1.0 Introduction

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 evaluation of several activities related to reducing, refining, and replacing the use of rabbits in the current *in vivo* eye irritation test method (announced in *Federal Register* [FR] notice 69 FR 13859, March 24, 2004). In response to this nomination, ICCVAM evaluated the validation status of the bovine corneal opacity and permeability (BCOP), hen's egg test–chorioallantoic membrane (HET-CAM), isolated chicken eye (ICE), and isolated rabbit eye (IRE) test methods. ICCVAM evaluated the test methods' ability to identify ocular corrosives and severe (irreversible) irritants using the EPA, United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS), and European Union (EU) classification systems.

ICCVAM considered the BCOP and ICE test methods to have sufficient performance to substantiate their use for regulatory hazard classification for some types of substances. The IRE and HET-CAM test methods lacked sufficient performance and/or sufficient data to substantiate their use for regulatory hazard classification. ICCVAM subsequently recommended that the BCOP and ICE test methods should be used in a tiered-testing strategy as part of a weight-of-evidence approach, where positive substances can be classified as ocular corrosives or severe irritants without the need for 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 the Organisation for Economic Co-operation and Development (OECD) adopted the BCOP and ICE test methods as OECD Test Guidelines 437 and 438, respectively (OECD 2009a, 2009b). When used in this manner, the BCOP and ICE test methods should reduce the number of animals needed for ocular safety testing and refine animal use by avoiding the pain and distress associated with testing severely irritating and corrosive substances.

Among these final recommendations was a charge to further evaluate the usefulness and limitations of the BCOP, HET-CAM, ICE, and IRE test methods for the identification of nonsevere ocular irritants (i.e., substances that induce reversible ocular damage) and substances not labeled as irritants. In addition, the Cytosensor® Microphysiometer (CM) test method was evaluated as proposed by the European Centre for the Validation of Alternative Methods (ECVAM) as a possible alternative test method for the identification of ocular corrosives and severe irritants and substances not labeled as irritants. For these current evaluations, ICCVAM used the EPA, EU, Federal Hazardous Substances Act (FHSA), and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007). The FHSA classification system, which is based on the testing guidelines and associated criteria included in 16 CFR 1500.42 (CPSC 2003), was not used in the original analyses (i.e., ability of the test methods to identify ocular corrosives and severe irritants) because the FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test (Draize et al. 1944) does not distinguish between ocular corrosives and severe irritants and less severe irritants. For this reason, an evaluation to identify ocular corrosives and severe irritants using the FHSA classification system is not possible.

The ICCVAM Ocular Toxicity Working Group (OTWG) was charged with working with the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) in reviewing these *in vitro* alternatives. Drs. João Barroso, Thomas Cole, and Valerie Zuang were ECVAM liaisons, and Dr. Hajime Kojima was the Japanese Center for the Validation of Alternative Methods (JaCVAM) liaison to the OTWG.

To facilitate peer review, the OTWG and NICEATM, which administers ICCVAM and provides scientific and operational support for ICCVAM activities, prepared comprehensive draft background review documents (BRDs) that provided information and data from validation studies and the scientific literature for the BCOP, HET-CAM, ICE, and IRE test methods. A redacted BRD (i.e., an abbreviated version that does not include confidential business information) for the CM test method was prepared by ECVAM and submitted to NICEATM–ICCVAM for review.

A June 7, 2007, *Federal Register* notice (72 FR 31582, available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_10966.pdf) requested data and information on these test methods. In addition, an April 4, 2008 *Federal Register* notice (73 FR 18535, available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR-E8-6969.pdf) requested nominations of individuals to serve on an independent international scientific peer review panel (Panel). These requests were also disseminated via the ICCVAM electronic mailing list and through direct requests to over 100 stakeholders. In response to these requests, 12 individuals or organizations nominated members to the Panel; however, no test method data were submitted (see **Section 7.0**).

The BRDs form the basis for the ICCVAM test method recommendations described herein. The ECVAM and JaCVAM liaisons to the OTWG provided input and contributed throughout the evaluation process. A detailed timeline of the evaluation is provided in **Appendix A**. The ICCVAM-recommended test method protocol and the BRD for each test method are provided in **Appendices B through F**.

On March 31, 2009, ICCVAM announced the availability of the ICCVAM draft documents and a public Panel meeting to review the validation status of the test methods (74 FR 14556[1]). The ICCVAM draft BRDs and draft test method recommendations were posted on the NICEATM–ICCVAM website. All of the information provided to the Panel and all public comments received before the Panel meeting were made available on the NICEATM–ICCVAM website.[2]

The Panel met in public session on May 19–21, 2009, to review the ICCVAM draft BRDs for completeness and accuracy. The Panel then evaluated (1) the extent to which the draft BRDs addressed established validation and acceptance criteria and (2) the extent to which the BRDs supported ICCVAM's draft test method recommendations. Interested stakeholders from the public were provided opportunities to comment at the Panel meeting. The Panel considered these comments as well as those submitted prior to the meeting before concluding their deliberations. On July 12, 2009, ICCVAM posted the final report of the Panel's recommendations[3] (**Appendix G**) on the NICEATM–ICCVAM website for public review and comment (announced in 74 FR 33444).

ICCVAM provided SACATM with the draft BRDs, the draft Panel report, and all public comments for discussion at their meeting on June 25–26, 2009, where public stakeholders were given another opportunity to comment.

After SACATM's meeting, ICCVAM and the OTWG considered the SACATM comments, the Panel report, and all public comments (**Appendix H**) before finalizing the ICCVAM test method evaluation report and the BRDs provided in this report. As required by the ICCVAM Authorization Act, ICCVAM will make this test method evaluation report and the accompanying final BRDs available to the public and to U.S. Federal agencies for consideration. The relevant U.S. Federal laws, regulations, guidelines, and recommendations for eye irritation/corrosion testing are summarized in **Appendix I**. Federal agencies must respond to ICCVAM within 180 days after receiving ICCVAM test method recommendations. Agency responses will be made available to the public on the NICEATM–ICCVAM website as they are received.

---

[1] http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/E9-7220.pdf
[2] http://iccvam.niehs.nih.gov/methods/ocutox/PeerPanel09.htm
[3] http://iccvam.niehs.nih.gov/docs/ocutox_docs/OcularPRPRept2009.pdf

2.0 The Bovine Corneal Opacity and Permeability Test Method

The BCOP test method is an *in vitro* eye irritation test method using isolated bovine corneas that are byproducts from processing plants. In the BCOP test method, changes in corneal opacity caused by chemical damage are determined by measuring decreases in light transmission through the cornea. Changes in permeability of the cornea resulting from chemical damage are determined by measuring increases in the quantity of sodium fluorescein dye that passes through all corneal cell layers. Both measurements are used to calculate an *in vitro* irritancy score (IVIS), which is used to predict the *in vivo* ocular irritation/corrosion potential of a test substance.

ICCVAM previously evaluated the validation status of the BCOP test method as an *in vitro* alternative to the Draize rabbit eye test (Draize et al. 1944) to identify ocular corrosives and severe irritants (i.e., those that induce irreversible ocular damage; EPA Category I, EU R41, GHS Category 1). ICCVAM determined that the reproducibility and accuracy were sufficient to support its use for this purpose for some types of substances (ICCVAM 2006e). U.S. agencies and international organizations (OECD 2009a) have adopted the BCOP test method for this purpose. In the current evaluation, ICCVAM evaluated the validation status of the BCOP test method as an *in vitro* alternative to the Draize rabbit eye test for identifying nonsevere ocular irritants (i.e., those that induce reversible ocular damage [EPA Category II and III, EU R36, GHS Category 2A and 2B]) and substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) according to the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007).

2.1 ICCVAM Recommendations

2.1.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

Evaluation as a Screening Test to Identify Substances Not Labeled as Irritants
ICCVAM concludes that the accuracy and reliability of the BCOP test method do **not** support its use as a screening test to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other hazard categories (i.e., EPA Category I, II, and III; EU R41 or R36; FHSA Irritant; GHS Category 1, 2A, or 2B) when results are to be used specifically for hazard classification and labeling purposes under the EPA, EU, FHSA, or GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007). For the validation database of 211 substances, false positive rates were high, ranging from 53% (24/45) to 70% (63/90) depending on the hazard classification system used. Accordingly, all positive results from these tests would require additional testing in a valid test system that can accurately characterize whether such substances require hazard labeling.

False negative rates were 0% for the EU (0/54) and GHS (0/97) classification systems, 5% (6/132) for the FHSA classification system, and 6% (8/142) for the EPA classification system. Among the EPA false negatives were three substances (3/8 [38%]) classified as EPA eye irritants based on at least one rabbit with corneal injuries and opacity that did not resolve until day 3 of the study. A fourth substance was classified as an EPA eye irritant based on all six rabbits with a conjunctival redness score of 2 (n = 4: *diffuse, crimson color of the conjunctiva, individual blood vessels not easily discernable*) or 3 (n = 2: *diffuse beefy red*). The conjunctival redness scores for two of these animals did not recover to a score of 1 (*some blood vessels definitely hyperemic*) until day 6 of the study. The conjunctival redness scores for the remaining four rabbits recovered to a score of 1 on day 2 of the study. These four EPA false negative substances were also false negatives for the FHSA classification system. Given the significant lesions associated with these false negative substances, the BCOP test method cannot be recommended as a screening test to identify substances not labeled as irritants (i.e., EPA Category IV, FHSA Not Labeled) for the EPA or FHSA classification systems (EPA 2003; FHSA 2005).

Furthermore, although the false negative rate is 0% (0/97) for the GHS classification system (UN 2007), the GHS does not classify as eye hazards substances that produce the corneal and conjunctival injuries described above. Such substances must be labeled as eye hazards according to the EPA and FHSA classification systems. These findings led NICEATM–ICCVAM to look more closely at the GHS eye hazard classification criteria. NICEATM evaluated results from rabbit eye test studies from two independent databases: (1) 149 studies obtained from a publicly available database (ECETOC 1998) and (2) 144 studies included in the Detailed Review Document (DRD) on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (OECD 1999). These data (**Appendix J**) confirmed that approximately 30% of the substances that require labeling as eye irritation hazards according to current U.S. hazard classification and labeling requirements (EPA and FHSA) are not labeled as eye irritation hazards by the GHS system. This includes at least 70% of currently labeled EPA Category III irritants (those causing eye injuries persisting for 24 hours to 7 days). The nature, severity, and duration of these eye injuries suggest the potential to cause human injury. The purpose of ocular toxicity labeling is to communicate potential hazards of chemicals and products to workers and consumers so that appropriate measures can be taken to avoid accidental or inadvertent contact with the eye. In addition, ocular safety labels provide the necessary first aid measures that should be taken in the event of accidental exposures.

Among the fundamental principles agreed upon by participants establishing the GHS was the assurance that "the level of protection offered to workers, consumers, the general public and the environment should not be reduced as a result of harmonizing the classification and labeling systems" (UN 2007). ICCVAM has conducted technical analyses to support the development of appropriate recommendations for GHS options that would continue to provide at least equivalent protection as current U.S. ocular hazard classification and labeling requirements. ICCVAM recommends that U.S. agencies consider the GHS ocular hazard classification criteria and categories and the level of protection they provide compared to current U.S. hazard classification systems.

Federal law requires agencies to determine that new test methods recommended by ICCVAM generate data that are at least equivalent to that generated by test methods they currently require or recommend for hazard identification purposes. Given that the BCOP test method does not identify eye irritation hazards when using the EU or GHS hazard classification systems that are currently identified using U.S. hazard identification and classification requirements (EPA and FHSA), ICCVAM cannot recommend using the BCOP test method as a screening test to identify substances not labeled as irritants for the GHS classification system. ICCVAM will revisit recommendations for the BCOP test method based on any updates to the GHS eye hazard classification criteria that may occur as a result of the NICEATM analyses.

Identification of Reversible Eye Irritation Hazard Categories

Based on an evaluation of available data and corresponding performance (accuracy and reliability), ICCVAM concludes that the BCOP test method is **not** recommended to identify moderate and mild ocular irritants as defined by the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[4]

[4] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between mild and moderate ocular irritants.

Table 2-1 *In Vivo* Ocular Lesions from False Negative Substances in the BCOP Test Method Using the EPA Classification System

Compound	N	*In Vivo* Scores[1]			
		Corneal Opacity: Score (day cleared)	Iris: Score (day cleared)	Conjunctival Redness: Score (day cleared)	Chemosis: Score (day cleared)
Dimethylbiquanide	3	N = 1 1(2) N = 1 1(3)	N = 1 1(2)	N = 2 2(3)	N = 2 2(1)
EDTA	3	N = 1 1(3)	N = 2 1(1)	N = 3 2(2)	N = 1 2(1) N = 1 2(2) N = 1 3(2)
Magnesium carbonate	3	N = 1 1(2) N = 1 1(3)	None	None	None
Polyalkenylsuccinate ester/amine salt	6	N = 2 1(2)	None	N = 1 2(6) N = 3 2(2) N = 1 3(2) N = 1 3(6)	N = 1 2(1) N = 1 2(2)
Compound I	6	N = 1 1(2)	None	None	None
Iminodibenzyl	3	N = 3 1(2)	None	None	None
Methylcyclopentane	6	None	None	N = 1 2(3)	None
Tween 20	4	None	None	N = 2 2(2)	None

Abbreviations: N = number of animals.

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

Evaluation as a Screening Test to Identify Ocular Corrosives and Severe Irritants

In the original ICCVAM evaluation of the BCOP test method as a screening test to identify substances as ocular corrosives and severe irritants, overall accuracy was 79% (113/143) to 81% (119/147), false positive rates were 19% (20/103) to 21% (22/103), and false negative rates were 16% (7/43) to 25% (10/40) depending on the hazard classification system (i.e., EPA Category I, EU R41, GHS Category 1).

