Appendix C

Background Review Document
Current Status of *In Vitro* Test Methods
for Identifying Mild/Moderate Ocular Irritants:
Bovine Corneal Opacity and Permeability Test Method

This page intentionally left blank

Background Review Document
Current Status of *In Vitro* Test Methods for Identifying Mild/Moderate Ocular Irritants: Bovine Corneal Opacity and Permeability Test Method

Interagency Coordinating Committee on the Validation of Alternative Methods

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

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

2010

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

This page intentionally left blank

Table of Contents

List of Tables **C-9**

List of Figures **C-13**

List of Acronyms and Abbreviations **C-14**

Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives **C-16**

Acknowledgements **C-17**

Preface **C-21**

Executive Summary **C-23**

1.0 Introduction **C-29**

1.1 Background C-29

1.2 Use of the BCOP Test Method in Overall Strategy of Hazard or Safety Assessment C-30

1.3 Validation of the BCOP Test Method C-30

1.4 Search Strategies and Selection of Citations for the BCOP BRD C-30

2.0 Bovine Corneal Opacity and Permeability Test Method Protocol Components **C-32**

2.1 Overview of How the BCOP Test Method is Conducted C-32

3.0 Substances Used for Validation of the Bovine Corneal Opacity and Permeability Test Method **C-34**

4.0 *In Vivo* Reference Data Used for an Assessment of Test Method Accuracy **C-37**

4.1 *In Vivo* Classification Criteria Used for BRD Analysis C-38

4.2 *In Vivo* Data Quality C-40

5.0 Bovine Corneal Opacity and Permeability Test Method Data and Results **C-41**

5.1 Availability of Copies of Original Data Used to Evaluate the Accuracy and Reliability C-41

5.2 Description of the Statistical Approaches Used to Evaluate the Resulting Data C-41

5.3 Summary of Results C-42

5.4 Use of Coded Chemicals and Compliance with GLP Guidelines.....C-42

6.0 Bovine Corneal Opacity and Permeability Test Method Accuracy C-43

6.1 Ability to Distinguish Ocular Corrosives and Severe Irritants from All Other Classes C-43

6.2 GHS Classification System: BCOP Test Method Accuracy.....C-44

6.2.1 Identification of Category 1 Substances (Ocular Corrosives/Severe Irritants).....C-44

6.2.2 Identification of Category 2A Substances (Moderate Ocular Irritants) C-44

6.2.3 Identification of Category 2B Substances (Mild Ocular Irritants) C-45

6.2.4 Identification of Substances Not Classified as Irritant C-45

6.2.5 Overall Correct Classification C-45

6.2.6 Ability to Distinguish Substances Not Classified as Irritant from All Other Classes C-45

6.2.7 Discordant Results According to the GHS Classification System C-48

6.3 EPA Classification System: BCOP Test Method Accuracy C-54

6.3.1 Identification of Category I Substances (Ocular Corrosives/Severe Irritants).....C-54

6.3.2 Identification of Category II Substances (Moderate Ocular Irritants) C-54

6.3.3 Identification of Category III Substances (Mild Ocular Irritants) C-54

6.3.4 Identification of Category IV Substances C-54

6.3.5 Ability to Distinguish Category IV from All Other Classes C-54

6.3.6 Discordant Results According to the EPA Classification System C-59

6.4 EU Classification System: BCOP Test Method Accuracy C-62

6.4.1 Identification of R41 Substances (Ocular Corrosives/Severe Irritants).....C-62

6.4.2 Identification of R36 Substances (Irritants) C-62

6.4.3 Identification of Not Labeled Substances C-64

6.4.4 Ability to Distinguish Not Labeled Substances from All Other Classes.....C-64

6.4.5 Discordant Results According to the EU Classification System C-67

6.5 FHSA Classification System: BCOP Test Method Accuracy C-72

6.5.1 Ability to Distinguish Not Labeled Substances from All Other Classes C-72

6.5.2 Discordant Results According to the FHSA Classification System C-74

7.0 Bovine Corneal Opacity and Permeability Test Method Reliability C-78

7.1 Interlaboratory Reproducibility of Hazard Classification Category Using the GHS Classification System C-78

7.2 Interlaboratory Reproducibility of Hazard Classification Category Using the EPA Classification System C-84

7.3 Interlaboratory Reproducibility of Hazard Classification Category Using the EU Classification System C-89

8.0 Bovine Corneal Opacity and Permeability Test Method Data Quality C-94

8.1 Adherence to National and International GLP Guidelines C-94

9.0 Reports in the Peer-Reviewed Literature C-95

10.0 Animal Welfare Considerations (Reduction, Refinement, and Replacement) C-96

10.1 How the BCOP Test Method Will Reduce, Refine, or Replace Animal Use C-96

10.2 Requirement for the Use of Animals C-96

11.0 Practical Considerations C-97

12.0 References C-98

13.0 Glossary C-101

Annex I Characterization of Substances Tested in the BCOP Test Method ... C-109

I-1 Chemical and Product Classes of Substances Tested in the BCOP Assay C-111

I-2 Components of Formulations Tested in Gettings et al. (1996) C-125

I-3 Components of Formulations Tested in Swanson et al. (1995) C-137

I-4 Components of Formulations Tested in Swanson and Harbell (2000).....C-143

I-5 Components of AMCP Formulations Tested.....C-147

Annex II *In Vitro* Data for Substances Tested in the BCOP Test Method C-161

II-1 BCOP Data Sorted by Reference.....C-163

II-2 BCOP Data Sorted by Substance Name C-215

Annex III Comparison of *In Vivo* and *In Vitro* Ocular Irritancy Classifications..... C-269

III-1 BCOP Data Sorted by Reference.....C-271

III-2 BCOP Data Sorted by Substance Name C-285

List of Tables

Table 1 Evaluation of the Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by GHS, EPA, and EU Classification Systems C-26

Table 2 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the GHS, EPA, EU, and FHSA Classification Systems C-27

Table 2-1 *In Vitro* Ocular Irritancy Classification Scheme for the BCOP Test Method (ICCVAM 2006b) C-32

Table 2-2 *In Vitro* Ocular Irritancy Classification Scheme for the BCOP Test Method (AMCP BRD 2008 Submission) C-33

Table 3-1 Chemical Classes Tested in the BCOP Test Method C-35

Table 3-2 Product Classes Tested in the BCOP Test Method..... C-36

Table 4-1 FHSA Classification System (16 CFR 1500.42) C-39

Table 4-2 Proposed FHSA "Proportionality" Criteria C-40

Table 6-1 Accuracy of the BCOP Test Method in Distinguishing Ocular Corrosives/Severe Irritants from All Other Categories, as Defined by the EPA, GHS, and EU Classification Systems..... C-44

Table 6-2 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by EPA, GHS, and EU Classification Systems..... C-46

Table 6-3 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the GHS Classification System, by Study and Overall..... C-47

Table 6-4 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the GHS Classification System, by Study and Overall..... C-48

Table 6-5 Under- and Overprediction of the BCOP Test Method Using the GHS Classification System in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property..... C-50

Table 6-6 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined

by the GHS Classification System, with Discordant Chemical and Physical Classes Excluded C-52

Table 6-7 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the GHS Classification System, with Discordant Chemical and Physical Classes Excluded C-53

Table 6-8 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System, by Study and Overall C-56

Table 6-9 Accuracy of the BCOP Test Method in Distinguishing Category IV Ocular Irritants from All Other Irritant Classes, as Defined by the EPA Classification System, by Study and Overall..... C-57

Table 6-10 BCOP False Negative Substances C-58

Table 6-11 Under- and Overprediction of the BCOP Test Method Using the EPA Classification System in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property..... C-60

Table 6-12 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System, with Discordant Chemical and Physical Classes Excluded..... C-63

Table 6-13 Accuracy of the BCOP Test Method in Distinguishing Category IV Ocular Irritants from All Other Irritant Classes, as Defined by the EPA Classification System, with Discordant Chemical and Physical Classes Excluded C-64

Table 6-14 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System, by Study and Overall C-65

Table 6-15 Accuracy of the BCOP Test Method in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the EU Classification System, by Study and Overall..... C-66

Table 6-16 Under- and Overprediction of the BCOP Test Method Using the EU Classification System in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property..... C-68

Table 6-17 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System, with Discordant Chemical and Physical Classes Excluded C-70

Table 6-18 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the EU Classification System, with Discordant Chemical and Physical Classes Excluded C-71

Table 6-19 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-20% Classification System, by Study and Overall C-73

Table 6-20 Accuracy of the BCOP Test Method in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the FHSA-67% Classification System, by Study and Overall C-75

Table 6-21 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-20% Classification System, with Discordant Chemical and Physical Classes Excluded C-76

Table 6-22 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-67% Classification System, with Discordant Chemical and Physical Classes Excluded C-77

Table 7-1 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate/Mild Irritants, as Defined by the GHS Classification System, by Study C-80

Table 7-2 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the GHS Classification System, by Study C-82

Table 7-3 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate/Mild Irritants, as Defined by the EPA Classification System, by Study C-85

Table 7-4 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System, by Study C-87

Table 7-5 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate Irritants, as Defined by the EU Classification System, by Study C-90

Table 7-6 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System, by Study..... C-92

List of Figures

Figure 1-1 GHS Testing Strategy for Serious Eye Damage and Eye Irritation..... C-31

List of Acronyms and Abbreviations

AMCP	Antimicrobial cleaning product
BCOP	Bovine corneal opacity and permeability
BRD	Background review document
CASRN	Chemical Abstracts Service Registry Number
CPSC	(U.S.) Consumer Product Safety Commission
CV	Coefficient of variation
EC	European Commission
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
EC/HO	European Commission/British Home Office
ECVAM	European Centre for the Validation of Alternative Methods
EEC	European Economic Community
EPA	(U.S.) Environmental Protection Agency
EU	European Union
FDA	(U.S.) Food and Drug Administration
FHSA	U.S. Federal Hazardous Substances Act
FR	*Federal Register*
g	Gram
GHS	Globally Harmonized System of Classification and Labelling of Chemicals
GLP	Good Laboratory Practice
HET-CAM	Hen's egg test – chorioallantoic membrane
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ICE	Isolated chicken eye
IRE	Isolated rabbit eye
IVIS	*In vitro* irritancy score
JaCVAM	Japanese Center for the Validation of Alternative Methods
μg	Microgram
μL	Microliter

μm	Micrometer
MAS	Maximum average score
MeSH	(National Library of Medicine) Medical Subject Headings
mL	Milliliter
MMAS	Modified maximum average score
NA	Not applicable
NC	Not Classified (as irritant)
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	National Institute of Environmental Health Sciences
NL	Not Labeled (as irritant)
NTP	(U.S.) National Toxicology Program
OD	Optical density
OECD	Organisation for Economic Co-operation and Development
OPPTS	Office of Prevention, Pesticides and Toxic Substances
OSHA	Occupational Safety and Health Administration
OTWG	Ocular Toxicity Working Group
r	rho (correlation coefficient)
SCNM	Study criteria not met
SD	Standard deviation
TG	Test Guideline
UN	United Nations

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

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

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

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

Department of Defense
* Robert E. Foster, Ph.D.
+ Patty Decot
Harry Salem, Ph.D.
Peter J. Schultheiss, D.V.M., DACLAM

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

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

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

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

Food and Drug Administration
Office of the Commissioner
* Suzanne Fitzpatrick, Ph.D., DABT
Center for Biologics Evaluation and Research
Richard McFarland, Ph.D., M.D.
Ying Huang, Ph.D.
Center for Devices and Radiological Health
Melvin E. Stratmeyer, Ph.D.
Vasant G. Malshet, Ph.D., DABT
Center for Drug Evaluation and Research
+ Abigail C. Jacobs, Ph.D.
Paul C. Brown, Ph.D.
Center for Food Safety and Applied Nutrition
David G. Hattan, Ph.D.
Robert L. Bronaugh, Ph.D.
Center for Veterinary Medicine
Devaraya Jagannath, Ph.D.
M. Cecilia Aguila, D.V.M.
National Center for Toxicological Research
Paul Howard, Ph.D.
Donna Mendrick, Ph.D.
William T. Allaben, Ph.D. (to Jan. 2009)
Office of Regulatory Affairs
Lawrence D'Hoostelaere, Ph.D.

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

National Institute of Environmental Health Sciences
* William S. Stokes, D.V.M., DACLAM
+ Raymond R. Tice, Ph.D.
Rajendra S. Chhabra, Ph.D., DABT
Jerrold J. Heindel, Ph.D.

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

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

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

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

* Principal agency representative
+ Alternate principal agency representative

Acknowledgements

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

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

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

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

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

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

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

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

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

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

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

National Institute of Environmental Health Sciences

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

Warren Casey, Ph.D., DABT
Deputy Director

Deborah McCarley
Special Assistant; Assistant Project Officer

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

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

Statistical Consultant for ILS, Inc.

Joseph Haseman, Ph.D.

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

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

Bob Guest
SafePharm Laboratories, Ltd.
Derby, United Kingdom

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

Joe Haseman, Ph.D.
Consultant
Raleigh, North Carolina

Penny Jones
Unilever Research
Sharnbrook, United Kingdom

Menk Prinsen
TNO Nutrition & Food Research Institute
The Netherlands

Horst Spielmann, Dr.med.
ZEBET
Berlin, Germany

Companies and Individuals that Provided *In Vitro* and/or *In Vivo* Data for the BCOP Background Review Document

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

Access Business Group
Luann Potts
Tom Truszkowski
Washington, DC

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

ECVAM
Chantra Eskes, Ph.D.
Ispra, Italy

ExxonMobil Biomedical Sciences, Inc.
James Freeman, Ph.D.
Annandale, NJ

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

JohnsonDiversey, Inc.
John Hamilton, Ph.D.
Sarah Willems, B.S.
Sturdivant, WI

Johnson & Johnson Pharmaceutical R&D
Freddy van Goethem, Ph.D.
Phillipe Vanparys, Ph.D.
Beerse, Belgium

L'Oreal
Christine Van den Berghe, Ph.D.
Paris, France

Merck
Joseph Sina, Ph.D.
West Point, PA

S.C. Johnson & Son, Inc./JohnsonDiversey, Inc.
Nicole Cuellar, M.S.
Judith Swanson, B.S./B.A.
Racine, WI

Preface

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

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

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

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

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

Accordingly, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the ICCVAM Ocular Toxicity Working Group (OTWG) prepared draft background review documents that summarize the current validation status of each test method based on published studies and other data and information submitted in response to a

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

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

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

We gratefully acknowledge the organizations and scientists who provided data and information for this document. We also acknowledge the efforts of those individuals who helped prepare this background review document, including the following staff from the NICEATM support contractor, Integrated Laboratory Systems, Inc.: David Allen, Jon Hamm, Nelson Johnson, Brett Jones, Elizabeth Lipscomb, Linda Litchfield, Steven Morefield, 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. (U.S. EPA) and Jill Merrill, Ph.D. (U.S. Food and Drug Administration), and ICCVAM representatives who reviewed and commented on draft versions. We also thank Valerie Zuang, Ph.D., and Hajime Kojima, Ph.D., 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., DACLAM
Rear Admiral/Assistant Surgeon General, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Executive Summary

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

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

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

- Because the FHSA classification system (2005) is based on a sequential testing strategy that uses up to 18 animals, only a small percentage of the substances in the BCOP database would be classifiable if the FHSA criteria were strictly applied. To maximize the number of substances included in these analyses, "proportionality" criteria were applied for the purpose of assigning an FHSA classification to test results that would require additional testing according to the FHSA sequential testing strategy. These "proportionality" criteria (FHSA-20% and FHSA-67%) are as follows:
- FHSA-20% is based on the proportion of positive animals needed to identify a substance as an irritant using the FHSA sequential testing strategy, where 20% of the animals must demonstrate a positive response for a substance to be identified as an irritant. A substance tested using 3 to 6 animals would not be labeled if $\leq 1/6$ animals were positive based on the FHSA criteria. The substance would be labeled as an irritant if there were ≥ 1 positive animal in a 3- to 5-animal test or ≥ 2 positive animals in a 6-animal test.
- FHSA-67% is based on the proportion of positive animals needed to identify a substance as an irritant using the "first test" of the FHSA sequential testing strategy, where 67% of the animals must demonstrate a positive response for a substance to be identified as an irritant. A substance tested using 3 to 6 animals would not be labeled as an irritant if $\leq 1/6$ animals were positive based on the FHSA criteria. The substance would be labeled as an irritant if there

were ≥2/3, 3/4, 4/5, or 4/6 positive animals. If 1/3, 1/4, 2/4, 1/5, 2/5, 3/5, 2/6, or 3/6 animals were positive, further testing would be required.

