of HET-CAM data are alcohols (n = 75), carboxylic acids (n= 51), and formulations (n = 53). For some of the test substances that were identified as formulations, components of the formulation and the relative concentrations of the components were available. The most common product classes tested are solvent, shampoo, surfactants, and cosmetics. Analyses of each of the multiple HET-CAM protocols indicate that the Irritation Score (A) (IS[A]) analysis method achieved the best performance when evaluating substances not labeled as irritants. A total of 63 test substances are included in the available IS(A) database, 60 of which had sufficient *in vivo* data to be assigned an ocular irritancy hazard classification. Among these 60 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 were represented by more than three substances).

Detailed *in vivo* data, consisting of cornea, iris and conjunctiva scores for each animal at 24, 48, and 72 hours and/or assessment of the presence or absence of lesions at 7, 14, and 21

603 days was necessary to calculate the appropriate EPA (1996), EU (2001), and GHS (UN 2003)
604 ocular irritancy hazard classification. Thus, some of the test substances for which there was
605 only limited *in vivo* data could not be used for evaluating test method accuracy and
606 reliability.

607 **HET-CAM Test Method Accuracy**

608 ***Identification of All Ocular Hazard Categories***

609 The ability of the HET-CAM test method to identify all categories of ocular irritation
610 potential, as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001;
611 UN 2003) was evaluated. As indicated in **Table 1**, the overall correct classification for the
612 HET-CAM test method ranged from 38% (23/60) to 41% (24/59), depending on the
613 classification system used.

614 Because a specific analysis method is the focus of the evaluation of HET-CAM for
615 identifying all hazard categories (the IS[A] analysis method), separate analyses were also
616 conducted for all chemical classes and specific physical properties of interest represented in
617 this database of 60 substances by at least five substances (i.e., surfactant based formulations,
618 oil/water emulsions, and alcohols). The results indicate that alcohols tend to be overpredicted
619 by HET-CAM; that is, 75% [6/8] to 88% [7/8] of alcohols classified as mild irritant or not
620 labeled based on Draize test results (and depending on the classification system used) were
621 overpredicted by HET-CAM by at least one hazard category. Similarly, approximately half
622 (44% [8/18] to 53% [9/17]) of the oil/water emulsions were overpredicted by HET-CAM by
623 at least one hazard category. By comparison, surfactants classified as ocular corrosives or
624 severe irritants based on Draize results tended to be underpredicted by HET-CAM (73%
625 [13/17] to 75% [12/16] ocular corrosives or severe irritants underpredicted by HET-CAM as
626 mild or moderate irritants). However, none of these substances were underpredicted as not
627 labeled.

628 Given the proportion of substances in the HET-CAM IS(A) database represented by these
629 chemical and product classes (i.e., 85% [51/60] of the substances are included in one of these
630 three categories), separate analyses without these discordant substances are not particularly
631 informative. However, because of the associated discordance with each type, overall

632 performance particularly for the ocular corrosive and severe irritant category can be
633 improved by excluding certain product types (i.e., surfactant based formulations).

634 **Table 1 Evaluation of the Performance of the HET-CAM Test Method In Predicting Ocular Irritant Classes Compared to**
635 **the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA, EU, or GHS Classification Systems**

Data Source	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)
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)
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)

636 Abbreviations: EPA = Environmental Protection Agency Hazard Classification System (EPA 1998); EU = European Union Hazard Classification System (EU
637 2007); GHS = Globally Harmonized System (UN 2007); HET-CAM = Hen's Egg Test - Chorioallantoic Membrane

It is apparent from **Table 1** that the number of substances (n = 0-2) in the moderate irritant category (i.e., EPA Category II, EU R36, and GHS Category 2A) that an adequate evaluation of HET-CAM performance for this category is not feasible. Similarly, while there are 18 substances classified as EPA Category III, there are only five substances classified as GHS Category 2B (the EU system does not distinguish mild irritants). This trend is also apparent when evaluating the correct classification for the corrosive/severe substances.

Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories

The ability of the HET-CAM test method to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) from all other ocular hazard categories (i.e., EPA Category I, II, or III; EU R41 or R36; GHS Category 1, 2A, or 2B) was also evaluated. Again, this same analysis was performed without specific chemical classes and/or physical properties.

As indicated in **Table 2**, overall accuracy ranged from 62% (41/59) to 78 (47/60)% depending on the hazard classification system used. Overall accuracy for the identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) from all other categories ranged from 58% (36/58) to 60% (47/60) depending on the hazard classification system used. False positive and false negative rates ranged from approximately 60% (9/15) to 69% (22/32) and 0% (0/26) to 9% (4/45), respectively. The lowest false negative rate (0% [0/26 or 0/31]) was noted for the EU and GHS systems, respectively followed by 9% (4/45) for the EPA system. For all three systems, the correctly identified substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) were cosmetic formulations that were either oil/water emulsions or surfactant containing formulations. Among the four false negatives for the EPA system, 100% (4/4, all oil/water emulsion cosmetic formulations) were EPA Category III substances based on a conjunctival redness score of two that required at least three days to resolve. For one of the substances, one out of the six rabbits tested had a conjunctival redness score of two that required 14 days to resolve. Four of the remaining five rabbits in this study had conjunctival redness scores of two that resolved within three days; the last rabbit did not have this lesion.

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

Data Source	N[2]	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.[3]	%	No.	%	No.	%	No.	%	No.
Overall (EPA)	60	78	47/60	91	41/45	40	6/15	60	9/15	9	4/45
Overall (EU)	58	62	36/58	100	26/26	31	10/32	69	22/32	0	0/26
Overall (GHS)	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31

Abbreviations: EPA = Environmental Protection Agency Hazard Classification System (EPA 1998); EU = European Union Hazard Classification System (EU 2007); GHS = Globally Harmonized System (UN 2007); HET-CAM = Hen's Egg Test – Chorioallantoic Membrane

[This Page Intentionally Left Blank]

HET-CAM Test Method Reliability

Quantitative and qualitative evaluations of HET-CAM test method reliability have been conducted previously (ICCVAM 2006a). Since 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.

Interlaboratory Reproducibility

However, additional qualitative analyses of interlaboratory reproducibility were conducted to evaluate the extent of agreement of HET-CAM hazard classifications among the five participating laboratories from the interlaboratory validation study (Hagino et al. 1999). As for the accuracy evaluation study, qualitative evaluations of reproducibility were conducted based on 1) the use of the HET-CAM test method for identifying all ocular hazard categories according to the EPA, EU, or GHS systems, and 2) the use of the HET-CAM test method to distinguish substances not labeled as irritants from all other irritant categories.

Using the first approach (i.e., identifying all ocular hazard categories), there was 100% agreement among the five laboratories for a majority of the Draize ocular corrosives/severe irritants correctly classified by HET-CAM based on all three classification systems (i.e., 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 HET-CAM. There also was 100% agreement for the mild ocular irritants (i.e., EPA Category III, GHS Category 2B; the EU does not have a mild irritant category), which were uniformly overpredicted. For the Hagino et al. (1999) database, all of the substances not classified as irritants (based on Draize results; i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) were overclassified by HET-CAM. There was 100% among the five laboratories for 86% (6/7) or 75% (3/4) of these substances for the EU and GHS systems, respectively. By comparison, for the two EPA Category IV substances tested, there was either 100% agreement or 80% among the five laboratories.

Using the second approach (i.e., distinguishing substances not labeled as irritants from all other ocular hazard categories), there was 100% agreement among the five laboratories for 82% (14/17), 76% (13/17), and 94% (16/17) for the 17 substances included in the Hagino et al. (1999) database for the EPA, EU, and GHS classification systems, respectively.

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

There was 100% agreement among the five laboratories for 63% (5/8) of the substances correctly identified as an irritant according to the EU system (i.e., 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 HET-CAM, there was 100% agreement among the five laboratories for 86% (6/7) of these substances that were overpredicted by HET-CAM.