Based on the current updated BCOP validation database, which has increased from 145 to 211 substances, overall accuracy of the BCOP test method as a screening test to identify substances as ocular corrosives and severe irritants is 77% (91/118) to 79% (148/187) depending on the hazard classification system (i.e., EPA Category I, EU R41, GHS Category 1).[5] The false positive rate is 23% (29/124) to 24% (29/122), and false negative rates are 15% (10/65) to 21% (7/33) depending on the hazard classification system used. Based on the similar performance statistics of the current and the original databases, the ICCVAM recommendation for the use of the BCOP test method to identify substances as ocular corrosives and severe irritants remains unchanged:

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

[5] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an evaluation of the BCOP test method as a screening test to identify ocular corrosives/severe irritants using the FHSA classification system is not possible.

testing strategy, as part of a weight-of-evidence approach. In a tiered-testing strategy, when a positive result is obtained in an appropriately validated *in vitro* test, a test substance may be classified as an ocular hazard without testing in rabbits. A substance that tests negative in the *in vitro* ocular toxicity test would need to be tested in the *in vivo* ocular test to identify possible *in vitro* false negatives and to identify moderate and mild ocular irritants (ICCVAM 2006e).

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the available data and test method performance support the ICCVAM draft recommendations that the BCOP test method is not recommended to distinguish substances from all hazard categories as defined by the EPA, EU, and GHS classification systems. The Panel agreed with ICCVAM that the BCOP test method continue to be recommended as a screening test for severe irritants. The Panel also concluded that the BCOP test method can be used as a screening test to distinguish substances not labeled as irritants from all other hazard categories when results are to be used for EU or GHS hazard classifications. *It should be noted that this recommendation preceded the NICEATM evaluation of the GHS classification system and, therefore, this information was not taken into consideration.* However, like ICCVAM, the Panel concluded that, because of the significant lesions associated with 50% (4/8) of the EPA Category III substances that tested as false negative, the BCOP test method cannot be recommended as a screening test to identify EPA Category IV substances.

2.1.2 ICCVAM Recommendations: BCOP Test Method Protocol

For use of the BCOP test method as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1), ICCVAM recommends using the updated ICCVAM BCOP test method protocol that is included as an appendix to this report (**Appendix B**). In addition, all future studies intended to further characterize the usefulness and limitations of the BCOP test method should be conducted using this protocol.

Independent Peer Review Panel Conclusions and Recommendations

While the BCOP test method protocol was previously reviewed for use in identifying ocular corrosives and severe irritants, the Panel emphasized the importance of protocol elements. They emphasized that use of this protocol to identify mild/moderate ocular irritants in future studies should include (1) methods for harvest and storage of eyes, (2) timeframe from harvest to use of eyes, (3) consistent animal age, (4) screening for existing corneal lesions prior to use, (5) concurrent positive and negative controls, (6) inclusion of an untreated negative control, and (7) refinement of histopathological methodology.

2.1.3 ICCVAM Recommendations: Future Studies for the BCOP Test Method

To further the use of this test method and to evaluate the use of the BCOP test method as a potential replacement for the Draize rabbit eye test or for the identification of mild and moderate ocular irritants (i.e., EPA Category II, III; EU R36; GHS Category 2A, 2B) and substances not labeled as irritants (i.e., EPA Category IV; EU Not Labeled; FHSA Not Labeled; GHS Not Classified), ICCVAM recommends additional studies be considered and undertaken.

- Additional optimization studies/evaluations should be conducted in an attempt to improve the correct classification of mild and moderate ocular irritants and substances not labeled as irritants. After optimization, additional studies to further assess the reliability and accuracy of the test method are recommended.
- ICCVAM recommends that a histopathological evaluation of the corneal tissue, using standardized procedures, be included when the BCOP test method is conducted. Such data will allow for development of decision criteria and future assessments on the usefulness of

this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.

- ICCVAM encourages users to provide all data that are generated from future studies. They could be used to further characterize the usefulness and limitations of the BCOP test method for the identification of all ocular hazard categories.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the available data support the ICCVAM draft recommendations. The Panel encouraged continued test method development and refinement of the protocol to achieve more accurate classification of mild and moderate irritants. The Panel also recommended that problematic chemical classes within these hazard categories be identified in order to determine if improved performance could be achieved by restricting the applicability domain.

2.1.4 ICCVAM Recommendations: Performance Standards for the BCOP Test Method

Based on the available data and associated performance described in the final ICCVAM BRD (**Appendix C**), ICCVAM recommends that the development of performance standards for the BCOP test method is not warranted at this time.

2.2 Validation Status of the BCOP Test Method

The following is a synopsis of the information in the final ICCVAM BRD (**Appendix C**), which reviews the available data and information for the BCOP test method. The ICCVAM BRD describes the current validation status of the BCOP test method, including what is known about its reliability and accuracy, the scope of substances tested, and standardized protocols for the validation study.

2.2.1 Test Method Description

The BCOP test method is an *in vitro* eye irritation test method using isolated bovine corneas that are byproducts from processing plants. In the BCOP test method, opacity is determined by the amount of light transmitted through the cornea, and permeability is determined by the amount of sodium fluorescein dye that passes through all corneal cell layers. Both measurements are used to calculate an IVIS, which is used to assign an *in vitro* irritancy classification to predict the *in vivo* ocular irritation potential of a test substance.

2.2.2 Validation Database

An online literature search conducted in support of the evaluation of the validation status of the BCOP test method identified four publications containing BCOP test method results. However, none of these publications included raw data or *in vivo* reference data, or they included data cited from earlier studies that were already included in the validation database. Accordingly, these were not added to the database. The results from the BCOP test method for 66 antimicrobial cleaning products (AMCPs) were obtained from a submission to ICCVAM that describes a non-animal approach for evaluating eye irritation potential and labeling requirements for AMCPs. Therefore, the previous validation database for the BCOP test method (ICCVAM 2006a) was updated to include BCOP data for the 66 AMCPs. The updated BCOP validation database contains 211 substances, representing a wide variety of chemical and product classes, and including 135 commercial products or formulations.

Detailed *in vivo* data were necessary to calculate the appropriate EPA, EU, FHSA, and GHS ocular hazard classifications (EPA 2003; EU 2001; FHSA 2005; UN 2007) (**Appendix C**). These data include 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 days. Thus, some of the test substances for which there was only limited *in vivo* data could not be used for evaluating test method accuracy and reliability. Additionally, because the FHSA classification system is based on a sequential testing

strategy that uses up to 18 animals, only a small percentage of the substances in the test method databases would be classifiable if the FHSA criteria were strictly applied. Therefore, to maximize the number of substances included in these analyses, two separate "proportionality" criteria were applied for the purpose of assigning an FHSA classification. Based on the minimum number of positive animals needed to identify a substance as an irritant using a single 6-animal test in the FHSA sequential testing strategy, a 67% threshold for positive responding animals was used (i.e., at least 2/3 or 4/6 positive animals) to assign an irritant classification. Alternatively, based on the minimum number of positive animals needed (4/18 [22%]) to identify a substance as an irritant when three 6-animal tests are required in the sequential testing strategy, a 20% threshold for positive responding animals was used to assign an irritant classification.

2.2.3 Test Method Accuracy

Identification of All Ocular Hazard Categories

The ability of the BCOP test method to identify all categories of ocular irritation potential was evaluated for the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[6] As indicated in **Table 2-2**, overall correct classifications ranged from 49% (91/187) to 55% (102/187), depending on the hazard classification system used when evaluating the entire database. Using alternative decision criteria to identify ocular corrosives and severe irritants (i.e., IVIS $\geq$ 75 [used in the AMCP submission protocol] instead of IVIS $\geq$ 55.1 [as per the ICCVAM-recommended BCOP protocol]) did not improve test method performance.

Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories

The ability of the BCOP test method to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other ocular hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B), as defined by the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007) was also evaluated.

As indicated in **Table 2-3**, overall accuracy for the identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other categories ranged from 64% (76/118) to 83% (148/179, 155/107, or 161/194) depending on the hazard classification system used. While false positive rates were high (53% [24/45 or 25/47] to 70% [63/90]) depending on the hazard classification system used), the false negative rates were low (0% [0/54 or 0/97] to 6% [8/142]) depending on the hazard classification system used. All eight of the false negative substances for the EPA classification system were EPA Category III (i.e., ocular injuries to the cornea and/or iris [inside the eye] and/or conjunctival injuries that persisted more than 24 hours but less than 7 days) based on Draize rabbit eye test data (**Table 2-1**). This included three substances (38% [3/8]) that were classified as EPA eye irritants based on at least one rabbit with corneal lesions and opacity that did not resolve until day 3 of the study. A fourth substance was classified as an EPA eye irritant based on all six rabbits with conjunctival redness scores of 3 (*producing diffuse, crimson color of the conjunctiva, individual blood vessels not easily discernable*). The conjunctival redness scores for two of these animals did not recover to a score of 1 (s*ome blood vessels definitely hyperemic*) until day 6 of the study. The conjunctival redness scores for the remaining four rabbits recovered to a score of 1 on day 2 of the study.

[6] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an evaluation of the BCOP test method to identify all ocular hazard categories using the FHSA classification system is not possible.

Table 2-2 Evaluation of the Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by EPA, EU, and GHS Classification Systems[1]

Severe using ≥55.1											
	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
GHS	49% (91/187)	85% (55/65)	15% (10/65)	62% (16/26)	27% (7/26)	11% (3/26)	67% (4/6)	33% (2/6)	0% (0/6)	70% (63/90)	30% (27/90)
EPA	55% (102/187)	84% (53/63)	16% (10/63)	50% (11/22)	32% (7/22)	18% (4/22)	50% (28/57)	36% (21/57)	14% (8/57)	53% (24/45)	47% (21/45)
EU	50% (59/118)	79% (26/33)	21% (7/33)	48% (10/21)	52% (11/21)	0% (0/21)	NA	NA	NA	66% (42/64)	34% (22/64)
Severe using ≥75											
		Severe		Moderate			Mild			Not Labeled	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
GHS	50% (94/187)	78% (51/65)	22% (14/65)	31% (8/26)	54% (14/26)	15% (4/26)	67% (4/6)	33% (2/6)	0% (0/6)	70% (63/90)	30% (27/90)
EPA	49% (92/187)	78% (49/63)	22% (14/63)	36% (8/22)	45% (10/22)	19% (4/22)	47% (27/57)	39% (22/57)	14% (8/57)	53% (24/45)	47% (21/45)
EU	51% (60/118)	73% (24/33)	27% (9/33)	29% (6/21)	67% (14/21)	4% (1/21)	NA	NA	NA	66% (42/64)	34% (22/64)

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency; EU = European Union; GHS = Globally Harmonized System; NA = not applicable.

[1] EPA classification system (EPA 2003); EU classification system (EU 2001); GHS classification system (UN 2007).

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

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

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

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

Table 2-3 Accuracy of the BCOP Test Method for Distinguishing Substances Not Labeled as Irritants[1] from All Other Irritant Classes

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
GHS	187	66	124/187	100	97/97	30	27/90	70	63/90	0	0/97
EPA	187	83	155/187	94	134/142	47	21/45	53	24/45	6	8/142
EU	118	64	76/118	100	54/54	34	22/64	66	42/64	0	0/54
FHSA-20%	194	83	161/194	95	139/147	47	22/47	53	25/47	5	8/147
FHSA-67%	179	83	148/179	95	126/132	47	22/47	53	25/47	5	6/132

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency; EU = European Union; FHSA = U.S. Federal Hazardous Substances Act; GHS = Globally Harmonized System; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Not Classified vs. Category 1/2A/2B.
EPA classification system (EPA 2003): Category IV vs. Category I/II/III.
EU classification system (EU 2001): Not Labeled vs. R41/R36.
FHSA classification system (FHSA 2005): Not Labeled vs. Irritant.

2.2.4 Test Method Reliability

Interlaboratory Reproducibility

Quantitative and qualitative evaluations of the BCOP test method reliability have been conducted previously (ICCVAM 2006a). However, additional qualitative analyses of interlaboratory reproducibility were conducted to evaluate the extent of agreement of BCOP hazard classifications among the laboratories participating in the three interlaboratory validation studies (Balls et al. 1995; Gautheron et al. 1994; Southee 1998). As was done for the accuracy evaluation, these qualitative evaluations of reproducibility were based on (1) the use of the BCOP test method for identifying all ocular hazard categories according to the EPA, EU, or GHS systems and (2) the use of the BCOP 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, III; EU R41, R36; GHS Category 1, 2A, 2B). Given that the performance of the BCOP test method was similar for the EPA and FHSA classification systems, additional reliability analyses were not conducted for the FHSA classification system.

Using the first approach (i.e., identifying all ocular hazard categories) among the three interlaboratory studies for the Balls et al. (1995) study, there was 100% agreement among the five laboratories for a majority of the Draize ocular corrosives and severe irritants based on all three classification systems, whether they were correctly identified or underclassified by the BCOP test method (e.g., for the GHS system, there was 100% agreement for 88% [15/17] of the correctly identified Category 1 substances). There was also 100% agreement among the five laboratories for 100% (10/10) of the overpredicted Not Labeled substances and for at least 50% (2/4) of the correctly identified Not Labeled substances.