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

BCOP Test Method Protocol

The BCOP test method is an *in vitro* eye irritation test method that uses isolated bovine eyes that are byproducts from processing plants. Changes in corneal opacity and permeability are assessed as a measure of test substance damage. To determine opacity, the amount of light transmitted through the cornea is measured with an opacitometer. To determine permeability, the amount of sodium fluorescein dye that passes through all corneal cell layers is measured with a visible light spectrophotometer. Both permeability and opacity are used to calculate an *in vitro* irritancy score (IVIS) that is used to assign an *in vitro* irritancy classification, which predicts the potential of a test substance to cause *in vivo* ocular irritation.

Validation Database

An online literature search was conducted to support the initial evaluation of the validation status of the BCOP test method. The search identified four publications containing BCOP test method results. However, none of these publications included raw data or referenced *in vivo* data. Some of these publications also included data from earlier studies that were already in the validation database. ICCVAM received the BCOP test results for 66 antimicrobial cleaning products (AMCPs) in a submission that describes a non-animal approach for evaluating eye irritation potential and labeling requirements for AMCPs. The previous validation database for the BCOP test method (ICCVAM 2006a) was updated to include these results.

The updated BCOP validation database contains a total of 211 substances, including 135 commercial products or formulations. The most commonly tested chemical classes are alcohols, carboxylic acids, esters, formulations, heterocyclic compounds, hydrocarbons, ketones, and onium compounds. The formulations tested include hair shampoos, personal care cleansers, detergents, bleaches, insect repellents, petroleum products, and fabric softeners. The most commonly tested product classes are chemical/synthetic intermediates, cleaners, drugs/pharmaceuticals/therapeutic agents, petroleum products, solvents, shampoos, and surfactants.

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

BCOP Test Method Accuracy

Identification of All Ocular Hazard Categories

ICCVAM evaluated how well the BCOP test method identified all categories of ocular irritation potential as defined by the EPA (2003a), GHS (UN 2007), and EU (2001) classification systems. Because the FHSA classification system does not distinguish between ocular corrosives and severe irritants and less severe irritants, an evaluation for all ocular hazard categories using the FHSA classification system was not possible.

As shown in **Table 1**, overall correct classifications ranged from 49% (91/187) to 55% (102/187) when using the entire database, depending on the hazard classification system used. Using different decision criteria to identify ocular corrosive/severe irritants (IVIS $\geq$ 75), based on the AMCP BRD (2008), instead of IVIS $\geq$ 55.1 as outlined in the ICCVAM BCOP BRD (2006a), does not improve test method performance.

Distinguishing Substances Not Labeled as Irritants from All Other Hazard Categories

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

As shown in **Table 2**, overall accuracy ranged from 64% (76/118) to 83% (148/179, 155/187, and 161/194), depending on the hazard classification system used. The lowest false negative rate (0% [0/97 and 0/54]) was noted for the GHS and EU classification systems, followed by 5% (8/147 and 6/132) for FHSA-20% and FHSA-67% criteria, and 6% (8/142) for the EPA classification system. Among the eight false negatives for the EPA classification system, all were EPA Category III substances based on Draize rabbit eye test data. For the FHSA-20% and FHSA-67% criteria, eight and six substances were false negatives, respectively. The lowest false positive rate (53% [24/45, 25/47, and 25/47]) was noted for the EPA, FHSA-20%, and FHSA-67% classification systems, followed by 66% (42/64) for the EU classification system, and 70% (63/90) for the GHS classification system. The exclusion of discordant classes had a minor effect or no effect on accuracy (ranged from 60% (39/65) to 82% (53/65) when discordant classes were removed versus 64% (76/118) to 83% (148/179, 155/187, and 161/194) for overall accuracy, depending on the hazard classification system used.

BCOP Test Method Reliability

Interlaboratory Reproducibility

Previous quantitative and qualitative evaluations of the reliability of the BCOP test method have been conducted (ICCVAM 2006a). Additional qualitative analyses of interlaboratory reproducibility were conducted to evaluate how well the BCOP hazard classifications agreed among the participating laboratories from the three different interlaboratory validation studies (Balls et al. 1995; Gautheron et al. 1994; Southee 1998). These evaluations were based on the use of the BCOP test method (1) to identify all ocular hazard categories according to the EPA, EU, or GHS systems, and (2) to distinguish substances not labeled as irritants (EPA Category IV, GHS Not Classified, EU Not Labeled) from all other ocular hazard categories (EPA Categories I, II, III; GHS Categories 1, 2A, 2B; EU R41, R36). Because the performance of the BCOP test method was similar for the EPA and FHSA hazard classification systems, additional reliability analyses were not conducted for the FHSA hazard classification system.

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

Severe using IVIS ≥55.1											
Hazard Classification System	**Overall Correct Classification**	**Severe[2]**		**Moderate[3]**			**Mild[4]**			**Not Labeled[5]**	
		Actual	**Under**	**Over**	**Actual**	**Under**	**Over**	**Actual**	**Under**	**Over**	**Actual**
Severe using IVIS ≥ 55.1 (ICCVAM BCOP BRD [2006a])											
GHS	49% (91/187)	85% (55/65)	15% (10/65)	62% (16/26)	27% (7/26)	11% (3/26)	67% (4/6)	33% (2/6)	0% (0/6)	70% (63/90)	30% (27/90)
EPA	55% (102/187)	84% (53/63)	16% (10/63)	50% (11/22)	32% (7/22)	18% (4/22)	50% (28/57)	36% (21/57)	14% (8/57)	53% (24/45)	47% (21/45)
EU	50% (59/118)	79% (26/33)	21% (7/33)	48% (10/21)	52% (11/21)	0% (0/21)	NA	NA	NA	66% (42/64)	34% (22/64)
Severe using IVIS ≥75 (AMCP BRD [2008])											
Hazard Classification System		**Severe**		**Moderate**			**Mild**			**Not Labeled**	
		Actual	**Under**	**Over**	**Actual**	**Under**	**Over**	**Actual**	**Under**	**Over**	**Actual**
GHS	50% (94/187)	78% (51/65)	22% (14/65)	31% (8/26)	54% (14/26)	15% (4/26)	67% (4/6)	33% (2/6)	0% (0/6)	70% (63/90)	30% (27/90)
EPA	49% (92/187)	78% (49/63)	22% (14/63)	36% (8/22)	45% (10/22)	19% (4/22)	47% (27/57)	39% (22/57)	14% (8/57)	53% (24/45)	47% (21/45)
EU	51% (60/118)	73% (24/33)	27% (9/33)	29% (6/21)	67% (14/21)	4% (1/21)	NA	NA	NA	66% (42/64)	34% (22/64)

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

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

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

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

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

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

Table 2 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the GHS, EPA, EU, and FHSA Classification Systems

Hazard Classification System	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall (GHS)[1]	187	66	124/187	100	97/97	30	27/90	70	63/90	0	0/97
Without Alcohols, Ketones, and Solids[2]	66	64	42/66	100	34/34	25	8/32	75	24/32	0	0/34
Overall (EPA)[3]	187	83	155/187	94	134/142	47	21/45	53	24/45	6	8/142
Without Alcohols, Ketones, and Solids	65	82	53/65	96	47/49	44	7/16	56	9/16	4	2/49
Overall (EU)[4]	118	64	76/118	100	54/54	34	22/64	66	42/64	0	0/54
Without Alcohols, Ketones, and Solids	65	60	39/65	100	31/31	24	8/34	76	26/34	0	0/31
Overall (FHSA-20%)[5]	194	83	161/194	95	139/147	47	22/47	53	25/47	5	8/147
Without Alcohols, Ketones, and Solids	132	81	107/132	98	94/96	36	13/36	64	23/36	2	2/96
Overall (FHSA-67%)[5]	179	83	148/179	95	126/132	47	22/47	53	25/47	5	6/132
Without Alcohols, Ketones, and Solids	120	80	96/120	99	83/84	36	13/36	64	23/36	1	1/84

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

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

[2] Alcohols, ketones, and solids were previously identified as discordant in the BCOP test method relative to the *in vivo* hazard classification (ICCVAM 2006a).

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

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

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

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

For the Gautheron et al. (1994) study, there was 100% agreement among the 11 laboratories for a majority of the Draize ocular corrosives and severe irritants based on all three classification systems, whether they were correctly identified or underclassified by the BCOP test method. For example, for the GHS system, there was 100% agreement for 67% [4/6] of the correctly identified Category I substances. There was also 100% agreement among the 11 laboratories for a majority of the overpredicted Not Labeled substances (for example, for the EU system, there was 100% agreement for 54% [7/13] of the correctly identified Not Labeled substances) and for a majority of the incorrectly identified Not Labeled substances (for example, for the EU system, there was 100% agreement for 91% [21/23] of the correctly identified substances).

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

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

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

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

1.0 Introduction

1.1 Background

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

In 2006, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) completed an evaluation of the bovine corneal opacity and permeability (BCOP) test method for its ability to identify ocular corrosives and severe irritants (ICCVAM 2006a). ICCVAM concluded that the BCOP test method could be used, in appropriate circumstances and with certain limitations, as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, European Union [EU] R41, GHS Category 1) (ICCVAM 2006b). While it was not considered valid as a complete replacement for the *in vivo* rabbit eye test, the BCOP test method was recommended for use as part of a tiered-testing strategy for regulatory classification and labeling within a specific applicability domain. Accordingly, substances that are positive in this test method can be classified as ocular corrosives or severe irritants without further testing in rabbits, while a substance that tests negative would need additional testing in rabbits using a sequential testing strategy as outlined in Organisation for Economic Co-operation and Development Test Guideline 405 (OECD 2002).

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

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

Parallel reviews of the isolated rabbit eye (IRE), hen’s egg test–chorioallantoic membrane (HET-CAM), and isolated chicken eye (ICE) test methods are being conducted. The expert panel report and the

analyses presented in the BRDs will be used to support ICCVAM recommendations on the proposed standardized test method protocols, proposed list of recommended reference substances, and additional optimization and/or validation studies that may be necessary to further develop and characterize the usefulness and limitations of these methods.

For a more detailed discussion of the background of the BCOP test method, including its scientific basis and regulatory rationale and applicability, see the ICCVAM *Background Review Document—Current Status of In Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability* (ICCVAM 2006a).

1.2 Use of the BCOP Test Method in Overall Strategy of Hazard or Safety Assessment

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

1.3 Validation of the BCOP Test Method

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

Validation is the process that establishes the reliability and relevance of a test method for a specific purpose (ICCVAM 2003). *Relevance* is defined as the extent to which a test method will correctly predict or measure the biological effect of interest (ICCVAM 2003). For the BCOP test method described in the BCOP BRD (ICCVAM 2006a), relevance is restricted to how well the test method identifies substances that are capable of producing corrosive or severe irritant effects on the eye. For the current BRD, relevance is based on how well the test method identifies (1) substances that are capable of producing nonsevere ocular irritation or (2) substances not labeled as irritants.

Reliability is defined as the reproducibility of a test method within and among laboratories. Reliability should be based on performance with a diverse set of substances that represent the types of chemical and product classes likely to be tested and that cover the range of responses that need to be identified. The validation process will provide data and information to allow U.S. Federal agencies to develop guidance on the development and use of the BCOP test method as part of a tiered-testing approach to evaluating substances' eye irritation potential.

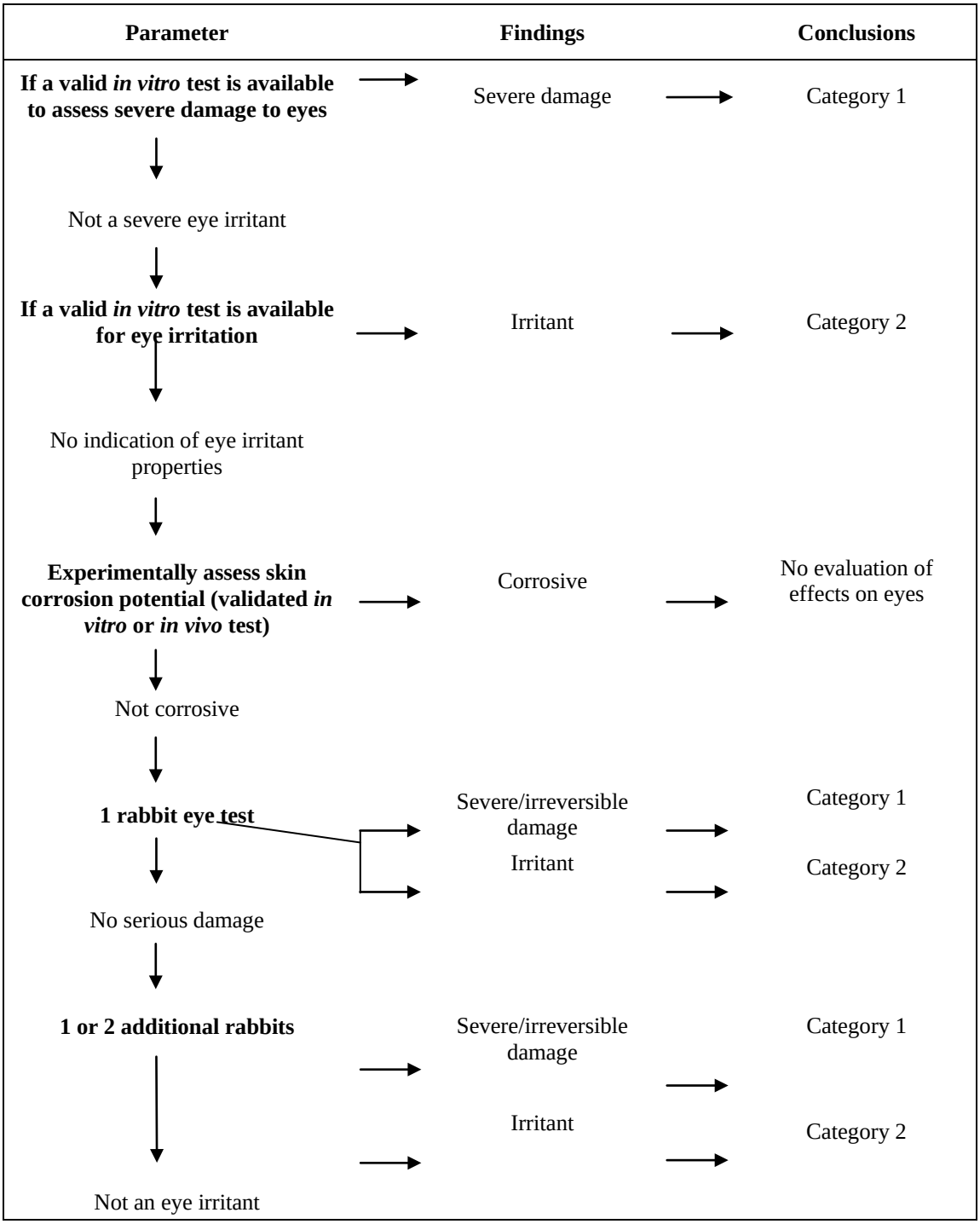
The first stage in this validation process is the preparation of a BRD that presents and evaluates the relevant data and information about the test method, including its mechanistic basis, proposed uses, reliability, and performance characteristics (ICCVAM 2003). This BRD summarizes the available information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

1.4 Search Strategies and Selection of Citations for the BCOP BRD

The BCOP test method data summarized in this BRD are based on information found in the peer-reviewed scientific literature as detailed in the *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). The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is currently evaluating a non-animal approach for assessing eye irritation potential and labeling requirements for antimicrobial cleaning

products (AMCPs). Three *in vitro* test methods, including the BCOP, are proposed in the testing strategy. The Institute for In Vitro Sciences gave the final AMCP BRD to NICEATM on July 21, 2008. Those substances in the AMCP validation database that had been tested in the BCOP test method were added to the BCOP validation database (ICCVAM 2006a). A subsequent literature search conducted in January 2009 revealed no new articles containing BCOP test method results.

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



Adapted from UN (2007).

2.0 Bovine Corneal Opacity and Permeability Test Method Protocol Components

2.1 Overview of How the BCOP Test Method is Conducted

The BCOP test method is an *in vitro* model that provides short-term maintenance of the physiological and biochemical function of the bovine cornea. In this test method, damage by the test substance is assessed by quantitative measurements of changes in corneal opacity and permeability with an opacitometer and a visible light spectrophotometer, respectively. Both measurements are used to calculate an *in vitro* irritancy score (IVIS), which is used to assign an *in vitro* irritancy hazard classification category for prediction of the *in vivo* ocular irritation potential of a test substance.

