Draft Summary Review Document

Strategy for U.S. Environmental Protection Agency Ocular Hazard Classification and Labeling of Antimicrobial Cleaning Products Using *In Vitro* Alternative Test Methods

Interagency Coordinating Committee on the Validation of Alternative Methods

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

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

April 1, 2009

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

[This Page Intentionally Left Blank]

Table of Contents

List of Tables

158 # List of Figures

163

164

List of Abbreviations and Acronyms

%CV	Percent coefficient of variation
AMCP	Antimicrobial cleaning product
ATWG	Alternative Testing Working Group
BCOP	Bovine corneal opacity and permeability test method
BRD	Background review document
CASRN	Chemical Abstracts Service Registry number
CM	Cytosensor microphyiometer test method
Colipa	European Cosmetic, Toiletry and Perfumery Association
Conc.	Concentration tested
CPSC	U.S. Consumer Product Safety Commission
CTFA	Cosmetic, Toiletry and Fragrance Association
CV	Coefficients of variation
ECVAM	European Centre for the Validation of Alternative Methods
EO	EpiOcular™ test method
EPA	U.S. Environmental Protection Agency
ESAC	European Centre for the Validation of Alternative Methods Scientific Advisory Committee
ET_{50}	Time needed to reduce cell viability by 50%
FDA	U.S. Food and Drug Administration
FR	*Federal Register*
GHS	United Nations Globally Harmonized System of Classification and Labelling of Chemicals
GLP	Good Laboratory Practice
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
IIVS	Institute for In Vitro Sciences
ILS	Integrated Laboratory Systems, Inc.
ISO	International Organization for Standardization
IVIS	*In vitro* irritancy score
JaCVAM	Japanese Center for the Validation of Alternative Methods
LVET	Low volume eye test
MAS	Maximum average score
MRD_{50}	Estimated concentration of a test substance needed to reduce the basal metabolic rate of L929 cells by 50%
NA	Not applicable
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

NIEHS	National Institute of Environmental Health Sciences
NTP	National Toxicology Program
OECD	Organisation for Economic Co-operation and Development
OPP	Office of Pesticide Programs
OPPTS	Office of Prevention, Pesticides and Toxic Substances
OTWG	Ocular Toxicity Working Group
PPDC	Pesticide Product Dialog Committee
REACH	Registration, Evaluation. Authorisation and Restriction of Chemicals (Regulation [EC] No 1907/2006)
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
SD	Standard deviation
SM	Silicon microphysiometer
TG	Test guideline
TNO	TNO Nutrition and Food Research Institute (Netherlands)
U.K.	United Kingdom
U.N.	United Nations
U.S.	United States
w/v	Weight-to-volume ratio

[This Page Intentionally Left Blank]

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

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

Consumer Product Safety Commission
• Marilyn L. Wind, Ph.D. (Chair)
◊ Kristina Hatlelid, Ph.D.
Joanna Matheson, Ph.D.

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

Department of Defense
• Robert E. Foster, Ph.D.
◊ Patty Decot
Peter J. Schultheiss, D.V.M., D.A.C.L.A.M.
Harry Salem, Ph.D.

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

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

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

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

• Principal agency representative
◊ Alternate principal agency representative

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

National Cancer Institute
• T. Kevin Howcroft, Ph.D.
◊ Chand Khanna, DVM, Ph.D.

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

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

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

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

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

[This Page Intentionally Left Blank]

Acknowledgements

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

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

Department of Defense
Harry Salem, Ph.D.

Department of Transportation
Steve Hwang, Ph.D.

U.S. Environmental Protection Agency
Office of Pesticide Programs
Meta Bonner, Ph.D.
Jonathan Chen, Ph.D.
Masih Hashim, D.V.M., Ph.D.
Karen Hicks
Marianne Lewis
Deborah McCall
Timothy McMahon, Ph.D.
Mark Perry, Ph.D.
John Redden, Ph.D.
Amy Rispin, Ph.D.
Jenny Tao, Ph.D.
Office of Research and Development
Andrew Geller, Ph.D.
Office of Science Coordination and Policy
Karen Hamernik, Ph.D. (OTWG Co-Chair)

U.S. Food and Drug Administration
Center for Drug Evaluation and Research
Paul C. Brown, Ph.D.
Abigail Jacobs, Ph.D.
Jill Merrill, Ph.D. (OTWG Co-Chair)
Center for Food Science and Nutrition
Robert Bronaugh, Ph.D.
Donnie Lowther
Office of Science and Health Coordination
Suzanne Fitzpatrick, Ph.D., D.A.B.T.

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

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

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

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

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

National Institute of Environmental Health Sciences

William Stokes, D.V.M., D.A.C.L.A.M.
Director; Project Officer

Deborah McCarley
Special Assistant; Assistant Project Officer

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

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

Jonathan Hamm, Ph.D.
Senior Toxicologist

Nelson Johnson
Senior Project Coordinator/Technical Writer

Elizabeth Lipscomb, Ph.D.
Staff Toxicologist

Linda Litchfield
Meeting Coordinator/Admin. Asst.

Greg Moyer, M.B.A.
Project Manager

Catherine Sprankle
Senior Communications Specialist

James Truax
Senior Project Coordinator/Technical Writer

Other Acknowledgements

ICCVAM and NICEATM gratefully acknowledge the following individuals and institutions that submitted data to NICEATM used for the evaluation of the Antimicrobial Cleaning Products Testing Strategy.

Clorox
Pleasanton, CA

Colgate-Palmolive Company
Piscataway, NJ

The Dial Corporation
Scottsdale, AZ

EcoLabs
St. Paul, MN

Institute for In Vitro Sciences
Rodger Curren, Greg Mun, and Hans Raabe
Gaithersburg, MD

MatTek Corporation
Patrick Hayden, Ph.D.
Ashland, MA

JohnsonDiversey Inc.
Sturdivant, WI

SC Johnson & Son, Inc.
Nicole Cuellar
Racine, WI

The Procter & Gamble Company
Len Sauers, Ph.D.
Cincinnati, OH

[This Page Intentionally Left Blank]

Preface

Commercial and household cleaning products require labeling to indicate if they are hazardous to the consumer and have the potential to cause injuries during handling or use, including possible ingestion by children. The Consumer Product Safety Commission typically regulates these products under the Federal Hazardous Substances Act (15 U.S.C. 1261 and 16 CFR 1500) and the Poison Prevention Packaging Act (16 CFR 1700). However, under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (7 U.S.C. 136-136y, 40 CFR 161), inclusion of an antimicrobial claim in such cleaning products necessitates their registration as antimicrobial pesticides with the U.S. Environmental Protection Agency (EPA) Office of Pesticide Products (OPP). Accordingly, to comply with EPA classification and labeling requirements for eye irritation (EPA 2003c), a product manufacturer must provide Draize rabbit eye test data (Draize et al. 1944) to the EPA (40 CFR 158; 40 CFR 161).

In June 2004, the Office of Pesticide Programs contacted the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) to seek the assistance of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in a technical assessment of a nonanimal approach that would meet their need to evaluate, categorize, and label antimicrobial cleaning products (AMCPs) for eye irritation. Subsequently, the Alternative Testing Working Group (ATWG) developed a nonanimal approach for this limited group of products. The ATWG comprises seven consumer product companies: Clorox, Colgate-Palmolive, Dial, EcoLabs, JohnsonDiversey, Procter & Gamble, and SC Johnson. The Institute for In Vitro Sciences, Inc. (IIVS), which coordinated the ATWG collaboration, performed additional testing to complete parallel sets of *in vivo* and *in vitro* data and described the final approach in a background review document. The EPA and the ATWG requested that NICEATM and ICCVAM use information in the background review document to conduct a technical review of the scientific validity of the proposed approach. The EPA and the ATWG sought to determine whether EPA could be assured with a reasonable degree of certainty that the approach would be useful for making hazard classification and labeling decisions for AMCPs in order to appropriately inform users. A *Federal Register* (*FR)* notice (70 FR 13512) issued on March 21, 2005, by NICEATM requested relevant alternative data and nominations for potential peer review panel members.

Three *in vitro* test methods are proposed in the testing strategy: the Cytosensor microphysiometer test method, the bovine corneal opacity and permeability test method, and the EpiOcular™ test method (MatTek Corporation, Ashland, MA). Representatives from the ATWG first presented an overview of the proposed AMCP testing strategy to the ICCVAM Ocular Toxicity Working Group (OTWG) on August 25, 2005, to solicit feedback and recommendations for a submission to ICCVAM. The AMCP team updated the OTWG on October 24, 2006. NICEATM received an initial draft of the AMCP BRD from IIVS on December 27, 2007; a formal transmittal letter followed on January 8, 2008. Representatives from the ATWG presented an overview of the AMCP submission to the OTWG on January 22, 2008. On March 28, 2008, following a preliminary review of the BRD, the OTWG requested additional information and data from IIVS. The additional data, which were necessary to complete an evaluation, were received on April 4, 2008.

An April 4, 2008, *FR* notice (73 FR 18535) requested relevant data and nominations for potential peer review panel members for the AMCP submission. On June 23-24, 2008, the OTWG and ICCVAM assigned this activity a high priority following consideration of comments from the public and ICCVAM's advisory committee, the Scientific Advisory Committee on Alternative Toxicological Methods. IIVS submitted to NICEATM the final AMCP background review document on July 21, 2008.

The OTWG and NICEATM prepared a draft summary review document (SRD) that summarizes the current validation status of the proposed testing strategy based on information in the AMCP BRD and other related information and data obtained by NICEATM following submission of the BRD. The draft ICCVAM SRD also provides similar information for an proposed alternate testing strategy based on the current validation database for each of the proposed test methods in the testing strategy. The SRD summarizes information from the BRD needed to evaluate the validation status of each of the three component test methods and both proposed testing strategies, and forms the basis for draft ICCVAM test method recommendations, which are provided in a separate document.

An international independent scientific peer review panel (Panel) will be convened in public forum on May 19–21, 2009, to develop conclusions and recommendations on the proposed AMCP testing strategies. The panel includes expert scientists nominated by ECVAM and JaCVAM. We anticipate that these organizations will be able to use the independent report of

the Panel for their deliberations and development of test method recommendations. The Panel will meet to consider this SRD and to evaluate the extent to which the available information supports the draft ICCVAM test method recommendations. ICCVAM will consider the conclusions and recommendations of the Panel, along with comments received from the public and SACATM, and then finalize the SRD and test method recommendations. These will be forwarded to Federal agencies for their consideration and acceptance decisions where appropriate.

We gratefully acknowledge the organizations and scientists who provided data and information for this document. We also acknowledge the efforts of those individuals contributing to the preparation of this summary review document, including the following staff from the NICEATM Support Contractor, Integrated Laboratory Systems, Inc.: David Allen, Jon Hamm, Nelson Johnson, Elizabeth Lipscomb, Linda Litchfield, Gregory Moyer, Catherine Sprankle, and Jim Truax. We also thank the members of the ICCVAM Ocular Toxicity Working Group, chaired by Karen Hamernik, Ph.D. (EPA) and Jill Merrill, Ph.D. (FDA), and ICCVAM representatives who reviewed and commented on draft versions. We also thank Valerie Zuang, Ph.D., and Dr. Hajime Kojima, Ph.D., the liaisons to the Ocular Toxicity Working Group from the European Centre for the Validation of Alternative Methods and the Japanese Center for the Validation of Alternative Methods, respectively, for their participation.

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

William S. Stokes, D.V.M., D.A.C.L.A.M.
Rear Admiral, U.S. Public Heath Service
Director, NICEATM
Executive Director, ICCVAM

April 1, 2009

Executive Summary

Background

In June 2004, the EPA Office of Pesticide Programs contacted the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) to seek the assistance of ICCVAM in a technical assessment of a nonanimal approach that would meet OPP's need to evaluate, categorize, and label antimicrobial cleaning products (AMCPs) for eye irritation. Subsequently, the Alternative Testing Working Group (ATWG) developed a nonanimal approach for this limited group of products. The ATWG is comprised of seven consumer product companies (Clorox, Colgate-Palmolive, Dial, EcoLabs, JohnsonDiversey, Procter & Gamble, and SC Johnson). The Institute for In Vitro Sciences, Inc. (IIVS), which coordinated the ATWG collaboration, performed additional testing to complete sets of comparative *in vivo* and *in vitro* data and prepared a background review document (BRD) describing the final approach. The EPA and the ATWG requested that NICEATM and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) use information within the BRD to conduct a technical review of the scientific validity of the proposed approach to determine whether EPA could be assured with a reasonable degree of certainty that the approach would be useful for making hazard classification and labeling decisions that appropriately inform the user of AMCPs.

After receiving the final AMCP BRD (**Appendix A**), ICCVAM and NICEATM compiled this summary review document , which summarizes the available data and information regarding the validity (usefulness and limitations) of each of the three individual test methods, the proposed ATWG testing strategy, and the proposed alternate strategy.

Test Method Protocols and the Proposed AMCP Testing Strategies

Bovine Corneal Opacity and Permeability, Cytosensor Microphysiometer, and EpiOcular™

In the AMCP BRD submission, three *in vitro* test methods were used to develop a proposed testing strategy: bovine corneal opacity and permeability (BCOP) test method, the Cytosensor microphysiometer (CM) test method, and the EpiOcular™ (EO) test method. Detailed protocols for each test method are provided in the AMCP BRD submission. These test methods use a variety of endpoints to predict ocular irritation potential. For each test method, decision criteria

have been developed to correspond to up to four of the different categories of ocular irritation defined by the EPA hazard classification system (i.e., EPA Categories I–IV). The endpoint for the CM is the estimated concentration of a test substance needed to reduce the basal metabolic rate of L929 cells by 50% (the MRD_{50}). $MRD_{50} < 2$ = EPA Category I; $MRD_{50} \geq 2$ mg/mL and ≤ 80 mg/mL = EPA Category III; $MRD_{50} > 80$ mg/mL = EPA Category IV. Decision criteria for the CM are not proposed in the AMCP BRD submission for Category II classification. The endpoint for the EO is the time needed to reduce cell viability by 50% (ET_{50}). Classification of the EO data is based on $ET_{50} < 4$ min = EPA Category I; $ET_{50} \geq 4$ min and ≤ 70 min = EPA Category III; $ET_{50} > 70$ mg/mL = EPA Category IV. Decision criteria for the EO are not proposed in the AMCP BRD submission for Category II classification. The BCOP includes two primary endpoints, the extent of corneal opacity and permeability (both measured quantitatively and used to calculate an *in vitro* irritancy score [IVIS]) and histopathology of the cornea, an optional endpoint that is still under development for use in the BCOP. An IVIS > 75 indicates a Category I, IVIS between 25 to 75 indicates a Category II, and IVIS < 25 indicates a Category III. Decision criteria for the BCOP are not proposed in the AMCP BRD submission for Category IV classification. The additional endpoint of histopathology is proposed for distinguishing between EPA Category I and II substances.