For the Gautheron et al. (1994) study, there was 100% agreement among the 11 laboratories for a majority of the Draize ocular corrosives and severe irritants based on all three classification systems, whether they were correctly identified or underclassified by the BCOP test method (e.g., for the GHS system, there was 100% agreement for 67% [4/6] of the correctly identified Category 1 substances). There was also 100% agreement among the 11 laboratories for a majority of the overpredicted Not Labeled substances (e.g., for the EU system, there was 100% agreement for 54% [7/13] of the

correctly identified Not Labeled substances) and for a majority of the incorrectly identified Not Labeled substances (e.g., for the EU system, there was 100% agreement for 91% [21/23] of the correctly identified substances).

For the Southee (1998) study, there was 100% agreement among the three laboratories for all of the ocular corrosives and severe irritants based on all three classification systems, whether they were correctly identified or underclassified by the BCOP test method (e.g., for the GHS system, there was 100% agreement for 100% [4/4] of the Draize ocular corrosives and severe irritants). There was also 100% agreement among the two correctly identified Not Labeled substances.

Using the second approach (i.e., distinguishing substances not labeled as irritants from all other ocular hazard categories) for the Balls et al. (1995) study, there was 100% agreement for 92% (55/60) to 93% (56/60) of the substances tested *in vitro*, depending on the classification system used. All five laboratories were in 100% agreement on the classification of 50% (2/4) of Not Labeled substances and 94% (32/34) to 96% (48/50) of all other irritant class substances, depending on the classification system used.

For the Gautheron et al. (1994) study, there was 100% agreement among the eleven laboratories for 65% (34/52) of the substances tested *in vitro*, for all classification systems. There was 100% agreement among the laboratories on the classification of 83% (10/12) to 87% (27/31) of all other irritant class substances, depending on the classification system used.

There was 100% agreement among the three laboratories in the Southee (1998) study for 88% (14/16) of the substances tested *in vitro*, for all classification systems. All three laboratories were in 100% agreement on the classification of 100% (2/2) Not Labeled substances and 90% (9/10) to 92% (11/12) of all other irritant class substances, depending on the classification system used.

As stated above, the final ICCVAM BRD (**Appendix C**) provides a comprehensive summary of the current validation status of the BCOP test method, including what is known about its reliability and accuracy, and the scope of substances tested. Raw data for the BCOP test method will be maintained for future use, so that these performance statistics may be updated as additional information becomes available.

2.2.5 Animal Welfare Considerations

The BCOP test method refines animal use. Because these animals are being humanely processed for nonlaboratory purposes, the testing procedure inflicts no additional pain or distress. Substances that are identified as corrosive or severe irritants *in vitro* are excluded from *in vivo* testing.

The BCOP test method can also reduce animal use because the test method utilizes animal species routinely raised as a food source in large numbers and thus replaces the need for laboratory animals.

3.0 The Cytosensor Microphysiometer Test Method

A number of *in vitro* test systems that have been proposed as alternatives to the Draize rabbit eye test rely on cell death as an endpoint. However, reversible cell changes may provide more appropriate endpoints for the assessment of ocular irritation potential. Good correlations have been reported between results obtained from the CM test method and *in vivo* eye irritancy data. The method is noninvasive, and thus allows the determination of recovery of the cells from the toxic insult. The CM test method measures the rate of extracellular acidification of populations of living cells maintained in flow chambers. After establishing a baseline acidification rate for each set of cells and measuring the new rates subsequent to each sample addition, the concentration of test material (w/v%) required to reduce the acidification rate to 50% is computed by interpolation between the rate data points spanning the 50% response level. This value is termed the MRD_{50} and is the endpoint for the test.

ICCVAM evaluated the validation status of the CM test method, which was not part of the ICCVAM 2006 evaluation, as an *in vitro* alternative to the Draize rabbit eye test for identifying ocular corrosives and severe irritants (i.e., those that induce irreversible ocular damage; EPA Category I, EU R41, GHS Category 1) and substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) according to the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007).

3.1 ICCVAM Recommendations

3.1.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

Evaluation as a Screening Test to Identify Substances Not Labeled as Irritants
ICCVAM concludes that the accuracy and reliability of the CM test method are sufficient to support its use as a screening test to distinguish water-soluble surfactant chemicals and certain types of surfactant-containing formulations (e.g., cosmetics and personal care product formulations, but not pesticide formulations) that are not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled) from all other hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant) when results are to be used specifically for hazard classification and labeling purposes under the EPA, EU and FHSA classification systems (EPA 2003; EU 2001; FHSA 2005). Until the issues associated with the GHS classification system are further discussed (see "BCOP Test Method Usefulness and Limitations"), ICCVAM is deferring final recommendations on the usefulness and limitations of using the CM test method as a screening test to identify substances not labeled as irritants according to the GHS classification system.

When the CM test method was used to distinguish substances not labeled as irritants among the database of 53 surfactant-containing substances, the false negative rate ranged from 0% (0/27) to 2% (1/47) depending on the hazard classification system used. The one false negative substance based on *in vivo* data was EPA Category III or FHSA Irritant. For this substance, six test animals were included. One test animal had no observable effects, three test animals had conjunctival redness (score = 1), and two test animals had corneal opacity (score = 1) that cleared after one day.

When the CM test method was used to distinguish substances not labeled as irritants among the database of 29 water-soluble nonsurfactant substances and formulations, the false negative rate ranged from 24% (5/21) to 40% (8/20) depending on the hazard classification system used. Because of these high false negative rates, the CM test method is **not** recommended as a screening test to distinguish substances not labeled as irritants among these types of substances.

Evaluation as a Screening Test to Identify Ocular Corrosives and Severe Irritants
ICCVAM recommends that the CM test method can be used as a screening test to identify water-soluble substances (i.e., water-soluble surfactants, surfactant-containing formulations, and nonsurfactants) as ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS

Category 1[7]) in a tiered-testing strategy, as part of a weight-of-evidence approach.[8] A substance that tests negative with the CM test method would need to be tested in another test method that is capable of identifying possible *in vitro* false negative ocular corrosives and severe irritants and to distinguish between moderate and mild ocular irritants. Currently, the Draize rabbit eye test is the only test method capable of making such a distinction.

Given that the CM test method (INVITTOX Protocol Number 102) is proposed for use as a screening test to identify both ocular corrosives and severe irritants and substances not labeled as irritants, users may want to consider using the CM test method prior to another *in vitro* ocular test method for testing these types of substances. However, testing in another test method would be necessary for certain substances not identified with the CM test method: (1) water-soluble substances that are not identified as ocular corrosives and severe irritants or (2) water-soluble surfactant chemicals and specific types of surfactant-containing formulations that are not identified as substances not labeled as irritants. The other test method must be capable of correctly classifying substances into each of the four hazard classification categories for the EPA or GHS classification systems. Currently, the only test method accepted for these purposes is the Draize rabbit eye test. Because of the high false positive rate (50% [3/6] to 69% [18/26] depending on the hazard classification system used) for substances not labeled as irritants, users may not want to use the CM test method if the intention is to identify substances not labeled as irritants first.

Independent Peer Review Panel Conclusions and Recommendations

The Panel concluded that the CM test method is recommended as a screening test to identify water-soluble surfactant substances as ocular corrosives and severe irritants and substances not labeled as irritants in a tiered-testing strategy, as part of a weight-of-evidence approach. However, major concerns are the continued availability of the instrument used to conduct the CM test method, and what new manufacturing processes, including the subsequent required revalidation, might mean to already existing CM test method data.

When using the CM test method to identify water-soluble surfactant substances as substances not labeled as irritants, the false negative rate for the EU, FHSA-67%, and GHS systems was 0% (0/27 or 0/28). For the EPA system it was 2% (1/46). Therefore, the CM test method was recommended for such testing purposes. The Panel recommended that further studies using the CM test method are needed, in particular for EPA Categories III and IV, and that the available data restrict the applicability domain of the CM test method to water-soluble surfactants and surfactant-based formulations (e.g., cosmetics and personal care products).

3.1.2 ICCVAM Recommendations: CM Test Method Protocol

For use of the CM test method as a screening test to identify water-soluble substances as ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1), or to identify substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified), ICCVAM recommends using the ICCVAM CM test method protocol that is included as an appendix to this report (**Appendix B**). In addition, all future studies intended to further characterize the usefulness and limitations of the CM test method should be conducted using this protocol.

---

[7] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an evaluation of the CM test method as a screening test to identify ocular corrosives/severe irritants using the FHSA classification system is not possible.

[8] The ECVAM Scientific Advisory Committee (ESAC) has also recommended the CM test method for this purpose and for this limited applicability domain (ESAC 2009; **Appendix K**).

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that critical aspects of the CM test method had been adequately justified and described and that the protocol was sufficiently detailed. The Panel supported the use of the recommended protocol for future studies to further characterize the usefulness and limitations of the CM test method. However, they expressed concern that the CM test method is unlikely to be widely used because manufacture of the instrument required to conduct the test method has been discontinued.

3.1.3 ICCVAM Recommendations: Future Studies for the CM Test Method

To expand the applicability domain of the CM test method for the identification of ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) and substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified), ICCVAM recommends additional studies be considered and undertaken.

- ICCVAM recommends that these substances be selected from the ICCVAM-recommended reference substances for validation of *in vitro* ocular safety test methods for the evaluation of ocular corrosives and severe irritants[9] in order to provide for a more direct assessment of the CM test method's utility as a screening test for identifying ocular corrosives and severe irritants. Similarly, a reference set could also be selected from this list for the purposes of assessing the utility of the CM test method as a screening test for identifying substances not labeled as irritants.
- ICCVAM recommends that future optimization studies be directed towards increasing the performance of the CM test method for identifying all categories of ocular irritancy hazard classification according to the EPA, EU, or GHS hazard classification systems. This will require that an increased number of substances in the moderate and mild ocular irritant categories (i.e., EPA Category II, III; EU R36; GHS Category 2A, 2B) be identified and tested.
- ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of the CM test method for the identification of all ocular hazard categories.

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that additional studies should be conducted to further characterize the usefulness and limitations of the CM test method for use as a screening test for identifying ocular corrosives and severe irritants or substances not labeled as irritants. Because the CM test method is limited to testing water-soluble surfactants and certain types of surfactant formulations, the Panel recommended that a range of surfactant concentrations should be tested because surfactants form micelles at higher concentrations, which reduce the number of surfactant molecules available to react with the target tissue.

3.1.4 ICCVAM Recommendations: Performance Standards for the CM Test Method

Based on the available data and associated performance described in the redacted ECVAM CM BRD (**Appendix D**), ICCVAM recommends that the development of performance standards for the CM test method is not warranted at this time.

3.2 Validation Status of the Cytosensor Microphysiometer Test Method

The following is a synopsis of the information for three of the peer-reviewed publications (Balls et al. 1995, Gettings et al. 1996, Brantom et al. 1997) referenced in the redacted ECVAM CM BRD (**Appendix D**) and utilized by ICCVAM in its review. The redacted ECVAM CM BRD describes the

[9] http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_tmer.htm

current validation status of the CM test method, including what is known about its reliability and accuracy, the scope of substances tested, and standardized protocols for the validation study.

3.2.1 Test Method Description

The CM test method estimates the metabolic rate (i.e., glucose utilization rate) of cells by measuring the rate of excretion of acid byproducts and resulting decrease in pH of the surrounding medium in an enclosed chamber. The rate of change in pH per unit time becomes the metabolic rate of the population. If a test material causes cytotoxicity to this population of cells it is assumed that the metabolic rate will fall. Although the metabolic rate is the physical parameter that is measured with the CM test method, the magnitude of the metabolic rate itself is not directly related to eye irritation potential. Rather, the reduction of the metabolic rate to 50% of its basal rate is the parameter used to measure the impact of the test article on the test system (L929 cells in almost all cases). The CM test method exposes a population of cells to increasing concentrations of the test article (diluted in medium). The exposure follows a 3-step process. The first step is the exposure to the diluted test article, the second is the test article rinse-out, and the third is the measurement of the metabolic activity. This means that the impact of the exposure is measured immediately, and then a subsequent exposure is performed until the highest testable concentration has been used or the population of cells is severely damaged and the metabolic rate has declined to effectively zero. From the concentration response curve, the concentration that leads to a 50% decline in the metabolic rate of the population (the MRD_{50}) is calculated. The MRD_{50} values are used to compare test materials and provide a measure of ocular irritancy potential.

3.2.2 Validation Database

Data on 53 water-soluble surfactant and surfactant-containing formulations were provided in the redacted ECVAM CM BRD (**Appendix D**), where data from the Draize rabbit eye test were also available to assess the accuracy of the CM test method. The database of 53 water-soluble surfactants tested in the CM test method included 21 surfactant chemicals and 32 surfactant-containing formulations tested across seven different laboratories.

The nonsurfactant substances database (n = 29) consisted of 27 water-soluble nonsurfactant chemicals, which included a range of chemical classes (e.g., acids, alcohols, alkalis, and ketones), and two nonsurfactant formulations (n = 2) tested in seven laboratories.

3.2.3 Test Method Accuracy

Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories
NICEATM evaluated the CM test method's ability to distinguish substances not labeled as irritants (EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other hazard categories (EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B), as defined by the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007) among the water-soluble surfactants and surfactant-containing formulations.

As indicated in **Table 3-1**, overall accuracy for the identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other categories for the water-soluble surfactants and surfactant-containing formulations ranged from 66% (35/53) to 93% (43/46) depending on the hazard classification system used. The false negative rate ranged from 0% (0/27, 0/28, or 0/40) to 2% (1/46 or 1/47) depending on the hazard classification system used. The one false negative in both the EPA and the FHSA-20% classification systems was classified as Category III and Irritant, respectively, based on Draize rabbit eye test data. For this substance, six rabbits were included in the *in vivo* test. One rabbit had no observable effects, three rabbits had conjunctival redness (score = 1), and two rabbits had corneal opacity (score = 1) that cleared after one day.