For a detailed description of how the BCOP test method is conducted, see the *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). Briefly, isolated corneas are obtained from the eyes of freshly slaughtered cattle. Test substances are applied to the epithelial surface of the cornea using different treatment methods depending on the physical nature and chemical characteristics (e.g., solids, semisolids [including creams and waxes], liquids, viscous [including gels] vs. nonviscous liquids) of the test substance. Liquids are tested undiluted, while surfactants are tested at a concentration of 10% in a 0.9% sodium chloride solution, distilled water, or other solvent demonstrated to have no adverse effects on the test system. Corneas are exposed to liquids and surfactants for 10 minutes. Nonsurfactant solids are typically tested as solutions or suspensions at a 20% concentration in a 0.9% sodium chloride solution, distilled water, or other solvent demonstrated to have no adverse effects on the test system. Solids may also be tested neat by direct application to the corneal surface. Corneas are exposed to solids for 4 hours.

Corneal opacity is quantified as the amount of light passing through the cornea, resulting in opacity values measured on a continuous scale. Permeability is quantified as the amount of sodium fluorescein dye that passes across the full thickness of the cornea, as detected in the posterior chamber medium. The mean opacity and mean permeability (OD_{490}) values for each treatment group are then used to calculate an *in vitro* score for each treatment group:

In vitro irritancy score = mean opacity value + (15 x mean OD490 value)

The *in vitro* irritation classification schemes used for this evaluation were based on two different predetermined ranges of *in vitro* scores. The differences between the two ranges are attributed to two different criteria used to identify ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41). One approach (**Table 2-1**) included the ICCVAM-recommended decision criteria for identifying an ocular corrosive/severe irritant (i.e., IVIS ≥ 55.1, ICCVAM 2006b).

Table 2-1 *In Vitro* Ocular Irritancy Classification Scheme for the BCOP Test Method (ICCVAM 2006b)

***In Vitro* Score Range**	***In Vitro* Classification**
0–3.0	Not Labeled
3.1–25	Mild Irritant
25.1–55	Moderate Irritant
≥55.1	Severe Irritant

The second approach (**Table 2-2**) included an alternative decision criterion for identifying an ocular corrosive/severe irritant in the AMCP BRD (2008) submission (i.e., IVIS ≥ 75).

Table 2-2 *In Vitro* Ocular Irritancy Classification Scheme for the BCOP Test Method (AMCP BRD 2008 Submission)

***In Vitro* Score Range**	***In Vitro* Classification**
0–3.0	Not Labeled
3.1–25	Mild Irritant
25.1–74.9	Moderate Irritant
≥75	Severe Irritant

For the purposes of this evaluation, Nonirritant = EPA Category IV, GHS Not Classified, EU Not Labeled, FHSA Not Labeled; Mild Irritant = EPA Category III, GHS Category 2B; Moderate Irritant = EPA Category II, GHS Category 2A; Severe Irritant = EPA Category I, GHS Category 1, EU Category R41. The Mild and Moderate Irritant categories were combined to generate EU Category R36. The Mild, Moderate, and Severe Irritant categories were combined to generate FHSA Irritant.

For this BRD, the *in vitro* classification was compared to the corresponding *in vivo* classification for each of the EPA, GHS, and EU classification systems (EPA 2003a; UN 2007; EU 2001). For the FHSA classification system, the *in vivo* classification was compared to the *in vitro* classification based on the EPA classification system. *In vitro* classifications of Mild, Moderate, and Severe Irritant were classified as FHSA Irritant and Nonirritant was classified as FHSA Not Labeled.

3.0 Substances Used for Validation of the Bovine Corneal Opacity and Permeability Test Method

In vitro ocular test method validation studies should evaluate an adequate sample of test substances and products from chemical and product classes that have also been evaluated using the *in vivo* rabbit eye test method. Test substances with a wide range of *in vivo* ocular responses (corrosive/severe irritant to Not Labeled) also should be assessed to determine limits to the range of responses that can be evaluated by the *in vitro* test method.

The substances tested in the BCOP test method and included in the AMCP BRD were added to BCOP data employed in the ICCVAM evaluation of the BCOP for identifying ocular corrosives and severe irritants (ICCVAM 2006a). Thus, the database in the current evaluation comprises substances from the AMCP BRD along with previously evaluated published reports (Bailey et al. 2004; Balls et al. 1995; Gautheron et al. 1994; Southee 1998; Swanson et al. 1995; Swanson and Harbell 2000).

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

As shown in **Table 3-1**, the chemical classes with the greatest amount of *in vitro* BCOP data are alcohols, carboxylic acids, esters, formulations, heterocyclic compounds, hydrocarbons, ketones, and onium compounds. Other chemical classes tested include amines, ethers/polyethers, inorganic and organic salts, and organic sulfur compounds. The formulations tested include hair shampoos, personal care cleansers, detergents, bleaches, insect repellents, petroleum products, and fabric softeners.

As shown in **Table 3-2**, the product classes tested most often in the BCOP test method are AMCPs, chemical/synthetic intermediates, cleaners, drugs/pharmaceuticals/therapeutic agents, petroleum products, shampoos, solvents, and surfactants. Other product classes tested include detergents, insect repellents, lubricants, personal care cleansers, pesticides, and plasticizers.

Table 3-1 Chemical Classes Tested in the BCOP Test Method

Chemical Class	# of Substances	Chemical Class	# of Substances
Acyl halide	3	Imide	2
Alcohol	22	Inorganic salt	6
Aldehyde	1	Ketone	12
Alkali	3	Lactone	3
Aluminum compound	1	Nitrile compound	1
Amide	2	Nitro compound	2
Amidine	6	Oil	1
Amine	10	Onium compound	12
Amino acid	4	Organic salt	3
Boron compound	1	Organic sulfur compound	5
Carboxylic acid	17	Organophosphate	1
Ester	12	Organosilicon compound	1
Ether/Polyether	9	Phenol	1
Formulation	69	Polycyclic compound	3
Heterocyclic compound	12	Terpene	1
Hydrocarbon	18	Wax	1

Table 3-2 Product Classes Tested in the BCOP Test Method

Product Class	# of Substances	Product Class	# of Substances
Adhesive	1	Fertilizer	1
Agricultural chemical	2	Flame retardant	1
Antifreeze agent	1	Flavor ingredient	3
Antimicrobial cleaning product	66	Food additive	1
Bactericide/Fungicide/ Disinfectant/Germicide	11	Herbicide	3
Beverage	1	Insect repellant	8
Bleach	3	Lubricant/lubricant additive	6
Chelating agent	2	Paint, lacquer, varnish (component)	1
Chemical/synthetic intermediate	28	Pesticide	8
Cleaner	15	Petroleum product	16
Cleanser (personal care)	13	Photographic chemical/ developing agent	2
Coupling agent	1	Plant growth regulator	2
Cutting fluid	2	Plasticizer	4
Degreaser	1	Preservative	2
Dessicant	1	Reagent	5
Detergent	11	Shampoo (hair)	14
Drug/Pharmaceutical/ Therapeutic agent and/or metabolite	17	Soap	3
Dry cleaning preparation	1	Solvent	34
Dye, in manufacture of	3	Surfactant	39
Emulsifier	1	Anionic surfactant	3
Etching and/or electroplating	2	Cationic surfactant	6
Explosive	1	Nonionic surfactant	5
Fabric softener	1	Thermometer fluid	1

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

The Draize rabbit eye test protocol used to generate the *in vivo* reference data is detailed in the ICCVAM *Test Method Evaluation Report: In Vitro Ocular Toxicity Methods for Identifying Severe Irritants and Corrosives* (2006b). A number of national and international test guidelines also describe this procedure (CPSC 2003; EPA 1998; EU 2004; OECD 2002). The subjective scoring system used to assign an ocular hazard classification is based on a discrete scale for grading the severity of ocular lesions on the cornea, iris, and conjunctiva.

Most of the BCOP studies evaluated in this BRD include *in vivo* reference data generated using the basic procedures for the Draize rabbit eye test method. NICEATM used these data to assign an ocular hazard classification according to the EPA (2003a), EU (2001), FHSA (2005), and the GHS (UN 2007) ocular irritancy classification systems (**Annex III**). Exceptions included the following:

For Gautheron et al. (1994), the *in vivo* reference data were obtained from concurrent *in vivo* studies performed by Dr. J. Giroux at the Agence du Medicament in Montpelier, France. Studies were performed according to European Economic Committee (EEC) (1984 and 1991) guidelines with a few modifications. Three rabbits were used per test substance, and a maximum average score (MAS) (Draize et al. 1944) was calculated. Only the MAS and Day 1 scores for the 52 compounds are presented in the Gautheron et al. publication. The substances were classified by the study authors according to both EEC (1984) and Kay and Calandra (1962) systems. Detailed *in vivo* data consisting of cornea, iris, and conjunctiva scores for each animal were provided by Dr. Philippe Vanparys in January 2005. Sufficient *in vivo* data were provided to allow 48 to 52 of these substances to be classified by NICEATM according to the EPA (EPA 2003a), EU (EU 2001) FHSA (2005), and GHS (UN 2007) ocular irritancy classification systems (**Annex III**).

For the European Commission/British Home Office validation study (Balls et al. 1995), modified maximum average scores (MMASs) were calculated for the 59 test substances from existing and concurrently run *in vivo* studies, all of which were performed according to OECD Test Guideline 405 and Good Laboratory Practices (GLP) guidelines. The data were generated since 1981 and met the following criteria:

- At least 3 New Zealand White rabbits were normally tested at the same time.
- A volume of 0.1 mL or the equivalent weight of substance was instilled into the conjunctival sac.
- Anesthesia was not used.
- Observations were made at least at 1, 2, and 3 days after instillation.

All 59 of these substances were classified by NICEATM according to the EU (2001) classification system, but due to lack of sufficient *in vivo* data, only 52, 55, 57, and 58 substances, respectively, were classified according to the FHSA-67% (2005), EPA (2003a), GHS (UN 2007), and the FHSA-20% ocular irritancy classification systems (**Annex III**).

For the Swanson et al. (1995) study, *in vivo* reference data were obtained from standard (100 µL of test material; 7 formulations) or modified (30 µL of test material; 13 formulations) Draize rabbit eye tests. An MAS(30) or an MAS(100) was reported for each test substance. *In vivo* categories reported in the publication are mild (2 substances), mild/moderate (2), moderate (4), moderate/severe (1), severe/corrosive (4), and corrosive (7). These categories are based on an internal classification scheme used at S.C. Johnson & Son, Inc. After publication of the study, the sponsor, S.C. Johnson & Son, Inc., assigned EPA (2003a) and GHS (UN 2007) classifications to the substances. The sponsor provided these classifications, along with detailed *in vivo* data for each test substance, to NICEATM. NICEATM verified the EPA and GHS ocular irritancy classifications for 13 of the substances and classified the

same 13 test substances based on the EU (2001) and FHSA (2005) ocular irritancy classification systems (**Annex III**). However, 11 of the test substances evaluated using a 30 μL test substance volume were not included in the accuracy analysis, because definitive classifications could not be assigned for the four regulatory ocular irritancy classification systems.

For the European Community prevalidation study of the BCOP test method (Southee 1998), cornea, iris, and conjunctiva scores for each animal for all substances were available in the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) Reference Chemicals data bank (ECETOC 1998). Fifteen of the substances have been classified by NICEATM according to the EU (2001) and FHSA-20% (2005) systems; 14 of the substances were classified according to the EPA (2003a, GHS (UN 2007) and the FHSA-67% (2005) ocular irritancy classification systems (**Annex III**).

S.C. Johnson and Son, Inc., provided detailed *in vivo* reference data for 9 of the 13 test substances evaluated in the Swanson and Harbell (2000) study of ethanol-containing insect repellent formulations. The standard Draize rabbit eye test protocol was used for these nine test substances. Each test included six animals.

ExxonMobil Biomedical Sciences, Inc., provided detailed *in vivo* reference data for the 16 petrochemical products evaluated by Bailey et al. (2004). All 16 substances had been tested previously using the standard Draize rabbit eye test protocol. Each test included either three or six animals.

4.1 *In Vivo* Classification Criteria Used for BRD Analysis

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

At least three rabbits were tested in the study unless a severe effect (e.g., corrosion of the cornea) was noted in a single rabbit. In such cases, substance classification could proceed based on the effects observed in less than three rabbits.

A volume of 0.1 mL or 0.1 g was tested in each rabbit. A study in which a lower quantity was applied to the eye was accepted for substance classification provided that a severe effect (e.g., corrosion of the cornea, lesion persistence) was observed in a rabbit.

Observations of the eye were made at least 24, 48, and 72 hours after test substance application if no severe effect was observed.

Observations of the eye were made until reversibility was assessed, typically meaning that all endpoint scores were cleared. Results from a study terminated early were not used unless the reason for the early termination was documented.

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

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

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

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

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

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

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

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

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

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

Table 4-2 Proposed FHSA "Proportionality" Criteria

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

Abbreviations: CPSC = U.S. Consumer Product Safety Commission; FHSA = Federal Hazardous Substances Act; NL = not labeled; No. = number

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

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

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

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

4.2 *In Vivo* Data Quality

Ideally, all data supporting the validity of a test method should be obtained and reported from studies conducted in accordance with GLP guidelines, which are nationally and internationally recognized rules designed to produce high-quality laboratory data and records (EPA 2003b, 2003c; FDA 2003; OECD 1998). To ensure the integrity, reliability, and accountability of a study, these guidelines provide an internationally standardized approach for the conduct of studies, reporting requirements, archival of study data and records, and information about the test protocol.

Although an attempt was made, original study records could not be obtained for the *in vivo* rabbit eye studies used to provide the comparative data in the published BCOP validation reports. Therefore, the extent to which they complied with GLP guidelines is based on the information provided in the reports themselves. Balls et al. (1995) and Southee (1998) explicitly state that GLP guidelines were followed. For the Bailey et al. (2004) report, approximately half of the *in vivo* studies were conducted according to GLP guidelines, while GLP compliance was not explicitly stated for the remaining substances. For Gautheron et al. (1994), the *in vivo* studies were conducted according to EEC test guidelines (1984 and 1991), predecessors of the current EU test guideline for eye irritation. However, this information alone does not give enough information about GLP compliance. For the remaining reports (Swanson et al. 1995 and Swanson and Harbell 2000), the extent of GLP compliance is not known.

5.0 Bovine Corneal Opacity and Permeability Test Method Data and Results

Eight reports, seven published and one unpublished, were obtained for this evaluation and used for an accuracy analysis. Test method data were extracted from seven publications, data submissions, or study reports, including Gautheron et al. (1994), Balls et al. (1995), Swanson et al. (1995), Southee (1998), Swanson and Harbell (2000), Bailey et al. (2004), and the AMCP BRD (2008). The data were sufficient for an accuracy analysis of the BCOP test method for the identification of all categories of ocular irritation. As detailed in **Section 6.0**, the data were evaluated collectively and on a per-study basis.

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

NICEATM staff made several attempts to obtain original *in vitro* and *in vivo* data from BCOP test method studies. In addition, NICEATM requested original BCOP data and *in vivo* reference data from authors of published BCOP studies. As a result of these efforts, some original BCOP test method data (i.e., corrected opacity and OD_{490} values for individual corneas) were obtained. The European Centre for the Validation of Alternative Methods (ECVAM) provided corrected opacity and OD_{490} values in a written report for 16 substances evaluated in the European Community Prevalidation Study of the BCOP (Southee 1998).

Dr. Joseph Sina also submitted corrected opacity and OD_{490} values electronically for 43 compounds; however, corresponding *in vivo* reference data was not obtained. ECVAM subsequently provided the mean opacity values, mean permeability values, and mean *in vitro* scores obtained for the 59 substances evaluated in the Balls et al. (1995) study. Dr. Freddy Van Goethem provided a summary table and individual cornea data for 52 compounds tested in the EEC validation study (Gautheron et al. 1994). S.C. Johnson & Son, Inc., provided transformed BCOP data (mean opacity, permeability, and *in vitro* scores) for the Swanson et al. (1995) and Swanson and Harbell (2000) studies. ExxonMobil Biomedical Sciences, Inc., provided detailed study reports for the Bailey et al. (2004) study.

The majority of other published BCOP reports, which are discussed in **Section 9.0**, did not contain sufficient *in vitro* or *in vivo* data with which to conduct an accuracy analysis.

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

The BCOP studies included in the accuracy analysis in this document (**Section 6.0**) evaluated variability in the BCOP test method by calculating the mean (± SD) for the opacity values and the OD_{490} values for each treatment group and control group. The mean opacity and mean permeability (OD_{490}) values for each treatment group were then used to calculate an *in vitro* irritancy score for each treatment group as follows:

In vitro irritancy score = mean opacity value + (15 x mean OD_{490} value)

Sina et al. (1995) reported that this formula was derived empirically during in-house and interlaboratory studies. The data generated for a series of 36 compounds in a multilaboratory study were subjected to a multivariate analysis to determine the equation of best fit between *in vivo* and *in vitro* data. Analysis performed by scientists at two separate companies generated nearly identical derived equations. The *in vitro* irritancy score provides a numerical value that can be used to compare the relative irritancy of test substances.