Combining the BCOP, CM, and EO into a Testing Strategy: AMCP Submission Proposal

As described in the AMCP BRD (see **Appendix A**) the first test method used in the proposed AMCP testing strategy reportedly depends on knowledge of the chemical properties of the test substance. If the substance is an oxidizer, which suggests that it will be an ocular corrosive or severe irritant, it is first tested in the BCOP. As noted above, test substances that produce an IVIS > 75 would be classified as EPA Category I. Test substances that produce an IVIS < 75 and do not meet the criteria for classification based on histopathology are judged to cause less than irreversible or severe ocular damage. They are subsequently tested in the CM or EO test methods to delineate the final ocular hazard category (EPA Cat II, III, or IV). Selection of the CM or EO depends on water solubility of the test substance; water-soluble substances would be tested in the CM and water-insoluble substances would be tested in the EO to determine the final hazard classification.

Combining the BCOP and the EO into a Testing Strategy: Alternate Strategy for Evaluation

After assessing the available data, an alternative testing strategy was also evaluated, which would include only the BCOP and the EO. The alternative strategy was to determine if results in the BCOP could be used to identify Category I or II substances, and if results in the EO could be used to identify Category III or IV substances. This alternative strategy was proposed for evaluation in part because only the BCOP and EO have included in their databases a list of the same AMCPs that were tested in both methods (see **Validation Database** below). Another reason for the alternative evaluation was the draft position by the OTWG regarding the validation status of the LVET (which is being reviewed separately by the Panel). Based on the available data, the draft OTWG position is that the LVET predictivity for the Draize test makes it inadequate to serve as a reference test method to support the validity of *in vitro* test methods. For this reason, the CM and some EO data could not be considered to support the testing strategy (see **Validation Database** below).

Validation Database

Substances Tested in the BCOP, CM, or EO

A total of 228 substances were included in the validation database of the AMCP BRD submission (**Appendix A**). These include 68 substances tested in the BCOP, 105 substances tested in CM, and 55 substances tested in EO. None of the 228 substances have been tested in all three of the proposed *in vitro* test methods (i.e., BCOP, CM, and EO). Twenty-eight AMCPs have been tested in both the BCOP and the EO. According to the submitter, "a minimum 28 of the materials are EPA registered anti-microbial cleaning products, with eight additional materials being in-use dilutions of concentrates which are EPA registered."

The distribution of product categories differed among the different validation databases. Most of the 105 substances tested in CM are surfactants (78% [82/105]) or solvents (18% [19/105]), while the substances tested in the BCOP (n=68) and EO (n=55) are relatively equally distributed among alkalis, oxidizers, solvents, and surfactants (approximately 20% to 30% each).

In Vivo Reference Data

The test method protocol used to generate the *in vivo* reference data varied among the 228 substances. For the 68 substances tested in the BCOP, 85% (58/68) were tested in the traditional Draize rabbit eye test protocol (i.e., OECD TG 405, OECD 1987). Another 12% (8/68) were tested in a nontraditional protocol (i.e., application volume of 30 μL instead of 100 μL, or application as an aerosol spray). The remaining 3% (2/68) were tested in the low volume eye test (LVET), a modification to the rabbit eye test that involves application of 10 μL of the test substance directly to the corneal surface instead of 100 μL of the test substance applied into the conjunctival sac. For the 55 substances tested in EO, 54% (29/54) were tested in the Draize rabbit eye test, while 46% (25/54) were tested in the LVET. All 105 of the substances tested in CM were tested in the LVET.

As noted above, the validation status of the LVET is being evaluated separately based on a comparison of the LVET to the Draize test, which is included in an ICCVAM summary review document (provided as a separate document to the Panel), a BRD submission to ECVAM for the LVET (**Appendix B**), and in the AMCP BRD submission (**Appendix A**). To date, the LVET has not been demonstrated as an adequately valid *in vivo* reference test method. Although the reported advantage of the LVET is that it underpredicts the Draize test and is less overpredictive of the human response than the Draize test, definitive data to support this claim are not available. Human data are generally a mix of clinical data from exposures to very mildly irritating or nonirritating products and accidental exposures where precise measures of amount and time of exposure are not known. The use of LVET as an *in vivo* reference test method is also restricted by the limited types of substances that have been tested (i.e. surfactant-based cleaning products). For this reason, concern exists regarding the feasibility of the LVET to accurately identify severe irritants and ocular corrosives. Therefore, it cannot be recommended as an acceptable *in vivo* reference test method against which to compare *in vitro* test method results

Test Method Accuracy

Cytosensor Microphysiometer

Test method accuracy for each of the three *in vitro* test methods included in the proposed strategy is provided in the AMCP BRD submission using EPA and United Nations Globally Harmonized System of Classification and Labeling of Chemicals (GHS) regulatory

674 classification systems (**Table 1)**. Based on the validation database of 105 substances tested in
675 both the CM and LVET test methods, the CM correctly classified 30% (32/108) of the test
676 substances (note that three substances were tested twice in LVET and resulted in different
677 hazard categories). The majority of Category II, III, and IV substances (based on LVET results)
678 included in the database were overclassified (100% [11/11 Category II AMCPs overclassified;
679 67% [40/60] Category III AMCPs overclassified; 89% [25/28] Category IV AMCPs
680 overclassified). Among the 25 LVET Category IV substances that were overclassified, 16%
681 (4/25 [all surfactants]) were classified by CM as Category I, and 84% (21/25 [6 solvents, 2
682 bases, and 13 surfactants]) were classified by CM as Category III. Because CM does not
683 include decision criteria for EPA Category II, all LVET Category II or III substances that were
684 overclassified by CM were as Category I. All but one of the 40 LVET Category III substances
685 that were overclassified by CM were surfactants; the remaining substance is a solvent. All 11 of
686 LVET Category II substances that were overclassified by CM were surfactants.

687 All nine of the Category I substances (all surfactants) were correctly identified. None of the
688 irritant categories (i.e., EPA Categories I, II, or III) were underpredicted by the CM results.

689 Additional data on 53 surfactant and surfactant-containing formulations were provided in a
690 BRD submitted to ECVAM for review of the validation status of the CM test method (see
691 **Appendix C**). These substances were not claimed as AMCPs, but they were surfactant-
692 containing formulations with similar composition to AMCPs. The database of 53 water-soluble
693 surfactants tested in CM includes 21 surfactant chemicals and 32 surfactant-containing
694 formulations tested across seven different laboratories. Based on the performance of CM using
695 these 53 substances, ICCVAM has proposed[1] that the CM test method can be used as a
696 screening test to identify water-soluble surfactant chemicals and certain types of surfactant-
697 containing formulations (e.g., cosmetics and personal care product formulations, but not
698 pesticide formulations) as either EPA Category I, GHS Category 1, or EU Category R41; or as
699 EPA Category IV, GHS Not Labeled, EU Not Classified in a tiered-testing strategy, as part of a
700 weight-of-evidence approach. A substance that is not classified into one of these two categories
701 would need to be tested in another test method that is capable of correctly identifying possible

[1] This evaluation is currently undergoing separate peer review by an ECVAM Scientific Advisory Committee Peer Review Panel, which includes two members of the ICCVAM Ocular Peer Review Panel (Drs. Hayes and Wilson).

in vitro false positives. Positives would also need to be additionally tested with methods that can correctly identify severe, moderate, and mild ocular irritants (for more detail, see ICCVAM Draft Proposed Recommendations on Cell Function-Based Assays for Identifying All Categories of Ocular Hazard). Analyses performed to identify the ocular hazard potential of these non-AMCP test substances based on Draize reference data suggest that the CM test method could be useful in a testing strategy.

Bovine Corneal Opacity and Permeability

Based on the validation database of 66 substances tested in both the BCOP and Draize test methods, the BCOP correctly classified 55% (36/66) of the substances among the four EPA categories. While only 60% (3/5) or 50% (6/12) of the Category II and III substances, respectively tested in both the BCOP and the Draize test, were correctly identified, 90% (27/30) of the Category I substances were correctly identified. Among the three Category I substances that were underpredicted by the BCOP as a Category II, two are classified as oxidizers and one as a base. It should be noted that one of these two substances (the base) would be correctly identified if the decision criteria was IVIS $\geq$ 55.1 (as recommended in the ICCVAM BCOP protocol) instead of IVIS > 75 (as is proposed in the AMCP submission). However, such a change would also result in two Category II substances (one oxidizer and one acid) and one Category III substance (a base) being overpredicted as Category I.

Among the Draize Category II substances that were incorrectly identified by the BCOP, one (a base) was underclassified as Category III and one (an oxidizer) was overclassified as Category I. Among the six Draize Category III substances that were incorrectly identified, three (one each of a solvent, a base, and a surfactant) was overclassified as Category II and three (two oxidizers and one base) was overclassified as Category I. Because the BCOP protocol followed in the submission does not propose decision criteria for Draize Category IV substances, all 19 were overpredicted; two as Category II (both solvents) and 17 (8 surfactants, 3 solvents, 3 acids, one base, one oxidizer, and one "other") as Category III.

EpiOcular™

As noted above, among the 54 substances tested in EO, 29 were also tested in the Draize test and 25 were tested in the LVET. Based on the database of 29 substances tested in both the EO and Draize test methods, the EO correctly classified 76% (22/29) of the substances. Among the four Draize Category III substances, 75% (3/4) were correctly identified. The one substance

733 incorrectly identified (a base) was overclassified as a Category I. Among the nine Draize
734 Category IV substances, 44% (4/9) were correctly identified. Four of the five incorrectly
735 identified substances were overclassified as Category III (2 solvents, 1 acid, and on surfactant),
736 and the remaining substance (a surfactant) was overclassified as a Category I. All of the Draize
737 Category I substances (15/15, including 12 bases, 2 solvents, and 1 "other") were correctly
738 identified.

739 Based on the database of 25 substances tested in both the EO and LVET test methods, the EO
740 correctly classified 44% (11/25) of the substances. Among the 12 LVET Category III
741 substances, 67% (8/12) were correctly identified. The four substances incorrectly identified
742 (two surfactants and two oxidizers) were overclassified as a Category I. Among the nine LVET
743 Category IV substances, 0% (0/9) were correctly identified; 44% (4/9, including three
744 surfactants and one solvent) were overclassified as Category III and 56% (5/9, including three
745 oxidizers and two solvents) were overclassified as Category I. All of the LVET Category I
746 substances (3/3, including two oxidizers and one surfactant) were correctly identified by the
747 EO.

Table 1 Performance of AMCP in the Cytosensor Microphysiometer, EpiOcular™, and Bovine Corneal Opacity and Permeability Test Methods Compared to the Low Volume Eye Test or the Draize Rabbit Eye Test as Reported in the AMCP BRD[1] Using the EPA Ocular Hazard Classification System

In Vitro Test Method	*In Vivo* Test Method	Overall Classification	Performance of the *In Vitro* Test Method Compared to the *In Vivo* Reference Test Method Using the EPA Ocular Hazard Classification System									
			I		II			III			IV	
			Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
CM[2]	LVET	30% (32/108)	100% (9/9)	0% (0/9)	100% (11/11)	0% (0/11)	0% (0/11)	67% (40/60)	33% (20/60)	0% (0/60)	89% (25/28)	11% (3/28)
BCOP[4]	Draize	55% (36/66)	90% (27/30)	10% (3/30)	20% (1/5)	60% (3/5)	20% (1/5)	50% (6/12)	50% (6/12)	0% (0/12)	100% (19/19)	0% (0/19)
EO[5]	Draize	76% (22/29)	100% (15/15)	0% (0/15)	0% (0/1)	0% (0/1)	100% (1/1)	25% (1/4)	75% (3/4)	0% (0/4)	56% (5/9)	44% (4/9)
EO[3]	LVET	44% (11/25)	100% (3/3)	0% (0/3)	100% (1/1)	0% (0/1)	0% (0/1)	33% (4/12)	67% (8/12)	0% (0/12)	100% (9/9)	0% (0/9)

Abbreviations: AMCP = Antimicrobial cleaning products; BCOP = Bovine corneal opacity and permeability test method; BRD = Background review document; Cat = Category; CM = Cytosensor microphysiometer; EO™ = EpiOcular™; EPA = U.S. Environmental Protection Agency; ET_{50} = Estimated time to decrease keratinocyte viability in the EO™ test method by 50%; IIVS = *in vitro* irritancy score; LVET = Low volume eye test; MRD_{50} = Concentration of test substance that decreases the metabolic rate by 50% determined by a plot of the concentration-response curve

[1]**Appendix A**.

[2]Classification of the CM data was based on $MRD_{50} < 2$ = EPA Category (Cat) I; $MRD_{50} \geq 2$ mg/mL and ≤ 80 mg/mL = EPA Cat III; $MRD_{50} > 80$ mg/mL = EPA Cat IV. The CM was not proposed to identify EPA Cat II moderate irritants. The database consisted of 108 substances tested in the CM and in the LVET (105 different substances because three duplicates were tested twice).

[3]Classification of the EO data was based on $ET_{50} < 4$ min = EPA Cat I; $ET_{50} \geq 4$ min and ≤ 70 min = EPA Cat III; $ET_{50} > 70$ mg/mL = EPA Cat IV. The CM was not proposed to identify EPA Cat II moderate irritants. The database consisted of 25 substances tested in the EO and in the LVET.

[4]Classification of the BCOP data using either the decision criteria in the AMCP BRD (**Appendix A**) (IIVS ≥ 75 to assign EPA Category 1) or in the BCOP BRD (ICCVAM 2006a) (IIVS ≥ 55 to assign EPA Category I) yields identical results. All BCOP classifications, including high-solvent substances, used a 10-minute exposure time. The BCOP was not proposed to identify EPA Cat IV. The database consisted of 66 substances tested in the BCOP and in the Draize test.

[5]Classification of the EO data was based on $ET_{50} < 4$ min = EPA Cat I; $ET_{50} \geq 4$ min and ≤ 70 min = EPA Cat III; $ET_{50} > 70$ mg/mL = EPA Cat IV. The CM was not proposed to identify EPA Cat II moderate irritants. The database consisted of 29 substances tested in the EO and in the Draize test.