The ability of the CM test method to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B), as defined by the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007) among the water-soluble nonsurfactant substances was also evaluated.

As indicated in **Table 3-2**, overall accuracy for the identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other categories for the water-soluble nonsurfactant substances ranged from 63% (15/24) to 76% (22/29) depending on the hazard classification system used. The false negative rate ranged from 24% (5/21) to 40% (8/20) depending on the hazard classification system used. Eight substances were false negative when using the EPA, GHS, and FHSA classification systems. In the EPA system, they were classified *in vivo* as Category 1 (n = 1), Category II (n = 3) and Category III (n = 4). In the GHS system, they were classified *in vivo* as Category 1 (n = 1) and Category 2A (n = 7). For the FHSA system, they were classified *in vivo* as Irritant.

Table 3-1 Accuracy of the CM Test Method for Distinguishing Substances Not Labeled as Irritants[1] from All Other Irritant Classes for Surfactant-Containing Substances

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
GHS	53	68	36/53	100	28/28	32	8/25	68	17/25	0	0/28
EPA	52	92	48/52	98	45/46	50	3/6	50	3/6	2	1/46
EU	53	66	35/53	100	27/27	31	8/26	69	18/26	0	0/27
FHSA-20%	53	92	49/53	98	46/47	50	3/6	50	3/6	2	1/47
FHSA-67%	46	93	43/46	100	40/40	50	3/6	50	3/6	0	0/40

Abbreviations: CM = Cytosensor Microphysiometer; EPA = U.S. Environmental Protection Agency; EU = European Union; FHSA = U.S. Federal Hazardous Substances Act; GHS = Globally Harmonized System; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Not Classified vs. Category 1/2A/2B.
EPA classification system (EPA 2003): Category IV vs. Category I/II/III.
EU classification system (EU 2001): Not Labeled vs. R41/R36.
FHSA classification system (FHSA 2005): Not Labeled vs. Irritant.

Table 3-2 Accuracy of the CM Test Method for Distinguishing Substances Not Labeled as Irritants[1] from All Other Irritant Classes for Nonsurfactant Substances

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
GHS	25	64	16/25	62	13/21	75	3/4	25	1/4	38	8/21
EPA	29	66	19/29	67	16/24	60	3/5	40	2/5	33	8/24
EU	29	76	22/29	76	16/21	75	6/8	25	2/8	24	5/21
FHSA-20%	25	64	16/25	62	13/21	75	3/4	25	1/4	38	8/21
FHSA-67%	24	63	15/24	60	12/20	75	3/4	25	1/4	40	8/20

Abbreviations: CM = Cytosensor Microphysiometer; EPA = U.S. Environmental Protection Agency; EU = European Union; FHSA = U.S. Federal Hazardous Substances Act; GHS = Globally Harmonized System; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Not Classified vs. Category 1/2A/2B.
EPA classification system (EPA 2003): Category IV vs. Category I/II/III.
EU classification system (EU 2001): Not Labeled vs. R41/R36.
FHSA classification system (FHSA 2005): Not Labeled vs. Irritant.

Distinguishing Ocular Corrosives and Severe Irritants from All Other Hazard Categories

The ability of the CM test method to distinguish ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1) from all other ocular hazard categories (i.e., EPA Category II, III, IV; EU R36, Not Labeled; GHS Category 2A, 2B, Not Classified) as defined by the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007)[10] was evaluated among the water-soluble surfactants and surfactant-containing formulations.

As indicated in **Table 3-3**, overall accuracy for the identification of ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1) from all other categories for surfactant-containing substances ranged from 85% (44/52) to 94% (50/53) depending on the hazard classification system used. The false positive rates ranged from 3% (1/30) to 10% (3/29) depending on the hazard classification system used. The three false positives when using the EPA classification system are classified as Category II (n = 2) or III (n = 1) based on Draize rabbit eye test data. The one false positive when using the GHS and EU classification systems is Not Classified and Not Labeled, respectively, based on Draize rabbit eye test data.

Table 3-3 Accuracy of the CM Test Method for Distinguishing Corrosives/Severe Irritants[1] from All Other Irritant Classes for Surfactant-Containing Substances

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
GHS	53	94	50/53	91	21/23	97	29/30	3	1/30	9	2/23
EPA	52	85	44/52	78	18/23	90	26/29	10	3/29	22	5/23
EU	53	89	47/53	81	21/26	96	26/27	4	1/27	19	5/26

Abbreviations: CM = Cytosensor Microphysiometer; EPA = U.S. Environmental Protection Agency; EU = European Union; GHS = Globally Harmonized System; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Category 1 vs. Category 2A/2B/NC.
EPA classification system (EPA 2003): Category I vs. Category II/III/IV.
EU classification system (EU 2001): R41 vs. R36/NL.

The ability of the CM test method to distinguish ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1) among the water-soluble nonsurfactant substances was evaluated for the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[11] As indicated in **Table 3-4**, overall accuracy ranged from 79% (23/29) to 92% (23/25) depending on the hazard classification system used. The false positive rate was 0% (0/17 or 0/18) for all classification systems used.

[10] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an evaluation of the CM test method as a screening test to identify ocular corrosives/severe irritants using the FHSA classification system is not possible.

[11] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an evaluation of the CM test method as a screening test to identify ocular corrosives/severe irritants using the FHSA classification system is not possible.

Table 3-4 Accuracy of the CM Test Method for Distinguishing Corrosives/Severe Irritants[1] from All Other Irritant Classes for Nonsurfactant Substances

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
GHS	29	83	24/29	55	6/11	100	18/18	0	0/18	45	5/11
EPA	25	92	23/25	71	5/7	100	18/18	0	0/18	29	2/7
EU	29	79	23/29	50	6/12	100	17/17	0	0/17	50	6/12

Abbreviations: CM = Cytosensor Microphysiometer; EPA = U.S. Environmental Protection Agency; EU = European Union; GHS = Globally Harmonized System; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Category 1 vs. Category 2A/2B/NC. EPA classification system (EPA 2003): Category I vs. Category II/III/IV. EU classification system (EU 2001): R41 vs. R36/NL.

3.2.4 Test Method Reliability

Interlaboratory Reproducibility

For the CM test method, intralaboratory reproducibility was assessed quantitatively based on calculated coefficients of variation (CVs) for MRD_{50} values for two different studies. Mean CVs ranged from 10% to 24% and tended to be slightly higher for surfactant substances than for nonsurfactant substances.

Interlaboratory reproducibility of the CM test method was also assessed using the data from the European Commission/Home Office (EC/HO; Balls et al. 1995) and European Cosmetic, Toiletry and Perfumery Association (COLIPA; Brantom et al. 1997) validation studies, which included four laboratories and two laboratories, respectively. Mean CVs in the EC/HO study ranged from 16% to 37% for surfactant substances and up to 51% for nonsurfactant substances. For surfactant materials, all four laboratories using the CM test method had 100% agreement for 55% (6/11) of the test substances; 75% of the laboratories had identical results for 27% (3/11) of the test substances; and 50% of the laboratories had agreement for 18% (2/11) of the test substances. For nonsurfactant substances, agreement among the laboratories was 100% for 48% (11/23) of the test substances, 75% for 22% (5/23) of the test substances, 67% for 4% (1/23) of the test substances, and 50% for 13% (3/23) of the test substances.

For the COLIPA study, substances were divided into surfactant materials, surfactant-based formulations and mixtures, and nonsurfactant substances. Two laboratories had mean between-laboratory CVs ranging from 16% to 23% for surfactant materials, approximately 16% for surfactant-based formulations and mixtures, and 32% to 51% for nonsurfactant substances. For surfactant materials, the laboratories had 100% agreement for 90% (9/10) of the test substances and 0% agreement for 10% (1/10) of them. The laboratories had 100% agreement for 100% (7/7) surfactant-based formulations and mixtures. For nonsurfactant substances, the laboratories had 100% agreement for 78% (7/9) of the test substances and 0% agreement for 22% (2/9) of them.

3.2.5 Animal Welfare Considerations

Except for the mice originally used to develop the L929 cell line, no animals are used for the CM test method.

4.0 The Hen's Egg Test–Chorioallantoic Membrane Test Method

The HET-CAM test method uses the chorioallantoic membrane, which is a vascular fetal membrane composed of the fused chorion and allantois. The acute effects induced by a test substance on the small blood vessels and proteins of this soft tissue membrane are used as an indicator of effects induced by the same test substance in the eye of a treated rabbit.

ICCVAM previously evaluated the validation status of the HET-CAM test method as an *in vitro* alternative to the Draize rabbit eye test to identify ocular corrosives and severe irritants (i.e., those that induce irreversible ocular damage; EPA Category I, EU R41, GHS Category 1) and determined that the reproducibility and accuracy was **not** sufficient to support its use for this purpose (ICCVAM 2006e). In the current evaluation, ICCVAM evaluated the validation status of the HET-CAM test as an *in vitro* alternative to the Draize rabbit eye test for identifying nonsevere ocular irritants (i.e., those that induce reversible ocular damage; EPA Category II and III, EU R36, GHS Category 2A and 2B) and substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) according to the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007).

4.1 ICCVAM Recommendations

4.1.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

Evaluation as a Screening Test to Identify Substances Not Labeled as Irritants
Based on the current evaluation, ICCVAM concludes that the scientific validity of the HET-CAM test method has been adequately evaluated and that the HET-CAM test method is **not** recommended as a screening test to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B) when results are to be used specifically for hazard classification and labeling purposes under the EPA, EU, FHSA, or GHS hazard classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007). There are too few surfactants or oil/water emulsions in the moderate irritant categories to allow sufficient confidence in the ability of the HET-CAM test method to distinguish them from the substances not labeled as irritants (i.e., there were no GHS Category 2A substances and only two EPA Category II or EU R36 substances).

Identification of Reversible Eye Irritation Hazard Categories
ICCVAM further concludes that, based on an evaluation of available data and corresponding performance (accuracy and reliability), the HET-CAM test method is **not** recommended to identify moderate and mild ocular irritants as defined by the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[12]

Evaluation as a Screening Test to Identify Ocular Corrosives and Severe Irritants
The available validation database for the HET-CAM test method has remained unchanged since the original ICCVAM evaluation (ICCVAM 2006b). Therefore, the original ICCVAM recommendation for the use of the HET-CAM test method to identify substances as ocular corrosives and severe irritants remains unchanged:

> The use of the HET-CAM test method for screening and identifying ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach, is **not** recommended (ICCVAM 2006e).

[12] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between mild and moderate ocular irritants.

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that the HET-CAM test method cannot identify substances from all hazard categories. The Panel also concluded (with one minority opinion) that the HET-CAM test method using the IS(A) analysis method cannot be used as a screening test to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) from all other hazard categories (i.e., EPA Category I, II, III; EU R41, R36; GHS Category 1, 2A, 2B) when results are to be used for EPA, EU or GHS hazard classifications. The Panel reached this conclusion because there were too few surfactants or oil/water emulsions in the mild to moderate irritant categories to have sufficient confidence in the ability of the test to distinguish them from substances not labeled as irritants.

4.1.2 ICCVAM Recommendations: HET-CAM Test Method Protocol

The updated ICCVAM-recommended HET-CAM test method protocol is included as an appendix to this report (**Appendix B**). The protocol has been modified from a generic description of the IS analysis method to include a more detailed IS(A) analysis method to be used for prospective studies. However, a description of the IS(B) method is included for retrospective analyses, where IS(B) analysis method data could be converted to fixed time points similar to those used for the IS(A) analysis method. All future studies intended to further characterize the usefulness and limitations of the HET-CAM test method should be conducted using this protocol.

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that the protocol is sufficiently detailed that it could be conducted reproducibly in other laboratories. However, they emphasized that the protocol should reflect any restrictions of the current applicability domain and it should also reflect details specific to the testing of certain types of substances. In addition, they noted that the protocol reflects the IS(A) analysis method, which was the subject of the ICCVAM draft recommendation, but additional data derived using the IS(B) analysis method could be collected and extrapolated to the IS(A) analysis method.

4.1.3 ICCVAM Recommendations: Future Studies for the HET-CAM Test Method

ICCVAM recommends that additional studies be conducted to further optimize the HET-CAM test method decision criteria that would be used to identify ocular corrosives and severe irritants (EPA Category I, EU R41, GHS Category 1), moderate and mild irritants (i.e., EPA Category II, III; EU R36; GHS Category 2A, 2B), and substances not labeled as irritants (i.e., EPA Category IV; EU Not Labeled; FHSA Not Labeled; GHS Not Classified) as defined by the EPA, EU, FHSA, and GHS classification systems. Such studies could potentially improve the usefulness of the HET-CAM test method for identifying these types of substances.

Additionally, in order to further optimize the protocol and more adequately characterize the usefulness of the HET-CAM test method for identifying substances not labeled as irritants:

- ICCVAM recommends that the applicability domain be expanded to include a broader range of chemical and product classes.
- ICCVAM encourages users to provide all data that are generated from future studies because they could be used to further characterize the usefulness and limitations of the HET-CAM test method for the identification of all ocular hazard categories.

Independent Peer Review Panel Conclusions and Recommendations
In order to further optimize the protocol and more adequately characterize the usefulness of the HET-CAM test method for identifying substances not labeled as irritants, the Panel recommended that additional data be collected on mild and moderate irritants and that the applicability domain be expanded to include a broader range of chemical and product classes.

They also noted that most of the single ingredients tested in the HET-CAM performed poorly, whereas formulations performed better. Hence, the effect of increasing the concentration of single ingredients on accuracy and sensitivity should be assessed in the HET-CAM to determine if there are test substance concentration limits for specific chemical classes.