The accuracy analysis in this document focused on evaluating the ability of the BCOP test method to identify ocular corrosives and severe irritants as defined by the EPA (2003a), EU (2001), and GHS (UN 2007) hazard classification schemes. The decision criteria applied to *in vitro* data to classify a test substance as a severe ocular irritant or a nonsevere ocular irritant (i.e., mild irritant, moderate irritant)

and/or Not Labeled are similar for the four BCOP test method protocols (Gautheron et al. 1994; Balls et al. 1995; Southee 1998; Bailey et al. 2004). The *in vitro* irritation classification scheme used in these studies is similar to the decision criteria first proposed by Gautheron et al. (1994), for which *in vitro* irritancy categories were based on predetermined ranges of *in vitro* scores (see **Section 2.0**).

5.3 Summary of Results

Where provided, the specific information extracted for each substance included its name, CASRN (if available), the concentration tested, the available BCOP data (e.g., mean opacity value, mean OD_{490} value, standard deviation, number of replicates, mean *in vitro* score), the *in vitro* irritation classification of the test substance (based on the *in vitro* irritation classification scheme applied or noted by the study author), and the literature reference. Other supporting information, such as the source, purity, and physicochemical characteristics of the test substances, was included to the extent this information was available. If not provided, the CASRN was obtained from various sources, including the National Library of Medicine's ChemIDplus® database. Chemical and product classes were assigned based on the MeSH classification system (available at http://www.nlm.nih.gov/mesh). **Annex I** provides information on the names, synonyms, CASRNs, and chemical/product class, where available, for each substance. **Annex II** contains the *in vitro* BCOP test method data sorted by reference and alphabetically by substance name.

5.4 Use of Coded Chemicals and Compliance with GLP Guidelines

Ideally, all data supporting the validity of a test method should be obtained and reported in accordance with GLP guidelines and with the use of coded chemicals (OECD 1998; EPA 2003b, 2003c; FDA 2003). The data quality was evaluated by reviewing the methods sections in literature references and submitted reports. The quality of data presented in the reviewed literature references can be evaluated to the extent this information was provided in the published reports. Based on the available information, the reports that stated that they had followed GLP guidelines or used data obtained according to GLP guidelines were Bailey et al. (2004), Balls et al. (1995), and Southee (1998). The extent of GLP compliance for Swanson et al. (1995) and Swanson and Harbell (2000) were not known. The reports that said they used coded chemicals were Gautheron et al. (1994), Balls et al. (1995), Swanson et al. (1995), Southee (1998), Swanson and Harbell (2000), and Bailey et al. (2004).

6.0 Bovine Corneal Opacity and Permeability Test Method Accuracy

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

Accuracy (concordance): the proportion of correct outcomes (positive and negative) of a test method

Sensitivity: the proportion of all positive substances that are classified as positive

Specificity: the proportion of all negative substances that are classified as negative

Positive predictivity: the proportion of correct positive responses among substances testing positive

Negative predictivity: the proportion of correct negative responses among substances testing negative

False positive rate: the proportion of all negative substances that are falsely identified as positive

False negative rate: the proportion of all positive substances that are falsely identified as negative

ICCVAM evaluated the ability of the BCOP test method to identify all categories of ocular irritation potential as defined by the EPA (EPA 2003), GHS (UN 2007), and EU (EU 2001) classification systems. Given that the FHSA classification system is used to identify eye irritants based on incidence and does not differentiate between irreversible (i.e., corrosive or severe) and reversible (i.e., nonsevere) ocular effects based on Draize rabbit eye test results, an evaluation for all ocular hazard categories using the FHSA classification system was not possible.

Analyses were also performed with specific chemical classes and/or physical properties excluded based on their previous identification as discordant in the BCOP test method (ICCVAM 2006a). These evaluations were conducted on the overall data set by combining results from the reports indicated in **Section 5.0** then assigning an overall ocular irritancy classification for each substance (**Annexes II** and **III**). When the same substance was evaluated in multiple laboratories, an overall BCOP classification was based on the majority classification among all of the studies. When there were equal numbers of different irritancy classifications for substances (e.g., two tests classified a substance as Not Labeled as Irritant, and two tests classified a substance as a mild irritant), the more severe irritancy classification was used for the overall classification of the substance (mild irritant, in this case).

The *in vitro* irritation classification schemes used for this evaluation were based on two different predetermined ranges of *in vitro* scores. The differences between the two ranges are attributed to two different criteria used to identify ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, EU R41). One approach (**Table 2-1**) included the ICCVAM-recommended decision criteria for identifying an ocular corrosive/severe irritant (i.e., IVIS $\geq$ 55.1, ICCVAM 2006b). The second approach (**Table 2-2**) included an alternative decision criteria for identifying an ocular corrosive/severe irritant in the AMCP BRD (2008) submission (i.e., IVIS $\geq$75).

6.1 Ability to Distinguish Ocular Corrosives and Severe Irritants from All Other Classes

The BCOP test method has been recommended previously for use in identifying ocular corrosives and severe irritants (i.e., EPA Category I, GHS Category 1, and EU R41; ICCVAM 2006b). The original ICCVAM evaluation of the BCOP test method was based on 145 substances. Overall accuracy rates were 79% (113/143) to 81% (119/147) depending on the hazard classification system evaluation (i.e., EPA, GHS, or EU). False positive rates were 19% (20/103) to 21% (22/103), and false negative rates were 16% (7/43) to 25% (10/40), also depending on the hazard classification system.

Because additional substances with sufficient BCOP and *in vivo* data were added to the BCOP test method validation database, this evaluation was repeated to verify similar performance. Based on the

current BCOP validation database, which has increased to 211 substances, overall accuracy is 77% (91/118) to 79% (147/186) depending on the hazard classification system evaluation (i.e., EPA, GHS, or EU). The false positive rate is 23% (29/124) to 24% (29/122), and false negative rates are 15% (10/65) to 21% (7/33) depending on the hazard classification system evaluation (**Table 6-1**).

Table 6-1 Accuracy of the BCOP Test Method in Distinguishing Ocular Corrosives/Severe Irritants from All Other Categories, as Defined by the EPA, GHS, and EU Classification Systems[1]

BCOP	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
GHS	187	79	148/187	85	55/65	76	93/122	24	29/122	15	10/65
EPA	187	79	148/187	84	53/63	77	95/124	23	29/124	16	10/63
EU	118	77	91/118	79	26/33	76	65/85	24	20/85	21	7/33

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

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

The following sections provide detailed analyses and results of the performance of the BCOP test method for each of the ocular hazard classification systems (i.e., EPA, GHS, and EU).

6.2 GHS Classification System: BCOP Test Method Accuracy

This accuracy evaluation used seven reports: Gautheron et al. (1994), Balls et al. (1995), Swanson et al. (1995), Southee (1998), Swanson and Harbell (2000), Bailey et al. (2004), and the AMCP BRD (2008) submission. These included BCOP data for 211 substances, 187 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the GHS classification system (UN 2007 (see **Annex III**). Among these studies, Gautheron et al. (1994), Balls et al. (1995), and Southee (1998) provided BCOP data for substances tested in multiple laboratories. Thus a consensus *in vitro* classification had to be assigned to each substance. Based on results from *in vivo* rabbit eye experiments, 35% (65/187) were classified as Category 1, 14% (26/187) were classified as Category 2A, 3% (6/187) were classified as Category 2B, and 48% (90/187) were classified as Not Classified as Irritant. Twenty-four substances could not be classified according to the GHS classification system due to the lack of adequate animal data.

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

The BCOP test method correctly identified 85% (55/65) and 78% (51/65) of the Category 1 substances using decision criteria of IVIS ≥55.1 and IVIS ≥75, respectively (**Table 6-2**). Among the Category 1 substances that were underpredicted by BCOP (based on IVIS ≥55.1), 9% (6/65) were classified as Category 2A, and 6% (4/65) were classified as Category 2B. Among the Category 1 substances that were underpredicted by the BCOP test method (based on IVIS ≥75), 15% (10/65) were classified as Category 2A and 6% (4/65) were classified as Category 2B.

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

Of the 26 substances that could be evaluated, the BCOP test method correctly identified 27% (7/26) as moderate irritants, overpredicted 62% (16/26), and underpredicted 11% (3/26) using decision criteria defining ocular corrosives/severe irritants ≥55.1 (**Table 6-2**). Using decision criteria defining ocular

corrosives/severe irritants ≥75, the BCOP test method correctly identified 54% (14/26) as moderate irritants, overpredicted 31% (8/26), and underpredicted 15% (4/26) (**Table 6-2**).

6.2.3 Identification of Category 2B Substances (Mild Ocular Irritants)

Regardless of the decision criteria used to define ocular corrosives/severe irritants, of the six substances that could be evaluated, the BCOP test method correctly identified 33% (2/6) as mild irritants while overpredicting 67% (4/6) (**Table 6-2**).

6.2.4 Identification of Substances Not Classified as Irritant

Regardless of the decision criteria used to define ocular corrosives/severe irritants, of the 90 substances that could be evaluated, the BCOP test method correctly identified 30% (27/90) as Not Classified as Irritant while overpredicting 70% (63/90) (**Table 6-2**).

6.2.5 Overall Correct Classification

As indicated in **Table 6-2**, the use of the alternative decision criteria proposed in the AMCP BRD (2008), in which ocular corrosives/severe irritants ≥75, did not improve the overall performance of BCOP hazard classification. Therefore, the remaining analyses will present results utilizing the ICCVAM-recommended decision criteria for ocular corrosives/severe irritants (≥55.1). Overall, correct classification for the entire database of 187 substances was 49% (91/187) but ranged from 25% (2/8) to 60% (6/10 or 9/15) when each of the eight individual validation databases was evaluated (**Table 6-3**).

6.2.6 Ability to Distinguish Substances Not Classified as Irritant from All Other Classes

In addition to evaluating the ability of the BCOP test method to identify each individual ocular hazard category according to the GHS classification system, ICCVAM evaluated the ability of the BCOP test method to distinguish substances not classified as irritants from all other irritant classes. Using this approach for the 187 substances considered, the BCOP test method has an accuracy of 66% (124/187), a sensitivity of 100% (97/97), a specificity of 30% (27/90), a false positive rate of 70% (63/90) and a false negative rate of 0% (0/97) (**Table 6-4**).

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

Gautheron et al. (1994): Based upon the *in vivo* rabbit data, 47 substances could be assigned a GHS classification. Based on these 47 substances, the BCOP test method has an accuracy of 55% (26/47), sensitivity of 100% (13/13), specificity of 38% (13/34), false positive rate of 62% (21/34), and a false negative rate of 0% (0/13) (**Table 6-4**).

Balls et al. (1995): Based upon the *in vivo* rabbit data, 54 substances could be assigned a GHS classification. Based on these 54 substances, the BCOP test method has an accuracy of 83% (45/54), sensitivity of 100% (40/40), specificity of 36% (5/14), false positive rate of 64% (9/14), and a false negative rate of 0% (0/40) (**Table 6-4**).

Swanson et al. (1995): Based upon the *in vivo* rabbit data, 10 substances could be assigned a GHS classification. Based on these 10 substances, the BCOP test method has an accuracy of 60% (6/10), sensitivity of 100% (6/6), specificity of 0% (0/4), false positive rate of 100% (4/4), and a false negative rate of 0% (0/6) (**Table 6-4**).

Table 6-2 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by EPA, GHS, and EU Classification Systems[1]

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

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

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

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

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

[4] Mild = EPA Category III, GHS Category 2B, EU R36.

[5] Not Labeled = Not Labeled or Classified as Irritant.

Table 6-3 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the GHS Classification System,[1] by Study and Overall

Data Source	Overall Correct Classification	Severe (Category 1)		Moderate (Category 2A)			Mild (Category 2B)			Not Classified[2]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Gautheron et al. (1994)	43% (20/47)	75% (6/8)	25% (2/8)	67% (2/3)	33% (1/3)	0% (0/3)	100% (2/2)	0% (0/2)	0% (0/2)	62% (21/34)	38% (13/34)
Balls et al. (1995)	50% (27/54)	73% (16/22)	27% (6/22)	57% (8/14)	29% (4/14)	14% (2/14)	50% (2/4)	50% (2/4)	0% (0/4)	64% (9/14)	36% (5/14)
Swanson et al. (1995)	60% (6/10)	100% (6/6)	0% (0/6)	0% (0/0)	0% (0/0)	0% (0/0)	0% (0/0)	0% (0/0)	0% (0/0)	100% (4/4)	0% (0/4)
Southee (1998)	60% (9/15)	57% (4/7)	43% (3/7)	33% (1/3)	67% (2/3)	0% (0/3)	50% (0/2)	50% (1/2)	0% (0/2)	33% (1/3)	67% (2/3)
Swanson and Harbell (2000)	25% (2/8)	100% (1/1)	0% (0/1)	50% (2/4)	25% (1/4)	25% (1/4)	0% (0/0)	0% (0/0)	0% (0/0)	100% (3/3)	0% (0/3)
Bailey et al. (2004)	43% (6/14)	67% (2/3)	33% (1/3)	0% (0/0)	0% (0/0)	0% (0/0)	100% (1/1)	0% (0/1)	0% (0/1)	60% (6/10)	40% (4/10)
AMCP BRD (2008)	51% (33/65)	93% (27/29)	7% (2/29)	86% (6/7)	14% (1/7)	0% (0/7)	0% (0/0)	0% (0/0)	0% (0/0)	83% (24/29)	17% (5/29)
Overall	49% (91/187)	85% (55/65)	15% (10/65)	62% (16/26)	27% (7/26)	11% (3/26)	67% (4/6)	33% (2/6)	0% (0/6)	70% (63/90)	30% (27/90)

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability; BRD = background review document; GHS = Globally Harmonized System.

[1] GHS classification system (UN 2007).

[2] Not Classified = Not Classified as Irritant.

Table 6-4 Accuracy of the BCOP Test Method in Distinguishing Substances Not Classified as Irritants from All Other Irritant Classes, as Defined by the GHS Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gautheron et al. (1994)	47	55	26/47	100	13/13	38	13/34	62	21/34	0	0/13
Balls et al. (1995)	54	83	45/54	100	40/40	36	5/14	64	9/14	0	0/40
Swanson et al. (1995)	10	60	6/10	100	6/6	0	0/4	100	4/4	0	0/6
Southee (1998)	15	93	14/15	100	12/12	67	2/3	33	1/3	0	0/12
Swanson and Harbell (2000)	8	63	5/8	100	5/5	0	0/3	100	3/3	0	0/5
Bailey et al. (2004)	14	57	8/14	100	4/4	40	4/10	60	6/10	0	0/4
AMCP BRD (2008)	65	63	41/65	100	36/36	17	5/29	83	24/29	0	0/36
Overall	187	66	124/187	100	97/97	30	27/90	70	63/90	0	0/97

Abbreviations: BCOP = bovine corneal opacity and permeability; GHS = Globally Harmonized System; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] GHS (UN 2007): NL vs. Categories 1/2A/2B.

Southee (1998): Based upon the *in vivo* rabbit data, 15 substances could be assigned a GHS classification. Based on these 15 substances, the BCOP test method has an accuracy of 93% (14/15), sensitivity of 100% (12/12), specificity of 67% (2/3), false positive rate of 33% (1/3), and a false negative rate of 0% (0/12) (**Table 6-4**).

Swanson and Harbell (2000): Based upon the *in vivo* rabbit data, eight substances could be assigned a GHS classification. Based on these eight substances, the BCOP test method has an accuracy of 63% (5/8), sensitivity of 100% (5/5), specificity of 0% (0/3), false positive rate of 100% (3/3), and a false negative rate of 0% (0/5) (**Table 6-4**).

Bailey et al. (2004): Based upon the *in vivo* rabbit data, 14 substances could be assigned a GHS classification. Based on these 14 substances, the BCOP test method has an accuracy of 57% (8/14), sensitivity of 100% (4/4), specificity of 40% (4/10), false positive rate of 60% (6/10), and a false negative rate of 0% (0/4) (**Table 6-4**).

AMCP BRD (2008): Based upon the *in vivo* rabbit data, 65 substances could be assigned a GHS classification. Based on these 65 substances, the BCOP test method has an accuracy of 63% (41/65), sensitivity of 100% (36/36), specificity of 17% (5/29), false positive rate of 83% (24/29), and a false negative rate of 0% (0/36) (**Table 6-4**).

6.2.7 Discordant Results According to the GHS Classification System

In order to evaluate BCOP test method responses that disagreed with the *in vivo* hazard classification, several accuracy subanalyses were performed. These included specific classes of chemicals and certain properties of interest considered relevant to ocular toxicity testing (e.g., surfactants, physical form) with sufficiently robust numbers of substances ($n \geq 5$).