There was 100% agreement among the five laboratories for 100% (11/11) of the substances correctly identified as an irritant according to the GHS system (i.e., Category 1, 2A, or 2B). While none of the GHS Not Classified substances were correctly identified by HET-CAM, there was 100% agreement among the five laboratories for 75% (3/4) of these substances that were overpredicted by HET-CAM.

As stated above, this BRD provides a comprehensive summary of the current validation status of the HET-CAM test method, including what is known about its reliability and accuracy, and the scope of the substances tested. Raw 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.

[This Page Intentionally Left Blank]

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 that allows for relative categorization of severity for reversible effects such as mild, moderate, or severe irritants (e.g., see U.S. Environmental Protection Agency [EPA] Ocular Classification System discussed below). Current EPA ocular testing guidelines and the United Nations (UN) Globally Harmonized System (GHS) of Classification and Labeling of Chemicals (UN 2003) indicate that if serious ocular damage is anticipated (e.g., irreversible adverse effects on day 21), then a test on a single animal may be considered. If serious damage is observed, then no further animal testing is necessary (EPA 1998; UN 2003). If serious damage is not observed, additional test animals (1 or 2 rabbits) may be evaluated sequentially until concordant irritant or not labeled responses are observed (UN 2003).

In 2006, ICCVAM completed an evaluation of the Hen's Egg Test – Chorioallantoic Membrane (HET-CAM) to identify ocular corrosives and severe irritants (ICCVAM 2006a). Following this review, ICCVAM concluded the HET-CAM test method was not suitable for identifying ocular corrosives and severe irritants (i.e., EPA Category I, UN GHS Category 1, EU R41) (ICCVAM 2006b), but this recommendation could be revised as additional data becomes available.

ICCVAM is now conducting an evaluation to further characterize the usefulness and limitations of the HET-CAM test method for identifying non-severe irritants and substances not labeled as irritants (i.e., EPA Category II, III, IV; UN GHS Category 2A, 2B, NL, EU R36, NL) (ICCVAM 2006b). As part of this 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 number and type of substances tested, and the availability of a standardized protocol. This BRD was prepared for use by an ICCVAM expert panel to aid in the review of HET-CAM as a method to identify all categories of ocular irritants and substances not labeled as irritants. Parallel reviews of the ICE, IRE, and BCOP test methods are being conducted. Results of the Peer Review Panel Report, combined with the analyses

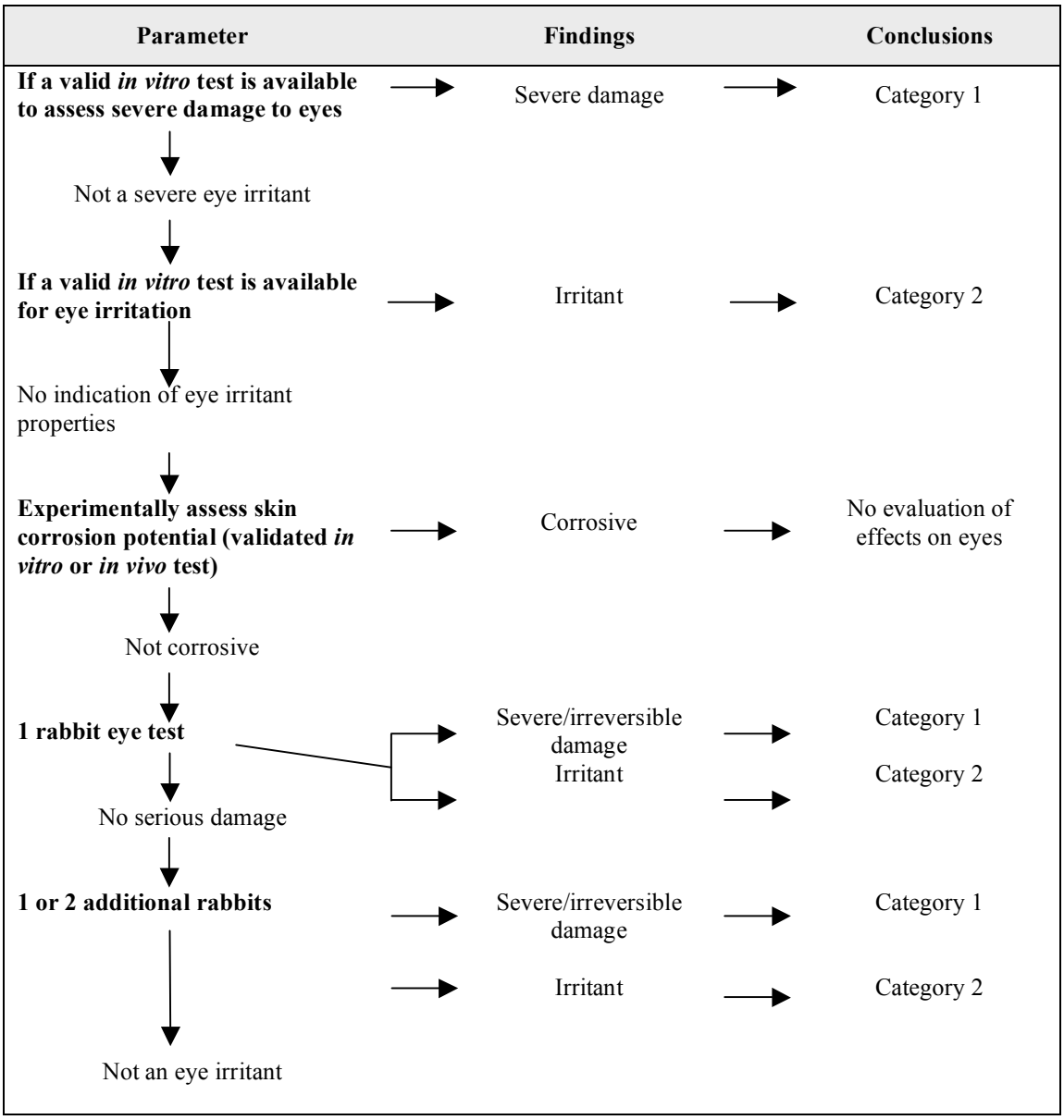
presented in the BRDs, will be used to support ICCVAM recommendations on the proposed standardized test method protocols, proposed list of recommended reference substances, and additional optimization and/or validation studies that may be necessary to further develop and characterize the usefulness and 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 BRD, *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 also allows for use of validated and accepted *in vitro* methods to identify severe ocular irritants/corrosives and ocular irritants without further testing. The HET-CAM test method is currently not recommended for use in identifying ocular corrosives and severe irritants in a tiered-testing strategy for regulatory classification and labeling for use in the GHS testing scheme (UN 2003). As indicated above, ICCVAM is now conducting an evaluation to further characterize the usefulness and limitations of the HET-CAM test method for identifying nonsevere irritants and substances not labeled as irritants.

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



803 Abbreviations: GHS = United Nations Globally Harmonized System of Classification and Labelling of
804 Chemicals

805 [1]Adapted from UN (2002).

806

807 ## 1.3 Validation of the HET-CAM Test Method

808 The ICCVAM Authorization Act (Sec. 4(c)) mandates that "[e]ach Federal Agency … shall
809 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].” (Public Law [P.L.] 106-545).

Validation is the process by which the reliability and relevance of an assay for a specific purpose are established (ICCVAM 2003). Relevance is defined as the extent to which an assay will correctly predict or measure the biological effect of interest (ICCVAM 2003). For the HET-CAM test method described in the ICCVAM 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. Reliability is defined as the reproducibility of a test method within and among laboratories and should be based on its performance with a diverse set of substances that are representative of the types of chemical and product classes that are expected to be tested and cover the range of responses that need to be identified. The validation process will provide data and information that will 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 the eye irritation potential of substances.