Combining the BCOP, CM, and EO into a Testing Strategy: AMCP Submission Proposal

None of the 228 substances included in the AMCP BRD were tested in all three *in vitro* test methods proposed for the testing strategy. Therefore, there are no data available for the proposed substances with which to characterize the actual performance of a testing strategy that includes the BCOP, CM, and EO.

Combining the BCOP and EO into a Testing Strategy: Alternate Strategy for Evaluation

However, 28 substances for which Draize eye test data are available were tested in both the BCOP and the EO. Therefore, an alternative testing strategy was evaluated to determine if BCOP results could be used to identify Category I or II substances and if EO results could be used to identify Category III or IV substances. The data were evaluated based on two approaches: (1) test in the BCOP first and then in the EO or (2) test in the EO first and then in the BCOP. For the first approach, the BCOP was evaluated for its ability to identify substances as either Category I or II. All substances that were classified as Category I or II in the BCOP (n=15) were removed from the database, and the remaining 13 substances were evaluated based on EO results for identifying Category III or IV substances. The reverse was done for the second approach: the EO was evaluated for its ability to identify substances as either Category III or IV, and all substances that were classified as Category III or IV in EO (n=13) were removed from the database. The remaining 15 substances were evaluated based on the BCOP results for identifying Category I or II substances.

Regardless of which approach was used, the performance of the proposed BCOP/EO testing strategy was the same (**Table 2**). The BCOP/EO testing strategy correctly classified 79% (22/28) of the substances, which includes identifying 100% (14/14) of the Category I substances, 100% (4/4) of the Category III substances, and 44% (4/9) of the Category IV substances. (There were no Category II substances among the 28 substances.) None of the irritant categories (i.e., Category I, II, or III) were underclassified as Category IV substances. However, it should be noted that, based on this database of 28 substances, the performance of the EO alone is the same as that of the proposed BCOP/EO testing strategy.

Because the AMCP BRD proposes different decision criteria to identify Category I substances (IVIS > 75) with the BCOP than those specified in the ICCVAM-recommended BCOP test method protocol (IVIS ≥ 55.1, ICCVAM 2006a), NICEATM also evaluated the testing strategy

using ICCVAM's decision criteria. Based on the limited database of 28 substances, this change did not affect the performance of the BCOP/EO testing strategy.

Table 2 AMCP Substances Tested in Both the BCOP and the EO: Performance Using an Alternate Testing Strategy

EPA	Overall Classification	Draize									
		I		II			III			IV	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Approach 1	79% (22/28)	100% (14/14)	0% (0/14)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/4)	100% (4/4)	0% (0/4)	56% (5/9)	44% (4/9)
EPA	**Overall Classification**	**Draize**									
		I		**II**			**III**			**IV**	
		Actual	**Under**	**Over**	**Actual**	**Under**	**Over**	**Actual**	**Under**	**Over**	**Actual**
Approach 2	79% (22/28)	100% (14/14)	0% (0/14)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/4)	100% (4/4)	0% (0/4)	56% (5/9)	44% (4/9)

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability test method; EO = EpiOcular™ test method; EPA = U.S. Environmental Protection Agency

Approach 1 = Test in the BCOP first to identify Category I or II and then in EO to identify Category III or IV; Approach 2 = Test in EO first to identify Category III or IV and then in the BCOP to identify Category I or II.

[1]Classification of the BCOP data using either the decision criteria in the AMCP BRD (**Appendix A**) (IIVS ≥ 75 to assign EPA Category 1) or in the BCOP BRD (ICCVAM 2006a) (IIVS ≥ 55 to assign EPA Category I) yields identical results. All BCOP classifications, including high-solvent substances, used a 10-minute exposure time.

[2]In the proposed testing strategy, the BCOP is intended to identify only Category I or II substances, and the EO is intended to identify only Category III or IV substances.

[3]When using 3-minute IIVS data for high solvents, the overall classification is 74% (17/23). Five high-solvent substances do not have 3-minute IIVS data and therefore cannot be considered in this analysis.

Test Method Reliability

Reliability of the test methods was determined using data provided in the AMCP BRD and data from other sources such as the European Commission/Home Office (EC/HO; Balls et al. 1995) and the European Cosmetic, Toiletry and Perfumery Association (Colipa; Brantom et al. 1997) validation studies in which the test methods were utilized. This additional data was in the form of non-AMCP data provided to ECVAM on the CM (**Appendix C**) and the EO (**Appendix D**), the ICCVAM *Background Review Document — Current Status of* In Vitro *Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method* (**Appendix E**), and in the Supplement to a Background Review Document of an *In Vitro* Approach for EPA Toxicity Labeling of Anti-Microbial Cleaning Products (**Appendix A**). The reliability evaluations were primarily based on measures of intra- and interlaboratory reproducibility. Reproducibility was evaluated quantitatively by comparing the percent coefficient of variation (%CV) values of each test method parameter and qualitatively as the percent concordance using either the EPA or United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS) classification systems based on the number of substances in agreement compared to the total number tested. Given the limited repeated testing of AMCPs, reliability was based largely on studies that tested substances other than AMCPs.

Test Method Reliability – Intralaboratory Reproducibility

For CM, 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.

For EO, intralaboratory reproducibility was assessed quantitatively based on calculated CVs for ET_{50} values from repeat testing of 0.3% Triton X-100 over a 9-year period in two different laboratories. Mean CVs between the two laboratories ranged from 21% to 22%.

For the BCOP, intralaboratory reproducibility was assessed quantitatively based on the calculated mean CV (20%) for IVIS values for five repeat tested AMCPs. Intralaboratory reproducibility was 20.3% for these five materials (2–6 values per material). Additionally, as noted in the ICCVAM Test Method Evaluation Report (ICCVAM 2006), calculated CVs of IVIS values from two studies ranged from 7% to 33%.

Test Method Reliability – Interlaboratory Reproducibility

Interlaboratory reproducibility of the CM 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 materials. For surfactant materials, all four laboratories using the CM 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 materials, 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 or mixtures; and nonsurfactants, ingredients, or mixtures. 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 no agreement for 10% (1/10) of them. The laboratories had 100% agreement for all (7/7) surfactant-based formulations and mixtures. For nonsurfactants, ingredients, and mixtures the laboratories had 100% agreement for 78% (7/9) of the test substances and no agreement for 22% (2/9) of them.

Interlaboratory reproducibility cannot be determined specifically for the AMCPs included in the AMCP submission because only one laboratory conducted the testing. A two-phased interlaboratory validation study for surfactants and surfactant-containing products was cited in the BRD. The protocol used in the validation study differed from the protocol in the BRD submission (e.g., in the two-phased validation study, surfactants were diluted to 20% before testing; the decision criteria were based on predicted Draize maximum average scores [MAS] and not on calculated ET_{50} values), but according to the BRD, “the vast majority of the manipulations were identical.” Other differences have not been specified. Based on the validation study, mean CVs ranged from 12% to 18%. Fifty-four pure surfactants and mixtures were tested by two laboratories in Phase II with a mean between-laboratories CV of 11.8%.

An interlaboratory validation study of the EO, which was included in a submission to ECVAM, is also cited as further evidence of interlaboratory reproducibility. It should be noted, however, that this reproducibility evaluation, which involves seven different laboratories, is based on an EO protocol that uses relative percent viability to assign an irritancy classification (i.e., irritant vs. nonirritant) and not on a calculated ET_{50} value to predict multiple ocular irritancy hazard categories (i.e., EPA Categories I–IV). The latter is the protocol included in the AMCP submission.

Interlaboratory reproducibility for the BCOP was evaluated using data obtained from three published reports (Gautheron et al. 1994; Balls et al. 1995, Southee 1998). The median %CVs ranged from 23% to 47%, whereas %CVs were higher due to increases at low BCOP IVIS values. Concordance of the EPA classifications was approximately 75% for 86% (44/51) of the test substances in the 11- to 12-laboratory study, > 80% agreement for 75% (44/59) of the test substances in the 5-laboratory study, and 100% agreement for 81% (13/16) of the substances tested in the 3-laboratory study using the BCOP.

Animal Welfare Considerations

The proposed testing strategy is a nonanimal approach for the classification and labeling of AMCP by the EPA (OPP). Bovine eyes used in the BCOP are obtained from animals that are being used for food and obtained post-mortem. The CM uses a mouse cell line that can be purchased. The EO uses primary human keratinocytes obtained from human donors during routine surgical procedures.

Test Method Transferability

The BCOP has been accepted internationally for use under certain circumstances and with specific limitations to classify substances as ocular corrosives and severe irritants. While it is not considered valid as a complete replacement for the *in vivo* rabbit eye test, the BCOP is recommended for use as part of a tiered-testing strategy for regulatory classification and labeling within a specific applicability domain. Use of the BCOP assay has been well documented in the ICCVAM BCOP BRD (ICCVAM 2006a), and uses and limitations are identified in the ICCVAM Test Method Evaluation Report, which includes an ICCVAM-recommended BCOP test method protocol (ICCVAM 2006e).

EO is commercially available from MatTek Corporation (Ashland, MA). The test method costs are in line with or less than those for a Draize rabbit eye test. The *in vitro* test methods may be run in less time than the *in vivo* Draize or LVET test methods, although it may take two weeks lead-time to procure tissue from the MatTek Corporation.

1.0 Introduction and Rationale for the Proposed Use of a Testing Strategy for U.S. Environmental Protection Agency Classification and Labeling of Antimicrobial Cleaning Products

1.1 Historical Background of *In Vitro* Ocular Corrosion and Irritation Test Methods and the Rationale for Their Development

Over the years, legislative statutes have been enacted that enable government agencies to regulate a variety of substances with the potential to pose a risk to ocular health. A synopsis of current U.S. regulatory laws that pertain to ocular corrosion and irritation is provided in **Table 1-1**.

Table 1-1 Summary of Current U.S. Legislation Related to Ocular Health[1]

Legislation (Year of Initial Enactment)	Agency	Substance
Food, Drug and Cosmetic Act (1938)	FDA	Pharmaceuticals and cosmetics
FIFRA (1947) and Federal Environmental Pesticide Control Act (1972)	EPA	Pesticides
FHSA (1964)	CPSC	Household products
FHSA (1964) and TSCA (1976)	Department of Agriculture and EPA	Agricultural and industrial chemicals
Occupational Safety and Health Act (1970)	OSHA	Occupational materials
Clean Air Act Amendments (1990)	Chemical Safety and Hazard Investigation Board and EPA	Accidentally released chemicals and air pollutants

[1]Adapted from Wilhelmus (2001).

Abbreviations: CPSC = U.S. Consumer Product Safety Commission; EPA = U.S. Environmental Protection Agency; FDA = U.S. Food and Drug Administration, FHSA = Federal Hazardous Substances Act; FIFRA = Federal Insecticide, Fungicide, and Rodenticide Act; OSHA = Occupational Safety and Health Administration; TSCA = Toxic Substances Control Act

Exposure of rabbit eyes to substances is the primary method for assessing the ocular hazard potential of substances that may come in contact with or be placed near the eye of a human. The test method currently accepted by U.S. Federal and international regulatory agencies (CPSC 1995; EPA 1998; OECD 2002) is the Draize rabbit eye test (Draize et. al. 1944), which involves placing a test substance into the lower conjunctival sac of one eye of a rabbit and comparing it to the contralateral eye, which serves as a negative control. The eyes of each rabbit are examined for adverse corneal (i.e., opacity and area of involvement), iridal, or conjunctival (i.e., redness, chemosis, and discharge) effects for a period up to 21 days after exposure to the test substance.

The current rabbit eye test method can identify both irreversible (corrosive) and reversible ocular effects. The wide ranges used for scoring a majority of these lesions permits

categorization of the severity of reversible effects as mild, moderate, or severe (see U.S. Environmental Protection Agency [EPA] Ocular Classification System discussed below). Current EPA ocular testing guidelines and the United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS; UN 2007) indicate that if serious ocular damage is anticipated (e.g., irreversible adverse effects on Day 21), then a test on a single animal may be considered. If serious damage is observed, then no further animal testing is necessary (EPA 1998; UN 2007). If serious damage is not observed, additional test animals (1 or 2 rabbits) may be evaluated sequentially until concordant irritant or nonirritant responses are observed (UN 2007).

Depending on the legislative mandate of various regulatory agencies and their goals for protecting human health, each agency's classification of irritant responses varies (**Table 1-2).** The EPA ocular irritation classification regulation and testing guidelines (EPA 1998, 2003c) are based on the most severe response in one animal in a group of 3 or more animals. This classification system takes into consideration the kinds of ocular effects produced, as well as the reversibility and the severity of the effects. The EPA classifies substances into four ocular irritant categories, ranging from I to IV (**Table 1-2**) (EPA 2003c). Category I substances are defined as corrosive or severe irritants, while classification in Categories II to IV is based on decreasing severity of ocular lesions, as well as the time required for the ocular lesions to clear. Irritation that clears in 8 to 21 days is classified as Category II, while irritation that clears within 7 days is classified as Category III. For Category IV substances, irritation clears within 24 hours. For the purpose of harmonizing the classification of ocular irritants internationally, the GHS (UN 2007) includes two categories (**Table 1-2**), one for irreversible effects on the eye/serious damage to the eye (Category 1) and one for reversible effects on the eye (Category 2) based on severity of the lesions and/or the duration of their persistence. Reversible effects are further classified based on the duration as Category 2A ("irritating to eyes" referring to an effect that reverses within 21 days) and Category 2B ("mildly irritating to eyes" referring to an effect that reverses within 7 days).