Table 6-5 shows some notable trends in the performance of the BCOP test method among these subgroups of substances. The chemical classes of substances that the BCOP test method most consistently overpredicted according to the GHS classification system were alcohols and hydrocarbons. Of the 53 overpredicted substances, eight were alcohols and eight were hydrocarbons. Additional chemical classes represented among the overpredicted substances were carboxylic acids (6), heterocyclic compounds (4), and esters (4). Among the 23 substances labeled as surfactants, the BCOP test method overpredicted 22% (5/23).

Forty-four of the substances overpredicted by the BCOP test method were liquids, and nine were solids. Considering the proportion of the total available database, the BCOP test method appears more likely to overpredict liquids (90/122 or 74%) than solids (32/122 or 26%).

Alcohols (2) and carboxylic acids (2) were most often underpredicted (i.e., false negatives[1]) by the BCOP test method according to the GHS classification system (see **Annex III**). As can be seen in **Table 6-5**, the 16 irritant substances labeled as surfactants were rarely underpredicted by the BCOP test method (7% [1/14] Category 1 substances was underpredicted; none of the Category 2A or 2B substances were underpredicted).

With regard to physical form, six of the substances underpredicted by the BCOP test method were liquids and five were solids. Given the proportion of the total available database, the BCOP test method appears more likely to underpredict solids (32/122 or 26%) than liquids (90/122 or 74%).

Table 6-6 shows the effects on the BCOP test method performance statistics of excluding from the data set problematic classes (i.e., those which gave the most discordant results according to the GHS classification system) identified in the BCOP BRD (ICCVAM 2006a). In general, exclusion of alcohols, ketones, or solids individually resulted in small changes in the performance statistics. Slight increases in the overall correct classification were noted with the exclusion of problematic classes, with the highest correct classification, 51% (49/97), noted when alcohols and ketones were both excluded. The exclusion of problematic classes had little impact on the ability of the BCOP test method to identify substances not labeled as irritants (see **Table 6-7**). Accuracy was 68% (83/122) with the entire database but ranged from 64% to 69% when problematic classes or combinations were excluded.

In **Table 6-5**, hydrocarbons are noted as discordant when the BCOP test method was evaluated for its ability to identify all hazard categories. Among the 11 hydrocarbons in the validation database, 73% (8/11) were overpredicted by the BCOP test method (**Table 6-5**). Compared to the entire database, exclusion of hydrocarbons resulted in only modest improvements in overall correct classification (50% [55/111] versus 48% (58/122]) and identification of Not Labeled substances (38% [19/50] versus 36% [22/61]) (**Table 6-6**). Exclusion of hydrocarbons also resulted in modest improvement in overall performance in identifying Not Labeled substances (see **Table 6-7**). Accuracy increased from 68% (83/112) to 72% (80/111). The false positive rate decreased from 64% (39/61) to 62% (31/50), while the false negative rate remained 0% (0/61 versus 0/61).

---

[1] *False negative* in this context refers to a substance that the BCOP test method classified as a nonsevere (mild or moderate) irritant or Not Labeled but that the *in vivo* data classified as a severe irritant.

Table 6-5 Under- and Overprediction of the BCOP Test Method Using the GHS Classification System[1] in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method by Chemical Class or Physical Property

Category	N	Underprediction (*In Vivo*/*In Vitro*)						Overprediction (*In Vivo*/*In Vitro*)					
		Severe (Category 1)			Moderate (Category 2A)		Mild (Cat 2B)	Moderate (Cat 2A)	Mild (Category 2B)		Not Classified as Irritant (NL)		
		2A	2B	NL	2B	NL	NL	1	1	2A	1	2A	2B
Overall	147	11% (4/36)	11% (4/36)	0% (0/36)	16% (3/19)	0% (0/19)	0% (0/6)	53% (10/19)	17% (1/6)	50% (3/6)	15% (9/61)	11% (7/61)	38% (23/61)
Chemical Class[2]													
Alcohol	18	33% (1/3)	33% (1/3)	0% (0/3)	0% (0/6)	0% (0/6)	0% (0/1)	67% (4/6)	0% (0/1)	100% (1/1)	43% (3/7)	0% (0/7)	0% (0/7)
Amine/Amidine	7	0% (0/5)	0% (0/5)	0% (0/5)	0% (0/2)	0% (0/2)	0/0	0/0	0/0	0/0	0% (0/4)	0% (0/4)	25% (1/4)
Carboxylic acid	14	0% (0/6)	33% (2/6)	0% (0/6)	0% (0/2)	0% (0/2)	0/0	50% (1/2)	0/0	0/0	33% (2/6)	33% (2/6)	17% (1/6)
Ester	10	0% (0/2)	0% (0/2)	0% (0/2)	33% (1/3)	0% (0/3)	0% (0/1)	33% (1/3)	0% (0/1)	0% (0/1)	0% (0/4)	5% (2/4)	25% (1/4)
Ether	6	0% (0/1)	0% (0/1)	0% (0/1)	0% (0/1)	0% (0/1)	0/0	100% (1/1)	0/0	0/0	25% (1/4)	0% (0/4)	0% (0/4)
Heterocyclic	13	0% (0/6)	17% (1/6)	0% (0/6)	0% (0/1)	0% (0/1)	0/0	0% (0/1)	0/0	0/0	17% (1/6)	0% (0/6)	50% (3/6)
Hydrocarbon	11	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	9% (1/11)	18% (2/11)	45% (5/11)
Inorganics	7	0% (0/4)	0% (0/4)	0% (0/4)	0/0	0/0	0% (0/1)	0/0	0% (0/1)	0% (0/1)	0% (0/2)	0% (0/2)	50% (1/2)
Ketone	9	0/0	0/0	0/0	0% (0/2)	0% (0/2)	0% (0/1)	0% (0/2)	0% (0/1)	0% (0/1)	33% (2/6)	0% (0/6)	17% (1/6)

Category	N	Underprediction (*In Vivo*/*In Vitro*)						Overprediction (*In Vivo*/*In Vitro*)					
		Severe (Category 1)			Moderate (Category 2A)		Mild (Cat 2B)	Moderate (Cat 2A)	Mild (Category 2B)		Not Classified as Irritant (NL)		
		2A	2B	NL	2B	NL	NL	1	1	2A	1	2A	2B
Onium compound	11	13% (1/8)	0% (0/8)	0% (0/8)	0/0	0/0	0% (0/1)	0/0	0% (0/1)	0% (0/1)	0% (0/2)	0% (0/2)	50% (1/2)
Polyether	2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0% (0/2)	0% (0/2)	0% (0/2)
Properties of Interest													
Liquids	90	8% (2/24)	4% (1/24)	0% (0/24)	18% (3/17)	0% (0/17)	0% (0/5)	53% (9/17)	20% (1/5)	60% (3/5)	16% (7/44)	16% (7/44)	39% (17/44)
Solids	32	17% (2/12)	25% (3/12)	0% (0/12)	0% (0/2)	0% (0/2)	0/0	50% (1/2)	0/0	0/0	12% (2/17)	0% (0/17)	35% (6/17)
Pesticide	8	20% (1/5)	20% (1/5)	0% (0/5)	0% (0/1)	0% (0/1)	0/0	100% (1/1)	0/0	0/0	50% (1/2)	0% (0/2)	50% (1/2)
Surfactant-Total	23	0% (0/14)	7% (1/14)	0% (0/14)	0% (0/1)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/1)	0% (0/1)	0% (0/7)	14% (1/7)	43% (3/7)
-nonionic	10	0% (0/5)	0% (0/5)	0% (0/5)	0% (0/1)	0% (0/1)	0/0	100% (1/1)	0/0	0/0	0% (0/4)	0% (0/4)	0% (0/4)
-anionic	9	0% (0/5)	20% (1/5)	0% (0/5)	0/0	0/0	0% (0/1)	0/0	0% (0/1)	0% (0/1)	0% (0/3)	33% (1/3)	67% (2/3)
-cationic	7	0% (0/6)	0% (0/6)	0% (0/6)	0/0	0/0	0/0	0/0	0/0	0/0	0% (0/1)	0% (0/1)	100% (1/1)

Abbreviations: BCOP = bovine corneal opacity and permeability; GHS = Globally Harmonized System.

[1] GHS classification system (UN 2007).

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

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

BCOP	Overall Correct Classification	Severe (Category 1)		Moderate (Category 2A)			Mild (Category 2B)			Not Classified[2]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall	48% (58/122)	78% (28/36)	22% (8/36)	53% (10/19)	32% (6/19)	15% (3/19)	67% (4/6)	33% (2/6)	0% (0/6)	64% (39/61)	36% (22/61)
Without Alcohols	49% (52/106)	82% (27/33)	18% (6/33)	46% (6/13)	31% (4/13)	23% (3/13)	60% (3/5)	40% (2/5)	0% (0/5)	65% (36/55)	35% (19/55)
Without Ketones	49% (55/113)	78% (28/36)	22% (8/36)	47% (8/17)	35% (6/17)	18% (3/17)	80% (4/5)	20% (1/5)	0% (0/5)	64% (35/55)	36% (20/55)
Without Solids	44% (40/90)	88% (21/24)	13% (3/24)	53% (9/17)	29% (5/17)	18% (3/17)	80% (4/5)	20% (1/5)	0% (0/5)	70% (31/44)	30% (13/44)
Without Alcohols and Ketones	51% (49/97)	82% (27/33)	18% (6/33)	36% (4/11)	36% (4/11)	27% (3/11)	75% (3/4)	25% (1/4)	0% (0/4)	65% (32/49)	35% (17/49)
Without Alcohols, Ketones, and Solids	47% (31/66)	91% (20/22)	9% (2/22)	33% (3/9)	34% (3/9)	33% (3/9)	100% (3/3)	0% (0/3)	0% (0/3)	75% (24/32)	25% (8/32)
Without Hydrocarbons	50% (55/111)	78% (28/36)	22% (8/36)	53% (10/19)	32% (6/19)	15% (3/19)	67% (4/6)	33% (2/6)	0% (0/6)	62% (31/50)	38% (19/50)

Abbreviations: BCOP = bovine corneal opacity and permeability; GHS = Globally Harmonized System.

[1] GHS classification system (UN 2007).

[2] Not Classified = Not Classified as Irritant.

Table 6-7 Accuracy of the BCOP Test Method in Distinguishing Substances Not Classified as Irritants from All Other Irritant Classes, as Defined by the GHS Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	187	66	124/187	100	97/97	30	27/90	70	63/90	0	0/97
Without Alcohols	106	66	70/106	100	51/51	35	19/55	65	36/55	0	0/51
Without Ketones	113	69	78/113	100	58/58	36	20/55	64	65/55	0	0/58
Without Solids	90	66	59/90	100	46/46	30	13/44	70	31/44	0	0/46
Without Alcohols and Ketones	97	67	65/97	100	48/48	35	17/49	65	32/49	0	0/48
Without Alcohols, Ketones, and Solids	66	64	42/66	100	34/34	25	8/32	75	24/32	0	0/34
Without Hydrocarbons	111	72	80/111	100	61/61	38	19/50	62	31/50	0	0/61

Abbreviations: BCOP = bovine corneal opacity and permeability; GHS = Globally Harmonized System; N = number of substances included in this analysis/the total number of substances in the study; NL = Not Labeled; No. = data used to calculate the percentage.

[1] GHS classification system (UN 2007): NL vs. Categories 1/2A/2B.

6.3 EPA Classification System: BCOP Test Method Accuracy

The seven reports used in the accuracy evaluation (Gautheron et al. 1994; Balls et al. 1995 ; Swanson et al. 1995 ; Southee 1998; Swanson and Harbell 2000; Bailey et al. 2004; and the AMCP BRD 2008) included BCOP data on 211 substances, 187 of which had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the EPA classification system (EPA 2003a) (see **Annex III**). Among these studies, Gautheron et al. (1994), Balls et al. (1995), and Southee (1998) provided BCOP data for substances tested in multiple laboratories and thus required that a consensus *in vitro* classification be assigned to each substance. Based on results from *in vivo* rabbit eye experiments, 35% (65/187) were classified as Category I, 14% (26/187) were classified as Category II, 3% (6/187) were classified as Category III, and 48% (90/187) were classified as Category IV. Twenty-four substances could not be classified according to the GHS classification system due to the lack of adequate animal data (noted in **Annex III)**.

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

The BCOP test method correctly identified 84% (53/63) and 78% (49/63) of the Category I substances using decision criteria defining ocular corrosives/severe irritants ≥55.1 and ocular corrosives/severe irritants ≥75, respectively (**Table 6-2**). Using decision criteria defining *in vitro* scores ≥55.1 as ocular corrosives/severe irritants, of the Category I substances that were underpredicted by the BCOP test method, 10% (6/63) were classified as Category II, and 6% (4/63) were classified as Category III. Using decision criteria defining *in vitro* scores ≥75 as ocular corrosives/severe irritants, of the Category I substances that were underpredicted by BCOP test method, 16% (10/63) were classified as Category II, and 6% (4/63) were classified as Category III.

6.3.2 Identification of Category II Substances (Moderate Ocular Irritants)

Of the 22 substances that could be evaluated, the BCOP test method correctly identified 32% (7/22) as moderate irritants, while 50% (11/22) were overpredicted and 18% (4/22) were underpredicted using decision criteria that defined *in vitro* scores ≥55.1 as ocular corrosives/severe irritants (**Table 6-8**). Using decision criteria defining *in vitro* scores ≥75 as ocular corrosives/severe irritants, the BCOP test method correctly identified 45% (10/22) as moderate irritants, while overpredicting 36% (8/22) and underpredicting 19% (4/22) (**Table 6-2**).

6.3.3 Identification of Category III Substances (Mild Ocular Irritants)

Using decision criteria defining *in vitro* scores ≥55.1 as ocular corrosives/severe irritants, for the 56 substances that could be evaluated, the BCOP test method correctly identified 36% (21/57) as mild irritants, while 50% (28/57) were overpredicted and 14% (8/57) were underpredicted (**Table 6-8**). Using decision criteria defining *in vitro* scores ≥75 as ocular corrosives/severe irritants, for the 57 substances that could be evaluated, the BCOP test method correctly identified 39% (22/57) as mild irritants, while 47% (27/57) were overpredicted and 14% (8/57) were underpredicted (**Table 6-2**).

6.3.4 Identification of Category IV Substances

Regardless of the decision criteria used to define *in vitro* scores as ocular corrosives/severe irritants, for the 45 substances that could be evaluated, the BCOP test method correctly identified 47% (21/45) as Category IV, while 53% (24/45) were overpredicted (**Tables 6-2 and 6-8**).

6.3.5 Ability to Distinguish Category IV from All Other Classes

In addition to evaluating the ability of the BCOP test method to identify each individual ocular hazard category according to the EPA classification system, ICCVAM also evaluated the ability of the BCOP

test method to distinguish Category IV from all other irritant classes. Using this approach for the 187 substances considered, the BCOP test method has an accuracy of 83% (155/187), a sensitivity of 94% (134/142), a specificity of 47% (21/45), a false positive rate of 53% (24/45), and a false negative rate of 6% (8/142) (**Table 6-9**).

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

Gautheron et al. (1994): Based upon the *in vivo* rabbit data, 48 substances could be assigned an EPA classification. Based on these 48 substances, the BCOP test method has an accuracy of 83% (40/48), sensitivity of 89% (31/35), specificity of 69% (9/13), false positive rate of 31% (4/13), and a false negative rate of 11% (4/35) (**Table 6-9**).

Balls et al. (1995): Based upon the *in vivo* rabbit data, 54 substances could be assigned an EPA classification. Based on these 54 substances, the BCOP test method has an accuracy of 93% (50/54), sensitivity of 92% (48/52), specificity of 100% (2/2), false positive rate of 0% (0/2), and a false negative rate of 8% (4/52) (**Table 6-9**).

Swanson et al. (1995): Based upon the *in vivo* rabbit data 10 substances could be assigned an EPA classification. Based on these 10 substances, the BCOP test method has an accuracy of 90% (9/10), sensitivity of 100% (9/9), specificity of 0% (0/1), false positive rate of 100% (1/1), and a false negative rate of 0% (0/9) (**Table 6-9**).

Southee (1998): Based upon the *in vivo* rabbit data, 15 substances could be assigned an EPA classification. Based on these 15 substances, the BCOP test method has an accuracy of 93% (14/15), sensitivity of 93% (13/14), specificity of 100% (1/1), false positive rate of 0% (0/1), and a false negative rate of 7% (0/14) (**Table 6-9**).

Swanson and Harbell (2000): Based upon the *in vivo* rabbit data, eight substances could be assigned an EPA classification. Based on these eight substances, the BCOP test method has an accuracy of 75% (6/8), sensitivity of 100% (6/6), specificity of 0% (0/2), false positive rate of 100% (2/2), and a false negative rate of 0% (0/6) (**Table 6-9**).