The first stage in this evaluation is the preparation of a BRD that presents and evaluates the relevant data and information about the assay, 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 performances of the assay 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 BRD, *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 published HET-CAM studies over the period from January 2005 to January 2009 using the same terminology and information databases as used in the 2006 ICCVAM BRD (ICCVAM 2006a) revealed four studies with available information on HET-CAM protocols or contained

839 data on test substances. While no *in vivo* reference data were included in any of the four
840 citations, *in vivo* data for six of nine substances included in one study were available from the
841 NICEATM database of Draize eye test results. However, these substances were already
842 included in the original analyses (and the HET-CAM results from the new study were in
843 agreement with the previous results), the database used in the HET-CAM performance
844 analysis is the same as the database used in ICCVAM (2006a).

845

[This Page Intentionally Left Blank]

2.0 HET-CAM 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) since 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 ICCVAM (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 multiple HET-CAM analysis methods indicates 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 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 totaled to give an 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 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 ocular irritancy classification scheme described in **Table 2-1** was used for the accuracy analysis presented in this BRD (see **Section 6.0**). Therefore, substances with an IS(A) or IS(B) value of nine 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 (1996), GHS (UN 2003), and EU (2001) classification systems is unknown.

914 **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

915 [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

In vitro ocular test method validation studies 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.

As noted in **Section 1.5**, although new HET-CAM data were identified among four studies published since the ICCVAM evaluation of HET-CAM for identifying 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., 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; Hagino et al. 1999). 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 are available for 60 compounds.

Table 3-1 and **Table 3-2** show **t**he 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 using the test substance are provided in **Appendix A**. However, if a product class was not assigned in the study report, this information was sought from other sources, including the National Library of Medicine's ChemID Plus 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) that 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.

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 3-1**, 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 **Appendix B**, 28 substances, including formulations and mixtures of unknown composition, could not be assigned a specific chemical class.

As shown in **Table 3-2**, the most common product classes tested in the HET-CAM assay are solvents (n = 13), hair shampoos (n = 13), surfactants (n = 17), and cosmetics (n = 14). Of the 504 substances included in **Appendix B,** 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 60 have available *in vivo* reference data. Among these 60 substances

959 are 43 cosmetic and personal care product formulations (including 25 surfactant based
960 formulations and 18 oil/water emulsions) and 17 individual substances (including seven
961 alcohols; no other classes represented by more than three substances).

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

Product Class	# of Substances
Aerosol formulation ingredient	1
Anti-freezing agent	1
Anti-infective agent, Anti-bacterial agent	2
Anti-perspirant	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
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

963

964

965

[This Page Intentionally Left Blank]

4.0 *In Vivo* Reference Data Used for an Assessment of 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 ICCVAM (2006). 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 (1996), the EU (2001), and the GHS (UN 2003) ocular irritancy classification systems (**Appendix D**). Exceptions included the following:

- *In vivo* data used by Gilleron et al. (1996) were obtained from the studies of Gautheron et al. (1994). According to the report, the studies were performed according to the French and European directives (EEC 1984, 1991). Substances were classified by the authors according to the EU (1993) classification system and used to assess the *in vitro* test method accuracy.

4.1 *In Vivo* Classification Criteria Used for BRD Analysis

As described in ICCVAM (2006a), the *in vivo* rabbit eye database used to conduct retrospective analyses of 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 currently devised to be applied to studies using no more than three rabbits. Thus, to maximize the amount of data used for the evaluation of HET-CAM, as well as for the three other *in vitro* test methods (ICE, IRE, BCOP) being evaluated, the decision criteria for each classification system were expanded to include studies that used more than three rabbits in their evaluation.

All classification systems require the scoring of rabbits using the Draize scoring system,

which occurs 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 this BRD, four criteria must apply. These criteria were:

- 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 less than three rabbits.
- A volume of 0.1 mL or 0.1 g was tested in each rabbit. A study in which a lower quantity was applied to the eye was 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 must have been made, at a minimum, at 24, 48, and 72 hours following test substance application if no severe effect was observed.
- Observations of the eye must have been 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 ICCVAM (2006a).

4.2 *In Vivo* Data Quality

Ideally, all data supporting the validity of a test method should be obtained and reported from studies conducted in accordance with Good Laboratory Practice (GLP) guidelines, which are nationally and internationally recognized rules designed to produce high-quality laboratory records (OECD 1998; EPA 2003a, 2003b; 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, in order to ensure the integrity, reliability, and accountability of a study.

The extent to which the *in vivo* rabbit eye studies, which were used to provide the comparative data in the published HET-CAM validation studies, were compliant 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).

[This Page Intentionally Left Blank]

5.0 HET-CAM Test Method Data and Results

A total of 12 published reports contained sufficient data for an accuracy analysis of the HET-CAM test method for the identification of all categories of ocular irritation. These reports are: 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

NICEATM requested original HET-CAM data for substances that also had been tested *in vivo* using the standard rabbit eye test by *FR* notice (69 FR 13589) published on March 24, 2004 and 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) and available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_05_3831.pdf. In addition, authors of selected published HET-CAM studies were contacted and asked to provide 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) were obtained for the 60 substances evaluated by Balls et al. (1995) from 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) were obtained from Drs. H. Spielmann and M. Liebsch. The data provided included the overall HET-CAM scores obtained by each laboratory for each substance evaluated. *In vitro* data for two control substances also were provided.
- Drs. Philippe Vanparys and Freddy Van Goethem provided individual endpoint scores for each egg evaluated for substances described in Gilleron et al. (1996, 1997). *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 (ITC) was determined. The ITC was typically combined with the IS for the test substance to evaluate ocular irritation or corrosivity potential of a substance.

The focus of the accuracy analysis in this BRD is on the ability of the HET-CAM test method to identify moderate and mild irritants, as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001; UN 2003). 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 1996), the EU (EU 2001), or the GHS (UN 2003). However, there have been attempts by some investigators (Gettings et al. 1991, 1994, and 1996; Spielmann et al. 1996) to correlate HET-CAM scores with the ocular irritation classification scheme described by the Federal Hazardous Substances Act classification system (CPSC 1988) and by the EU classification system (EU 1992), respectively.

To evaluate the ability of HET-CAM to identify all ocular hazard categories, as defined by the EPA (1996), GHS (UN 2003), and EU (2001) classification systems, HET-CAM results obtained using each of the different analysis methods were assigned 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, ranging 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.

A preliminary evaluation conducted by NICEATM using the various analysis methods (see **Section 6.1** and **Appendix E**) indicated that only the IS(A) analysis method had adequate accuracy with which 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 is shown in **Appendix D**. This table, sorted by reference, provides the CASRN (if available), the concentration tested, the calculated *in vitro* score, the *in vitro* irritation classification of the test substance (based on the irritation classification schemes in **Section 5.3**), the *in vivo* reference classifications (i.e., GHS, EPA, EU), and the 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. When provided, the specific information extracted for each substance included its name, CASRN (if available), chemical class, product class, concentration tested, form tested, *in vitro* classification, and reference. If not provided, the CASRN was obtained from various sources, including the National Library of Medicine's ChemID 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 based on the classification of the National Library of Medicine's Medical Subject Heading (MeSH; available at http://www.nlm.nih.gov/mesh). **Appendix B** provides information on the names, synonyms, CASRN, and chemical/product class, where available, for each substance

while **Appendix C** contains 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 2003a, 2003b; FDA 2003). The data quality was evaluated by a review of the methods section in literature references and the submitted reports. Thus, data quality presented in the reviewed literature references can only be evaluated to the extent such information was 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). Detailed information on coding procedures used in different studies is provided in Section 3.4 of the ICCVAM BRD (2006a).

6.0 HET-CAM Test Method Accuracy

6.1 Accuracy of the HET-CAM Test Method

A critical component of an ICCVAM evaluation of the validation status of a test method is an assessment of the accuracy of the proposed test method when compared to 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

The ability of the HET-CAM test method to identify all categories of ocular irritation potential, as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001; UN 2003), was evaluated. This same analysis was also performed with specific chemical classes and/or physical properties excluded based on their previously being identified as discordant in HET-CAM (ICCVAM 2006a).