958 **Table 1-2 *In Vivo* Ocular Irritancy Classification Systems**

Regulatory Agency (Authorizing Act)	Number of Animals	Minimum Observation Times (after treatment)	Mean Score Taken?	Positive Response	Irritant/Nonirritant Classification
EPA (FIFRA; TSCA; and The Federal Environmental Pesticide Control Act)	At least 3	1 hour, 1, 2, 3, 7, 14, and 21 days	No	- Maximum score in an animal used for classification - Opacity or Iritis ≥ 1 or Redness or Chemosis ≥ 2	One or more positive animals needed for classification in categories below. Category: I = Corrosive, corneal involvement, or irritation persisting more than 21 days II= Corneal involvement or irritation clearing in 8-21 days III = Corneal involvement or irritation clearing in 7 days or less IV = Minimal effects clearing in less than 24 hours
GHS-Irreversible Eye Effects	3	1, 2, 3 days (observation until Day 21)	Yes	Mean animal values (over Days 1, 2, and 3) of: Opacity ≥ 3 and/or Iritis ≥ 1.5	- At least 2 positive response animals = Eye Irritant Category 1 - At least 1 animal where Opacity, Chemosis, Redness, or Iritis > 0 on Day 21 = Eye Irritant Category 1
GHS-Reversible Eye Effects	3	1, 2, 3 days (observation until Day 21)	Yes	Mean animal values (over Days 1, 2, and 3) of: Opacity or Iritis ≥ 1 or Redness or Chemosis ≥ 2 and the effect fully reverses in 7 or 21 days	- At least 2 positive response animals and the effect fully reverses in 21 days = Eye Irritant Category 2A - At least 2 positive response animals and effect fully reverses in 7 days = Eye Irritant Category 2B

959 Abbreviations: CPSC = U.S. Consumer Product Safety Commission; EPA = U.S. Environmental Protection Agency; FDA = U.S. Food and Drug Administration;
960 FIFRA = Federal Insecticide, Fungicide, and Rodenticide Act; GHS = United Nations Globally Harmonized System of Classification and Labelling of
961 Chemicals; OSHA = Occupational Safety and Health Administration; TSCA = Toxic Substances Control Act

The GHS (UN 2007) categories are based on severity of the lesions and/or the duration of persistence. **Section 4.1.3** describes the GHS and the U.S. *in vivo* ocular irritancy classification systems in greater detail.

The U.S. Federal Hazardous Substances Act (FHSA) (CPSC 1995) and the European Union (EU; EU 2001) also have classification criteria for ocular irritation. However, because this evaluation focuses on ocular hazard classification according to the EPA and GHS systems, we will not discuss the FHSA and EU criteria. Additional details on these systems can be found in ICCVAM 2006a.

Recently, the EPA requested the evaluation of a nonanimal strategy to classify and label antimicrobial cleaning products (AMCPs). This strategy was developed by the Alternative Testing Working Group (ATWG), which is composed of seven consumer product companies (Clorox, Colgate-Palmolive, Dial, EcoLabs, JohnsonDiversey, Procter & Gamble, and SC Johnson). The *in vitro* test methods used to develop this strategy were the bovine corneal opacity and permeability (BCOP) test method, the Cytosensor microphysiometer (CM) test method, and the EpiOcular™ (EO) test method (MatTek Corporation, Ashland, MA). *In vitro* data were paired with *in vivo* data obtained in either the standard Draize rabbit eye test data or the low volume eye test (LVET).

On behalf of the ATWG, the Institute for In Vitro Sciences (IIVS) submitted a comprehensive background review document (BRD) to the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) for review of the validation status of the proposed strategy. The EPA and the ATWG requested that the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and ICCVAM use information within the BRD to conduct a technical review of the proposed approach to determine whether ICCVAM could assure the EPA with a reasonable degree of certainty that the approach would help the EPA determine AMCP labeling that would appropriately inform users.

This ICCVAM summary review summarizes the available data and information regarding the usefulness and limitations of one of the proposed testing strategies, as well as a proposed alternate strategy that uses only two of the three *in vitro* test methods (the BCOP and the EO).

1.2 Regulatory Rationale and Applicability

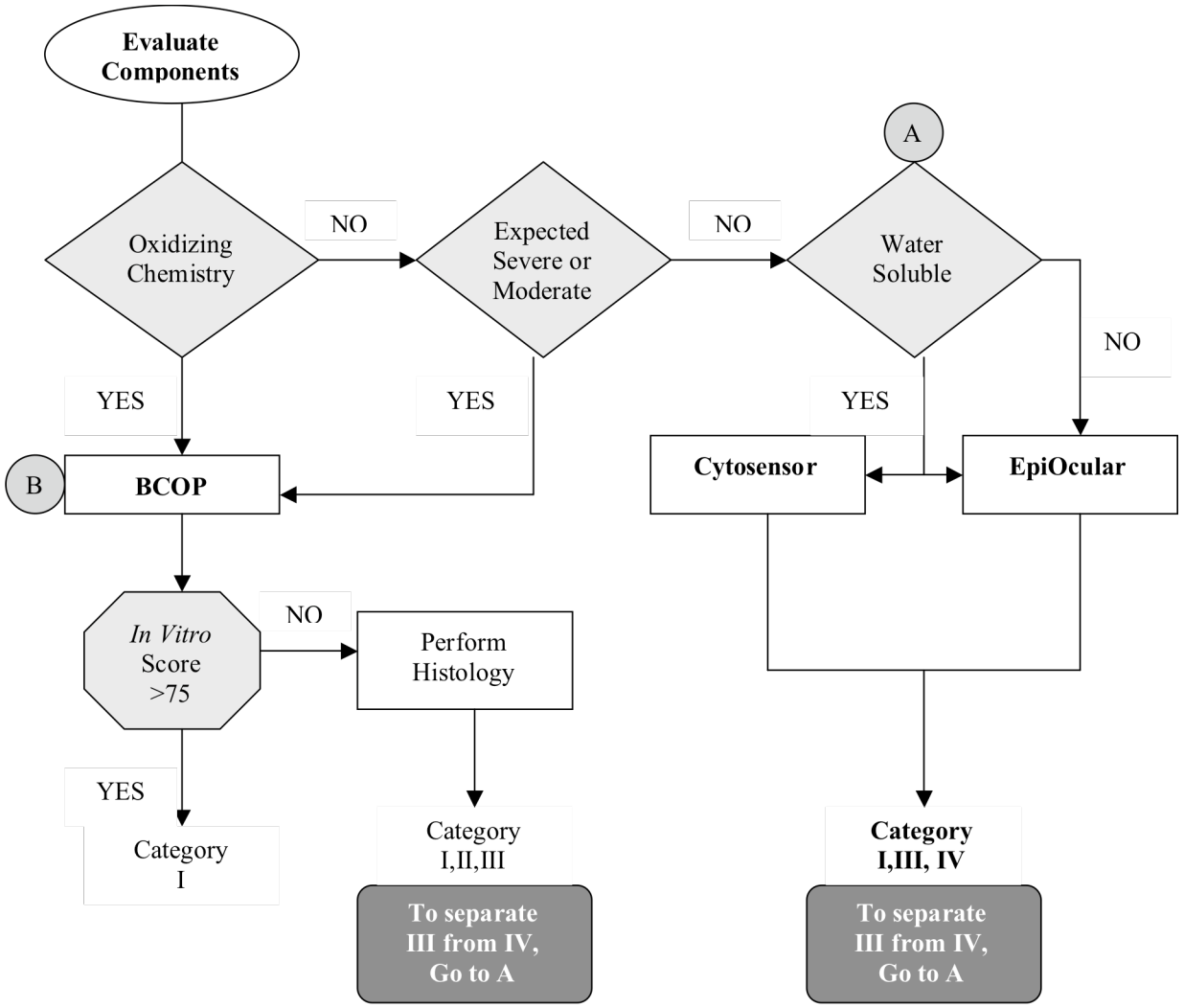
Methods to determine the ocular hazard potential of a cleaning product are currently regulated by the Consumer Product Safety Commission (CPSC) unless the manufacturer intends to label the product as an AMCP. In that case, jurisdiction for the regulation of the product shifts to the EPA. The producer must register the AMCP with the EPA as a pesticide. Currently, the EPA requires AMCPs to be tested in the Draize rabbit eye test in order to adequately characterize their ocular hazard potential.

2.0 Testing Strategies for Ocular Hazard Classification and Labeling of Antimicrobial Cleaning Products

2.1 Original Testing Strategy Proposed in the AMCP BRD Submission

The testing strategy (**Figure 2-1**) proposed in the AMCP BRD submission (**Appendix A**) is based on the use of three test methods: the bovine corneal opacity and permeability test method (BCOP), the Cytosensor microphysiometer test method (CM), and the EpiOcular™ test method (EO). For each test method, decision criteria have been developed to correspond to up to four of the different categories of ocular irritation defined by the Environmental Protection Agency (EPA) hazard classification system (i.e., EPA Categories I–IV). The endpoint for the CM is the estimated concentration of a test substance needed to reduce the basal metabolic rate of L929 cells by 50% (the MRD_{50}). $MRD_{50} < 2$ = EPA Cat I; $MRD_{50} \geq 2$mg/mL and ≤ 80 mg/mL = EPA Cat III; $MRD_{50} > 80$ mg/mL = EPA Cat IV. Decision criteria for the CM are not proposed in the AMCP BRD submission for Category II classification. The endpoint for the EO is the time needed to reduce cell viability by 50% (ET_{50}). Classification of the EO data is based on $ET_{50} < 4$ min = EPA Cat I; $ET_{50} \geq 4$ min and ≤ 70 min = EPA Cat III; $ET_{50} > 70$ mg/mL = EPA Cat IV. Decision criteria for the EO are not proposed in the AMCP BRD submission for Category II classification. The BCOP includes two primary endpoints, the extent of corneal opacity and permeability (both measured quantitatively and used to calculate an *in vitro* irritancy score[2] [IVIS]) and an optional endpoint that is still under development for use in BCOP, histopathology of the cornea. An IVIS > 75 indicates a Category I, IVIS between 25 to 75 indicates a Category II, and IVIS < 25 indicates a Category III. Decision criteria for the BCOP are not proposed in the AMCP BRD submission for Category IV classification. The additional endpoint of histopathology is proposed for distinguishing between EPA Category I and II substances. Detailed protocols for each test method are provided in the AMCP BRD submission (Annex A1-A4).

[2]The *in vitro* irritancy score (IVIS) is calculated as sum of the mean corrected opacity value (± standard deviation [SD]) and 15 times the mean corrected permeability value (OD_{490} units ± SD). Generally, an IVIS from 0 to 25 is considered a mild irritant, from 25.1 to 75 (or to 55 in early studies with pharmaceutical intermediates) is considered a moderate irritant, and above 75 is considered a severe irritant or corrosive.

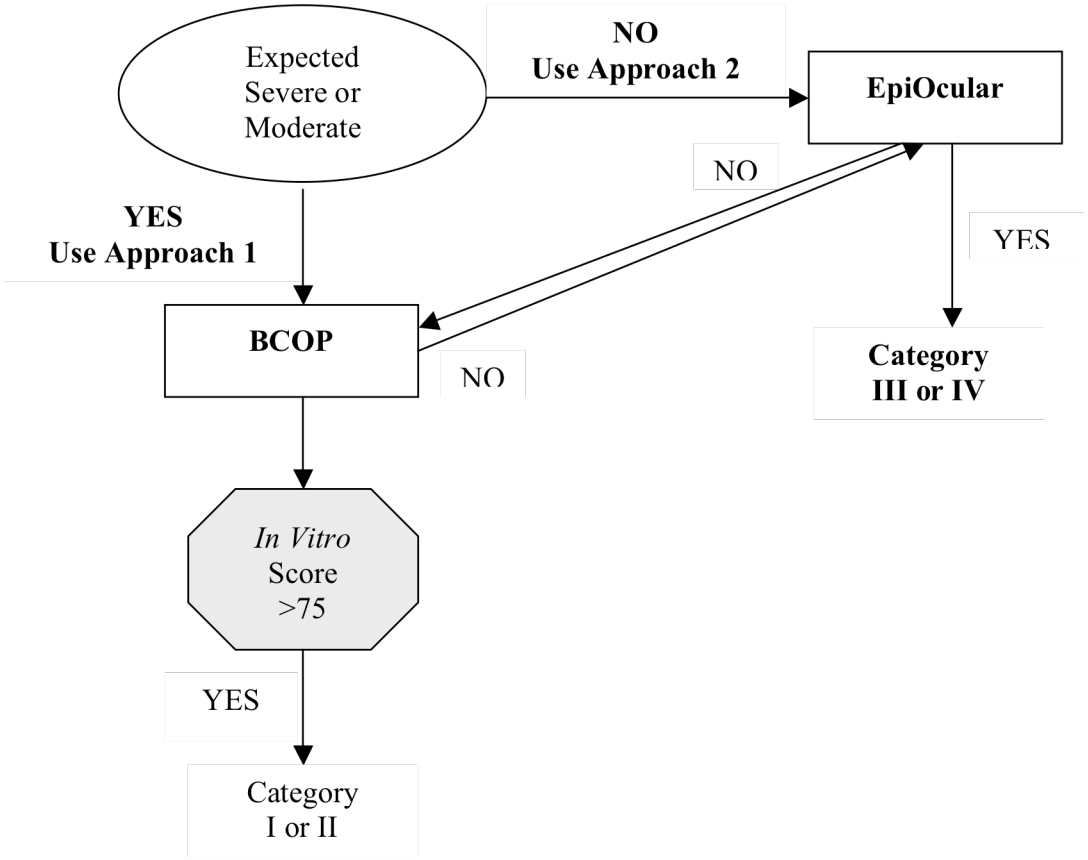


1026

1027 **Figure 2-1 Combining the BCOP, CM, and EO into a Testing Strategy: AMCP Submission Proposal (from Appendix A)**

This testing strategy is based on examination of the predictive capacity of each ocular test method relative to the test substance classifications obtained in either the Draize rabbit eye test or the low volume eye test (LVET), a modification to the Draize rabbit eye test that involves application of 10 μL of the test substance directly to the corneal surface instead of 100 μL of the test substance applied into the conjunctival sac. The physicochemical and other known properties or information on AMCP or of the components in the AMCP formulation (e.g., structure-activity relationships, pH extremes, chemical class, water solubility, physical form [e.g., solid, liquid, gel, paste]) are initially evaluated for their likelihood to produce ocular damage or for their relationship to other similar chemicals or products that are known to produce ocular damage. The first test method used in the proposed AMCP testing strategy depends on knowledge of the chemical properties of the test substance. If the substance is an oxidizer, which suggests that it will be an ocular corrosive or severe irritant, it is first tested in the BCOP, according to scheme B in **Figure 2-1**. AMCPs that are not expected to be moderate or severe irritants or corrosives are tested using scheme A in **Figure 2-1,** using the CM if they are water soluble or the EO if they are not water soluble. Based on these results, AMCPs would be classified as either Category I, III, or IV. AMCPs identified as Category II irritants would then be tested in scheme A illustrated in **Figure 2-1**.

Expected severe irritants/corrosives or oxidizing substances with the potential to be moderate or severe irritants/corrosives are tested in the BCOP test method (scheme B). If the IVIS in the BCOP test method is greater than 75, an EPA Category I classification is assigned. Substances than do not produce an IVIS greater than 75 would be subjected to a histopathology assessment to determine if they qualified as either an EPA Category II, III, or IV ocular irritant. To distinguish EPA Category III from Category IV, the AMCP would have to be tested in scheme A. Companies requiring separation of IIIs and IVs in scheme A or I and II in scheme B would require additional testing according to either scheme A or scheme B.