The Panel did not support additional studies for using the HET-CAM test method to identify all categories of ocular irritants, given that it has been extensively evaluated and proven incapable for this task. However, as noted above, ICCVAM considers such studies valuable because they could improve the usefulness of the HET-CAM test method for identifying these types of substances. Furthermore, it is essential that the full range of ocular irritancy potential be tested in order to establish whether false negatives in the HET-CAM test method present a significant public health risk (e.g., EPA Category I substances predicted as Category IV in the HET-CAM test method).

4.1.4 ICCVAM Recommendations: Performance Standards for the HET-CAM Test Method

Based on the available data and associated performance described in the final ICCVAM BRD (**Appendix E**), ICCVAM recommends that the development of performance standards for the HET-CAM test method is not warranted at this time.

4.2 Validation Status of the HET-CAM Test Method

The following is a synopsis of the information in the final ICCVAM BRD (**Appendix E**), which reviews the available data and information for the HET-CAM test method. The ICCVAM BRD describes the current validation status of the HET-CAM test method, including what is known about its reliability and accuracy, the scope of the substances tested, and standardized protocols for the validation study.

4.2.1 Test Method Description

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

4.2.2 Validation Database

No new HET-CAM data were obtained since the ICCVAM evaluation of the HET-CAM test method for identifying ocular corrosives and severe irritants (ICCVAM 2006b). Therefore, the same database was used in the current evaluation. The database is composed of 260 substances representing a wide variety of chemical and product classes. It includes more than 50 commercial products or formulations. However, of the 260 substances, 167 could not be classified within a product class.

Analyses of each of the multiple HET-CAM protocols indicate that the IS(A) analysis method achieved the best performance when evaluating substances not labeled as irritants. The available IS(A) database includes a total of 63 test substances, 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 were necessary to calculate the appropriate EPA, EU, FHSA, and GHS ocular hazard classifications (EPA 2003; EU 2001; FHSA 2005; UN 2007) (**Appendix E**). These data

consist 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 days. Thus, some of the test substances for which there was only limited *in vivo* data could not be used to evaluate test method accuracy and reliability. In order to maximize the number of substances included in these analyses, "proportionality" criteria (i.e., FHSA-20% and FHSA-67%) were applied for the purpose of assigning an FHSA classification for test results that would require additional testing according to the FHSA sequential testing strategy (see **Section 2.2.2**).

4.2.3 Test Method Accuracy

Identification of All Ocular Hazard Categories

The ability of the HET-CAM test method to identify all categories of ocular irritation potential was evaluated for the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[13] As indicated in **Table 4-1**, overall correct classifications ranged from 38% (23/60) to 41% (24/59) depending on the classification system used when evaluating the entire database.

It is apparent from **Table 4-1** that the limited number of substances (n = 0–2) in the moderate irritant category (i.e., EPA Category II, EU R36, and GHS Category 2A) prevents an adequate evaluation of HET-CAM performance for this category. 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).

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, FHSA Not Labeled, GHS Not Classified) from all other ocular hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B) as defined by the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007) was also evaluated.

As indicated in **Table 4-2**, overall accuracy for the identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other categories ranged from 62% (36/58) to 80% (44/55) depending on the hazard classification system used. False positive and false negative rates ranged from 60% (9/15) to 69% (22/32) and 0% (0/26 or 0/36) to 9% (4/45 or 4/47), respectively. Among the four false negatives for the EPA and FHSA-20% systems, 100% (4/4, all oil/water emulsion cosmetic formulations) were EPA Category III substances based on a conjunctival redness score of 2 that required at least 3 days to resolve. For one of these substances, one of the six rabbits tested had a conjunctival redness score of 2 that required 14 days to resolve. Four of the remaining five rabbits in this study had conjunctival redness scores of 2 that resolved within three days. The fifth rabbit did not have this lesion.

[13] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an evaluation of the HET-CAM test method to identify all ocular hazard categories using the FHSA classification system is not possible.

Table 4-1 Evaluation of the Performance of the HET-CAM Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA, EU, and GHS Classification Systems[1]

	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
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)
GHS	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)
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)

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

[1] EPA classification system (EPA 2003); EU classification system (EU 2001); GHS classification system (UN 2007).

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

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

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

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

Table 4-2 Accuracy of the HET-CAM Test Method for Distinguishing Substances Not Labeled as Irritants[1] from All Other Irritant Classes

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
EPA	60	78	47/60	91	41/45	40	6/15	60	9/15	9	4/45
GHS	59	69	41/59	100	31/31	36	10/28	64	18/28	0	0/31
EU	58	62	36/58	100	26/26	31	10/32	69	22/32	0	0/26
FHSA-20%	63	78	49/63	91	43/47	38	6/16	63	10/16	9	4/47
FHSA-67%	55	80	44/55	97	38/39	38	6/16	63	10/16	3	1/39

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

[1] GHS classification system (UN 2007): Not Classified vs. Category 1/2A/2B.
EPA classification system (EPA 2003): Category IV vs. Category I/II/III.
EU classification system (EU 2001): Not Labeled vs. R41/R36.
FHSA classification system (FHSA 2005): Not Labeled vs. Irritant.

4.2.4 Test Method Reliability

Interlaboratory Reproducibility

Quantitative and qualitative evaluations of HET-CAM test method reliability have been conducted previously (ICCVAM 2006b). Because the database used for the current evaluation of the HET-CAM test method has not changed, the quantitative evaluation of test method reliability remains unchanged. However, additional qualitative analyses of interlaboratory reproducibility were conducted to evaluate the extent of agreement of HET-CAM hazard classifications among the five laboratories participating in the interlaboratory validation study (Hagino et al. 1999). As was done for the accuracy evaluation, these qualitative evaluations of reproducibility were based on (1) the use of the HET-CAM test method for identifying all ocular hazard categories according to the EPA, EU, and GHS systems, and (2) the use 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 irritant categories (i.e., EPA Category I, II, III; EU R41, R36; GHS Category 1, 2A, 2B). Given that the performance of the HET-CAM test method was similar for the EPA and FHSA classification systems, additional reliability analyses were not conducted for the FHSA classification system.

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 and severe irritants correctly classified by the HET-CAM test method based on all three classification systems. (There was 100% agreement for 63% [5/8] of the correctly identified EPA Category I substances and 100% agreement for 71% [5/7] of the correctly identified GHS Category 1 or EU R41 substances.) There was 100% agreement among the five laboratories for the one moderate irritant in the database (EPA Category II or EU R36; no GHS Category 2A substances were included), which was overpredicted. There 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 the HET-CAM test method. There was 100% agreement among the five laboratories for 86% (6/7) or 75% (3/4) of these substances for the EU and GHS classification systems, respectively. By comparison, for the two EPA Category IV substances tested, there was either 100% or 80% agreement among the five laboratories.

Using the second approach (i.e., identifying substances not labeled as irritants), 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 irritants according to the EPA classification system (i.e., Category I, II, or III). While neither of the EPA Category IV substances was correctly identified by the HET-CAM test method, there was 60% agreement among the five laboratories for 100% (2/2) of the EPA Category IV substances that were overpredicted by the HET-CAM test method.

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

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

As stated above, this review 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.

4.2.5 Animal Welfare Considerations

The HET-CAM test method has the potential to reduce and refine animal use in eye irritation testing. It would refine animal use by the *in vitro* identification of ocular corrosives/severe irritants, nonsevere irritants, or substances not labeled as irritants in a tiered-testing strategy.

5.0 The Isolated Chicken Eye Test Method

The ICE test method is an *in vitro* eye irritation test method using chicken eyes that are byproducts from processing plants. In the ICE test method, damage by the test substance is assessed by determining corneal swelling, opacity, and fluorescein retention. These endpoints are used collectively as an indicator of effects induced by the test substance in the eye of a treated rabbit.

ICCVAM previously evaluated the validation status of the ICE test method as an *in vitro* alternative to the Draize rabbit eye test to identify ocular corrosives/severe irritants (i.e., those that induce irreversible ocular damage; EPA Category I, EU R41, GHS Category 1) and determined that the reproducibility and accuracy was sufficient to support its use for this purpose for some types of substances (ICCVAM 2006e). U.S. agencies and international organizations (OECD 2009b) have adopted the ICE test method for this purpose. Following this initial evaluation, ICCVAM evaluated the validation status of the ICE test method as an *in vitro* alternative to the Draize rabbit eye test for identifying nonsevere ocular irritants (i.e., those that induce reversible ocular damage [EPA Category II and III, EU R36, GHS Category 2A and 2B]) and substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) according to the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007).

5.1 ICCVAM Recommendations

5.1.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

Evaluation as a Screening Test to Identify Substances Not Labeled as Irritants
ICCVAM concludes that the ICE test method is **not** recommended as a screening test to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B) when results are to be used specifically for hazard classification and labeling purposes under the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007).

Identification of Reversible Eye Irritation Hazard Categories
Based on an evaluation of available data and test method performance (accuracy and reliability), ICCVAM concludes that the ICE test method is **not** recommended to identify moderate and mild ocular irritants as defined by the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[14]

Evaluation as a Screening Test to Identify Ocular Corrosives and Severe Irritants
The available validation database for the ICE test method has not changed since the original ICCVAM evaluation (ICCVAM 2006c). Therefore, the original ICCVAM recommendation for the use of the ICE test method to identify substances as ocular corrosives and severe irritants remains unchanged:

> There are sufficient data to support the use of the ICE test method, in appropriate circumstances and with certain limitations, as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, UN GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach. In a tiered-testing strategy, when a positive result is obtained in an appropriately validated *in vitro* test, a test substance may be classified as an ocular hazard without testing in rabbits. A substance that tests negative in the *in vitro* ocular toxicity test would need to be tested in the *in vivo* ocular test to identify possible *in vitro* false negatives and to identify moderate and mild ocular irritants (ICCVAM 2006e).

[14] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between mild and moderate ocular irritants.

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that the available data and test method performance (accuracy and reliability) supported the ICCVAM draft recommendation that the ICE test method is not recommended to identify substances from all hazard categories as defined by the EPA, EU, and GHS classification systems. The Panel further concluded that the ICE test method is not recommended as a screening test to distinguish substances as not labeled as irritants from all other hazard categories as defined by the EPA, EU, and GHS classification systems.

5.1.2 ICCVAM Recommendations: ICE Test Method Protocol

For use of the ICE test method as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1), ICCVAM recommends using the updated ICCVAM ICE test method protocol that is included as an appendix to this report (**Appendix B**). In addition, all future studies intended to further characterize the usefulness and limitations of the ICE test method should be conducted using this protocol.

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that the protocol is sufficiently detailed but noted that the protocol could be improved by adding objective endpoints for corneal opacity and fluorescein staining.

5.1.3 ICCVAM Recommendations: Future Studies for the ICE Test Method

To further the use of this test method and to evaluate its use as a potential replacement for the Draize rabbit eye test or for the identification of mild and moderate ocular irritants (i.e., EPA Category II, III; EU R36; GHS Category 2A, 2B) and substances not labeled as irritants (i.e., EPA Category IV; EU Not Labeled; FHSA Not Labeled; GHS Not Classified), ICCVAM recommends additional studies be considered and undertaken.

- Additional optimization studies/evaluations should be conducted in an attempt to improve the correct classification of mild and moderate ocular irritants and substances not labeled as irritants. After optimization, additional studies to further assess the reliability and accuracy of the test method are recommended.
- ICCVAM recommends that a histopathological evaluation of the corneal tissue, using standardized procedures, be included when the ICE test method is conducted. Such data will allow for development of decision criteria and future assessments on the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.
- ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of the ICE test method for the identification of all ocular hazard categories.

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that additional optimization studies should be required to validate the test method for the identification of all ocular irritation hazard categories. The Panel also noted that the use of histopathology to evaluate corneal tissue might improve test method accuracy.

5.1.4 ICCVAM Recommendations: Performance Standards for the ICE Test Method

Based on the available data and associated performance described in the final ICCVAM BRD (**Appendix F**), ICCVAM recommends that the development of performance standards for the ICE test method is not warranted at this time.

5.2 Validation Status of the ICE Test Method

The following is a synopsis of the information in the final ICCVAM BRD (**Appendix F**), which reviews the available data and information for the ICE test method. The ICCVAM BRD describes the

current validation status of the ICE test method, including what is known about its reliability and accuracy, the scope of the substances tested, and standardized protocols for the validation study.

5.2.1 Test Method Description

The ICE test method is an *in vitro* model that provides short-term maintenance of the chicken eye. In the ICE test method, damage by the test substance is assessed by determination of corneal swelling, opacity, and fluorescein retention. While the latter two parameters involve a qualitative assessment, analysis of corneal swelling provides for a quantitative assessment. Each measurement is either converted into a quantitative score used to calculate an overall irritation index, or assigned a qualitative categorization that is used to assign an *in vitro* ocular irritation classification. Either of these outcomes can then be used to predict the *in vivo* ocular irritation potential of a test substance.

5.2.2 Validation Database

No new ICE data were obtained after the ICCVAM evaluation of the ICE test method for identifying ocular corrosives and severe irritants (ICCVAM 2006c). Therefore, the same database was used in the current evaluation. It is composed of 175 substances representing a wide variety of chemical and product classes. However, of the 175 substances, 85 (including formulations of unidentified composition) could not be assigned a specific chemical class.

Detailed *in vivo* data were necessary to calculate the appropriate EPA, EU, FHSA, and GHS ocular hazard classifications (EPA 2003; EU 2001; FHSA 2005; UN 2007) (**Appendix F**). These data consist 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 days. Thus, some of the test substances for which there was only limited *in vivo* data could not be used to evaluate test method accuracy and reliability. In order to maximize the number of substances included in these analyses, "proportionality" criteria (i.e., FHSA-20% and FHSA-67%) were applied for the purpose of assigning an FHSA classification for test results that would require additional testing according to the FHSA sequential testing strategy (see **Section 2.2.2**).