These evaluations were conducted on the overall data set by combining results from the reports detailed in **Section 5.0,** then assigning an overall ocular irritancy classification for

each substance (**Appendix B** and **C**). 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 was an equal number of differing irritancy classifications for substances (e.g., two tests classified a substance as Not Labeled and two tests classified a substance as a mild irritant), the more severe irritancy classification was used for the overall classification for the substance (mild irritant, in this case).

HET-CAM performance analyses compared to the Draize rabbit eye test were performed for each classification system (i.e., GHS, EPA, EU) each of the six HET-CAM protocols (i.e., IS [A], IS [B], Q-Score, S-Score, IS, and ITC protocols, see **Appendix E**). 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, EU Not Labeled, GHS Not Classified). 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 Classified substances). Therefore, more extensive analyses of HET-CAM 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 on 60 substances that were assigned GHS ocular irritant classifications (UN 2003) (see **Appendix C**). Performance was evaluated for three individual studies (Gettings et al.1994; Gettings et al. 1996; Hagino et al. 1999). Individual analyses were not conducted on the other two studies (Bagley et al. 1992; Kojima et al. 1995) because they contained data on one and two substances, respectively. Based on *in vivo* rabbit eye test data, 45% (27/60) of substances were classified as Category 1, none were classified as Category 2A, and 47% (28/60) were classified as Not Labeled. The remaining 5% (3/60) could not be classified due to lack of adequate animal data and are so noted in **Appendix C**.

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

The HET-CAM test method correctly identified 48% (113/27) of the Category 1 substances (**Table 6-1**). Among the remaining 52% (14/27) of Category 1 substances underpredicted by

HET-CAM, 42% (11/26) were classified as Category 2A and 7.6% (2/26) were classified as Category 2B.

6.1.1.2 Identification of Category 2A Substances (Moderate Ocular Irritants)

The HET-CAM test method did not identify any substances as moderate ocular irritants (i.e., GHS Cat 2A (**Table 6-1**).

Table 6-1 Evaluation of the 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[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		actual	under	over	actual	under	over	actual	under	over	actual
Gettings et al. (1994)	50% (9/18)	100% (1/1)	0% (0/0)	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% (0/0)	0% (0/0)	0% (0/0	0% (0/0	100% (3/3)	0% (0/0)	0% (0/0	100% (4/4)	0% (0/0)
Overall[6]	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 2003)

[2]Severe = Category 1.

[3]Moderate = Category 2A.

[4]Mild = Category 2B.

[5]Not Labeled = Not Labeled.

[6]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.

[This Page Intentionally Left Blank]

6.1.1.3 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**).

6.1.1.4 Identification of Not Labeled Substances

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

6.1.1.5 Ability to Distinguish Substances Not Labeled 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 labeled as irritants from all irritant classes[1]. Using this approach of identifying substances not labeled as irritants from all other classes for the 59 substances considered, the HET-CAM test method has 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**).

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

Table 6-2 Accuracy of the HET-CAM Test Method (IS[A]) for Distinguishing Not Classified Substances from All Other Irritant Classes as Defined by the GHS 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)	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	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31

[1]GHS = Globally Harmonized System (UN2003); NC vs Cat 1/2A/2B.
[2]N= Number of substances included in this analysis/the total number of substances in the study.
[3]No. = Data used to calculate the percentage.
[46]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.

[This Page Intentionally Left Blank]

Gettings et al. (1996): Based on the *in vivo* rabbit data, 24 substances could be assigned a GHS classification. Based on 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).

Hagino et al. (1999): Based upon the *in vivo* rabbit data, 15 substances could be assigned a GHS classification. Based on 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).

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

Because a specific analysis method is the focus of the evaluation of HET-CAM for identifying all hazard categories (the IS[A] analysis method), 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). The results indicate that alcohols tend to be overpredicted by HET-CAM (i.e., 75% [4/6] of alcohols classified as Category 2B or Not Classified based on Draize test results [and depending on the classification system used] 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 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 were underpredicted as Not Classified.

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 Category 1 substances can be improved by excluding certain product types (i.e., surfactant based formulations, see **Table 6-3**).

When the ability of the HET-CAM test method to distinguish Not Classified substances from all 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**).

Table 6-3 Evaluation of the 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 Exclusion of Discordant Chemical and Physical Classes

HET-CAM Database	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		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)
w/o Alcohols	47% (24/51)	46% (11/24)	54% (13/24)	-	-	-	67% (2/3)	33% (1/3)	33% (1/3)	58% (14/24)	42% (10/24)
w/o Surfactant Formulations	49% (17/35)	90% (9/10)	10% (1/10)	-	-	-	100% (3/3)	0% (0/3)	0% (0/3)	64% (14/22)	36% (8/22)
w/o Oil/Water Emulsions	41% (15/41)	48% (12/25)	52% (13/25)	-	-	-	80% (4/5)	20% (1/5)	0% (0/5)	82% (9/11)	18% (2/11)
w/o Alcohols and Surfactant Formulations	56% (15/27)	87% (7/8)	12% (1/8)	-	-	-	100% (1/1)	0% (0/1)	0% (0/1)	56% (10/18)	44% (8/18)
w/o Alcohols and Oil/Water Emulsions	39% (13/33)	44% (10/23)	56% (13/23)	-	-	-	67% (2/3)	33% (1/3)	0% (0/3)	71% (5/7)	29% (2/7)
w/o Alcohols, Surfactant Formulations, and Oil/Water Emulsions	67% (6/9)	86% (6/7)	14% (1/7)	-	-	-	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). [2]Severe = Category 1. [3]Moderate = Category 2A. [4]Mild = Category 2B. [5]Not Labeled = Not Classified.

Table 6-4 Accuracy of the HET-CAM Test Method (IS[A]) for Distinguishing Substances not labeled as irritants from All Other Irritant Classes as Defined by the GHS Classification System[1], with Exclusion of Discordant Chemical and Physical Classes

HET-CAM Database	N[2]	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.[3]	%	No.	%	No.	%	No.	%	No.
Overall	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31
w/o Alcohols	51	73	37/51	100	27/27	0	0/3	58	14/24	0	0/27
w/o Surfactant Formulations	35	60	21/35	100	13/13	36	8/22	64	14/22	0	0/13
w/o Oil/Water Emulsions	41	78	32/41	100	30/30	18	2/11	82	9/11	0	0/30
w/o Alcohols and Surfactant Formulations	27	63	17/27	100	9/9	44	8/18/	56	10/18	0	0/9
w/o Alcohols and Oil/Water Emulsions	33	85	28/33	100	26/26	29	2/7	71	5/7	0	0/26
w/o Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	67	6/9	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

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 Category 1 surfactants, 75% (12/16) were underpredicted (**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 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**).

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 from 63 substances, 60 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the EPA classification system (EPA 1996) (see **Appendix C**). Based on results from *in vivo* rabbit eye experiments, 41% (26/63) were classified as severe irritants (i.e., Category 1), 3% (2/61) were classified as moderate irritants (i.e., Category II), 29% (18/63) were classified as mild irritants (i.e., Category III), and 24% (15/63) were classified as not labeled (i.e., Category IV). The remaining 3% (2/63) of substances could not be classified according to the EPA classification system due to the lack of adequate animal data and are so noted in **Appendix C**.

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

The HET-CAM test method correctly identified 48% (12/25) of the Category 1 substances (**Table 6-6**). Among the remaining 52% (13/25) Category 1 substances that were underpredicted by HET-CAM, 40% (10/25) were classified as Category 2A, and 12% (3/25) were classified as Category 2B.