1053

1054

1055 **2.2 Combining the BCOP and the EO into a Testing Strategy: Proposed Alternate Strategy for Evaluation**

The CM has been evaluated in an ECVAM BRD submission for additional types of water-soluble substances that are not identified as AMCPs (see **Appendix C**). However, because there are no comparative data for substances tested in all three *in vitro* methods included in the proposed testing strategy (see **Section 3.0**), concerns regarding the validation status of the LVET (see **Section 4.0** and the ICCVAM LVET Summary Review Document), which was used as the reference test method for all of the CM AMCP data, as well as lack of commercial availability of the instrumentation for the CM (see **Section 11.0**), an alternate testing strategy was evaluated that would include only the BCOP and the EO. In this proposed strategy, the BCOP would be used to identify EPA Category I or II substances, and the EO would be used to identify Category III or IV substances.

Testing in the alternate strategy could proceed in one of two approaches: (1) test in the BCOP first and then in the EO or (2) test in the EO first and then in the BCOP. Using the first approach, the BCOP would first classify all Category I and II substances. All other substances would then be tested in the EO and classified as either Category III or IV. Using the second approach, substances would first be tested in the EO, which would classify all Category III and IV results. All other substances would then be tested in the BCOP and classified as either Category I or II.

3.0 Substances Used for Validation of the Testing Strategies for EPA Classification of Antimicrobial Cleaning Products

3.1 Rationale for the Substances or Products Included in the Proposed AMCP Testing Strategy

A total of 228 substances were included in the validation database of the AMCP BRD submission (**Appendix A**). These include 68 substances tested in the BCOP, 105 substances tested in the CM, and 55 substances tested in the EO. None of the 228 substances have been tested in all three of the proposed *in vitro* test methods (i.e., BCOP, CM, and EO). Data analyses in the CM were based on an n=108 because three substances were included that were tested twice, each with a different result. Of 29 substances tested in both the BCOP and the EO, 28 met the criteria to assign an EPA hazard classification.

In the AMCP BRD, test substances were divided into chemical "buckets." These buckets were termed solvents, oxidizers, surfactants, acids, bases, and others. The distribution of these buckets by test method is presented in **Table 3-1**. Among the 105 substances tested in the CM, 17% (18/105) were solvents and 78% (82/105) were surfactants. Of 55 substances tested in the EO, 18% (10/55) were solvents, 24% (13/55) were oxidizers, 31% (17/55) were surfactants, and 20% (11/55) were bases. Among the 68 substances tested in the BCOP, 18% (12/68) were solvents, 24% (16/68) were oxidizers, 33% (18/55) were surfactants, and 21% (14/68) were bases.

Table 3-1 Distribution of Product Categories Evaluated in the Proposed AMCP Testing Strategy

Product Categories	Number of Substances Tested Per Test Method			
	Cytosensor Microphysiometer	EpiOcular™	BCOP	Total
Solvents	18	10	12	39
Oxidizers	0	13	16	33
Surfactants	82	17	18	114
Acids	1	2	7	10
Bases	4	11	14	29
Others	0	2	1	3
Total	105	55	68	228

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability test method

It should be noted that, according to the submitter, "a minimum 28 of the materials are EPA registered anti-microbial cleaning products, with eight additional materials being in-use dilutions of concentrates which are EPA registered."

As reported in the AMCP BRD submission (**Appendix A**), all 105 substances tested in the CM were tested *in vitro* and the results compared to *in vivo* LVET data. No *in vivo* Draize rabbit eye test data were available for comparison to *in vitro* data obtained in the CM. Of the 55 substances tested in the EO, 25 were tested in the LVET data and 30 were tested in the Draize rabbit eye test. Of the 30 substances tested in the Draize rabbit eye test, 29 qualified for EPA hazard classification (i.e., one substance with Draize scores greater than 1 was not evaluated through Day 21 as required by EPA). For the BCOP, 85% (58/68) were tested in the Draize rabbit eye test, 12% (8/68) were tested in a nontraditional Draize rabbit eye test[3], and the remaining 3% (2/68) were tested in the LVET.

3.2 Rationale for the Substances or Products Included in the Proposed Alternate Testing Strategy

NICEATM requested additional ocular data on substances tested in the BCOP, the EO, and Draize rabbit eye tests. Additional EpiOcular™ data for which BCOP and Draize test data were available were provided by MatTek Corporation (Ashland, MA), but it was determined that these data were generated using a different protocol or prediction model than those used for all of the performance analyses described in the AMCP BRDs. No other data were found.

The evaluation of the proposed alternate AMCP testing strategy was limited to 28 substances that were tested in both the EO and BCOP and which were also tested in the Draize rabbit eye test. The product categories of these 28 substances included five surfactants, two acids, ten alkalis, four oxidizers, six solvents, and one "other" (or nonspecified) as shown in **Table 3-2**.

[3]The nontraditional Draize test data included seven substances tested with 30 □L rather than the traditional 100 □L instilled in the conjunctival sac of the rabbit and one substance that was tested as an aerosol sprayed directly on the cornea.

Table 3-2 Product Categories of AMCPs Tested in Both the BCOP and the EO

Product Category	Number of Products Tested	*In Vivo* Draize Classification - EPA			
		I	II	III	IV
Surfactant	5	0	0	2	3
Acid	2	0	0	1	1
Alkali	10	9	1	0	0
Oxidizer	4	3	0	0	1
Solvent	6	2	0	1	3
Other	1	0	0	0	1
Total	28	14	1	4	9

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability test method; EO = EpiOcular™ test method; EPA = U.S. Environmental Protection Agency

4.0 *In Vivo* Reference Data

4.1 Consideration of LVET Data

As reported in the AMCP BRD submission (**Appendix A**), all 105 substances tested in the CM were tested *in vivo* in the LVET. No *in vivo* Draize rabbit eye test data were available for comparison to *in vitro* data obtained in the CM. For the 55 substances tested in the EO, 25 were tested in the LVET and 30 were tested in the Draize rabbit eye test. Of those tested in the BCOP, 85% (58/68) were tested in the Draize rabbit eye test, 12% (8/68) were tested in a nontraditional Draize rabbit eye test[4], and the remaining 3% (2/68) were tested in the LVET. The proposed alternate AMCP testing strategy is based on the results for the 28 substances that were tested in both the BCOP and the EO, and that were also tested in the Draize rabbit eye test and qualified for assignment of an EPA hazard classification.

The Draize rabbit eye test (Draize et al. 1944) is the standard test method accepted by U.S. regulatory agencies such as the EPA for ocular irritation testing and for the classification and labeling of chemicals and products. Test guidelines describing the procedure have been published by the EPA (OPPTS 870.2400 [EPA 1998]) and the Organisation for Economic Co-operation and Development (Test Guideline 405 [OECD 2002]). The original reference data are summarized in Section 4.2 of the AMCP BRD submission, and the individual animal data are provided in Annex C of the AMCP BRD submission.

The LVET is an *in vivo* rabbit eye test developed by Griffith et al. (1980) that differs from the Draize rabbit eye test by applying 10 μL (instead of 100 μL) of a test substance directly on the cornea (instead of the conjunctival sac). Scoring of corneal, iridal, and conjunctival lesions in the LVET is identical to that of the Draize rabbit eye test. Background information on the LVET and comparison of the LVET to the Draize test is available in an ICCVAM summary review document (provided to the Panel as a separate document), a BRD submission to ECVAM for the LVET (**Appendix B**), and in the AMCP BRD submission (**Appendix A**). To date, the LVET has not demonstrated adequate validity as an *in vivo* reference test method. Although the reported advantage of the LVET is that it underpredicts the Draize test and overpredicts the human response less than the Draize test, definitive data to support this claim

[4]The nontraditional Draize test data included seven substances tested with 30 μL rather than the traditional 100 μL instilled in the conjunctival sac of the rabbit and one aerosol test substance that was sprayed directly on the cornea.

1150 are not available. Human data are generally a mix of clinical data from exposures to very mildly
1151 irritating or nonirritating products and from accidental exposures where precise measures of
1152 amount and duration of exposure are not known. The use of the LVET as an *in vivo* reference
1153 test method is also restricted by the limited types of substances that have been tested (i.e.,
1154 primarily surfactant-based cleaning products).

1155 Although at least one personal-care products company has used LVET data to support
1156 submission of AMCP data to the EPA, these results were used in a weight-of-evidence
1157 approach with supporting Draize rabbit eye test data. Searches for additional LVET data in the
1158 literature did not provide any additional data.

1159

5.0 Test Method Data and Results

5.1 Original Testing Strategy Proposed in the AMCP BRD Submission

The database in the original AMCP BRD (**Appendix A**) includes, where available, the following specific information for each test substance: name, Chemical Abstracts Service Registry Number (CASRN), physicochemical properties (e.g., purity, form tested), study reference, formulation ingredients, and chemical class (Annex B1). Test concentrations, individual and mean opacity scores, individual and mean permeability scores, ET_{50} or MRD_{50} values, and hazard classification information are provided in Annex B2. No attempt was made to identify the source or purity of the test substance if this information was missing.

5.2 Bovine Corneal Opacity and Permeability

In the AMCP BRD (see Annex D of **Appendix A**), data were available for a total of 68 substances that were tested in the BCOP test method and had corresponding Draize rabbit eye test data. These included 18 surfactants, 7 acids, 14 alkalis, 16 oxidizers, 12 solvents, and one nonclassified material.

Participating companies provided data on 38 substances that had formulations similar to those found in AMCPs. *In vivo* data for 30 of these substances were available for comparison to the BCOP data as shown in Table 5-10 of the AMCP BRD (**Appendix A**). However, two substances were tested only in the LVET. In addition, Gettings et al. (1996) evaluated 25 surfactant or surfactant-containing materials in the BCOP as part of the Cosmetic, Toiletry and Fragrance Association Phase III study, and the raw data for these studies were available for inclusion in the BRD. Although not AMCPs, these surfactant-based substances contain formulations similar to those used in many AMCPs. *In vivo* data from the Draize and LVET test methods were available for these 25 test substances. Raw data were also available for a wide range of materials including 15 surfactants tested in the BCOP in the Balls et al. (1995) European Commission/Home Office (EC/HO) validation study. These *in vitro* data were paired with Draize test results. Thus, 68 substances were tested in the BCOP with available Draize rabbit eye test data.

All of the materials evaluated in the BCOP test method were coded to prevent the possibility of bias in the interpretation of test results and to insure that individual companies were not associated with specific products or formulations.

5.3 Cytosensor Microphysiometer

Participating companies provided CM data on 105 unique substances generated using at least two protocols. One protocol was based on the silicon microphysiometer (SM) test method, the predecessor of the CM, that used a 500 second exposure to L929 cells grown on a cover slip, compared to the cells grown using a patented Transwell™ membrane system in the CM protocol (IIVS and Proctor & Gamble) with an 810-second exposure. An algorithm was derived and used to convert the SM data to be consistent with the CM data.

CM data were also obtained on 25 substances paired with Draize data and 25 substances paired with LVET data from the CTFA Phase III validation study of surfactant-based formulations (Gettings et al. 1996) using the SM method. CM data were also available for 20 unique materials from the Colipa Eye Irritation Validation study (Brantom et al. 1997) that were not tested in any other test method using an 810-second CM protocol developed by IIVS, Microbiological Associates, and Proctor & Gamble. The CM test method data are available in Annex E of the AMCP BRD (**Appendix A**).

5.4 EpiOcular™

Participating companies submitted EO data for 61 test substances having formulations similar to those found in typical cleaning product formulations, but sufficient *in vivo* data to determine the *in vivo* EPA hazard classification were available for only 55 of these. The raw animal data can be found in Annex C1 of the AMCP BRD (**Appendix A**). EO data (i.e., ET_{50} values) and corresponding *in vivo* reference data were available for 55 of these test substances (30 with Draize data, 25 with LVET data).

Data from another set of studies conducted to validate the EO were also submitted for the AMCP BRD. Seventy-three surfactants or surfactant-based materials (or dilutions of materials) were tested in these studies. However, the EO protocol used in those studies differs from the protocol being proposed in this BRD in that the test material was diluted before testing; therefore, these studies will be presented only as supporting information for interlaboratory reproducibility (**Section 7.2.3**).

5.5 Combining the BCOP, the CM, and the EO into a Testing Strategy: AMCP Submission Proposal

None of the 228 substances included in the AMCP BRD were tested in all three *in vitro* test methods proposed for the testing strategy. Therefore, there are no data available for the

proposed substances with which to characterize the actual performance of a testing strategy that includes the BCOP, the CM, and the EO.

5.6 Combining the BCOP and the EO into a Testing Strategy: Proposed Alternate Strategy for Evaluation

There were 28 substances tested in both the BCOP and the EO for which Draize reference data were available. The composition of each of the 28 formulations evaluated in the proposed alternate testing strategy is provided in **Appendix F**. The BCOP IVIS and the EO ET_{50} values for each of the 28 substances tested, along with the associated in vitro and in vivo EPA hazard classification are provided in **Appendix G**.

6.0 Test Method Accuracy

6.1 Original Testing Strategy Proposed in the AMCP BRD Submission

The performance of each of the test methods included in the proposed testing strategy is detailed in the AMCP BRD submission (**Appendix A**) according to either the EPA (EPA 2003c) or GHS (UN 2007) regulatory classifications systems. Therefore, we only briefly summarize performance in this report. Additionally, because the results for EPA or GHS classification systems are similar, we discuss only the EPA results. The data from the original submission are summarized in **Table 6-1**.

6.1.1 Bovine Corneal Opacity and Permeability

Based on the validation database of 66 substances tested in both the BCOP and Draize test methods, the BCOP correctly classified 55% (36/66) of the substances overall (see **Table 6-1**). However, while only 60% (3/5) and 50% (6/12) of the Category II and III substances, respectively, tested in both BCOP and the Draize test were correctly identified, 90% (27/30) of the Category I substances were correctly identified. Among the three Category I substances that were underpredicted by the BCOP as a Category II, two were classified as oxidizers and one as a base. It should be noted that one of these two substances (the base) would be correctly identified if the decision criteria was IVIS $\geq$ 55.1, as recommended in the ICCVAM BCOP protocol, instead of IVIS > 75 as proposed in the AMCP submission. However, such a change would also result in two Category II substances (one oxidizer and one acid) and one Category III substance (a base) being overpredicted as Category I.