Bailey et al. (2004): Based upon the *in vivo* rabbit data, 13 substances could be assigned an EPA classification. Based on these 13 substances, the BCOP test method has an accuracy of 62% (8/13), sensitivity of 75% (3/4), specificity of 44% (4/9), false positive rate of 56% (5/9), and a false negative rate of 25% (1/4) (**Table 6-9**).

AMCP BRD (2008): Based upon the *in vivo* rabbit data, 66 substances could be assigned an EPA classification. Based on these 66 substances, the BCOP test method has an accuracy of 79% (52/66), sensitivity of 98% (47/48), specificity of 28% (5/18), false positive rate of 72% (13/18), and a false negative rate of 2% (1/48) (**Table 6-9**).

Table 6-8 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System,[1] by Study and Overall

Data Source	Overall Correct Classification	Severe (Category I)		Moderate (Category II)			Mild (Category III)			Not Labeled[2] (Category IV)	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Gautheron et al. (1994)	52% (25/48)	75% (6/8)	25% (2/8)	75% (3/4)	25% (1/4)	0% (0/4)	44% (10/23)	39% (9/23)	17% (4/23)	31% (4/13)	69% (9/13)
Balls et al. (1995)	46% (25/54)	68% (13/19)	32% (6/19)	50% (6/12)	33% (4/12)	17% (2/12)	52% (11/21)	29% (6/21)	19% (4/21)	0% (0/2)	100% (2/2)
Swanson et al. (1995)	60% (6/10)	100% (6/6)	0% (0/6)	0% (0/0)	0% (0/0)	0% (0/0)	100% (3/3)	0% (0/3)	0% (0/3)	100% (0/1)	0% (0/1)
Southee (1998)	47% (7/15)	50% (3/6)	50% (3/6)	50% (1/2)	50% (1/2)	0% (0/2)	50% (3/6)	33% (2/6)	17% (1/6)	0% (0/1)	100% (1/1)
Swanson and Harbell (2000)	50% (4/8)	100% (3/3)	0% (0/3)	0% (0/2)	50% (1/2)	50% (1/2)	100% (1/1)	0% (0/1)	0% (0/1)	100% (2/2)	0% (0/2)
Bailey et al. (2004)	38% (5/13)	0% (0/1)	100% (1/1)	0% (0/0)	0% (0/0)	0% (0/0)	33% (1/3)	33% (1/3)	33% (1/3)	56% (5/9)	44% (4/9)
AMCP BRD (2008)	62% (41/66)	94% (29/31)	6% (2/31)	60% (3/5)	20% (1/5)	20% (1/5)	42% (5/12)	50% (6/12)	8% (1/12)	72% (13/18)	28% (5/18)
Overall	55% (102/187)	84% (53/63)	16% (10/63)	50% (11/22)	32% (7/22)	18% (4/22)	50% (28/57)	36% (21/57)	14% (8/57)	53% (24/45)	47% (21/45)

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability; BRD = background review document; EPA = U.S. Environmental Protection Agency.

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

[2] Not Labeled = Category IV.

Table 6-9 Accuracy of the BCOP Test Method in Distinguishing Category IV Ocular Irritants from All Other Irritant Classes, as Defined by the EPA Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gautheron et al. (1994)	48	83	40/48	89	31/35	69	9/13	31	4/13	11	4/35
Balls et al. (1995)	54	93	50/54	92	48/52	100	2/2	0	0/2	8	4/52
Swanson et al. (1995)	10	90	9/10	100	9/9	0	0/1	100	1/1	0	0/9
Southee (1998)	15	93	14/15	93	13/14	100	1/1	0	0/1	7	0/14
Swanson and Harbell (2000)	8	75	6/8	100	6/6	0	0/2	100	2/2	0	0/6
Bailey et al. (2004)	13	62	8/13	75	3/4	44	4/9	56	5/9	25	1/4
AMCP BRD (2008)	66	79	52/66	98	47/48	28	5/18	72	13/18	2	1/48
Overall	187	83	155/187	94	134/142	47	21/45	53	24/45	6	8/142

Abbreviations: AMCP = antimicrobial cleaning products; BCOP = bovine corneal opacity and permeability; BRD = background review document; EPA = U.S. Environmental Protection Agency; N = number of substances included in this analysis; No. = data used to calculate the percentage.

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

Among the eight false negatives for the EPA system, 100% (8/8) were EPA Category III substances based on Draize test data. For 38% (3/8) of these substances, the categorization was based on at least one rabbit with a corneal opacity score of 1 that was not resolved until Day 3 of the study. Another substance was categorized based on all six rabbits with a conjunctival redness score of 3 that was not resolved until Day 7 of the study. Among the seven false negative substances for which chemical class and/or physical properties could be assigned, 71% (5/7) were from discordant classes that have previously been identified for the BCOP test method (i.e., either ketones or solids; see also ICCVAM 2006a). Chemical class information was unavailable for the one substance that was from the AMCP BRD 2008 (**Table 6-10**).

Table 6-10 BCOP False Negative Substances[1]

Substance (Discordant Class Y/N)	*In Vivo* Classification					*In Vivo* Scores		
	EPA	GHS	EU	FHSA-20%	FHSA-67%	N	Corneal Opacity: Score (Day Cleared)	Conjunctival Redness: Score (Day Cleared)
Dimethylbiquanide (Y)	III	NC	NL	Irr	Irr	3	N=1 1(2) N=1 1(3)	N=2 2(3)
EDTA (Y)	III	NC	NL	Irr	Irr	3	N=1 1(3)	N=3 2(2)
Iminodibenzyl (Y)	III	NC	NL	Irr	Irr	3	N=3 1(2)	-
Magnesium Carbonate (Y)	III	NC	NL	Irr	Irr	3	N=1 1(2) N=1 1(3)	-
Methylcyclopentane (Y)	III	NC	NL	NL	NL	6	-	N=1 2(3)
Polyalkenylsuccinate ester/amine salt (N)	III	SCNM	SCNM	Irr	Irr	6	N=2 1(2)	N=1 2(6), N=3 2(2) N=1 3(2) N=1 3(6)
Tween 20 (N)	III	NC	NL	Irr	FTR	4	-	N=2 2(2)
Compound I (Disinfectant/ Cleaner; Unknown)	III	SCNM	SCNM	NI	NI	6	N=1 1(2)	N=2 1(2)
L-Aspartic acid (Y)	SCNM	SCNM	SCNM	Irr	Irr	3	N=1 1(3), N=1 1(>3) N=1 3(>3)	N=3 3(2)
DL-Glutamic acid (Y)	SCNM	SCNM	SCNM	Irr	FTR	3	N=1 1(2)	-

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency; EU = European Union; FTR = further testing required; GHS = Globally Harmonized System; Irr = irritant; N = number of animals; NC = Not Classified (as irritant); NL = Not Labeled (as irritant); SCNM = study criteria not met.

For the purposes of this evaluation, *clearing* is defined in the EPA hazard classification system as corneal opacity or iritis scores = 0 or redness or chemosis scores = 1; in the GHS and EU hazard classification systems as corneal opacity, iritis, redness, or chemosis scores = 0.

[1] False negative compounds (shaded here) are those that test as nonirritants *in vitro* but are mild, moderate, or severe ocular irritants/corrosive *in vivo*, i.e., EPA Categories I, II, and III; GHS Categories 1, 2A, and 2B; and EU R41 and R36.

6.3.6 Discordant Results According to the EPA Classification System

In order to evaluate discordant responses of the BCOP test method relative to the *in vivo* hazard classification, several accuracy subanalyses were performed. These included specific classes of chemicals with sufficiently robust numbers of substances ($n \geq 5$), as well as certain properties of interest considered relevant to ocular toxicity testing (e.g., pesticides, surfactants, pH, physical form). **Table 6-11** shows some notable trends in the performance of the BCOP test method among these subgroups of substances. According to the EPA classification system, alcohols are the chemical class most consistently overpredicted by the BCOP test method. Nine of the 41 overpredicted substances were alcohols. Additional chemical classes represented among the overpredicted substances were hydrocarbons (6), carboxylic acids (5), ketones (4), esters (4), ethers (3), inorganic (1), and onium compounds (1). Among the substances labeled as surfactants, the BCOP test method overpredicted 32% (7/22).

Thirty-seven of the substances overpredicted by the BCOP test method were liquids and four were solids. Considering the proportion of the total available database, liquids (89/121 or 74%) appear more likely than solids (32/121 or 26%) to be overpredicted by the BCOP test method. Among the 22 substances labeled as surfactants, the BCOP test method overpredicted 32% (7/22).

According to the EPA classification system (see **Annex III**), the BCOP test method underpredicted relatively few substances (i.e., false negatives). Alcohols (2), esters (2), and heterocyclic compounds were most often underpredicted. As can be seen in **Table 6-11**, the 19 irritant substances labeled as surfactants were rarely underpredicted by the BCOP test method (9% [1/11] Category 1 substances were underpredicted; no Category II were underpredicted and 17% [1/6] Category III substances were underpredicted).

Nine of the substances underpredicted by the BCOP test method were solids, and nine were liquids. Given the proportion of the total available database, the BCOP test method appears more likely to underpredict solids (32/121 or 26%) than liquids (89/121 or 74%).

Table 6-12 shows the effects on the BCOP test method performance statistics of excluding from the data set problematic classes (i.e., those that gave the most discordant results according to the EPA classification system) identified in the BCOP BRD (ICCVAM 2006a). In general, the exclusion of alcohols, ketones, or solids individually resulted in small changes in the performance statistics. Exclusion of both alcohols and ketones improved the overall classification rate: 56% (54/96) versus 51% (62/121) for all compounds in the database. The classification of ocular corrosives/severe irritants was most improved by the exclusion of problematic classes. Using the entire database, 75% (24/32) of severe ocular corrosives/severe irritants were accurately classified. Removal of solids resulted in 86% (18/21) correct classification. Removal of alcohols, ketones, and solids resulted in correct classification of 90% (18/20) of Category I substances.

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

Category	N	Underprediction (*In Vivo*/*In Vitro*)						Overprediction (*In Vivo*/*In Vitro*)					
		Severe (Category I)			Moderate (Category II)		Mild (Cat III)	Moderate (Cat II)	Mild (Category III)		Not Labeled (Category IV)		
		II	III	IV	III	IV	IV	I	I	II	I	II	III
Overall	121	13% (4/32)	13% (4/32)	0% (0/32)	18% (3/17)	0% (0/17)	16% (7/45)	47% (8/17)	29% (13/45)	20% (9/45)	4% (1/27)	0% (0/27)	37% (10/27)
Chemical Class[2]													
Alcohol	17	50% (1/2)	50% (1/2)	0% (0/2)	0% (0/6)	0% (0/6)	0% (0/5)	67% (4/6)	80% (4/5)	20% (1/5)	0% (0/4)	0% (0/4)	0% (0/4)
Amine\Amidine	7	0% (0/2)	0% (0/2)	0% (0/2)	0/0	0/0	50% (2/4)	0/0	0% (0/4)	25% (1/4)	0% (0/1)	0% (0/1)	0% (0/1)
Carboxylic Acid	15	0% (0/7)	0% (0/7)	0% (0/7)	0% (0/2)	0% (0/2)	20% (1/5)	50% (1/2)	20% (1/5)	40% (2/5)	100% (1/1)	0% (0/1)	0% (0/1)
Ester	10	0% (0/1)	0% (0/1)	0% (0/1)	25% (1/4)	0% (0/4)	20% (1/5)	50% (2/4)	0% (0/5)	40% (2/5)	0/0	0/0	0/0
Ether	6	0/0	0/0	0/0	0% (0/4)	0% (0/4)	0% (0/2)	100% (1/1)	67% (2/3)	0% (0/3)	0% (0/4)	0% (0/4)	0% (0/4)
Heterocyclic	12	0% (0/5)	20% (1/5)	0% (0/5)	0/0	0/0	25% (1/4)	0% (0/1)	20% (1/5)	0% (0/5)	0% (0/1)	0% (0/1)	0% (0/1)
Hydrocarbon	11	0/0	0/0	0/0	0% (0/4)	0% (0/4)	0% (0/2)	0/0	20% (1/5)	40% (2/5)	0% (0/6)	0% (0/6)	50% (3/6)
Inorganics	7	0% (0/3)	0% (0/3)	0% (0/3)	0% (0/1)	0% (0/1)	33% (1/3)	100% (1/1)	0% (0/3)	0% (0/3)	0/0	0/0	0/0
Ketone	10	0/0	0/0	0/0	0% (0/1)	0% (0/1)	14% (1/7)	100% (1/1)	43% (3/7)	0% (0/7)	0% (0/1)	0% (0/1)	0% (0/1)

Category	N	Underprediction (*In Vivo*/*In Vitro*)						Overprediction (*In Vivo*/*In Vitro*)					
		Severe (Category I)			Moderate (Category II)		Mild (Cat III)	Moderate (Cat II)	Mild (Category III)		Not Labeled (Category IV)		
		II	III	IV	III	IV	IV	I	I	II	I	II	III
Onium Compound	10	17% (1/6)	0% (0/6)	0% (0/6)	0% (0/1)	0% (0/1)	0% (0/2)	100% (1/1)	0% (0/2)	0% (0/2)	0% (0/1)	0% (0/1)	0% (0/1)
Polyether	2	0/0	0/0	0/0	0/0	0/0	100% (1/1)	0/0	0% (0/1)	0% (0/1)	0% (0/1)	0% (0/1)	0% (0/1)
Properties of Interest													
Liquids	89	10% (2/21)	5% (1/21)	0% (0/21)	20% (3/15)	0% (0/15)	9% (3/33)	47% (7/15)	36% (12/33)	27% (9/33)	0% (0/20)	0% (0/20)	45 (9/20)
Solids	32	18% (2/11)	27% (3/11)	0% (0/11)	0% (0/2)	0% (0/2)	36% (4/11)	50% (1/2)	9% (1/11)	0% (0/11)	14% (1/7)	0% (0/7)	14% (1/7)
Pesticide	9	20% (1/5)	20% (1/5)	0% (0/5)	0/0	0/0	0% (0/4)	0/0	67% (2/3)	0% (0/3)	0/0	0/0	0/0
Surfactant–Total	22	0% (0/11)	9% (1/11)	0% (0/11)	0% (0/2)	0% (0/2)	17% (1/6)	100% (2/2)	33% (2/6)	33% (2/6)	0% (0/3)	0% (0/3)	33% (1/3)
-nonionic	11	0% (0/4)	0% (0/4)	0% (0/4)	0% (0/1)	0% (0/1)	33% (1/3)	100% (1/1)	67% (2/3)	0% (0/3)	0% (0/3)	0% (0/3)	33% (1/3)
-anionic	8	0% (0/5)	20% (1/5)	0% (0/5)	0/0	0/0	0% (0/2)	0/0	0% (0/2)	100% (2/2)	0% (0/1)	0% (0/1)	100% (1/1)
-cationic	6	0% (0/4)	0% (0/4)	0% (0/4)	0% (0/1)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/1)	0% (0/1)	0/0	0/0	0/0

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency.

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

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

As shown in **Table 6-11**, hydrocarbons were also noted as discordant when the BCOP test method was evaluated for its ability to identify all hazard categories. Among the 11 hydrocarbons in the validation database, the BCOP test method overpredicted 55% (6/11) (**Table 6-11**). Compared to the entire database, exclusion of hydrocarbons resulted in only modest improvement in overall correct classification (52% [57/110] versus 51% [62/121]) and identification of Category IV substances (62% [13/21] versus 59% [16/27]) (**Table 6-12**). Accuracy increased from 85% (103/121) to 86% (95/110), and the false positive rate decreased from 41% (11/27) to 38% (8/21). However, exclusion of hydrocarbons slightly increased the false negative rate from 7% (7/94) to 8% (7/89).

Table 6-13 shows how the ability of the BCOP test method to distinguish Category IV substances was affected by exclusion of problematic classes from the data set. Exclusion of problematic classes individually or in combination had little effect on accuracy (85% versus 82% to 87%), sensitivity (91% to 96%), or specificity (44% to 63%). The overall false positive rate of 7% (7/94) showed the largest decrease following the exclusion of solids, the false positive rate dropping to 4% (3/69).

6.4 EU Classification System: BCOP Test Method Accuracy

The six reports used in the accuracy evaluation (Gautheron et al. 1994, Balls et al. 1995, Swanson et al. 1995, Southee 1998, Swanson and Harbell 2000, and Bailey et al. 2004) included BCOP data on 118 substances that had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the EU classification system (EU 2004) (see **Annex III**). Among these studies, Gautheron et al. (1994), Balls et al. (1995), and Southee (1998) provided BCOP data for substances tested in multiple laboratories and thus required that a consensus *in vitro* classification be assigned to each substance. Based on results from *in vivo* rabbit eye experiments, 28% (33/118) were classified as R41, 14% (21/118) were classified as R36, and 54% (64/118) were classified as Not Labeled.