Table 6-5 Evaluation of the Performance of the HET-CAM Test Method Using the GHS[1] Classification System 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*)					
		1 (Severe)[2]			2A (Moderate)[3]		2B (Mild)[4]	2A (Mod)	2B (Mild)		NL (Not Labeled)		
		NL	2B	2A	NL	2B	NL	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[6]													
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	--	-	-	-	-	-	-	-	-	-	-	-	-
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

[1]GHS classification system (UN 2003) [2]Severe = Category 1 (GHS) [3]Moderate = Category 2A (GHS) [4]Mild = Category 2B (GHS).

Table 6-6 Evaluation of the 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[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		actual	under	over	actual	under	over	actual	under	over	actual
Gettings et al. (1994)	78% (14/18)	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% (0/0)	50% (2/4)`	50% (2/4)
Hagino et al. (1999)	47% (7/15)	100% (7/7)	0% (0/0)	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[6]	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 2003)

[2]Severe = Category I.

[3]Moderate = Category II.

[4]Mild = Category III.

[5]Not Labeled = Category IV.

[4]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 was therefore could not be used in the analysis.

6.1.2.2 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 2A while 50% (1/2) were overpredicted and 0% (0/2) were underpredicted (**Table 6-6**).

6.1.2.3 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 2B while 56% (10/18) were overpredicted and 22% (4/18) were underpredicted (**Table 6-6**).

6.1.2.4 Identification of Category IV 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 6-6**).

6.1.2.5 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]. Using this approach of identifying substances not labeled as irritants from all other classes for the 60 substances considered, the HET-CAM test method has 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] ICCVAM (2006) provides an evaluation of the HET-CAM test method for distinguishing ocular corrosives and severe irritants from all other classes. Since the database of HET-CAM test method results has not changed, this analysis has not been repeated here.

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).

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

Because a specific analysis method is the focus of the evaluation of HET-CAM for identifying all hazard categories (the IS[A] analysis method), 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 excluded certain product types (see **Table 6-8**). The results indicate that alcohols tend to be overpredicted by HET-CAM (i.e., 100% [7/7] of alcohols classified as Category III or IV based on Draize test results [and depending on the classification system used] were overpredicted by HET-CAM by at least one hazard category). Similarly, 47% (8/17) of the oil/water emulsions 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 were underpredicted as Category IV.

When the ability of the HET-CAM test method to distinguish Category IV substances from all 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**).

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 6-10**).

According to the EPA classification system, the most overpredicted substances (false positives) were alcohols, which overpredicted 100% (7/7) of alcohols. Because the entire HET-CAM IS(A) database is made up of liquid substances, the physical form of underpredicted substances was liquids. Three of the surfactants tested in HET-CAM were overpredicted (**Table 6-10**).

Table 6-7 Accuracy of the HET-CAM Test Method for Distinguishing Substances not labeled as irritants from All Other Irritant Classes as Defined by the EPA 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)	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	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

[1] EPA classification system (EPA 1996). Cat IV vs. Cat I/II/III.

[2]N = Number of substances included in this analysis/the total number of substances in the study.

[3]No. = Data used to calculate the percentage.

[4]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. Data on one substance from Kojima et al. (1995) was classified as a GHS Category 1/2A and, therefore, was also not used in the analysis.

Table 6-8 Evaluation of the 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 Exclusion of Discordant Chemical and Physical Classes

HET-CAM Database	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		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)
w/o 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)
w/o 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)
w/o 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)
w/o 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)
w/o 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)
w/o 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 1996) [2]Severe = Category 1 [3]Moderate = Category II [4]Mild = Category III [5]Not Labeled = Category IV.

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

HET-CAM Database	N[2]	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
w/o Alcohols	52	87	45/52	100	39/39	46	6/13	54	7/13	10	4/39
w/o Surfactant Formulations	35	80	28/35	100	24/24	29	4/14	82	9/11	17	4/24
w/o Oil/Water Emulsions	41	78	32/41	100	30/30	18	2/11	82	9/11	0	0/30
w/o Alcohols and Surfactant Formulations	27	81	22/27	100	18/18	44	4/9	56	5/9	44	4/18
w/o Alcohols and Oil/Water Emulsions	34	94	32/34	100	30/30	50	2/4	50	2/4	0	0/30
w/o Alcohols, Surfactant Formulations, and Oil/Water Emulsions	9	78	7/9	100	9/9	-	-	-	-	-	-

Abbreviations: EPA = U.S. Environmental Protection Agency; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane

[1] EPA classification system (EPA 1996). Cat IV vs. Cat I/II/III.

[2] N = Number of substances included in this analysis/the total number of substances in the study.

[3] No. = Data used to calculate the percentage.

Table 6-10 Evaluation of Under and Overprediction of the HET-CAM Test Method Using the EPA[1] Classification System 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*)					
		I (Severe)[2]			II (Moderate)[3]		III (Mild)[4]	II (Mod)	III (Mild)		IV (Not labeled)		
		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)	22% (4/18)	50% (1/2)	28% (5/18)	28% (5/18)	40% (6/15)	0% (0/18)	20% 3/18)
Chemical Class[6]													
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	--	-	-	-	-	-	-	-	-	-	-	-	-
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 classification system (EPA 1996); HET-CAM = Hen's Egg Test - Chorioallantoic Membrane

[1] EPA classification system (EPA 1996)

[2]Severe = Category I.

[3]Moderate = Category II.

[4]Mild = Category III.

[5]Non-rritant = Category IV.

[6]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 MeSH categories (www.nlm.nih.gov/mesh) as defined in Appendix A.

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, 59 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the EU classification system (EU 2001) (see **Appendix C**). Based on results from *in vivo* rabbit eye experiments, 38% (24/63) were classified as Category R41 (i.e., severe irritants), 3% (2/63) were classified as Category R36 (i.e., moderate irritants), 51% (32/63) were classified as Not Labeled. The remaining 6% (4/63) of substances could not be classified according to the EU classification system due to the lack of adequate animal data and are so noted in **Appendix C**.

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

The HET-CAM test method correctly identified 50% (12/24) of the Category R41 substances (**Table 6-11**). Among the remaining 50% (12/24) R41 substances that were underpredicted by HET-CAM, 42% (10/24) were classified as Category R36, and 8% (2/24) were classified as not labeled.

6.1.3.2 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 overpredicted and 0% (0/2) were underpredicted (**Table 6-11**).

6.1.3.3 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 6-11**).

6.1.3.4 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 irritant classes[3]. Using this approach of identifying substances not labeled as irritants from all other classes for the 60 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 6-12**).

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

Gettings et al. (1994): Based upon the *in vivo* rabbit data, 14 substances were assigned an EU classification. The HET-CAM test method, by comparison, has an accuracy of 64% (9/14), sensitivity of 100% (1/1), specificity of 62% (8/14), false positive rate of 33% (1/3), and a false negative rate of 0% (0/1) (**Table 6-12**).

Gettings et al. (1996): Based upon the *in vivo* rabbit data, 17 substances were assigned a EU classification. The HET-CAM test method, by comparison, has an accuracy of 82% (14/17), sensitivity of 100% (14/14), specificity of 67% (2/3), false positive rate of 50% (2/4), and a false negative rate of 0% (0/8) (**Table 6-12**).

Hagino et al. (1999): Based upon the *in vivo* rabbit data, 14 substances were assigned a EU classification. The HET-CAM test method, by comparison, has an accuracy of 50% (8/14), sensitivity of 100% (8/8), specificity of 0% (0/6), false positive rate of 100% (6/6), and a false negative rate of 0% (0/26) (**Table 6-12**).

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

Because a specific analysis method is the focus of the evaluation of HET-CAM for identifying all hazard categories (the IS[A] analysis method), 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

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

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 6-13**). The results indicate that alcohols tend to be overpredicted by HET-CAM (i.e., 83% [5/6] of alcohols classified as Not Labeled based on Draize test results [and depending on the classification system used] 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 Category R36). However, none of these substances were underpredicted as Not Labeled.

When the ability of the HET-CAM test method to distinguish Not Labeled substances from all 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 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) that were excluded (**Table 6-14**).

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 6-10**).