Among the Draize Category II substances that were incorrectly identified by the BCOP, one (a base) was underclassified as Category III, and one (an oxidizer) was overclassified as Category I. Among the six Draize Category III substances that were incorrectly identified, three (a solvent, a base, and a surfactant) were overclassified as Category II, and three (two oxidizers and one base) were overclassified as Category I. Because the BCOP protocol followed in the submission does not propose decision criteria for Draize Category IV substances, all 19 were overpredicted; two as Category II (both solvents) and 17 as Category III (8 surfactants, 3 solvents, 3 acids, one base, one oxidizer, and one “other”).

Table 6-1 Performance of AMCP in the Cytosensor Microphysiometer, EpiOcular™, and Bovine Corneal Opacity and Permeability Test Methods Compared to the Low Volume Eye Test or the Draize Rabbit Eye Test as Reported in the AMCP BRD[1] Using the EPA Ocular Hazard Classification System

In Vitro Test Method	*In Vivo* Test Method	Overall Classification	Performance of the *In Vitro* Test Method Compared to the *In Vivo* Reference Test Method Using the EPA Ocular Hazard Classification System									
			I		II			III			IV	
			Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
CM[2]	LVET	30% (32/108)	100% (9/9)	0% (0/9)	100% (11/11)	0% (0/11)	0% (0/11)	67% (40/60)	33% (20/60)	0% (0/60)	89% (25/28)	11% (3/28)
BCOP[4]	Draize	55% (36/66)	90% (27/30)	10% (3/30)	20% (1/5)	60% (3/5)	20% (1/5)	50% (6/12)	50% (6/12)	0% (0/12)	100% (19/19)	0% (0/19)
EO[5]	Draize	76% (22/29)	100% (15/15)	0% (0/15)	0% (0/1)	0% (0/1)	100% (1/1)	25% (1/4)	75% (3/4)	0% (0/4)	56% (5/9)	44% (4/9)
EO[3]	LVET	44% (11/25	100% (3/3)	0% (0/3)	100% (1/1)	0% (0/1)	0% (0/1)	33% (4/12)	67% (8/12)	0% (0/12)	100% (9/9)	0% (0/9)

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability; Cat = Category; CM = Cytosensor Microphysiometer test method; EO™ = EpiOcular™test method; EPA = U.S. Environmental Protection Agency; ET_{50} = Estimated time to decrease keratinocyte viability in the EO test method by 50%; LVET = Low volume eye test; MRD_{50} = Concentration of test substance that decreases the metabolic rate by 50% determined by a plot of the concentration-response curve; IVIS = *in vitro* irritancy score

[1]**Appendix A** of AMCP BRD.

[2]Classification of the CM data was based on $MRD_{50} < 2$ = EPA Cat I; $MRD_{50} \geq 2$mg/mL and ≤ 80 mg/mL = EPA Cat III; $MRD_{50} > 80$ mg/mL = EPA Cat IV. The CM was not proposed to identify a EPA Cat II moderate irritants. The database consisted of 108 substances tested in the CM and in the LVET (105 different substances since three duplicates were tested twice).

[3]Classification of the EO data was based on $ET_{50} < 4$ min = EPA Cat I; $ET_{50} \geq 4$ min and ≤ 70 min = EPA Cat III; $ET_{50} > 70$ mg/mL = EPA Cat IV. The CM was not proposed to identify EPA Cat II moderate irritants. The database consisted of 25 substances tested in the EO™ and in the LVET.

[4]Classification of the BCOP data using either the decision criteria in the AMCP BRD (**Appendix A**) (IVIS ≥ 75 to assign EPA Category 1) or in the BCOP BRD (ICCVAM 2006a) (IVIS ≥ 55 to assign EPA Category I) yields identical results. All BCOP classifications, including high solvent substances, used a 10-minute exposure time. The BCOP test method was not proposed to identify EPA Cat IV. The database consisted of 66 substances tested in the BCOP and the Draize test.

[5]Classification of the EO data was based on $ET_{50} < 4$ min = EPA Cat I; $ET_{50} \geq 4$ min and ≤ 70 min = EPA Cat III; $ET_{50} > 70$ mg/mL = EPA Cat IV. The CM was not proposed to identify EPA Cat II moderate irritants. The database consisted of 29 substances tested in the EO and the Draize test.

BCOP IVIS scores were also considered with histopathology data in an attempt to distinguish between Category I and Category II substances. There were 17 substances for which IVIS scores and histopathology data were available. As noted in **Table 6-2**, accuracy of the overall EPA classification (i.e., Cat I, II, III, IV) was reduced from 41% (7/17) to 35% (6/17). Although using histopathology removed one of the Category I false negatives, it added three Category II false positives. Therefore, based on this limited database of 17 test substances, accuracy in the BCOP did not improve with the inclusion of histopathology as an additional endpoint.

6.1.2 Cytosensor Microphysiometer

An evaluation of the CM was based on a comparison to LVET data using the EPA regulatory classification system for 105 test substances (**Table 6-1**). The results of the performance analysis indicated that the majority of Category II, III, and IV substances (based on LVET results) included in the database were overclassified (100% [11/11] Category II AMCPs overclassified; 67% [40/60] Category III AMCPs overclassified; 89% [25/28] Category IV AMCPs overclassified). Among the 25 LVET Category IV substances that were overclassified, 16% (4/25 [all surfactants]) were classified by the CM as Category I, and 84% (21/25 [6 solvents, 2 bases, and 13 surfactants]) were classified by the CM as Category III. Because the CM does not include decision criteria for EPA Category II, all LVET Category II or III substances that were overclassified by the CM were classified as Category I. All but one of the 40 LVET Category III substances that were overclassified by the CM were surfactants; the remaining substance is a solvent. All 11 of the LVET Category II substances that were overclassified by the CM were surfactants.

Additional analyses were performed using data on 25 surfactant-based formulations from the Phase III surfactant study of Gettings et al. (1996) using either Draize or LVET reference data. The results for CM identification of Draize or LVET Category I calls were 80% (8/10) vs 100% (3/3) concordance, respectively. For identification of Category III calls ,concordance was 63% (10/16) for LVET and 91% (10/11) for Draize data. For identification of Category IV calls, accuracy was 17% (1/6) vs 25% (1/4) with false positive rates of 83% (5/6) and 75% (3/4), respectively. None of the CTFA substances were underpredicted by the CM.

Table 6-2 Comparison of the BCOP and the BCOP Using Histopathology

EPA	Overall Classification	Draize Test									
		I		II			III			IV[1]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
BCOP Only	41% (7/17)	50% (3/6)	50% (3/6)	0% (0/4)	25% (1/4)	75% (3/4)	25% (1/4)	25% (1/4)	0% (0/4)	100% (3/3)	0% (0/3)
BCOP with Histology	35% (6/17)	67% (4/6)	33% (2/6)	75% (3/4)	0% (0/4)	75% (3/4)	25% (1/4)	25% (1/4)	0% (0/4)	100% (3/3)	0% (0/3)

Abbreviations: BCOP = bovine corneal opacity and permeability test method

[1]The BCOP test method decision criteria do not propose to identify EPA Category IV substances.

[2]The BCOP test method was based on the use of AMCP decision criteria with a cutoff for corrosives or severe irritants of ≥ 75 tested with a 10 min exposure time.

In the Colipa Eye Irritation Validation study (Brantom et al. 1997), 129 surfactants and surfactant-containing materials were evaluated (data not shown). The CM correctly identified 78% (7/9) for identification of Draize Category I substances, and 100% (6/6) for Category III or Category IV (2/2) substances. LVET results were identical to those for the Draize test. Twenty-two percent of the Category I substances were underpredicted (2/9 as EPA Category III). Of the surfactants, 11% (2/19) were either under- or overpredicted.

6.1.3 EpiOcular™

EO results are summarized in **Table 6-1**. As indicated in **Section 5.4**, EO data (i.e., ET_{50} values) and corresponding *in vivo* reference data were available for 55 test substances (30 with Draize data, 25 with LVET data). Among the 29 substances that were classified based on Draize data using the EPA hazard classification system, all Category I substances (15/15, including 12 bases, 2 solvents, and 1 "other") were correctly identified by the EO. Among the four Draize Category III substances, 75% (3/4) were correctly identified. The one substance incorrectly identified (a base) was overclassified as a Category I. Among the nine Draize Category IV substances, 44% (4/9) were correctly identified. Four of the five incorrectly identified substances were overclassified as Category III (two solvents, one acid, and one surfactant), and the remaining substance (a surfactant) was overclassified as a Category I.

Among the 25 substances classified based on LVET data, all of the Category I substances (3/3, including two oxidizers and one surfactant) were correctly identified by EO. Among the 12 LVET Category III substances, 67% (8/12) were correctly identified. The four substances incorrectly identified (two surfactants and two oxidizers) were overclassified as a Category I. Among the nine LVET Category IV substances, 0% (0/9) were correctly identified; 44% (4/9, including 3 surfactants and 1 solvent) were overclassified as Category III; and 56% (5/9, including 3 oxidizers and 2 solvents) were overclassified as Category I.

6.1.4 Combining the BCOP, the CM, and the EO into a Testing Strategy: AMCP Submission Proposal

None of the 228 substances included in the AMCP BRD were tested in all three *in vitro* test methods proposed for the testing strategy. Therefore, no data are available for the proposed substances with which to characterize the actual performance of a testing strategy that includes the BCOP, the CM, and the EO.

6.2 Combining the BCOP and the EO into a Testing Strategy: Proposed Alternate Strategy for Evaluation

A number of different analyses were conducted to determine an optimal alternate testing strategy that would include the BCOP and the EO. For the BCOP, one set of performance calculations was based on Draize data that were available for 210 substances (i.e., AMCP and non-AMCP) tested in the BCOP. However, overall accuracy using this large set of substances was low (47% [99/210]) when a 10-minute exposure time for all substances including high solvents was used with a cutoff value of ≥ 55.1 to identify an ocular corrosive or severe irritant as recommended in the ICCVAM BCOP protocol (ICCVAM 2006e). Overall accuracy was 56% (37/66) using the higher cutoff value of ≥ 75 as proposed in the AMCP BRD (**Appendix A**). When only BCOP AMCP data with corresponding Draize data (n=66) were evaluated using these two decision criteria, overall accuracy was still low (58% [38/66]) and (56% [37/66], respectively).

By comparison, when only the EO AMCP data with corresponding Draize data (n=29 substances with EPA classification assigned) were evaluated, overall accuracy was higher (76% [22/29]) than the BCOP (55% [36/66]). However, while all Category I substances in the Draize test were correctly predicted by the EO (100%; [15/15]), the one Category II substance that was tested was underpredicted by the EO as a Category III. Of the four Category III substances, 75% (3/4) were correctly predicted and 25% (1/4) was overpredicted. Of nine substances identified as Category IV, 44% (4/9) were correctly predicted whereas 56% (5/9) were overpredicted.

The final set of performance calculations was based on the 28 substances that were tested in both the BCOP and the EO, with Draize reference data for each data set. As noted in **Section 2.0**, these data were evaluated based on two approaches: test in the BCOP first and then in the EO, or test in the EO first and then in the BCOP. Using the first approach, substances would first be tested in the BCOP, and all Category I and II results would be classified. All other substances would subsequently be tested in the EO and classified as either Category III or IV. Using the second approach, substances would first be tested in the EO, and all Category III and IV results would be classified. All other substances would subsequently be tested in the BCOP and classified as either Category I or II.

6.2.1 Approach 1: Test in the BCOP Followed by the EO

Using Approach 1 (i.e., test in the BCOP first to identify Category I or II substances, then in the EO to identify Category III or IV substances) and using either the ≥ 55.1 or ≥ 75 cutoff values to identify Category I substances, the overall correct classification was 78% (22/28) (**Table 6-3**). The boxes in **Table 6-3** represent the correct calls for the BCOP test method (bolded numbers) or for the EO test method (numbers in parentheses). All of the substances classified as EPA Category I by the Draize test were correctly identified by the BCOP-EO testing strategy using Approach 1 (100% [14/14]). Similarly, the EO test method correctly predicted (100%; 4/4) all of the Category III substances and 44% (4/9) of the Category IV substances. Thus, 56% (5/9) were overpredicted as Category III irritants.

6.2.2 Approach 2: Test in the EO Followed by the BCOP

Using Approach 2 (i.e., test in the EO first to identify Category III or IV substances, then in the BCOP to identify Category I or II substances) and using either the ≥ 55.1 or ≥ 75 cutoff values to identify Category I substances, the overall correct classification was also 78% (22/28) (**Table 6-4**). The boxes in **Table 6-4** represent the correct calls for the BCOP test method (bolded numbers) or for the EO test method (numbers in parentheses). The EO test method correctly identified all (100%; 4/4) of the Category III substances and 44% (4/9) of the Category IV substances. Five Category IV substances (56% [5/9]) were overclassified by the EO as Category III. All of the substances classified as Category I by the Draize test were correctly identified by the BCOP-EO testing strategy using Approach 2 (100% [14/14]). No Draize substances were underpredicted. The BCOP overpredicted one Category IV nonirritant as a Category II moderate irritant.

Table 6-3 Performance of AMCP Substances Tested in Both the BCOP and the EO[1] Using Approach 1

EPA		Classification (BCOP→EO)[3] Using Approach 1				
		I	**II**	**III**	**IV**	**Totals**
Draize Classification	I	**14** (0)	**0** (0)	**0** (0)	**0** (0)	14
	II	**0** (0)	**0** (0)	**0** (1)	**0** (0)	1
	III	**0** (0)	**0** (0)	**0** (4)	**0** (0)	4
	IV	**0** (1)	**1** (0)	**0** (3)	**0** (4)	9
	Totals	**14** (1)	**1** (0)	**0** (8)	**0** (4)	28

EPA	Overall Classification	Draize									
		I		II			III			IV	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Approach to Identify Ocular Corrosives and Severe Irritants	79% (22/28)	100% (14/14)	0% (0/14)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/4)	100% (4/4)	0% (0/4)	56% (5/9)	44% (4/9)

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability; EO = EpiOcular™; EPA = U.S. Environmental Protection Agency; IVIS = *in vitro* irritancy score

[1]Classification of the BCOP data using either the decision criteria in the AMCP BRD (**Appendix A**) (IVIS ≥ 75 to assign EPA Category 1) or in the BCOP BRD (ICCVAM 2006a) (IVIS ≥ 55 to assign EPA Category I) yields identical results. All BCOP classifications, including high solvent substances, used a 10 min exposure time.

[2]Bolded numbers indicate the BCOP classification and numbers in parentheses indicate EO classification when using the proposed strategy.