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

The BCOP test method correctly identified 79% (26/33) and 73% (24/33) of the R41 substances using decision criteria that defined *in vitro* scores ≥55.1 as R41 and *in vitro* scores ≥75 as R41, respectively (**Table 6-2**). Using decision criteria that defined *in vitro* scores ≥55.1 as R41, all seven substances that were underpredicted by the BCOP test method were classified as R36. Using decision criteria that defined *in vitro* scores ≥75 as R41, all nine substances that were underpredicted by the BCOP test method were classified as R36.

6.4.2 Identification of R36 Substances (Irritants)

For the 21 substances that could be evaluated, the BCOP test method correctly identified 52% (11/21) as R36, while 48% (10/21) were overpredicted using decision criteria defining *in vitro* scores ≥55.1 as R41 (**Table 6-14**). Using decision criteria that defined *in vitro* scores ≥75 as R41, the BCOP test method correctly identified 67% (14/21) as R36, while 29% (6/21) were overpredicted and 4% (1/21) were underpredicted (**Table 6-2**).

Table 6-12 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	Overall Correct Classification	Severe (Category I)		Moderate (Category II)			Mild (Category III)			Not Labeled (Category IV)	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall	51% (62/121)	75% (24/32)	25% (8/32)	47% (8/17)	35% (6/17)	18% (3/17)	49% (22/45)	36% (16/45)	15% (7/45)	41% (11/27)	59% (16/27)
Without Alcohols	54% (57/105)	73% (24/33)	43% (14/33)	36% (4/11)	36% (4/11)	27% (3/11)	43% (17/40)	40% (16/40)	18% (7/40)	46% (11/24)	54% (13/24)
Without Ketones	53% (59/112)	75% (24/32)	25% (8/32)	44% (7/16)	38% (6/16)	19% (3/16)	47% (18/38)	37% (14/38)	16% (6/38)	42% (11/26)	58% (15/26)
Without Solids	48% (43/89)	86% (18/21)	14% (3/21)	47% (7/15)	33% (5/15)	20% (3/15)	64% (21/33)	27% (9/33)	9% (3/33)	45% (9/20)	55% (11/20)
Without Alcohols and Ketones	56% (54/96)	80% (24/30)	20% (6/30)	30% (3/10)	40% (4/10)	30% (3/10)	39% (13/33)	42% (14/33)	18% (6/33)	48% (11/23)	52% (12/23)
Without Alcohols, Ketones, and Solids	54% (35/65)	90% (18/20)	10% (2/20)	25% (2/8)	38% (3/8)	37% (3/8)	57% (12/21)	33% (7/21)	10% (2/21)	56% (9/16)	44% (7/16)
Without Hydrocarbons	52% (57/110)	75% (24/32)	25% (8/32)	47% (8/17)	35% (6/17)	18% (3/17)	48% (19/40)	35% (14/40)	17% (7/40)	38% (8/21)	62% (13/21)

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency.

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

Table 6-13 Accuracy of the BCOP Test Method in Distinguishing Category IV Ocular Irritants from All Other Irritant Classes, as Defined by the EPA Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	187	83	155/187	94	134/142	47	21/45	53	24/45	6	8/142
Without Alcohols	105	83	87/105	91	74/81	63	13/24	46	11/24	9	7/81
Without Ketones	112	85	95/112	93	80/86	58	15/26	42	11/26	7	6/86
Without Solids	89	87	77/89	96	66/69	55	11/20	45	9/20	4	3/69
Without Alcohols and Ketones	96	82	79/96	92	67/73	52	12/23	48	11/23	8	6/73
Without Alcohols, Ketones, and Solids	65	82	53/65	96	47/49	44	7/16	56	9/16	4	2/49
Without Hydrocarbons	110	86	95/110	92	82/89	62	13/21	38	8/21	8	7/89

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency;.N = number of substances included in this analysis/total number of substances in the study; No. = data used to calculate the percentage.

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

6.4.3 Identification of Not Labeled Substances

Regardless of the decision criteria used to define R41, for the 64 substances that could be evaluated, the BCOP test method correctly identified 34% (22/64) as Not Labeled, while 66% (42/64) were overpredicted (**Table 6-14**).

6.4.4 Ability to Distinguish Not Labeled Substances from All Other Classes

In addition to evaluating the ability of the BCOP test method to identify each individual ocular hazard category according to the EU classification system, ICCVAM also evaluated the ability of the BCOP test method to distinguish Not Labeled substances from all other irritant classes. Using this approach for the 118 substances considered, the BCOP test method has an accuracy of 64% (76/118), a sensitivity of 100% (54/54), a specificity of 34% (22/64), a false positive rate of 66% (42/64), and a false negative rate of 0% (0/54) (**Table 6-15**).

Table 6-14 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System,[1] by Study and Overall

Data Source	Overall Correct Classification	Severe (R41)		Moderate (R36)			Mild			Not Labeled[2]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Gautheron et al. (1994)	42% (18/43)	75% (6/8)	25% (2/8)	50% (2/4)	50% (2/4)	0% (0/4)	NA	NA	NA	68% (21/31)	32% (10/31)
Balls et al. (1995)	54% (27/50)	74% (14/19)	26% (5/19)	47% (7/15)	53% (8/15)	0% (0/15)	NA	NA	NA	69% (11/16)	31% (5/16)
Swanson et al. (1995)	50% (6/12)	100% (6/6)	0% (0/6)	0% (0/0)	0% (0/0)	0% (0/0)	NA	NA	NA	100% (6/6)	0% (0/6)
Southee (1998)	60% (9/15)	67% (4/6)	33% (2/6)	40% (2/5)	60% (3/5)	0% (0/5)	NA	NA	NA	50% (2/4)	50% (2/4)
Swanson and Harbell (2000)	38% (3/8)	100% (1/1)	0% (0/1)	50% (2/4)	50% (2/4)	0% (0/4)	NA	NA	NA	100% (3/3)	0% (0/3)
Bailey et al. (2004)	46% (6/13)	67% (2/3)	33% (1/3)	0% (0/0)	0% (0/0)	0% (0/0)	NA	NA	NA	60% (6/10)	40% (4/10)
AMCP BRD (2008)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Overall	50% (59/118)	79% (26/33)	21% (7/33)	48% (10/21)	52% (11/21)	0% (0/21)	NA	NA	NA	66% (42/64)	34% (22/64)

Abbreviations: AMCP = antimicrobial cleaning product; BCOP = bovine corneal opacity and permeability; BRD = background review document; EU = European Union; NA = not applicable.

[1] EU classification system (EU 2001).

[2] Not Labeled = Not Labeled as Irritant.

Table 6-15 Accuracy of the BCOP Test Method in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the EU Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gautheron et al. (1994)	43	51	22/43	100	12/12	32	10/31	68	21/31	0	0/12
Balls et al. (1995)	50	78	39/50	100	34/34	31	5/16	69	11/16	0	0/34
Swanson et al. (1995)	12	50	6/12	100	6/6	0	0/6	100	6/6	0	0/6
Southee (1998)	15	87	13/15	100	11/11	50	2/4	50	2/4	0	0/11
Swanson and Harbell (2000)	8	63	5/8	100	5/5	0	0/3	100	3/3	0	0/5
Bailey et al. (2004)	13	54	7/13	100	3/3	40	4/10	60	6/10	0	0/3
AMCP BRD (2008)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Overall	118	64	76/118	100	54/54	34	22/64	66	42/64	0	0/54

Abbreviations: BCOP = bovine corneal opacity and permeability; BRD = background review document; EU = European Union; N = number of substances included in this analysis; No. = data used to calculate the percentage.

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

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

Gautheron et al. (1994): Based upon the *in vivo* rabbit data, 43 substances could be assigned EU classifications. Based on these 43 substances, the BCOP test method has an accuracy of 51% (22/43), sensitivity of 100% (12/12), specificity of 32% (10/31), false positive rate of 68% (21/31), and a false negative rate of 0% (0/12) (**Table 6-15**).

Balls et al. (1995): Based upon the *in vivo* rabbit data, 50 substances could be assigned EU classifications. Based on these 50 substances, the BCOP test method has an accuracy of 78% (39/50), sensitivity of 100% (34/34), specificity of 31% (5/16), false positive rate of 69% (11/16), and a false negative rate of 0% (0/34) (**Table 6-15**).

Swanson et al. (1995): Based upon the *in vivo* rabbit data, 12 substances could be assigned EU classifications. Based on these 12 substances, the BCOP test method has an accuracy of 50% (6/12), sensitivity of 100% (6/6), specificity of 0% (0/6), false positive rate of 100% (6/6), and a false negative rate of 0% (0/6) (**Table 6-15**).

Southee (1998): Based upon the *in vivo* rabbit data, 15 substances could be assigned EU classifications. Based on these 15 substances, the BCOP test method has an accuracy of 87% (13/15), sensitivity of 100% (11/11), specificity of 50% (2/4), false positive rate of 50% (2/4), and a false negative rate of 0% (0/11) (**Table 6-15**).

Swanson and Harbell (2000): Based upon the *in vivo* rabbit data, eight substances could be assigned EU classifications. Based on these eight substances, the BCOP test method has an accuracy of 63% (5/8), sensitivity of 100% (5/5), specificity of 0% (0/3), false positive rate of 100% (3/3), and a false negative rate of 0% (0/5) (**Table 6-15**).

Bailey et al. (2004): Based upon the *in vivo* rabbit data, 13 substances could be assigned EU classifications. Based on these 13 substances, the BCOP test method has an accuracy of 54% (7/13), sensitivity of 100% (3/3), specificity of 40% (4/10), false positive rate of 60% (6/10), and a false negative rate of 0% (0/3) (**Table 6-15**).

6.4.5 Discordant Results According to the EU Classification System

In order to evaluate discordant responses of the BCOP test method relative to the *in vivo* hazard classification, several accuracy subanalyses were performed. These included specific classes of chemicals with sufficiently robust numbers of substances (n ≥5), as well as certain properties of interest considered relevant to ocular toxicity testing (e.g., surfactants and physical form, respectively).

Table 6-16 shows some notable trends in the performance of the BCOP test method among these subgroups of substances. The chemical class of substances that was most consistently overpredicted according to the EU classification system by the BCOP test method was hydrocarbons. Seven of the 42 overpredicted substances were hydrocarbons. Additional chemical classes represented among the overpredicted substances were ketones (5), esters (5), carboxylic acids (4), alcohols (3), and heterocyclic compounds (3). Among the 24 substances labeled as surfactants, the BCOP test method overpredicted 25% (6/24).

The BCOP test method overpredicted 35 liquids and 7 solids. Considering the proportion of the total available database, the BCOP test method appears more likely to overpredict liquids (88/118 or 75%) than solids (30/118 or 25%).

According to the EU classification system (see **Annex III**), alcohols (2) were most often underpredicted (i.e., false negatives) by the BCOP test method. As can be seen in **Table 6-16**, none of the 24 substances labeled as surfactants was underpredicted by the BCOP test method (0% [0/24]).

The BCOP test method underpredicted five solids and one liquid. As a proportion of the total available database, solids (30/118 or 25%) appear more likely than liquids (88/118 or 75%) to be underpredicted by the BCOP test method.

Table 6-17 shows how the BCOP test method performance statistics were affected by excluding from the data set problematic classes (i.e., those that gave the most discordant results, according to the EU classification system) identified in the BCOP BRD (ICCVAM 2006a). In general, the exclusion of alcohols, ketones, or solids individually resulted in small changes in the performance statistics. Exclusion of both alcohols and ketones improved the overall classification rate: 53% (50/94) versus 50% (59/118) for all compounds in the database. The classification of ocular corrosives/severe irritants was most improved by the exclusion of problematic classes. Using the entire database, 79% (26/33) of severe ocular corrosives/severe irritants were accurately classified, while removal of solids resulted in 91% (21/23) correct classification. Removal of alcohols, ketones, and solids resulted in correct classification of 95% (20/21) ocular corrosives/severe irritants. Evaluation of overpredicted substances shows 64% (7/11) of hydrocarbons were overpredicted (**Table 6-16**). Compared to the entire database, exclusion of hydrocarbons improved overall correct classification (52% [56/107)] versus 50% [62/121]) and slightly improved identification of substances Not Labeled as Irritants (36% [19/53] versus 34% [22/64]) (**Table 6-17**).

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

Category	N	Underprediction (*In Vivo/In Vitro*)			Overprediction (*In Vivo/In Vitro*)		
		Severe (R41)		Moderate (R36)	Moderate (R36)	Not Labeled (NL)[2]	
		R36	NL	NL	R41	R41	R36
Overall	118	21% (7/33)	0% (0/33)	0% (0/21)	48% (10/21)	13% (8/64)	38% (24/64)
Chemical Class[3]							
Alcohol	16	67% (2/3)	0% (0/3)	0% (0/6)	50% (3/6)	0% (0/7)	0% (0/7)
Amine/Amidine	6	0% (0/2)	0% (0/2)	0/0	0/0	0% (0/4)	25% (1/4)
Carboxylic acid	13	25% (1/4)	0% (0/4)	0% (0/3)	33% (1/3)	33% (2/6)	17% (1/6)
Ester	10	0% (0/2)	0% (0/2)	0% (0/3)	33% (1/3)	40% (2/5)	40% (2/5)
Ether	6	0% (0/1)	0% (0/1)	0% (0/1)	100% (1/1)	0% (0/2)	0% (0/2)
Heterocyclic	13	17% (1/6)	0% (0/6)	0% (0/1)	0% (0/1)	0% (0/6)	50% (3/6)
Hydrocarbon	11	0/0	0/0	0/0	0/0	18% (2/11)	45% (5/11)
Inorganics	7	0% (0/5)	0% (0/5)	0% (0/1)	0% (0/1)	0% (0/2)	50% (1/2)
Ketone	9	0/0	0/0	0% (0/2)	100% (2/2)	14% (1/7)	28% (2/7)
Onium compound	11	13% (1/8)	0% (0/8)	0% (0/1)	0% (0/1)	0% (0/2)	50% (1/2)
Polyether	2	0/0	0/0	0/0	0/0	0% (0/2)	0% (0/2)
Properties of Interest							
Liquids	88	4% (1/23)	0% (0/23)	0% (0/18)	50% (9/18)	17% (8/47)	38% (18/47)
Solids	30	50% (5/10)	0% (0/10)	0% (0/2)	50% (1/2)	0% (0/17)	35% (6/17)
Pesticide	7	50% (2/4)	0% (0/4)	0% (0/1)	100% (1/1)	0% (0/2)	50% (1/2)
Surfactants: total	24	0% (0/13)	0% (0/13)	0% (0/2)	50% (1/2)	22% (2/9)	33% (3/9)
Surfactants: nonionic	11	0% (0/5)	0% (0/5)	0% (0/1)	100% (1/1)	0% (0/5)	20% (1/5)

Category	N	Underprediction (*In Vivo/In Vitro*)			Overprediction (*In Vivo/In Vitro*)		
		Severe (R41)		Moderate (R36)	Moderate (R36)	Not Labeled (NL)[2]	
		R36	NL	NL	R41	R41	R36
Surfactants: anionic	9	0% (0/4)	0% (0/4)	0% (0/1)	0% (0/1)	50% (2/4)	50% (2/4)
Surfactants: cationic	7	0% (0/6)	0% (0/6)	0/0	0/0	0% (0/1)	100% (1/1)

Abbreviations: BCOP= bovine corneal opacity and permeability; EU = European Union; N = number of substances used in this analysis/total number of substances in the study.

[1] EU classification system (EU 2001).

[2] Not Labeled (NL) = Not Labeled as Irritant.

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

Table 6-18 shows how the ability of the BCOP test method to distinguish substances not labeled as irritants was affected by exclusion of problematic classes from the data set. Exclusion of problematic classes individually or in combination had a minimal effect on accuracy (64% versus 60% to 66%) and specificity (24% to 35%). Sensitivity was 100% using the overall database and therefore remained unchanged. None of the R41 substances was classified by the BCOP test method as not labeled as an irritant. Exclusion of hydrocarbons resulted in modest improvement in overall performance in identifying substances not labeled as irritants (see **Table 6-18**). Accuracy increased from 64% (76/118) to 68% (73/107). The false positive rate decreased from 66% (42/64) to 64% (34/53), while the false negative rate remained 0% (0/54 versus 0/54).