5.2.3 Test Method Accuracy

Identification of All Ocular Hazard Categories

The ability of the ICE test method to identify all categories of ocular irritation potential was evaluated for the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).[15] As indicated in **Table 5-1**, overall correct classifications ranged from 59% (83/141) to 77% (118/153) depending on the classification system used when evaluating the entire database.

[15] The FHSA ocular hazard category that is assigned based on results from the Draize rabbit eye test does not distinguish between ocular corrosives/severe irritants and less severe irritants. For this reason, an avaluation of the ICE test method to identify all ocular hazard categories using the FHSA classification system is not possible.

Table 5-1 Evaluation of the Performance of the ICE Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA, EU, and GHS Classification Systems[1]

	Overall Correct Classification	Severe[2]		Moderate[3]			Mild[4]			Not Labeled[5]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
EPA	62% (87/140)	48% (13/27)	52% (14/27)	31% (5/16)	50% (8/16)	19% (3/16)	29% (11/38)	53% (20/38)	18% (7/38)	22% (13/59)	78% (46/59)
EU	77% (118/153)	59% (19/32)	41% (13/32)	18% (5/28)	57% (16/28)	25% (7/28)	NA	NA	NA	11% (10/93)	89% (83/93)
GHS	59% (83/141)	52% (15/29)	48% (14/29)	36% (8/22)	36% (8/22)	28% (6/22)	18% (2/11)	73% (8/11)	9% (1/11)	34% (27/79)	66% (52/79)

Abbreviations: EPA = U.S. Environmental Protection Agency; EU = European Union; GHS = Globally Harmonized System; ICE = isolated chicken eye; NA = not applicable.

[1] EPA classification system (EPA 2003); GHS classification system (UN 2007); EU classification system (EU 2001).

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

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

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

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

Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories
The ability of the ICE test method to distinguish substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other ocular hazard categories (i.e., EPA Category I, II, III; EU R41, R36; FHSA Irritant; GHS Category 1, 2A, 2B) as defined by the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007) was also evaluated.

As indicated in **Table 5-2**, overall accuracy for the identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, FHSA Not Labeled, GHS Not Classified) from all other categories ranged from 78% (110/141) to 85% (130/153) depending on the hazard classification system used. The false negative rates ranged from 6% (4/62) to 22% (13/60) depending on the hazard classification system used. The lowest false negative rate (6%) was noted for the GHS system, followed by 9% (7/76) for the FHSA-67% system, 12% (10/84) for the FHSA-20% system, 14% (11/81) for the EPA system, and 22% (13/60) for the EU system. However, at least one of these false negatives is classified as an ocular corrosive and severe irritant based on Draize rabbit eye test data (n = 1 each for the EPA and GHS systems and n = 6 for the EU system).

Table 5-2 Accuracy of the ICE Test Method for Distinguishing Substances Not Labeled as Irritants[1] from All Other Irritant Classes

	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
EPA	140	83	116/140	86	70/81	78	46/59	22	13/59	14	11/81
GHS	141	78	110/141	94	58/62	66	52/79	34	27/79	6	4/62
EU	153	85	130/153	78	47/60	89	83/93	11	10/93	22	13/60
FHSA-20%	146	83	121/146	88	74/84	76	47/62	24	15/62	12	10/84
FHSA-67%	138	84	116/138	91	69/76	76	47/62	24	15/62	9	7/76

Abbreviations: EPA = U.S. Environmental Protection Agency; EU = European Union; FHSA = U.S. Federal Hazardous Substances Act; GHS = Globally Harmonized System; ICE = isolated chicken eye; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): Not Classified vs. Category 1/2A/2B.
EPA classification system (EPA 2003): Category IV vs. Category I/II/III.
EU classification system (EU 2001): Not Labeled vs. R41/R36.
FHSA classification system (FHSA 2005): Not Labeled vs. Irritant.

5.2.4 Test Method Reliability

Interlaboratory Reproducibility
Quantitative and qualitative evaluations of ICE test method reliability have been conducted previously (ICCVAM 2006c). However, additional qualitative analyses of interlaboratory reproducibility were conducted to evaluate the extent of agreement of ICE hazard classifications among the four laboratories participating in the interlaboratory validation study (Balls et al. 1995). As was done for the accuracy evaluation, these qualitative evaluations of reproducibility were based on (1) the use of the ICE test method for identifying all ocular hazard categories according to the EPA, EU, or GHS systems, and (2) the use of the ICE 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). Given that the performance of the ICE test method was similar for the EPA and FHSA classification systems, additional reliability analyses were not conducted for the FHSA classification system.

Using the first approach (i.e., identifying all ocular hazard categories), there was 100% agreement among the four laboratories for a majority of the Draize ocular corrosives and severe irritants for all three classification systems, whether the substances were correctly identified or underclassified by the ICE test method. (For example, for the EPA system, there was 100% agreement for 70% [7/10] of the correctly identified Category I substances.) There was 100% agreement among the four laboratories for at least 50% (3/6 to 3/5) of the correctly identified moderate ocular irritants (EPA Category II, GHS Category 2A, EU R36). For the mild ocular irritants (EPA Category III, GHS Category 2B), there was 100% agreement among the four laboratories for 0% (0/2) to 13% (1/8) of the correctly identified substances. The four laboratories had only 50% agreement for 50% (4/8 or 1/2) of these substances for the EPA and GHS classification systems.

Among the four laboratories, a majority 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 the ICE test method. The four laboratories had at least 75% agreement for all but two of these substances. For example, there was at least 75% agreement for 85% (11/13) of the GHS Not Labeled substances overclassified by the ICE test method. The four laboratories had at least 75% agreement for 76% (13/17) of the EU Not Labeled substances, whether they were correctly identified or overclassified by the ICE test method. For example, there was at least 75% agreement for 77% (7/9) of the EU Not Labeled substances that were correctly identified and 75% (6/8) of those overclassified by the ICE test method.

Using the second approach (i.e., distinguishing substances not labeled as irritants from all other ocular hazard categories), there was 100% agreement among the four laboratories for 61% (36/59) to 75% (44/59) of the substances included in the Balls et al. (1995) study.

There was 100% agreement among the four laboratories for 81% (38/47) of the substances correctly identified as irritants according to the EPA system (i.e., Category I, II, or III). While none of the EPA Category IV substances was correctly identified by the ICE test method, there was 75% agreement among the four laboratories for both of the Category IV substances that were overpredicted by the ICE test method.

The four laboratories had 100% agreement for 87% (33/38) of the substances correctly identified as irritants according to the GHS system (i.e., Category 1, 2A, or 2B). While only one of the GHS substances not labeled as irritants was correctly identified by the ICE test method (for which there was 75% agreement among the laboratories), there was at least 75% agreement among the four laboratories for 85% (11/13) of the GHS substances not labeled as irritants that were overpredicted by the ICE test method.

There was 100% agreement among the four laboratories for 85% (22/26) of the substances correctly identified as irritants according to the EU system (i.e., R36 or R41). The laboratories had at least 75% agreement for 77% (7/9) of the substances correctly identified as Not Labeled.

The final ICCVAM BRD (**Appendix F**) provides a comprehensive summary of the current validation status of the ICE test method, including what is known about its reliability and accuracy, and the scope of the substances tested. Raw data for the ICE test method will be maintained for future use, so that these performance statistics may be updated as additional information becomes available.

5.2.5 Animal Welfare Considerations

The ICE test method refines animal use. Because these animals are being humanely processed for nonlaboratory purposes, the testing procedure inflicts no additional pain or distress on animals. Substances that are identified as corrosive or severe irritants *in vitro* are excluded from *in vivo* testing.

The ICE test method can also reduce animal use. The test method utilizes animal species routinely raised as a food source in large numbers to replace the need for laboratory animals.

6.0 The Isolated Rabbit Eye Test Method

The IRE test method is an *in vitro* eye irritation test method using eyes from rabbits that have been euthanized for other research purposes or are byproducts from processing plants. In the IRE test method, the treated eye may be evaluated for corneal opacity, corneal swelling, fluorescein penetration, and effects on the corneal epithelium at various times over a four-hour observation period.

ICCVAM previously evaluated the validation status of the IRE test method as an *in vitro* alternative to the Draize rabbit eye test to identify ocular corrosives and severe irritants (i.e., those that induce irreversible ocular damage; EPA Category I, EU R41, GHS Category 1) and determined that the reproducibility and accuracy was **not** sufficient to support its use for this purpose (ICCVAM 2006e). In the current evaluation, ICCVAM evaluated the validation status of the IRE test method as an *in vitro* alternative to the Draize rabbit eye test for identifying nonsevere ocular irritants (i.e., those that induce reversible ocular damage [EPA Category II and III; EU R36; GHS Category 2A and 2B]) and substances not labeled as irritants (i.e., EPA Category IV; EU Not Labeled; FHSA Not Labeled; GHS Not Classified) according to the EPA, EU, FHSA, and GHS classification systems (EPA 2003; EU 2001; FHSA 2005; UN 2007).

6.1 ICCVAM Recommendations

6.1.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

Evaluation as a Screening Test to Identify All Ocular Hazard Categories
There are insufficient data using all four recommended IRE endpoints (i.e., corneal opacity, corneal swelling, fluorescein penetration, and effects on the corneal epithelium) in a single study to assess test method accuracy and reliability. Among the studies that included all four recommended IRE endpoints, decision criteria are focused on distinguishing ocular corrosives and severe irritants from all other ocular hazard categories (i.e., moderate and mild irritants and substances not labeled as irritants) and do not specify decision criteria for each ocular hazard category. For these reasons, an adequate evaluation of the IRE test method for its ability to identify all ocular hazard categories is not feasible at this time.

Evaluation as a Screening Test to Identify Ocular Corrosives and Severe Irritants
The available validation database for the IRE test method has not changed since the original ICCVAM evaluation (ICCVAM 2006d). Therefore, the original ICCVAM recommendation for the use of the IRE test method to identify substances as ocular corrosives and severe irritants remains unchanged:

> The use of the IRE test method for screening and identifying ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach, is **not** recommended (ICCVAM 2006e).

Independent Peer Review Panel Conclusions and Recommendations
The Panel concluded that there are insufficient data from all four recommended IRE test method endpoints (i.e., corneal opacity, fluorescein penetration, corneal swelling, and effects on the corneal epithelium) to evaluate the accuracy and reliability of the test method when all four endpoints are evaluated in a single study. Therefore, the Panel recommended that additional optimization and validation studies be conducted to further evaluate the relevance and reliability of the IRE test method, and in turn develop more definitive recommendations.

6.1.2 ICCVAM Recommendations: IRE Test Method Protocol

An ICCVAM-recommended test method protocol for the IRE test method that should be used for all future studies is included as an appendix to this report (**Appendix B**). The recommended protocol

remains unchanged from the previous ICCVAM evaluation (ICCVAM 2006e) and includes all four recommended IRE endpoints that should be measured: maximal corneal opacity (opacity × area), maximal corneal swelling, fluorescein penetration (intensity × area) and assessment of epithelial integrity (0.5, 1, 2, 3, and 4 hours).

Independent Peer Review Panel Conclusions and Recommendations
The Panel recommended that there should be rigid criteria specifying the handling and storage of the eyes. The Panel emphasized the need for control of the length of time between death and study initiation to account for any postmortem effects on the eye and criteria for appropriate inclusion/exclusion of ocular tissue. The Panel further emphasized the importance of criteria on test article administration/washout (e.g., viscous substances).

6.1.3 ICCVAM Recommendations: Future Studies for the IRE Test Method

To further the use of this test method and to evaluate the use of the IRE test method as a potential replacement for the Draize rabbit eye test or for the identification of all ocular hazard categories (i.e., EPA Category I, II, III, IV; GHS Category 1, 2A, 2B, Not Classified; FHSA Irritant, Not Labeled; EU R41, R36, Not Labeled), ICCVAM recommends additional studies be considered and undertaken.

- Additional evaluation studies should be conducted to increase the current IRE database and optimize the IRE test method decision criteria. Once these studies are conducted, ICCVAM recommends that additional validation studies be conducted to further evaluate the relevance and reliability of the IRE test method.
- ICCVAM recommends that a histopathological evaluation of the corneal tissue, using standardized procedures, be included when the IRE test method is conducted. Such data will allow for development of decision criteria and future assessments on the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.
- ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of the IRE test method for the identification of all ocular hazard categories.

Independent Peer Review Panel Conclusions and Recommendations
The Panel recommended a validation study to compare the utility of shipped rabbit eyes versus freshly collected rabbit eyes. Specifically, the Panel recommended inclusion of the study of shipping effects on ocular tissues into the planned validation study by GlaxoSmithKline and SafePharm. The Panel reiterated its concerns that there should be rigid criteria on the handling and storage of the eyes. In addition, the Panel recommended development of appropriate inclusion/exclusion criteria for eyes. Finally, the Panel recommended that criteria on test article administration/washout (e.g., viscous substances) should be developed.

6.1.4 ICCVAM Recommendations: Performance Standards for the IRE Test Method

Given that there are insufficient data using all four recommended IRE endpoints (i.e., corneal opacity, corneal swelling, fluorescein penetration, and effects on the corneal epithelium) in a single study to assess test method accuracy and reliability, ICCVAM recommends that the development of performance standards for the IRE test method is not warranted at this time.

6.2 Validation Status of the IRE Test Method

The IRE BRD (ICCVAM 2006d) describes the current validation status of the IRE test method, including what is known about its reliability and accuracy, the scope of the substances tested, and standardized protocols for the validation study.