[3]In the proposed testing strategy, BCOP is only intended to identify Category I or II substances and EO is intended to identify only Category III or IV substances.

[4]When using three-minute *In Vitro* Irritancy Score data for high solvents, the overall classification is 74% (17/23). Five high solvent substances do not have three-minute *In Vitro* Irritancy Score data and thus cannot be considered in this analysis.

Table 6-4 Performance of AMCP Substances Tested in Both the BCOP and the EO[1] Using Approach 2

EPA		Classification (EO→BCOP)[3] Approach 2				
		I	**II**	**III**	**IV**	**Totals**
Draize Classification	I	**14** (0)	**0** (0)	**0** (0)	**0** (0)	14
	II	**0** (0)	**0** (0)	**0** (1)	**0** (0)	1
	III	**0** (0)	**0** (0)	**0** (4)	**0** (0)	4
	IV	**0** (0)	**0** (0)	**1** (4)	**0** (4)	9
	Totals	**14** (1)	**0** (0)	**1** (9)	**0** (4)	28

EPA	Overall Classification	Draize									
		I		II			III			IV	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Approach to Identify Category IV	79% (22/28)	100% (14/14)	0% (0/14)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/4)	100% (4/4)	0% (0/4)	56% (5/9)	44% (4/9)

Abbreviations: AMCP = antimicrobial cleaning product; BRD = background review document; BCOP = bovine corneal opacity and permeability test method; EO = EpiOcular™ test method; EPA = U.S. Environmental Protection Agency; IIVS = *in vitro* irritancy score

[1]Classification of the BCOP data using either the decision criteria in the AMCP BRD (**Appendix A**) (IIVS ≥ 75 to assign EPA Category 1) or in the BCOP BRD (ICCVAM 2006a) (IIVS ≥ 55 to assign EPA Category I) yields identical results. All BCOP classifications, including high-solvent substances, used a 10-minute exposure time.

[2]Bolded numbers indicate the BCOP classification and numbers in parentheses indicate EO classification when using the proposed strategy.

[3]In the proposed testing strategy, the BCOP is intended to identify only Category I or II substances, and the EO is intended to identify only Category III or IV substances.

[4]When using 3-minute IIVS data for high solvents, the overall classification is 74% (17/23). Five high-solvent substances do not have 3-minute IIVS data and thus cannot be considered in this analysis.

7.0 Reliability of the Test Methods Used in the Antimicrobial Cleaning Product Testing Strategies

An assessment of test method reliability is an essential element of any evaluation of the performance of an alternative test method (ICCVAM 2003). Test method reliability was assessed by analysis of intralaboratory repeatability (multiple runs of a substance in a test method conducted by a single laboratory over a short period of time (i.e., days), intralaboratory reproducibility (multiple runs of a substance in a test method conducted by a single laboratory over an extended period of time under similar conditions using identical protocols), and interlaboratory reproducibility (multiple runs of a substance in a test method conducted among several laboratories over an extended period of time under similar conditions using identical protocols). While some measures of repeatability and reproducibility were conducted using data sets presented in the AMCP BRD, there were insufficient data to accurately determine the reliability of the test methods. Therefore, information and data from other sources were used to establish reliability of the test methods used in the AMCP BRD. These include non-AMCP data provided in BRDs submitted to ECVAM on the CM (**Appendix C**) and EO (**Appendix D**) test methods and in ICCVAM's *Background Review Document—Current Status of In Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method* (ICCVAM 2006a). Additional information on test method reliability for the CM, the EO, and the BCOP was provided in IIVS' Supplement to a Background Review Document of an *In Vitro* Approach for EPA Toxicity Labeling of Anti-Microbial Cleaning Products (**Appendix A**). Reproducibility is typically measured as the coefficient of variation expressed as a percentage (%CV) of the MRD_{50}, the ET_{50}, or the IVIS in the BCOP and EO, respectively, using the prediction models outlined in **Section 6.0**.

7.1 Bovine Corneal Opacity and Permeability

7.1.1 Intralaboratory Repeatability

Intralaboratory repeatability (i.e., comparison of within-experiment runs of a test substance) was determined in the AMCP BRD (**Appendix A**) as the %CV of the opacity or permeability score and of the IVIS for each cornea (n=3 to 5) treated with a test substance. The data is shown in Table 7-27 of the AMCP BRD. Because the %CV was significantly impacted by the magnitude of the scores or IVIS, the table was prepared so that the %CVs for IVIS > 10 were separated from those IVIS ≤ 10. The mean %CVs for opacity score, permeability score, and

IVIS when the IVIS was ≤ 10 were 266%, 167.1%, and 66.4%, respectively. However, when the IVIS was > 10, the mean %CVs for opacity score, permeability score, and IVIS were 27.9%, 24.1%, and 18.3%, respectively.

Intralaboratory repeatability data included in the ICCVAM BCOP BRD (ICCVAM 2006a) were also referenced. Intralaboratory repeatability of IVIS was assessed by analyzing two studies (IVIS ≥ 55.1). For substances of varying irritancy in one study (three laboratories evaluated), the median coefficient of variation (CV) for IVISs for replicate corneas (n=3) ranged from 11.8% to 14.2%. In a second study, mean and median CV values for IVISs for replicate corneas (n=4) were 71% to 35%, respectively.

Intralaboratory repeatability of the BCOP was also determined in IIVS' BRD Supplement (**Appendix A**) as the concordance of EPA or GHS classifications among the three to five individual corneas run per test substance for a total of 75 substances. For the EPA classifications, there was 100% agreement among the corneas in a test for 63 of the 75 test substances (84%), 67% agreement for 11 of 75 substances (15%) and 60% agreement for 1 of 75 substances (1.3%). Of the 12 substances for which the test corneas were not in 100% agreement, seven had reactive chemistries, two were alkalis, and one was an acid. For the GHS classification system, there was 100% agreement for 63 of 75 test substances (84%), 67% agreement for 11 of 75 substances (15%), and 60% agreement for 1 of 75 substances (1.3%). The same 12 substances for which the test corneas were not in 100% agreement were noted for the GHS classification system.

7.1.2 Intralaboratory Reproducibility

Intralaboratory reproducibility of the BCOP (i.e., comparison of between-experiment runs of a test substance) was determined by comparing individual test runs of a substance within a single laboratory from different experiments under identical conditions using the same protocol and reported as the %CV. There was 100% agreement among repeat runs for five different test substances each tested twice or for one substance tested six times. Additionally, as noted in the ICCVAM (2006), a CV analysis of intralaboratory data (IVISs) from two studies was conducted. In one study, the between experiment (n=3) mean and median CV values for permeability values were 33.4% and 29.0%, respectively, for 25 surfactant-based personal care cleaning formulations. In the second study, the between-experiment mean CV values of *in vitro*

irritancy scores for 16 substances that had been tested two or more times in three laboratories ranged from 12.6% to 14.8%, while the median CV values ranged from 6.7% to 12.4%.

7.1.3 Interlaboratory Reproducibility

An analysis of interlaboratory reproducibility (i.e., comparison of between-laboratory runs of a test substance) based on data provided in the AMCP BRD (**Appendix A**) was precluded because a single laboratory generated these data. However, as noted in ICCVAM 2006, comparable BCOP data were available for multiple laboratories within each of three comparative validation studies, which allowed for an evaluation of the interlaboratory reproducibility of the BCOP. For these studies, interlaboratory reproducibility was evaluated qualitatively based on the ocular irritancy classification assigned to each substance by each laboratory and quantitatively using IVISs. In the qualitative assessment of interlaboratory reproducibility of hazard classification category, 67% to 94% of the substances were classified the same by the participating laboratories. Substances with less than complete agreement in the testing laboratories include those representing such chemical classes as alcohols, ketones, and heterocyclic compounds and such product classes as solvents, surfactants, chemical intermediates, and pesticides.

A quantitative evaluation of interlaboratory reproducibility also was conducted for these three studies by performing a CV analysis of IVISs obtained for substances tested in multiple laboratories. In one study, the 17 substances predicted as severe in the BCOP had mean and median CV values of 36% and 17%, respectively, for results obtained in either 11 or 12 laboratories. In a second study, the 32 substances predicted as severe in the BCOP assay had mean and median CV values of 25% and 22%, respectively, for results obtained in 5 laboratories. In a third study, the mean and median IVIS CV values for the 16 tested substances were 32.4% and 22.8%, respectively for results obtained in 3 laboratories. Finally, the interlaboratory correlation between the BCOP endpoint data generated by each laboratory was determined for 60 substances, as well as for various subsets of test substances (water soluble, water insoluble, surfactants, solids, solutions, and liquids). This analysis yielded a range of correlation coefficients for the subsets of test substances. Interlaboratory IVIS correlation coefficients generally spanned a range of 0.867 to 0.958 depending on the specific subsets of substances being evaluated.

7.2 EpiOcular™

7.2.1 Intralaboratory Repeatability

Intralaboratory repeatability data for 15 product formulations are provided in Table 7-20 of the AMCP BRD (**Appendix A**). Each test substance was tested in two tissues at four exposure times. The %CVs ranged from 0 to 49.5%.

7.2.2 Intralaboratory Reproducibility

Intralaboratory reproducibility data were provided in Table 7-22 of the AMCP BRD (**Appendix A**) from repeat testing of a single material, 0.3% Triton X-100. Data were presented as combined data over a number of years from MatTek Corporation and IIVS and also as data from IIVS only through October 2004. The %CV by either measure was 20.7 and 22.2%, respectively.

Three substances that were tested more than once at IIVS were also evaluated for their concordance using the EPA or GHS ocular hazard classification. There was 100% agreement for all three test substances for both EPA and GHS classifications.

7.2.3 Interlaboratory Reproducibility

Interlaboratory reproducibility of EO cannot be determined specifically for the AMCPs included in the submission (**Appendix A**) because only one laboratory conducted the testing. A two-phased interlaboratory validation study for surfactants and surfactant -containing products was cited in the BRD. The protocol used in the validation study differed from that in the BRD submission (e.g., in the two-phased validation study, surfactants were diluted to 20% before testing, the decision criteria are based on predicted Draize MAS scores and not on calculated ET_{50} values), but according to the BRD, "the vast majority of the manipulations were identical." Other differences have not been specified. From this study, two examples of interlaboratory reproducibility data were provided in Tables 7-24 and 7-25 of the AMCP BRD (**Appendix A**). These data were obtained from two phases of a validation study conducted for the Colgate-Palmolive Company using a different prediction model than those described in the AMCP BRD. The mean %CV for 19 surfactant-based formulations tested in four laboratories was 18.1% in Phase II and 11.8% for 54 surfactant-based formulations tested in two laboratories in Phase III.

However, it should be noted that this evaluation of reproducibility is based on an EO protocol that uses relative percent viability to assign an irritancy classification (i.e., irritant vs.

nonirritant) and not on a calculated ET_{50} value to predict multiple ocular irritancy hazard categories (i.e., EPA Categories I-IV), the latter which is the protocol included in the AMCP BRD submission.

These test substances were also evaluated for their concordance using the EPA and GHS ocular hazard classification systems. The data is presented in the AMCP BRD Supplement (**Appendix A**) in Tables 2-3, 2-4, 2-5, and 2-6. Using either the EPA or GHS classification systems in the Colgate-Palmolive Phase II validation study, there was 100% agreement for 14/19 (74%) substances, 75% agreement for 2/19 (11%) substances, and 50% agreement for 3/19 (16%) substances among four laboratories. In the Phase III validation study using either the EPA or GHS classification systems, there was 100% agreement for 51/54 (94%) substances and 0% agreement for 3/54 (6%) substances in two laboratories.

8.0 Data Quality: Antimicrobial Cleaning Product Background Review Document

8.1 Adherence to National and International Good Laboratory Practice Guidelines

The extent to which the studies included in the AMCP submission complied with national and international Good Laboratory Practice (GLP) guidelines (OECD 1998; EPA 2003a, 2003b, FDA 2003) is based on the information provided in the BRD. While it could not be ascertained that all of the *in vitro* data provided in the AMCP BRD were GLP compliant, those data that were generated in compliance with GLP guidelines were noted in the Excel® spreadsheets that contain the study data. All of the laboratories that contributed data for these studies have experience in conducting GLP-compliant studies. All of the new data generated for the studies in the AMCP BRD were collected under full GLP compliance.

8.2 Data Quality Audits

Formal assessments of data quality, such as quality assurance audits, generally involve a systematic and critical comparison of the data provided in a study report to the laboratory records generated for a study. No data quality audits were specifically conducted in the preparation of the AMCP BRD. However, those studies that were conducted according to GLP guidelines would have included such an audit.

8.3 Impact of Deviations from GLP Guidelines

The impact of deviations from GLP guidelines cannot be evaluated for the data reviewed in this BRD, because no information on data quality audits was obtained.

8.4 Availability of Laboratory Notebooks or Other Records

The original study notebooks, final reports, and other background information were available for the majority of the studies reported in the AMCP BRD. These materials are considered confidential by the companies who contributed data to the AMCP BRD, and they have asked that the individual companies not be associated with any particular product. However, it has been noted that the study materials will be available for inspection upon request by NICEATM or the EPA but that company identifiers would be removed to ensure compliance with this request.

9.0 Other Scientific Reports and Reviews

Individual BRDs for the CM and the EO have been submitted to ECVAM for review of their validation status in Europe. To date, these BRDs have not been made publicly available. A BRD for the BCOP was compiled by NICEATM and the ICCVAM Ocular Toxicity Working Group and published in March 2006.

NICEATM issued *Federal Register* notices on March 18, 2005, and April 4, 2008, requesting additional data for test methods used to evaluate AMCPs. No additional data were received in response to these requests.

10.0 Animal Welfare Considerations

10.1 How the AMCP Testing Strategy and *In Vitro* Methods will Refine, Reduce, or Replace Animal Use

Draize rabbit eye test data are currently used to classify and label AMCPs. The original testing strategy proposed in the AMCP BRD submission or the alternate testing strategy would provide a nonanimal approach to EPA classification and labeling of AMCPs and could thereby eliminate the use of rabbits for this type of testing.

10.2 Requirements for the Use of Animals

The EPA Office of Pesticide Programs currently requires a Draize rabbit eye test to be used for classification and labeling of AMCPs. The Draize eye irritation test method protocol is provided in the EPA Health Effects Test Guideline (OPPTS 87.2440 [EPA 1998] and in the OECD Test Guideline 405 (OECD 2002). The Draize rabbit eye test requires only one animal if the test substance is shown to be corrosive or a severe (irreversible) eye irritant and three animals per test substance for nonsevere irritants or nonirritants. This is in addition to similar sets of animals for both the positive and negative control groups within a study of multiple test substances. More animals may be required if the test results are equivocal with respect to an EPA classification category.