According to the EU classification system, the most overpredicted substances (false positives) were alcohols, which overpredicted 83% (5/6) alcohols. 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 6-10**).

Table 6-11 Evaluation of the 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[2]		Moderate[3]			Mild			Not Labeled[4]	
		actual	under	over	actual	under	over	actual	under	over	actual
Gettings et al. (1994)	64% (9/14)	100% (1/1)	0% (0/1)	0% (0/0)	0% (0/0)	0% (0/0)	NA	NA	NA	38% (5/13)	62% (8/13)
Gettings et al. (1996)	35% (6/17)	29% (4/14)	71% (10/14)	0% (0/0)	0% (0/0)	0% (0/0)	NA	NA	NA	33% (1/3)	67% (2/3)
Hagino et al. (1999)	50% (7/14)	100% (7/7)	0% (0/0)	100% (1/1)	0% (0/1)	0% (0/1)	NA	NA	NA	100% (6/6)	0% (0/6)
Overall[5]	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]Severe = R41.
[3]Moderate = Category 2A (GHS), Category II (EPA), and R36 (EU).
[4]Not Labeled = Not Labeled.
[5]Overall data set includes one additional test substance from Bagley et al. (1992) and two from Kojima et al. (1995) that were not included as individual data sources.

Table 6-12 Accuracy of the HET-CAM Test Method (IS[A]) for Distinguishing Substances not labeled as irritants from All Other Irritant Classes as Defined by the EU 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)	14	64	9/14	100	1/1	62	8/13	38	5/13	0	0/1
Gettings et al. (1996)	17	82	14/17	100	14/14	67	2/3	33	1/3	0	0/14
Hagino et al. (1999)	14	50	8/14	100	8/8	0	0/6	100	6/6	0	0/8
Overall[4]	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

[1]EU classification system (EU 2001). NL vs. R41/R36.

[2]N = Number of substances included in this analysis/the total number of substances in the study.

[3]No. = Data used to calculate the percentage.

[4]Overall data set includes one additional test substance from Bagley et al. (1992) and two from Kojima et al. (1995) that were not included as individual data sources.

Table 6-13 Evaluation of the 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 Exclusion of Discordant Chemical and Physical Classes

HET-CAM Database	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		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)
w/o Alcohols	42% (21/50)	45% (10/22)	55% (12/22)	50% (1/2)	0% (0/2)	50% (1/2)	NA	NA	NA	62% (16/26)	38% (10/26)
w/o 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)
w/o Oil/Water Emulsions	36% (14/39)	48% (11/23)	52 (12/23)	0% (0/1)	100% (1/1)	0% (0/1)	NA	NA	NA	87% (13/15)	0% (0/24)
w/o 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)
w/o 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)
w/o Alcohols, Surfactant Formulations, and Oil/Water Emulsions	62% (5/8)	100% (5/5)	0% (0/5)	100% (0/1)	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) [2]Severe = R41 [3]Moderate = R36 [4]Mild = NA [5]Not Labeled = Not Classified.

Table 6-14 Accuracy of the HET-CAM Test Method (IS[A]) for Distinguishing Substances not labeled as irritants from All Other Irritant Classes as Defined by the EU Classification System[1], with Exclusion of Discordant Chemical and Physical Classes

HET-CAM Database	N[2]	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.[3]	%	No.	%	No.	%	No.	%	No.
Overall	58	62	36/58	100	26/26	31	10/32	69	22/32	0	0/26
w/o Alcohols	50	42	21/50	100	24/24	38	10/26	62	16/26	0	0/24
w/o Surfactant Formulations	34	64	16/25	100	9/9	32	8/25	68	17/25	0	0/9
w/o Oil/Water Emulsions	39	67	26/39	100	2/24	13	2/15	87	13/15	0	0/24
w/o Alcohols and Surfactant Formulations	26	65	17/26	100	7/7	42	8/19	58	11/19	0	0/7
w/o Alcohols and Oil/Water Emulsions	32	78	25/32	100	23/23	22	2/9	78	7/9	0	0/23
w/o Alcohols, Surfactant Formulations, and Oil/Water Emulsions	8	62	5/8	100	6/6	0	0/2	100	2/2	0	0/6

Abbreviations: EU = European Union; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane

[1] EU classification system (EU 2001). NV vs. R41/R36.

[2] N = Number of substances included in this analysis/the total number of substances in the study.

[3] No. = Data used to calculate the percentage.

Table 6-15 Evaluation of the Performance of the HET-CAM Test Method Using the EU[1] Classification System 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*)		
		R41 (Severe)[2]		R36 (Moderate)[3]	R36 (Mod)[3]	NL (Not Labeled)[4]	
		NL	R36	NL	R41	R36	R41
Overall	61	7% (2/28)	43% (12/28)	50% (1/2)	0% (0/2)	16% (5/31)	23% (7/31)
Chemical Class[5]							
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	-	-	-	-	-	-	-
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: GHS = Globally Harmonized System; HET-CAM = Hen's Egg Test – Chorioallantoic Membrane

[1]EU classification system (EU 2001)

[2]Severe = Category R41 (EU).

[3]Moderate = Category R36 (EU).

[4]NL = Category NL (EU).

7.0 HET-CAM Test Method Reliability

An assessment of test method reliability (intralaboratory repeatability and intra- and inter-laboratory reproducibility) is an essential element of 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). Since 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 of HET-CAM hazard classifications among the laboratories.

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 2003). Of four not labeled and three Category 2B substances, none (0%; 0/4 and 0%; 0/3, respectively) were correctly identified by HET-CAM. None of the 15 GHS-classified substances tested were classified Category 2A by HET-CAM. However, eight substances classified as GHS Category 1 were correctly identified by HET-CAM (100%; 8/8).

The extent of agreement of calls among laboratories between irritants (i.e., Category 1, 2A, and 2B = "+" and Not Labeled = "-") regardless of the individual hazard classification was evaluated by comparison of *in vivo* and *in vitro* data (**Table 7-1**).

- For 11 substances, there was 100% agreement among the *in vivo* and *in vitro* calls (i.e., "+/+").
- For four substances that were overpredicted *in vitro* (i.e., "+/-"), there was 100% agreement for 3/4 (75%) of the substances and 80% agreement for 1/4 (25%) of the substances.
- For two substances that could not be assigned GHS classification, there was 100% agreement on the *in vitro* classifications (i.e., "?/+).
- An assessment of the agreement between laboratories for substances not labeled as irritants compared to all other classes could not be made because

there were no not labeled calls obtained using HET-CAM. However, overall, there was 100% agreement for 16/17 (94%) of substances and 80% agreement for 1/17 (6%) of substances [4].

• [4] Because the database of HET-CAM test method results has not changed, the qualitative evaluation of reproducibility presented in ICCVAM (2006) is not repeated here.

Table 7-1 Interlaboratory Variability of Hagino et al. (1999) In Predicting Not Labeled Ocular Substances or Corrosive/Severe/Moderate/Mild Irritants as Defined by the GHS Classification System

Report	Anal[1]	Classification (*In Vivo*/*In Vitro*)[2]	# of Labs	N[3]	Substances with 100% Agreement among Labs	Substances with 80% Agreement among Labs
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%)

Abbreviation: GHS = Globally Harmonized System (UN 2007)

[1]Anal = 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).

[2]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*.

[3]N indicates number of substances.

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

[This Page Intentionally Left Blank]

The extent of agreement among the five laboratories for a test substance was also evaluated based on prediction of the individual GHS hazard category (**Table 7-2**).

- Of four not labeled substances, all were overpredicted with 100% agreement by 3/4 (75%) laboratories and 80% agreement by 1/4 (25%) laboratories.
- Of the three Category III substances, all 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 substances with 100% agreement for 5/8 (63%) substances, 80% agreement for 1/8 (13%) substances, and 60% agreement for 2/8 (25%) substances.