6.2.1 Test Method Description

The IRE test method was developed by Burton and his colleagues at Unilever Research Laboratory, Colworth, United Kingdom, as an *in vitro* alternative to the Draize rabbit eye test for the assessment of eye irritation (Burton et al. 1981). In the IRE test method, liquid test substances are spread using a syringe, and solids are pulverized and applied as a powder over the corneas of enucleated rabbit eyes. The principal advantages of this test method are that the animals are euthanized prior to ocular irritancy testing (i.e., eyes from animals used for other toxicological purposes or from the food chain can be used) and testing is performed on the cornea, which is the part of the eye that is generally given the highest weight for scoring ocular irritancy in the Draize rabbit eye test. The effects of the test substance on the cornea of the isolated eye are measured quantitatively as an increase in thickness (swelling); subjectively as scores for corneal opacity, the area of corneal involvement, and fluorescein penetration; and descriptively as morphological changes to the corneal epithelium. However, the number of ocular parameters and the number of time points measured varies from study to study.

Two additional refinements of the IRE test method may be incorporated into the protocol or used *ad hoc* to supplement existing data. One is histopathological evaluation to confirm or identify the extent of irritancy at the cellular level, especially when the degree of irritancy falls between moderate and severe. Another is confocal microscopy to determine the extent and depth of ocular injury (Maurer et al. 2002). Many studies using the IRE test method evaluate single or multiple ocular endpoints at various times and then assign irritancy classifications to the substances tested (CEC 1991; Köeter and Prinsen 1985; Cooper et al. 2001; Jones et al. 2001), while others use mean data from one or more ocular endpoints assessed at various times after application of the test substance, typically 0.5 to 4 hours (Balls et al. 1995; Gettings et al. 1996). One protocol for the IRE test method was designed to specifically identify severe eye irritants (Guerriero et al. 2004). In this study, cut-off values for each ocular parameter tested were predetermined. If these cut-off values were achieved or exceeded in any single parameter over a period of 0.5 to 4 hours, including a significant change in the corneal epithelium, the test substance was classified as a severe eye irritant with potential to cause serious or irreversible damage to the human eye.

6.2.2 Validation Database

The available validation database for the IRE test method has remained unchanged since the original ICCVAM evaluation (ICCVAM 2006d). A total of 149 substances were evaluated in five studies, of which 25 were commercial products or formulations (ICCVAM 2006d). The chemical classes tested included but were not limited to alcohols, amides, amines, carboxylic acids, esters, ethers, formulations, heterocyclic, ketones, onium compounds, and sulfur compounds. The commercial products or formulations tested were skin cleansers, soaps, shampoos, conditioners, surfactants, and solvents.

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 days were necessary to calculate the appropriate EPA, EU, and GHS ocular hazard classifications (EPA 2003; EU 2001; UN 2007) (ICCVAM 2006d). Thus, some of the test substances for which there were only limited *in vivo* data could not be used to evaluate test method accuracy and reliability.

6.2.3 Test Method Accuracy

There are insufficient data using all four recommended IRE endpoints (i.e., corneal opacity, corneal swelling, fluorescein penetration, and effects on the corneal epithelium) in a single study to assess test method accuracy and reliability. Among the studies that included all four recommended IRE endpoints, decision criteria are focused on distinguishing ocular corrosives and severe irritants from all other ocular hazard categories (i.e., moderate and mild irritants and substances not labeled as irritants) and do not specify decision criteria for each ocular hazard category. For these reasons, an

adequate evaluation of the IRE test method for its ability to identify all ocular hazard categories is not feasible at this time.

6.2.4 Test Method Reliability

Due to the lack of quantitative IRE test method data for replicate experiments within an individual laboratory, an evaluation of the intralaboratory repeatability and reproducibility of the IRE test method could not be conducted. However, multilaboratory qualitative and quantitative IRE test method data were available for a collaborative study by the CEC (1991) and a validation study conducted by Balls et al. (1995). Three laboratories participated in the CEC (1991) collaborative study and four laboratories participated in the Balls et al. (1995) validation effort. In the CEC (1991) study, each substance tested was assigned a EU classification (R41, R36, or Not Labeled) based on *in vivo* rabbit eye test results. However, due to the lack of individual Draize rabbit eye test scores, a reliability assessment for the CEC (1991) study using the EPA (EPA 2003) or GHS (UN 2007) classification systems was not possible. The Balls et al. (1995) data were used for an evaluation of the interlaboratory reproducibility of the IRE test method according to the EPA, EU, and GHS classification systems (EPA 2003; EU 2001; UN 2007).

6.2.5 Animal Welfare Considerations

The IRE test method reduces animal use by obtaining eyes from rabbits raised for food or rabbits sacrificed after use in other laboratory procedures that do not adversely affect the eye. The IRE test method is a refinement of the *in vivo* rabbit eye test in that the animals are sacrificed prior to application of the test substance and, therefore, the animals do not experience pain and suffering when an ocular irritant is directly applied to the eye. Furthermore, because the IRE test method was adapted from the Draize rabbit eye test specifically to reduce the need for live animals for ocular irritation testing, pain and suffering of the animals is eliminated and the overall number of animals needed for ocular toxicity screening is reduced.

Although rabbits are required as a source of corneas for the IRE test method, only rabbits sacrificed for food or used for other laboratory purposes are used as eye donors (i.e., no live animals are specifically sacrificed for use in this test method).

7.0 ICCVAM Consideration of Public, SACATM, and ICATM Comments

The ICCVAM evaluation process incorporates a high level of transparency. This process is designed to provide numerous opportunities for stakeholder involvement, including submitting written public comments and providing oral comments at ICCVAM independent peer review panel meetings and SACATM meetings. **Table 7-1** lists the nine opportunities for public comments that were provided during the ICCVAM evaluation of the validation status of alternative ocular safety testing methods and approaches. The number of public comments received in response to each of the opportunities is also indicated. Thirty-seven comments were submitted. Comments received in response to or related to *Federal Register* notices (**Appendix H**) are also available on the NICEATM-ICCVAM website.[16] The following sections, delineated by *Federal Register* notice, briefly discuss the public comments received.

Table 7-1 Opportunities for Public Comments

Opportunities for Public Comments	Date	Number of Public Comments Received
70 FR 13512: Request for Data on Non-Animal Methods and Approaches for Determining Skin and Eye Irritation Potential of Antimicrobial Cleaning Product Formulations; Request for Nominations for an Independent Expert Panel	March 21, 2005	0
72 FR 26396: Request for Data on the Use of Topical Anesthetics and Systemic Analgesics for *In Vivo* Eye Irritation Testing	May 9, 2007	1
72 FR 31582: Request for Ocular Irritancy Test Data From Human, Rabbit, and *In Vitro* Studies Using Standardized Testing Methods	June 7, 2007	0
73 FR 18535: Non-Animal Methods and Approach for Evaluating Eye Irritation Potential for Antimicrobial Cleaning Products (AMCPs): Request for Nominations for an Independent Expert Panel and Submission of Relevant Data	April 4, 2008	12
74 FR 14556: Announcement of an Independent Scientific Peer Review Panel on Alternative Ocular Safety Testing Methods; Availability of Draft Background Review Documents (BRD); Request for Comments	March 31, 2009	8
74 FR 19562: Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)	April 29, 2009	2
Independent Scientific Peer Review Panel Meeting: Alternative Ocular Safety Testing Methods	May 19–21, 2009	12
SACATM Meeting, Arlington Hilton, Arlington, VA	June 25–26, 2009	2
74 FR 33444: Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of Alternative Ocular Safety Testing Methods and Approaches; Notice of Availability and Request for Public Comments	July 13, 2009	0

[16] Available at http://ntp-apps.niehs.nih.gov/iccvambp/searchPubCom.cfm

7.1 Public Comments in Response to 70 FR 13512 (March 21, 2005) Request for Data on Non-Animal Methods and Approaches for Determining Skin and Eye Irritation Potential of Antimicrobial Cleaning Product Formulations; Request for Nominations for an Independent Expert Panel

NICEATM requested (1) submission of data that would assist in evaluating the validation status of non-animal methods and approaches used for determining the skin and eye irritation potential of AMCP formulations to meet regulatory hazard classification and labeling purposes and (2) nominations of expert scientists to serve as members of an independent peer review panel.

No data or nominations were received in response to this *Federal Register* notice.

7.2 Public Comments in Response to 72 FR 26396 (May 9, 2007) Request for Data on the Use of Topical Anesthetics and Systemic Analgesics for *In Vivo* Eye Irritation Testing

NICEATM requested submission of (1) data and information on the use of topical anesthetics and systemic analgesics for alleviating pain and distress in rabbits during eye irritation testing and (2) information about other procedures and strategies that may reduce or eliminate pain and distress associated with *in vivo* eye irritation methods.

NICEATM received 1 comment in response to this *Federal Register* notice. This comment was not relevant to *in vitro* ocular safety test methods.

7.3 Public Comments in Response to 72 FR 31582 (June 7, 2007) Request for Ocular Irritancy Test Data From Human, Rabbit, and *In Vitro* Studies Using Standardized Testing Methods

NICEATM requested data on substances tested for ocular irritancy in humans, rabbits, and/or *in vitro* to be used to:

- Review the state of the science in regard to the availability of accurate and reliable *in vitro* test methods for assessing the range of potential ocular irritation activity, including whether ocular damage is reversible or not
- Expand NICEATM's high-quality ocular toxicity database. *In vitro* test methods for which data are sought include but are not limited to (1) the bovine corneal opacity and permeability test, (2) the isolated rabbit eye test, (3) the isolated chicken eye test, and (4) the hen's egg test–chorioallantoic membrane.

No data or information were received in response to this *Federal Register* notice.

7.4 Public Comments in Response to 73 FR 18535 (April 4, 2008) Non-Animal Methods and Approach for Evaluating Eye Irritation Potential for Antimicrobial Cleaning Products (AMCPs): Request for Nominations for an Independent Expert Panel and Submission of Relevant Data

NICEATM requested the following:

- Nominations of expert scientists to serve as members of an independent peer review panel
- Submission of relevant data and information on AMCPs or related substances obtained from (1) human testing or experience, including reports from accidental exposures, and (2) rabbit testing using the standard eye test or the LVET

- *In vitro* ocular safety test methods such as the bovine corneal opacity and permeability test method, the Cytosensor Microphysiometer test method, and the EpiOcular test method, including data supporting the accuracy and reproducibility of these methods

In response to this *Federal Register* notice, NICEATM received 12 comments, including nominations of 20 potential panelists. The nominees were included in the database of experts from which the Panel was selected. No additional data were received.

7.5 Public Comments in Response to 74 FR 14556 (March 31, 2009) Announcement of an Independent Scientific Peer Review Panel on Alternative Ocular Safety Testing Methods; Availability of Draft Background Review Documents (BRD); Request for Comments

NICEATM requested public comments on the draft BRDs, SRDs, and draft ICCVAM test method recommendations that were provided to an independent scientific peer review panel meeting (May 19–21, 2009). These documents summarized the current validation status of several test methods and testing strategies for identifying potential ocular irritants. The test methods and testing strategies included the following:

- A testing strategy that proposes the use of three *in vitro* test methods to assess the eye irritation potential of AMCPs
- Four *in vitro* test methods for identifying moderate (EPA Category II, UN Globally Harmonized System of Classification and Labelling of Chemicals [GHS] Category 2A) and mild (EPA Category III, GHS Category 2B) ocular irritants and substances not classified as ocular irritants (EPA Category IV, GHS Not Classified)
- The *in vivo* LVET
- A proposal for the routine use of topical anesthetics, systemic analgesics, and earlier humane endpoints to avoid and minimize pain and distress during *in vivo* ocular irritation testing

NICEATM received 20 comments in response to this *Federal Register* notice. Eight written comments were received before the Panel meeting, and 12 oral comments were provided at the Panel meeting. Of these comments, 10 were relevant to *in vitro* ocular safety test methods.

Public Responses (written)

HET-CAM—Two written comments were relevant to the HET-CAM test method.

Comment:
One commenter emphasized the importance of establishing one specific protocol and specific endpoints to be used for the HET-CAM test method. Based on a database of 145 substances tested in both the HET-CAM and Draize test methods, the commenter reported that the HET-CAM test method was not useful to identify water-soluble substances as severe irritants in a tiered-testing strategy. However, the HET-CAM might be applicable for excluding severe ocular irritants among water-insoluble substances.

ICCVAM Response:
ICCVAM recommends that additional studies should be conducted to further optimize the HET-CAM prediction models and the decision criteria that would be used to identify ocular corrosives and severe irritants (EPA Category I, EU R41, GHS Category 1). Such studies could potentially improve the usefulness of the HET-CAM test method for identifying these types of substances.

Comment:
Another commenter made the following comments on the HET-CAM draft BRD: (1) the terminology describing the endpoints used in the HET-CAM test method needs to be clarified, (2) the HET-CAM validation database appears to have some inconsistencies with regard to chemical class, (3) the HET-CAM validation database needs to be evaluated by solubility, and (4) the method for counting the days of embryonic development needs to be clarified.

ICCVAM Response:
The text was clarified and supporting references were provided in the final BRD (**Appendix E**) for the endpoints used in the HET-CAM test method. With regard to chemical class, the classifications in the BRD are based on the National Library of Medicine MeSH chemical classification system. Additional details for counting the days of embryonic development were provided in the final BRD.

ICE— One written comment was relevant to the ICE test method.

Comment:
The main point provided by the commenter is that the selection criteria set forth by ICCVAM for classification purposes of substances tested with the ICE test method are inappropriate. The commenter states that additional data from studies that were terminated earlier than 21 days after treatment or compounds lacking an *in vivo* eye irritation study because of proven *in vivo* skin corrosivity should be considered. The commenter also addressed the variability of the Draize rabbit eye test and expressed concern that the Draize rabbit eye test (i.e., OECD TG 405) has no standardized exposure regimen.