While the BCOP uses bovine tissue obtained from animals that are being slaughtered for food at the time the ocular tissue is procured. Cattle are not subject to pain and suffering during the harvest of corneal tissue, because it is obtained post-mortem and would otherwise be discarded as waste by the meatpacker.

No animals are used for the CM, except for the mice used to establish the original murine cell line used to establish the cell culture.

Primary human keratinocytes are used to generate the 3-dimensional corneal construct used in the EO. These cells are obtained during routine surgical procedures and their procurement to initiate a cell culture does not subject the donor to any pain or suffering.

11.0 Practical Considerations

Several issues in addition to performance evaluations must be taken into account when assessing the practicality of an alternative test method in comparison to the existing test method:

- Assessments of the laboratory equipment and supplies needed to conduct the alternative test method
- Level of personnel training
- Labor costs
- Time required to complete the test method

The time, personnel cost, and effort required to conduct the proposed test method(s) must be considered reasonable in comparison to those of the test method it is intended to replace.

11.1 Transferability of the Test Methods Included in the Testing Strategy

Test method transferability addresses the ability of a method to be performed accurately and reliably by multiple laboratories (ICCVAM 2003), including those experienced in the particular type of procedure as well as laboratories with less or no experience in the particular procedure. The degree of transferability of a test method can be evaluated based on interlaboratory reproducibility (see **Section 7.0**). The transferability of the test methods included in the strategy is discussed in detail in the AMCP BRD submission (**Appendix A**).

One important consideration regarding the transferability of the CM is the fact that the microphysiometer instrument is not currently available commercially (it has been discontinued). Therefore, a user would be required to either obtain a used instrument, or they would need to have one manufactured.

11.2 Training Considerations

The level of training and expertise needed to conduct the test methods used in the ICCVAM alternate strategy and the training requirements needed to demonstrate proficiency based on the ICCVAM test method submission guidelines (ICCVAM 2003) have been presented in detail in the AMCP BRD submission (**Appendix A**).

11.3 Cost Considerations

The cost for running a GLP-compliant Draize rabbit eye test ranges from $1160 to $14,500 depending on the lab and the maximum number of days the animals have to remain in the study (i.e., 21 days or less). A GLP-compliant CM test method will cost approximately $2050 for each of a minimum of two test materials, but the cost could be reduced to $1375 per test substance for five or more materials run concurrently (IIVS, Gaithersburg, MD). IIVS is reportedly the only commercial laboratory that performs the CM. The EO will cost $3700 per test substance at IIVS if tested individually, but the cost is reduced to $2750 per test substance for five or more materials run concurrently. For the EO, MB Research Laboratories (Spinnerstown, PA) charges $2200 per test substance for each test substance with two replicates at each of three time points or $3225 for inclusion of four time points. A GLP-compliant BCOP test method at IIVS will cost approximately $1850 for a single test substance, including positive and negative controls. Histopathology of the corneas used in that study will cost an additional $4750. Running multiple materials concurrently can reduce the cost of the BCOP with histopathology. For example, a single substance would cost $6600 compared to $3300 per substance for four substances run concurrently. MB Research Laboratories charges $1000 per test substance for the BCOP and $1900 for the BCOP with histopathology.

11.4 Time Considerations

The CM, including multiple runs of the test material, can be completed in a single workday. The EO requires a 2-week lead time to procure the tissue from MatTek Corporation (Ashland, MA) and up to 2 days for testing. The BCOP can be completed in one day but including a histopathology evaluation could add up to 4 weeks. The Draize or LVET *in vivo* test methods could require up to 21 days, in addition to several pretest days required to acclimatize the animals.

12.0 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, de Silva O, Dupuis J, Earl LK, Lovell DP, Pape WJW, Uttley M, Bagley DM, Baker FW, Bracher, M, Courtellemont P, Declercq L, Freeman S, Steiling W, Walker AP, Carr GJ, Dami N, Thomas G, Harbell J, Jones PA, Pfannenbecker U, Southee JA, Tcheng M, Argembeaux H, Castelli D, Clothier R, Esdaile DJ, Itigaki H, Jung, Y. Kasai, H. Kojima, U. Kristen, l.M. Larnico, R.W. Lewis, K. Marenus, Moreno O, Peterson KA, Rasmussen ES, Robles C and Stern MA. 1997. Summary report of the Colipa international validation study on alternatives to the Draize rabbit eye irritation test. Toxicol In Vitro 11:141-179.

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

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.

EPA. 1998. Health Effects Test Guideline. OPPTS 870.2440 Acute Eye Irritation. EPA 712-C-98-195. U.S. Environmental Protection Agency:Washington, DC.

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

EPA. 2003b. Good laboratory practice standards. 40CFR160.

EPA. 2003c. EPA Label Review Manual, 3rd Edition. U.S. Environmental Protection Agency, Office of Prevention, Pesticides & Toxic Substances (OPPTS). EPA 735-B-03-001. August 2003. Washington, DC:U.S. Environmental Protection Agency.

FDA. 2003. Good laboratory practice for nonclinical laboratory studies. 21CFR58.

Federal Insecticide, Fungicide and Rodenticide Act. 1947. Public Law 80-102.

FHSA. 1964. Federal Hazardous Substances Act. Public Law 86-613.

Gautheron et al 1994

Gettings S, Lordo R, Hintze K, Bagley D, Casterton P, Chudkowski M, Curren RD,

1711 Demetrulias JL, Dipasquale LC, Earl LK, Feder PI, Galli CL, Glaza SM, Gordon VC, Janus
1712 J, Kurtz PJ, Marenus KD, Moral J, Pape WJW, Renskers KJ, Rheins LA, Roddy MT, Rozen
1713 MG, Tedeschi JP, Zyracki J. 1996. The CTFA evaluation of alternatives program: An
1714 evaluation of in vitro alternatives to the Draize primary eye irritation test. (Phase III)
1715 surfactant-based formulations. Food Chem Toxic 34:79-117.

1716 Griffith et al 1980

1717 ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised,
1718 and Alternative Test Methods. NIH Publication No: 03-4508. Research Triangle
1719 Park:National Institute of Environmental Health Sciences.

1720 ICCVAM. 2006a. Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives
1721 and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method. Vol. 1/2. NIH
1722 Publication No. 06-4512. Research Triangle Park, NC:National Institute of Environmental
1723 Health Sciences.

1724 ICCVAM. 2006e. ICCVAM Test Method Evaluation Report. *In vitro* Ocular Toxicity Test
1725 Methods for Identifying Severe Irritants and Corrosives. NIH Publication No. 07-4517.
1726 Research Triangle Park, NC:National Institute of Environmental Health Sciences.

1727 Occupational Safety and Health Act of 1970 (OSHA). 1970. Public Law 91-596.

1728 OECD. 1998. OECD Series on Principles of Good Laboratory Practice and Compliance
1729 Monitoring Number 1: OECD principles on Good Laboratory Practice (as revised in 1997).
1730 ENV/MC/CHEM(98)17. OECD, Paris.

1731 OECD. 2002. Test guideline 405. Acute eye irritation/corrosion, adopted April 24, 2002. In:
1732 OECD Guidelines for Testing of Chemicals. OECD, Paris.

1733 Southee JA. 1998. Evaluation of the Prevalidation Process. Part 2, final report. Volume 2.
1734 The Bovine Corneal Opcity and Permeability (BCOP) Assay. European Community Contract
1735 No. 11279-95-10F 1ED ISB GB.

1736 Toxic Substances Control Act of 1976. 1976. Public Law 94-469.

1737 UN. 2007. Globally Harmonised System of Classification and Labelling of Chemicals (GHS)
1738 – "The Purple Book." New York & Geneva:United Nations Publications.

1739 Wilhelmus KR. 2001. The Draize eye test. Surv Ophthalmol 45(6):493-515.

13.0 Glossary[5]

Accuracy:[6] (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of *relevance*. The term is often used interchangeably with *concordance* (see also *two-by-two table*). Accuracy is highly dependent on the prevalence of positives in the population being examined.

Antimicrobial cleaning product (AMCP): Commercially available household cleaning products are regulated by the CPSC. However, when an antimicrobial claim is made, these products must be registered as pesticides with the U.S. EPA to carry the antimicrobial claim on their label.

Blepharitis: Inflammation of the eyelid.

Chemosis: A form of eye irritation in which the membranes that line the eyelids and surface of the eye (*conjunctivae*) become swollen.

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

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

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

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

Concordance:[28] The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance and one aspect of *relevance*. The term is often used interchangeably with *accuracy* (see also *two-by-two table*). Concordance is highly dependent on the prevalence of positives in the population being examined.

Conjunctiva: The mucous membrane that lines the inner surfaces of the eyelids and folds back to cover the front surface of the eyeball, except for the central clear portion of the outer

5 The definitions in this glossary are restricted to their uses with respect to the AMCP test methods and testing strategy.

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

eye (the cornea). The conjunctiva is composed of three sections: palpebral conjunctiva, bulbar conjunctiva, and fornix.

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

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

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

Corneal permeability: Quantitative measurement of damage to the corneal epithelium by a determination of the amount of sodium fluorescein dye that passes through all corneal cell layers.

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

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

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

Essential test method component:[28] Structural, functional, and procedural elements of a test method that are used to develop the test method protocol. These components include unique characteristics of the test method, critical procedural details, and quality control measures. Adherence to essential test method components is necessary when the acceptability of a proposed test method is being evaluated based on performance standards derived from mechanistically and functionally similar validated test method. [Note: Previously referred to as *minimum procedural standards*]

False negative:[28] A substance incorrectly identified as negative by a test method.

False negative rate:[28] The proportion of all positive substances falsely identified by a test method as negative (see *two-by-two table*). It is one indicator of test method accuracy.

False positive:[28] A substance incorrectly identified as positive by a test method.

False positive rate:[28] The proportion of all negative substances that are falsely identified by a test method as positive (see *two-by-two table*). It is one indicator of test method accuracy.

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

Good Laboratory Practice (GLP):[28] Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the OECD and Japanese authorities, which describe record keeping and quality assurance procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.

Hazard:[28] The potential for an adverse health or ecological effect. Hazard potential results only if an exposure occurs that leads to the possibility of an adverse effect being manifested.

Interlaboratory reproducibility:[28] A measure of whether different qualified laboratories using the same protocol and test substances can produce qualitatively and quantitatively similar results. Interlaboratory reproducibility is determined during the prevalidation and validation processes and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory repeatability:[28] The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Intralaboratory reproducibility:[28] The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.

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

***In vitro* score (IVS):** An empirically derived formula used in the BCOP test method whereby the mean opacity and mean permeability values for each treatment group are combined into a single *in vitro* score for each treatment group. The *in vitro* irritancy score (IIVS) = mean opacity value + (15 x mean permeability value).

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

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

Negative predictivity:[28] The proportion of correct negative responses among substances testing negative by a test method (see *two-by-two table*). It is one indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

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

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

Ocular: Of or relating to the eye.

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

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

Opacitometer: An instrument used to measure "corneal opacity" by quantitatively evaluating light transmission through the cornea. The instrument has two compartments, each with its own light source and photocell. One compartment is used for the treated cornea, while the other is used to calibrate and zero the instrument. The difference between photocell signals in the two compartments is measured electronically as a change in voltage, and is displayed digitally, generating numerical opacity values with arbitrary units.

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

Performance:[28] The accuracy and reliability characteristics of a test method (see *accuracy*, *reliability*).

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

Positive control: A substance known to induce a positive response used to demonstrate the sensitivity of the test method and to allow for an assessment of variability in the conduct of the test method over time. For most test methods, the positive-control substance is tested concurrently with the test substance and the vehicle/solvent control. However, for some *in vivo* test methods, periodic studies using a positive-control substance is considered adequate by the OECD.

Positive predictivity:[28] The proportion of correct positive responses among substances testing positive by a test method (see *two-by-two table*). It is one indicator of test method accuracy. Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.

Prevalence:[28] The proportion of positives in the population of substances tested (see *two-by-two table*).

Protocol:[28] The precise, step-by-step description of a test, including the listing of all necessary reagents, criteria, and procedures for the evaluation of the test data.

Quality assurance:[28] A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures is assessed independently by individuals other than those performing the testing.

Reduction alternative:[28] A new or modified test method that reduces the number of animals required.

Reference test method:[28] The accepted *in vivo* test method used for regulatory purposes to evaluate the potential of a test substance to be hazardous to the species of interest.

Refinement alternative:[28] A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal wellbeing.

Relevance:[28] The extent to which a test method correctly predicts or measures the biological effect of interest in humans or another species of interest. Relevance incorporates consideration of the *accuracy* or *concordance* of a test method.

Reliability:[28] A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. It is assessed by calculating intra- and interlaboratory reproducibility and intralaboratory repeatability.

Replacement alternative:[28] A new or modified test method that replaces animals with nonanimal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate).

Reproducibility:[28] The consistency of individual test results obtained in a single laboratory (*intralaboratory reproducibility*) or in different laboratories (*interlaboratory reproducibility*) using the same protocol and test substances (see intra- and *interlaboratory reproducibility*).

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

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

Sensitivity:[28] The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy (see *two-by-two table*).

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

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

Specificity:[28] The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy (see *two-by-two table*).

Test:[28] The experimental system used; used interchangeably with *test method* and *test method*.

Test method:[28] A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with *test* and *test method*. See also *validated test method* and *reference test*.

Tiered testing: A testing strategy where all existing information on a test substance is reviewed, in a specified order, prior to *in vivo* testing. If the irritancy potential of a test substance can be assigned, based on the existing information, no additional testing is required. If the irritancy potential of a test substance cannot be assigned, based on the existing

information, a step-wise animal testing procedure is performed until an unequivocal classification can be made.

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

Transferability:[28] The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Two-by-two table:[28] The two-by-two table can be used for calculating accuracy (concordance) ([c+d]/[a+b+c+d]), negative predictivity (d/[c+d]), positive predictivity (a/[a+b]), prevalence ([a+c]/[a+b+c+d]), sensitivity (a/[a+c]), specificity (d/[b+d]), false positive rate (b/[b+d]), and false negative rate (c/[a+c]).

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

Validated test method:[28] An accepted test method for which validation studies have been completed to determine the relevance and reliability of this method for a specific proposed use.

Validation:[28] The process by which the reliability and relevance of a procedure are established for a specific purpose.

Vehicle control: An untreated sample containing all components of a test system, including the vehicle that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the same vehicle.

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