None (0/8 [0%]), of the Category 1 substances were incorrectly identified. However, all four not labeled 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 to incorrectly classify any 0/8 (0%) of the GHS Category 1 substances, since all were correctly classified. There was 100% agreement to overclassify 3/3 (100%) of the GHS Category 2B substances, 100% agreement to overclassify 3/4 substances and 80% agreement to overclassify 1/4 of the not labeled 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 2003). Of two not labeled, five Category III, and one Category II substances, none (0%, 0/2, 0%, 0/5, and 0%, 0/1, respectively) were correctly identified by HET-CAM. However, seven substances classified as EPA Category I were correctly identified by HET-CAM (100%; 7/7).

Table 7-2 Evaluation of the Interlaboratory Variability of Hagino et al. (1999) in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method as Defined by the GHS Classification System

***In vivo* Classification (No.)[1]**	**Classification *(in vitro)***	**Number of Substances**	**Number of Testing Laboratories**	**Substances with 100% Agreement Among Laboratories (%)**	**Substances with 80% Agreement Among Laboratories (%)**	**Substances with 60% Agreement Among Laboratories (%)**
NL (4)	Actual	0	5	0	0	0
	Over	4	5	3 (75%)	1 (25%)	0
2B (3)	Under	0	5	0	0	0
	Actual	0	5	0	0	0
	Over	3	5	3 (100%)	0	0
2A (0)	Under	0	5	0	0	0
	Actual	0	5	0	0	0
	Over	0	5	0	0	0
1 (8)	Under	0	5	0	0	0
	Actual	8	5	5 (63%)	1 (13%)	2 (25%)

Abbreviations: GHS = United Nations Globally Harmonized System of Classification and Labelling of Chemicals

[1]Due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects), a GHS classification (UN 2007) 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*.

[This Page Intentionally Left Blank]

The extent of agreement of calls among laboratories between irritants (i.e., Category I, II, and III = "+" and Category IV = "-") regardless of the individual hazard classification was evaluated by comparison of *in vivo* and *in vitro* data (**Table 7-3**).

- For 13 substances, there was 100% agreement among the *in vivo* and *in vitro* calls (i.e., "+/+").
- For two substances that were overpredicted *in vitro* (i.e., "+/-"), there was 60% agreement for 2/2 (100%) of the substances.
- For two substances that could not be assigned an EPA classification, there was 100% agreement on the *in vitro* classifications (i.e., "?/+).
- An assessment of the agreement between laboratories for substances not labeled as irritants compared to all other classes could not be made because there were no not labeled calls obtained using HET-CAM. However, overall, there was 100% agreement for 14/17 (82%) of substances and 60% agreement for 3/17 (18%) of substances[5].

The extent of agreement among the five laboratories for a test substance was also evaluated based on prediction of the individual EPA hazard category (**Table 7-4**).

- Of two not labeled substances, all were overpredicted with 100% agreement by 1/2 (50%) of the laboratories and 80% agreement by 1/2 (50%) of the laboratories.
- Of the three Category III substances, all were overpredicted with 100% agreement among the five laboratories.
- One Category II substances was overpredicted with 100% agreement among the five laboratories.
- All seven substances were correctly predicted as Category I substances with 100% agreement for 5/7 (71%) of the substances and 80% agreement for 2/7 (29%) of the substances.

---

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

Table 7-3 Interlaboratory Variability of Hagino et al. (1999) In Predicting Not Labeled Ocular Substances or Corrosive/Severe/Moderate/Mild Irritants as Defined by the EPA Classification System

Report	Anal[1]	Classification (*In Vivo*/*In Vitro*)[2]	# of Labs	N[3]	Substances with 100% Agreement among Labs	Substances with 60% Agreement among Labs
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%)

Abbreviation: EPA = U.S. Environmental Protection Agency (EPA 2003)

[1]Anal = 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).

[2]A "+" indicates that the substance was assigned an overall classification of corrosive or a severe irritant (Category I); a "-" indicates that the substance was assigned an overall classification of nonsevere irritant (Category II or II) 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 EPA 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*.

[3]N indicates number of substances.

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

Table 7-4 Evaluation of the Interlaboratory Variability of Hagino et al. (1999) In Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method as Defined by the EPA Classification System

***In vivo* Classification (No.)[1]**	**Classification *(in vitro)***	**Number of Substances**	**Number of Testing Laboratories**	**Substances with 100% Agreement Among Laboratories (%)**	**Substances with 80% Agreement Among Laboratories (%)**
IV (2)	Actual	0	5	0	0
	Over	2	5	1 (50%)	1 (50%)
III (5)	Under	0	5	0	0
	Actual	0	5	0	0
	Over	5	5	5 (100%)	0
II (1)	Under	0	5	0	0
	Actual	0	5	0	0
	Over	1	5	1 (100%)	0
1 (7)	Under	0	5	0	0
	Actual	7	5	5 (71%)	2 (29%)

Abbreviation: EPA = U.S. Environmental Protection Agency (EPA 2003)

[1]Due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects), an EPA classification (EPA 2003) 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*.

None (0/7 [0%]), of the Category 1 substances were incorrectly identified. However, all two not labeled, five Category III, and one Category II substances (i.e., 2/2 [100%], 5/5 [100%], and 1/1 [100%], respectively, were incorrectly identified by HET-CAM (**Table 7-4**).

- There was no agreement among the five participating laboratories to incorrectly classify any 0/7 (0%) of the EPA Category I substances, since all were correctly classified. There was 100% agreement to overclassify 1/2 (50%) and 80% agreement to overclassify 1/2 (50%) of the EPA substances not labeled as irritants. For Category III substances, there was 100% agreement to overclassify 5/5 substances. For the Category II substance, there was 100% agreement to overclassify it.

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, one Category R36, and one Category R36 substances, none (0%, 0/7, 0%, 0/1, and 0%, 0/2, respectively) were correctly identified by HET-CAM. However, seven substances classified as EU Category I were correctly identified by HET-CAM (100%; 7/7).

The extent of agreement of calls among laboratories between irritants (i.e., Category R41, R36 = "+" and not labeled = "-") regardless of the individual hazard classification was evaluated by comparison of *in vivo* and *in vitro* data (**Table 7-5**).

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

- An assessment of the agreement between laboratories for substances not labeled as irritants compared to all other classes could not be made because there were no not labeled calls obtained using HET-CAM.

Table 7-5 Interlaboratory Variability of Hagino et al. (1999) In Predicting Not Labeled Ocular Substances or Corrosive/Severe/Moderate/Mild Irritants as Defined by the EU Classification System

Report	Anal[1]	Classification (*In Vivo*/*In Vitro*)[2]	# of Labs	N[3]	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)

Abbreviation: EU = European Union (EU 2001).

[1]Anal = 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).

[2]A "+" indicates that the substance was assigned an overall classification of corrosive or a severe irritant (Category R41); a "-" indicates that the substance was assigned an overall classification of nonsevere irritant (Category 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), a 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*.

[3]N indicates number of substances.

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

Table 7-6 Evaluation of the Interlaboratory Variability of Hagino et al. (1999) In Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method as Defined by the EU Classification System

***In vivo* Classification (No.)**[1]	**Classification *(in vitro)***	**Number of Substances**	**Number of Testing Laboratories**	**Substances with 100% Agreement Among Laboratories (%)**	**Substances with 80% Agreement Among Laboratories (%)**	**Substances with 60% Agreement Among Laboratories (%)**
NL (7)	Actual	0	5	0	0	0
	Over	7	5	6 (86%)	1 (14%)	0
R36 (1)	Under	0	5	0	0	0
	Actual	0	5	0	0	0
	Over	1	5	1 (100%)	0	0
R41 (7)	Under	0	5	0	0	0
	Actual	7	5[2]	5 (71%)	1 (14%)	1 (14%)

Abbreviation: EU = European Union (EU 2001).

[1]Due to the lack of appropriate *in vivo* data (e.g., studies were terminated too early to assess reversibility of effects), a EU classification (EU 2001) 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*.