ICCVAM Response:
The performance of the ICE test method was reevaluated after including skin corrosivity test results (n = 8) and corrections to the classification of specific test substances, where appropriate based on the additional data and information provided by the commenter. These changes are reflected in the ICE final BRD (**Appendix F**). However, the addition and modification of these test results did not significantly impact the performance of the ICE test method or the conclusions and recommendations of the Panel or ICCVAM.

HCE— One written comment was specific to the human corneal epithelial cell (HCE) model.

Comment:
The commenter provided comments on an ATLA journal article (Eskes et al. 2005) that described the human corneal epithelial cell (HCE) model and a list of key references omitted from that review.

ICCVAM Response:
Although the HCE model was not part of the current evaluation, ICCVAM welcomes comments on alternative *in vitro* test methods at any time. ICCVAM encourages the submission of data for the HCE model for future evaluation of its validation status.

Public Responses, Oral
Twelve oral public comments were provided at the Panel meeting (May 19-21, 2009).

HET-CAM— One commenter remarked specifically on the HET-CAM test method.

Comment:
One commenter indicated that the false negatives using the EPA classification system, which are Not Classified using the GHS classification system, result because the EPA classification system categorizes substances based upon the most severe category observed among the test rabbits (i.e., hazard classification is not based on the majority classification among rabbits tested). The commenter noted that because the types of formulations regulated by EPA are not present in the database, the EPA classification system should not be given too much weight.

ICCVAM Response:
Until the GHS classification system is formally adopted, ICCVAM will continue to consider all relevant hazard classification systems (i.e., EPA, EU, and GHS) when evaluating the usefulness and limitations of an *in vitro* test method.

ICE— One commenter remarked specifically on the ICE test method.

Comment:
One commenter indicated that the variability of the ICE test method was similar to that of the Draize rabbit eye test. The commenter stated that the ICE test method should not be held to a higher standard than the Draize rabbit eye test and noted that the concordance among laboratories was reasonable.

ICCVAM Response:
All *in vitro* test methods are evaluated according to the *ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods* (ICCVAM 2003), including a comparison to the currently accepted regulatory test method as a reference.

BCOP— Three commenters remarked specifically on the BCOP test method.

Comment:
One commenter indicated that the performance of the BCOP test method was unlikely to improve based on the lack of reproducibility with the Draize rabbit eye test in the mild and moderate categories. The commenter stated that results from Weil and Scala (1971) show that the extremes (i.e., corrosives/severe irritants and substances not labeled as irritants) are reproducible, but the mild and moderate levels of ocular irritation are highly variable. The commenter referenced the AMCP BRD that includes an analysis of the impact on the ocular hazard category when the results of a six-rabbit Draize test are randomly sampled for a three-rabbit test.

ICCVAM Response:
The Draize rabbit eye test (Draize et al. 1944) has a long history of demonstrated protection of public health and therefore, U.S. and international regulatory agencies currently use this test to identify potential ocular hazards. Alternatives are accepted only when they demonstrate the ability to provide equal or better protection than the reference test method. Given the uncertainty of the results associated with the BCOP test method for substances in the mild/moderate irritancy range, the BCOP test method cannot be considered a complete replacement at this time.

Comment:
A second commenter stated that damaged eyes are quickly removed and excluded from the BCOP test method and that Gautheron et al. (1992) used both fresh eyes and eyes maintained at 4ºC and found no differences in results. The commenter also asked the Panel to reconsider the use of a histopathology evaluation in the BCOP test method.

ICCVAM Response:
ICCVAM previously evaluated and recommended the BCOP test method for the identification of corrosive/severe ocular irritants (ICCVAM 2006e). In the current evaluation, ICCVAM, along with the Panel, recommends the BCOP test method for the identification of substances not labeled as irritants. Furthermore, the final ICCVAM recommendations state that a histopathological evaluation of the corneal tissue, using standardized procedures, should be included when the BCOP test method is conducted. Such data will allow for development of decision criteria and future assessments on the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.

Comment:
A third commenter discussed the “top-down” (i.e., screening for corrosives/severe irritants) and “bottom-up” (i.e., screening for substances not labeled as irritants) approaches using the ICE and

BCOP test methods. The commenter stated that ECVAM is developing a paper to recommend the use of these proposed testing strategies for both ICE and BCOP, where substances could be tested in the BCOP or ICE test methods in order to identify corrosives/severe irritants or substances not labeled as irritants without using an animal test.

ICCVAM Response:
ICCVAM currently recommends the ICE and BCOP test methods for use in a tiered-testing strategy, where positive substances can be classified as ocular corrosives and severe irritants without the need for animal testing (ICCVAM 2006e). However, identification of nonsevere ocular irritants and substances not labeled as irritants would require another test that has been demonstrated as scientifically validated for identifying such substances.

Comment:
One commenter questioned the need for performance standards for the CM test method, given that the Panel did not recommend performance standards for the BCOP and ICE test methods.

ICCVAM Response:
The final ICCAM recommendations state that the development of performance standards for the CM test method is not warranted at this time.

7.6 Public Comments in Response to 74 FR 19562 (April 29, 2009) Meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)

NICEATM announced the SACATM meeting (June 25–26, 2009) and requested written and public oral comments on the agenda topics.

NICEATM received four comments. Two written comments were received before the meeting, and two oral comments were provided at the SACATM meeting. None of these comments were relevant to *in vitro* ocular safety test methods.

SACATM Response:
In general, SACATM was pleased with the Panel report. One SACATM member expressed the need for harmonization in the assessment of performance standards. Another SACATM member said the focus should be on the GHS system because it will ultimately be adopted. Another SACATM member expressed concern regarding the availability of the Cytosensor Microphysiometer.

7.7 Public Comments in Response to 74 FR 33444 (July 13, 2009) Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of Alternative Ocular Safety Testing Methods and Approaches; Notice of Availability and Request for Public Comments

NICEATM requested submission of written public comments on the independent scientific peer review panel report.

No public comments were received.

7.8 Comments Received from ICATM Validation Organizations

In accordance with the International Cooperation on Alternative Test Methods (ICATM), each participating organization (i.e., ECVAM, Health Canada, and JaCVAM) was given the opportunity to comment prior to finalizing these recommendations. All ICATM partners agreed with these recommendations with one exception noted from ECVAM regarding the usefulness and limitations of the BCOP test method (see **Appendix H4**). ECVAM agreed with ICCVAM that the BCOP test method should not be recommended for the identification of substances not labeled as irritants under the EPA and FHSA classification system. However, they did not agree with ICCVAM's concerns

about the underprediction that would occur with the use of the BCOP test method under the GHS classification system compared to the EPA and FHSA systems currently used in the United States. Rather, ECVAM expressed the view that the predictivity of a test method should be calculated independently for each classification, without regard to the nature, severity, and duration of eye injuries that were the basis for ICCVAM's concerns, and that serve as the basis for classification as eye hazards using current U.S. classification criteria. Thus, ECVAM considered the calculated predictivity of the BCOP test method for available test data (0% [0/97] false negatives) to support its use for identifying substances not labeled as irritants when compared strictly to the eye hazard criteria in the current GHS classification system. While ECVAM also noted that their recommendations would be in line with the ICCVAM Peer Review Panel, as stated in **Section 2.1.1**, the Peer Panel deliberations preceded the NICEATM evaluation of the GHS classification system that indicates an estimated 30% or more of substances requiring labeling for eye irritation hazard according to current U.S. hazard classification requirements will not be labeled as eye irritation hazards by the current GHS criteria. ECVAM disagrees with the ICCVAM concern that, "*the nature, severity, and duration of these eye injuries suggest the potential to cause human injury,"* because they state that there is no empirical evidence to substantiate that there should in fact be such a concern. While ICCVAM was not able to find any human accidental exposure data for the chemicals that will no longer be labeled as eye hazards, there also are no human exposure or test data to suggest that these chemicals, which produced significant eye injuries in rabbits, will not produce significant injuries to human eyes.

NICEATM has searched numerous databases (i.e., OSHA, CPSC, EPA) for human eye injury data for these chemicals, but no relevant data have been identified. U.S. Federal law requires agencies to determine that new test methods recommended by ICCVAM generate data that are at least equivalent to data generated by current test methods required or recommended by each agency for hazard identification purposes. Therefore, until the issues associated with the GHS system as outlined above are further discussed, ICCVAM is deferring final recommendations on the usefulness and limitations of using BCOP as a screening test to identify substances not labeled as irritants according to the GHS classification system. At such time that relevant human testing or exposure data becomes available, ICCVAM will revisit this issue. In the meantime, the Scientific Advisory Committee on Alternative Toxicological Methods will be asked to comment at their June 2010 meeting about the extent that reversible chemically induced eye injuries in rabbits might indicate the potential for injury to humans. Minutes of this meeting will be available on the NTP website at: http://ntp.niehs.nih.gov/index.cfm?objectid=720165EC-BDB7-CEBA-F517D1DEE4D7D129.

8.0 References

Balls M, Botham PA, Bruner LH, Spielmann H. 1995. The EC/HO international validation study on alternatives to the Draize eye irritation test. Toxicol In Vitro 9:871–929.

Brantom PG, Bruner LH, Chamberlain M, DeSilva O, Dupuis J, Earl LK, et al. 1997. A summary report of the COLIPA International Validation Study on Alternatives to the Draize Rabbit Eye Irritation Test. Toxicol In Vitro 11:141–179.

Burton AB, York M, Lawrence RS. 1981. The *in vitro* assessment of severe eye irritants. Food Cosmet Toxicol 19:471–480.

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. Brussels:Commission of the European Communities.

Cooper KJ, Earl LK, Harbell J, Raabe H. 2001. Prediction of ocular irritancy of prototype shampoo formulations by the isolated rabbit eye (IRE) test and bovine corneal opacity and permeability (BCOP) assay. Toxicol In Vitro 15:95–103.

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

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

ECETOC. 1998. Eye Irritation – Reference Chemicals Data Bank. Technical Report No. 48(2). Brussels:European Centre for Ecotoxicology and Toxicology of Chemicals.

EPA. 2003. Label Review Manual. 3rd ed. EPA737-B-96-001. Washington, DC:U.S. Environmental Protection Agency.

ESAC. 2009. Statement on the scientific validity of cytotoxicity/cell-function based in vitro assays for eye irritation testing. Ispra, Italy:European Commission. Available: http://ecvam.jrc.it/index.htm.

Eskes et al. 2005. Eye irritation. Altern Lab Anim 33(Suppl. 1):47-81.

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.

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

Gautheron P, Dukic M, Alix D, Sina JF. 1992. Bovine corneal opacity and permeability test: An in vitro assay of ocular irritancy. Fundam Appl Toxicol 18:442-449.

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 SD, Lordo RA, Demetrulias J, Feder PI, Hintze KL. 1996. Comparison of low-volume, Draize and in vitro eye irritation test data. I. Hydroalcoholic formulations. Food Chem Toxicol 34:737–749.

Guerriero FJ, Seaman CW, Olson MJ, Guest R, Whittingham A. 2004. Retrospective assessment of the rabbit enucleated eye test (REET) as a screen to refine worker safety studies [Abstract No. 1282]. Toxicol Sci (supplement) 78(1-S).

Hagino S, Kinoshita S, Tani N, Nakamura T, Ono N, Konishi K, Iimura H, Kojima H, Ohno Y. 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, NC: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: Bovine Corneal Opacity and Permeability Test Method. NIH Publication No. 06-4512. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

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

ICCVAM. 2006c. Background Review Document: Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Isolated Chicken Eye Test Method. NIH Publication No. 06-4513. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

ICCVAM. 2006d. Background Review Document: Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Isolated Rabbit Eye Test Method. NIH Publication No. 06-4514. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

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

Jones PA, Budynsky E, Cooper KJ, Decker D, Griffiths HA, Fentem JH. 2001. Comparative evaluation of five *in vitro* tests for assessing the eye irritation potential of hair-care products. ATLA 29:669–692.

Koëter BWM, Prinsen MK. 1985. Comparison of *in vivo* and *in vitro* eye irritancy test systems: a study with 34 substances. In: Alternative Methods in Toxicology (Goldberg AM, ed). Vol 3. New York:Mary Ann Liebert, 571–579.

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

Maurer JK, Parker RD, Jester JV. 2002. Extent of corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. Reg Tox Pharmacol 36:106-117.

OECD. 1999. Detailed Review Document Number 14. Classification systems for eye irritation/corrosion in OECD member countries. ENV/JM/MONO(99)4. In: OECD Series on Testing and Assessment. Available: http://www.olis.oecd.org/olis/1999doc.nsf/LinkTo/env-jm-mono%2899%294.

OECD. 2009a. Test Guideline 437. Bovine corneal opacity and permeability test method for identifying ocular corrosives and severe irritants, adopted September 7, 2009. In: OECD Guidelines for Testing of Chemicals. Paris:Organisation for Economic Co-operation and Development.

OECD. 2009b. Test Guideline 438. Isolated chicken eye test method for identifying ocular corrosives and severe irritants, adopted September 7, 2009. In: OECD Guidelines for Testing of Chemicals. Paris:Organisation for Economic Co-operation and Development.

Southee JA. 1998. Evaluation of the Prevalidation Process. Part 2, Final Report. Vol. 2. The Bovine Corneal Opacity and Permeability (BCOP) Assay. European Community Contract No. 11279-95-10F 1ED ISP GB.

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

Weil CS, Scala RA. 1971. Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. Toxicol Appl Pharmacol.19:276-360.

This page intentionally left blank