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

- Of seven not labeled substances, all were overpredicted with 100% agreement by 6/7 (86%) of the laboratories and 80% agreement by 1/7 (14%) of the laboratories.
- The one R36 substance was overpredicted with 100% agreement among the five laboratories.
- Seven Category R41 substances were overpredicted with 100% agreement among the five laboratories for 5/7 (71%), 80% agreement for 1/7 (14%), and 60% agreement for 1/7 (14%) of the substances.

None, 0/7 (0%), of the Category R41 substances were incorrectly identified. However, all seven not labeled, one Category R36, and seven Category R41 substances (i.e., 7/7 (100%), 1/1 (100%), and 7/7 (100%), respectively, were incorrectly identified by HET-CAM (**Table 7-6**).

- There was no agreement among the five participating laboratories to incorrectly classify any 0/7 (0%) of the EU Category R41 substances, since all were correctly classified. There was 100% agreement to overclassify 6/7 (86%) and 80% agreement to overclassify 1/7 (14%) of the EPA substances not labeled as irritants. For Category R36 substances, there was 100% agreement to overclassify 1/1 substances.

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.

8.0 Test Method Data Quality

The database used in this assessment did not change from that used in the previous assessment of the ability of the HET-CAM test method to identify ocular corrosives and severe irritants. The evaluation of HET-CAM test method data quality is detailed in ICCVAM (2006a).

9.0 Other Scientific Reports and Reviews

Four additional studies were identified in the peer-reviewed literature over the period 2005 to 2009 that contained HET-CAM data (Debbasch et al. 2005; Vinardell and Mitjans, 2006; Mehling et al. 2007; Mancebo et al. 2008). 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 previously included in the accuracy analyses reported in the ICCVAM BRD (2006a). As such, they have in turn already been considered in the current evaluation.

In Debbasch et al. (2005), 12 coded make-up removers were tested in the HET-CAM, BCOP, and the corneal epithelial cell line (CEPI) test methods, and a clinical in-use test under ophthalmological control after their application to the external eyelid. Three hundred microliter of undiluted test product was applied to the chorioallantoic membrane of nine-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 OD490 nm. *In vitro* scores were classified according to Gautheron et al. (1994) and Harbell and Curren (1998), but 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. Using fertile eggs (Leghorn SA31, six per solvent), the substances to be tested were applied on the membrane in a constant volume of 0.3 ml at 37^{o}C. Following application of the test substances, 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 previously included in the accuracy analyses reported in the ICCVAM BRD (2006a). As such, they have in turn already been considered in the current evaluation.

In Mehling et al. (2007), 18 proprietary surfactants were tested using the red blood cell test (RBC), HET-CAM and the SkinEthic™ ocular tissue model. Following the standard operating procedure (SOP) of the COLIPA project (INVITTOX protocol No. 96), three

hundred microliter 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 semi-quantitatively 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 using the acute dermal toxicity and irritation/corrosion tests, 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 a period of five minutes. The three endpoints for this study were hemorrhage, vessel lyses, and coagulation. Although *in vivo* rabbit eye data were reported in the study, the available raw data was not and as such the study was not included in the HET-CAM performance analyses detailed in this BRD.

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

10.1 How the HET-CAM Test Method Will Reduce, Refine, and 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 nonanimal 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 and use of this 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.

10.2 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 have provisions for the protection of animals used for experimental or other scientific purposes. Some provisions indicate the time in which a test method using an animal embryo or fetus is considered an animal, and therefore protected by the regulations. According to some of these regulations, a bird is considered a protected animal (and therefore the test is considered an *in vivo* and not *in vitro* test) when greater 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 nine, 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 nine in the avian embryo and that there is significant development of the sensory nerve ending 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 nine there is little to no pain perceived by the developing embryo during the conduct of the HET-CAM test method.

11.0 References

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

Bagley D, Bruner L, DeSilva O, Cottin M, O'Brien K, Uttley M, Walker AP. 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.

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

CPSC. 1988. Commercial practices. Hazardous Substances and Articles; Administration and Enforcement Regulations. 16 CFR 1500.42.

CPSC. 1995. Test for eye irritants. 16CFR1500.42.

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

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 and 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. Acute Toxicity – Eye Irritation. In Directive 67/548 (6^{th} adaption); Annex V, Part B: Methods for the Determination of Toxicity. Official Journal of the European Communities 27 L251, 109.

EEC. 1991. Classification of irritant substances and preparations. Official Journal of the European Communities. L180:52.

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

EPA. 1998. Health Effects Test Guideline, OPPTS 870.2400 Acute Eye Irritation. EPA 712-C-98-195. Office of Prevention, Pesticides & Toxic Substances (OPPTS). Washington, DC: U.S. Environmental Protection Agency.

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

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

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

EU. 1986. Council Directive 86/69/EEC. Official Journal of the European Economic Community. 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 eighteenth 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:1-86.

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

FDA. 2003. Good laboratory practice for nonclinical laboratory studies. 21 CFR 58.

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, 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.

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.

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.

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: 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: National Toxicology Program.

ICCVAM Authorization Act of 2000. 2000. Public Law 106-545. [114 Stat. 2721]. Available: http://iccvam.niehs.nih.gov/about/overview.htm [accessed 21 October 2004].

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

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.

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.

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.

2042 Russell and Burch. 1992. The Principles of Humane Experimental Technique. 14th ed. South
2043 Mimms, Potters Bar, Herts, England: Universities Federation for Animal Welfare.

2044 Spielmann H, Liebsch M, Kalweit S, Moldenhauer F, Wirnsberger T, Holzhütter H, et al.
2045 1996. Results of a validation study in Germany on two in vitro alternatives to the Draize eye
2046 irritation test, HET-CAM test and the 3T3 NRU cytotoxicity test. ATLA 24:741-858.

2047 UN. 2002. OECD Guideline for the Testing of Chemicals. Acute Eye Irritation/Corrosion.
2048 Adopted: 24th April 2002. New York & Geneve:United Nations Publications.

2049 UN. 2003. Globally Harmonised System of Classification and Labelling of Chemicals
2050 (GHS). New York & Geneva:United Nations Publications.

2051 Vinardell M, Macián M. 1994. Comparative study of the HET-CAM test and the Draize eye
2052 test for assessment of irritancy potential. Toxic In Vitro 8:467-470.

2053 Vinardell MP, Mitjans M. 2006. The chorioallantoic membrane test as a model to predict the
2054 potential human eye irritation induced by commonly used laboratory solvents. Toxicol In
2055 Vitro 20:1066-1070.

12.0 Glossary[6]

Accuracy[7]: (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[2]: The experimental system used. Often used interchangeably with "test" and "test method."

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

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.

Blepharitis: Inflammation of the eyelids.

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

[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.

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[2]: 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.

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

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[2]: The biological process, response, or effect assessed by a test method.

False negative[2]: A substance incorrectly identified as negative by a test method.

False negative rate[2]: 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[2]: A substance incorrectly identified as positive by a test method.

False positive rate[2]: 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 Harmonized System (GHS): A classification system presented by the United Nations that provides (a) 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)[2]: 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[2]: 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[2]: 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[2]: 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[2]: 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.

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[2]: 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.

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[2]: 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[2]: 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[2]: The proportion of positives in the population of substances tested (see "two-by-two" table).

Protocol[2]: 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[2]: 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[2]: A new or modified test method that reduces the number of animals required.

Reference test method[2]: 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[2]: A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being.

Relevance[2]: 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[2]: 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 inter-laboratory reproducibility and intralaboratory repeatability.

Replacement alternative[2]: 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[2]: 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 inter-laboratory reproducibility).

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

Sensitivity[2]: 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[2]: 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 non-transparent test substances.

Test[2]: The experimental system used; used interchangeably with "test method" and "assay."

Test method[2]: 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 component: 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[2]: The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Two-by-two table[2]: 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 ([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[2]: 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[2]: 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.