Table 6-17 Performance of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	Overall Correct Classification	Severe (R41)		Moderate (R36)			Mild			Not Labeled[2]	
		Actual	Under	Over	Actual	Under	Over	Actual	Under	Over	Actual
Overall	50% (59/118)	79% (26/33)	21% (7/33)	48% (10/21)	52% (11/21)	0% (0/21)	NA	NA	NA	66% (42/64)	34% (22/64)
Without Alcohols	50% (52/103)	83% (25/30)	17% (5/30)	47% (7/15)	53% (8/15)	0% (0/15)	NA	NA	NA	67% (39/58)	33% (19/58)
Without Ketones	52% (59/109)	79% (26/33)	21% (7/33)	42% (8/19)	58% (11/19)	0% (0/19)	NA	NA	NA	65% (37/57)	35% (20/57)
Without Solids	49% (43/88)	91% (21/23)	9% (2/23)	50% (9/18)	50% (9/18)	0% (0/18)	NA	NA	NA	72% (34/47)	28% (13/47)
Without Alcohols and Ketones	53% (50/94)	83% (25/30)	17% (5/30)	38% (5/13)	62% (8/13)	0% (0/13)	NA	NA	NA	67% (34/51)	33% (17/51)
Without Alcohols, Ketones, and Solids	52% (34/65)	95% (20/21)	5% (1/21)	40% (4/10)	60% (6/10)	0% (0/10)	NA	NA	NA	76% (26/34)	24% (8/34)
Without Hydrocarbons	52% (56/107)	79% (26/33)	21% (7/33)	48% (10/21)	52% (11/21)	0% (0/21)	NA	NA	NA	64% (34/53)	36% (19/53)

Abbreviations: BCOP = bovine corneal opacity and permeability; EU = European Union; NA = not applicable.

[1] EU classification system (EU 2001).

[2] Not Labeled = Not Labeled as Irritant.

Table 6-18 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the EU Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	118	64	76/118	100	54/54	34	22/64	66	42/64	0	0/54
Without Alcohols	103	62	64/103	100	45/45	33	19/58	67	39/58	0	0/45
Without Ketones	109	66	72/109	100	52/52	35	20/57	65	37/57	0	0/52
Without Solids	88	61	54/88	100	41/41	28	13/47	72	34/47	0	0/41
Without Alcohols and Ketones	94	64	60/94	100	43/43	33	17/51	67	34/51	0	0/43
Without Alcohols, Ketones, and Solids	65	60	39/65	100	31/31	24	8/34	76	26/34	0	0/31
Without Hydrocarbons	107	68	73/107	100	54/54	36	19/53	64	34/53	0	0/54

Abbreviations: BCOP = bovine corneal opacity and permeability; EU = European Union; N = number of substances included in this analysis/the total number of substances in the study; No. = data used to calculate the percentage.

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

6.5 FHSA Classification System: BCOP Test Method Accuracy

The six reports used in the accuracy evaluation (Gautheron et al. 1994, Balls et al. 1995, Swanson et al. 1995, Southee 1998, Swanson and Harbell 2000, and Bailey et al. 2004) included BCOP data on 194 and 179 substances that had sufficient *in vivo* data to be assigned an ocular irritancy classification according to the FHSA-20% and FHSA-67% classification systems, respectively (FHSA 2005) (see **Annex III**). Among these studies, Gautheron et al. (1994), Balls et al. (1995), and Southee (1998) provided BCOP data for substances tested in multiple laboratories and thus required that a consensus *in vitro* classification be assigned to each substance. Based on results from *in vivo* rabbit eye experiments, 76% (147/194) and 74% (132/179) were classified as irritants in FHSA-20% and FHSA-67%, respectively, while 24% (47/194) and 26% (47/179) were classified as Not Labeled in FHSA-20% and FHSA-67%, respectively.

6.5.1 Ability to Distinguish Not Labeled Substances from All Other Classes

ICCVAM also evaluated the ability of the BCOP test method to distinguish Not Labeled substances from irritants using the FHSA-20% and FHSA-67% classification systems.

Ability to Distinguish Not Labeled Substances from All Other Classes using the FHSA-20% Classification System

ICCVAM evaluated the ability of the BCOP test method to distinguish Not Labeled substances from irritants using the FHSA-20% classification system. Using this approach for the 194 substances, the BCOP test method has an overall accuracy of 83% (161/194), a sensitivity of 95% (139/147), a specificity of 47% (22/47), a false positive rate of 53% (25/47), and a false negative rate of 5% (8/147) (**Table 6-19**).

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

Gautheron et al. (1994): Based upon the *in vivo* rabbit data, 52 substances could be assigned an FHSA-20% classification. Based on these 52 substances, the BCOP test method has an accuracy of 83% (43/52), sensitivity of 88% (35/40), specificity of 67% (8/12), false positive rate of 33% (4/12), and a false negative rate of 13% (5/40) (**Table 6-19**).

Balls et al. (1995): Based upon the *in vivo* rabbit data, 58 substances could be assigned an FHSA-20% classification. Based on these 58 substances, the BCOP test method has an accuracy of 91% (53/58), sensitivity of 93% (50/54), specificity of 75% (3/4), false positive rate of 25% (1/4), and a false negative rate of 7% (4/54) (**Table 6-19**).

Swanson et al. (1995): Based upon the *in vivo* rabbit data, 9 substances could be assigned an FHSA-20% classification. Based on these 9 substances, the BCOP test method has an accuracy of 89% (8/9), sensitivity of 100% (8/8), specificity of 0% (0/1), false positive rate of 100% (1/1), and a false negative rate of 0% (0/8) (**Table 6-19**).

Southee (1998): Based upon the *in vivo* rabbit data, 15 substances could be assigned an FHSA-20% classification. Based on these 15 substances, the BCOP test method has an accuracy of 93% (14/15), sensitivity of 93% (13/14), specificity of 100% (1/1), false positive rate of 0% (0/1), and a false negative rate of 7% (1/14) (**Table 6-19**).

Swanson and Harbell (2000): Based upon the *in vivo* rabbit data, 8 substances could be assigned an FHSA-20% classification. Based on these 8 substances, the BCOP test method has an accuracy of 75% (6/8), sensitivity of 100% (6/6), specificity of 0% (0/2), false positive rate of 100% (2/2), and a false negative rate of 0% (0/6) (**Table 6-19**).

Bailey et al. (2004): Based upon the *in vivo* rabbit data, 15 substances could be assigned an FHSA-20% classification. Based on these 15 substances, the BCOP test method has an accuracy of 73% (11/15),

sensitivity of 88% (7/8), specificity of 57% (4/7), false positive rate of 43% (3/7), and a false negative rate of 13% (1/8) (**Table 6-19**).

AMCP BRD (2008): Based upon the *in vivo* rabbit data, 63 substances could be assigned an FHSA-20% classification. Based on these 63 substances, the BCOP test method has an accuracy of 67% (42/63), sensitivity of 100% (42/42), specificity of 0% (0/21), false positive rate of 100% (21/21), and a false negative rate of 0% (0/42) (**Table 6-19**).

Table 6-19 Accuracy of the BCOP Test Method in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the FHSA-20% Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gautheron et al. (1994)	52	83	43/52	88	35/40	67	8/12	33	4/12	13	5/40
Balls et al. (1995)	58	91	53/58	93	50/54	75	3/4	25	1/4	7	4/54
Swanson et al. (1995)	9	89	8/9	100	8/8	0	0/1	100	1/1	0	0/8
Southee (1998)	15	93	14/15	93	13/14	100	1/1	0	0/1	7	1/14
Swanson and Harbell (2000)	8	75	6/8	100	6/6	0	0/2	100	2/2	0	0/6
Bailey et al. (2004)	15	73	11/15	88	7/8	57	4/7	43	3/7	13	1/8
AMCP BRD (2008)	63	67	42/63	100	42/42	0	0/21	100	21/21	0	0/42
Overall	194	83	161/194	95	139/147	47	22/47	53	25/47	5	8/147

Abbreviations: BCOP = bovine corneal opacity and permeability; FHSA = Federal Hazardous Substances Act; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] FHSA-20% classification system (2005): Not Labeled vs. Irritant.

Ability to Distinguish Not Labeled Substances from All Other Classes using the FHSA-67% Classification System

ICCVAM evaluated the ability of the BCOP test method to distinguish Not Labeled substances from irritants using the FHSA-67% classification system. Using this approach for the 179 substances, the BCOP test method has an overall accuracy of 83% (148/179), a sensitivity of 95% (126/132), a specificity of 47% (22/47), a false positive rate of 53% (25/47), and a false negative rate of 5% (6/132) (**Table 6-20**).

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

Gautheron et al. (1994): Based upon the *in vivo* rabbit data, 48 substances could be assigned an FHSA-67% classification. Based on these 48 substances, the BCOP test method has an accuracy of 83% (40/48), sensitivity of 89% (32/36), specificity of 67% (8/12), false positive rate of 33% (4/12), and a false negative rate of 11% (4/36) (**Table 6-20**).

Balls et al. (1995): Based upon the *in vivo* rabbit data, 52 substances could be assigned an FHSA-67% classification. Based on these 52 substances, the BCOP test method has an accuracy of 94% (49/52), sensitivity of 96% (46/48), specificity of 75% (3/4), false positive rate of 25% (1/4), and a false negative rate of 4% (2/48) (**Table 6-20**).

Swanson et al. (1995): Based upon the *in vivo* rabbit data, eight substances could be assigned an FHSA-67% classification. Based on these 8 substances, the BCOP test method has an accuracy of 88% (7/8), sensitivity of 100% (7/7), specificity of 0% (0/1), false positive rate of 100% (1/1), and a false negative rate of 0% (0/7) (**Table 6-20**).

Southee (1998): Based upon the *in vivo* rabbit data, 14 substances could be assigned an FHSA-67% classification. Based on these 14 substances, the BCOP test method has an accuracy of 100% (14/14), sensitivity of 100% (13/13), specificity of 100% (1/1), false positive rate of 0% (0/1), and a false negative rate of 0% (0/13) (**Table 6-20**).

Swanson and Harbell (2000): Based upon the *in vivo* rabbit data, seven substances could be assigned an FHSA-67% classification. Based on these 7 substances, the BCOP test method has an accuracy of 71% (5/7), sensitivity of 100% (5/5), specificity of 0% (0/2), false positive rate of 100% (2/2), and a false negative rate of 0% (0/5) (**Table 6-20**).

Bailey et al. (2004): Based upon the *in vivo* rabbit data, 14 substances could be assigned an FHSA-67% classification. Based on these 14 substances, the BCOP test method has an accuracy of 71% (10/14), sensitivity of 86% (6/7), specificity of 57% (4/7), false positive rate of 43% (3/7), and a false negative rate of 14% (1/7) (**Table 6-20**).

AMCP BRD (2008): Based upon the *in vivo* rabbit data, 63 substances could be assigned an FHSA-67% classification. Based on these 63 substances, the BCOP test method has an accuracy of 67% (42/63), sensitivity of 100% (42/42), specificity of 0% (0/21), false positive rate of 100% (21/21), and a false negative rate of 0% (0/42) (Table 6-20).

6.5.2 Discordant Results According to the FHSA Classification System

In order to evaluate discordant responses of the BCOP test method relative to the *in vivo* hazard classification, several accuracy subanalyses were performed. These included specific classes of chemicals with sufficiently robust numbers of substances (n ≥5), as well as certain properties of interest considered relevant to ocular toxicity testing (e.g., surfactants and physical form, respectively).

Discordant Results According to the FHSA-20% Classification System

Table 6-21 shows how the ability of the BCOP test method to distinguish substances not labeled as irritants was affected by exclusion of problematic classes from the data set. Exclusion of problematic classes individually or in combination had a minimal or no effect on accuracy (83% versus 80% to 84%), specificity (94% to 98%) and specificity (36% to 47%). Exclusion of hydrocarbons also resulted no significant improvement in overall performance in identifying substances not labeled as irritants. However, a slightly higher false positive rate and slightly lower false negative rate occurred with exclusion of discordant classes (see **Table 6-21**).

Discordant Results According to the FHSA-67% Classification System

Table 6-22 shows how the ability of the BCOP test method to distinguish substances not labeled as irritants was affected by exclusion of problematic classes from the data set. Exclusion of problematic classes individually or in combination had a minimal or no effect on accuracy (83% versus 80% to 83%), specificity (95% to 99%) and specificity (36% to 47%). Exclusion of hydrocarbons also resulted no significant improvement in overall performance in identifying substances not labeled as irritants. However, a slightly higher false positive rate and slightly lower false negative rate occurred with exclusion of discordant classes (see **Table 6-22**).

Table 6-20 Accuracy of the BCOP Test Method in Distinguishing Not Labeled Substances from All Other Irritant Classes, as Defined by the FHSA-67% Classification System,[1] by Study and Overall

Data Source	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Gautheron et al. (1994)	48	83	40/48	89	32/36	67	8/12	33	4/12	11	4/36
Balls et al. (1995)	52	94	49/52	96	46/48	75	3/4	25	1/4	4	2/48
Swanson et al. (1995)	8	88	7/8	100	7/7	0	0/1	100	1/1	0	0/7
Southee (1998)	14	100	14/14	100	13/13	100	1/1	0	0/1	0	0/13
Swanson and Harbell (2000)	7	71	5/7	100	5/5	0	0/2	100	2/2	0	0/5
Bailey et al. (2004)	14	71	10/14	86	6/7	57	4/7	43	3/7	14	1/7
AMCP BRD (2008)	59	64	38/59	100	38/38	0	0/21	100	21/21	0	0/38
Overall	179	83	148/179	95	126/132	47	22/47	53	25/47	5	6/132

Abbreviations: BCOP = bovine corneal opacity and permeability; FHSA = Federal Hazardous Substances Act; N = number of substances included in this analysis; No. = data used to calculate the percentage.

[1] FHSA-67% classification system (2005): Not Labeled vs. Irritant.

Table 6-21 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-20% Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	194	83	161/194	95	139/147	47	22/47	53	25/47	5	8/147
Without Alcohols	177	81	144/177	94	125/133	43	19/44	57	25/44	6	8/133
Without Ketones	184	82	151/184	94	131/139	44	20/45	56	25/45	6	8/139
Without Solids	157	84	132/157	98	114/116	44	18/41	56	23/41	2	2/116
Without Alcohols and Ketones	168	80	135/168	94	118/126	40	17/42	60	25/42	6	8/126
Without Alcohols, Ketones, and Solids	132	81	107/132	98	94/96	36	13/36	64	23/36	2	2/96
Without Hydrocarbons	184	83	153/184	94	133/141	47	20/43	53	23/43	6	8/141

Abbreviations: BCOP = bovine corneal opacity and permeability; FHSA = Federal Hazardous Substances Act; N = number of substances included in this analysis/the total number of substances in the study; No. = data used to calculate the percentage.

[1] FHSA-20% classification system (2005): Not Labeled vs. Irritant.

Table 6-22 Accuracy of the BCOP Test Method in Distinguishing Substances Not Labeled as Irritants from All Other Irritant Classes, as Defined by the FHSA-67% Classification System,[1] with Discordant Chemical and Physical Classes Excluded

BCOP	N	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate	
		%	No.	%	No.	%	No.	%	No.	%	No.
Overall	179	83	148/179	95	126/132	47	22/47	53	25/47	5	6/132
Without Alcohols	162	81	131/162	95	112/118	43	19/44	57	25/44	5	6/118
Without Ketones	170	82	139/170	95	119/125	44	20/45	56	25/45	5	6/125
Without Solids	144	83	120/144	99	102/103	44	18/41	56	23/41	1	1/103
Without Alcohols and Ketones	154	80	123/154	95	106/112	40	17/42	60	25/42	5	6/112
Without Alcohols, Ketones, and Solids	120	80	96/120	99	83/84	36	13/36	64	23/36	1	1/84
Without Hydrocarbons	170	83	141/170	95	121/127	47	20/43	53	23/43	5	6/127

Abbreviations: BCOP = bovine corneal opacity and permeability; FHSA = Federal Hazardous Substances Act; N = number of substances included in this analysis/the total number of substances in the study; No. = data used to calculate the percentage.

[1] FHSA-67% classification system (2005): Not Labeled vs. Irritant.

7.0 Bovine Corneal Opacity and Permeability Test Method Reliability

Assessment of test method reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) is essential to any evaluation of the performance of an alternative test method (ICCVAM 2003). Quantitative and qualitative evaluations of BCOP test method reliability have been conducted previously (ICCVAM 2006a).

However, additional qualitative analyses of test method reproducibility evaluated the extent of agreement of BCOP hazard classifications among the laboratories. Given that the performance of the BCOP test method was similar for the EPA and FHSA classification systems, additional reliability analyses were not conducted for the FHSA classification system.

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

Reliability analyses for the BCOP test method were evaluated for the following three studies: Balls et al. (1995), Gautheron et al. (1994), and Southee (1998).

Balls et al. (1995): Of 14 substances classified by the GHS as Not Labeled, 29% (4/14) were correctly identified, while two of four GHS Category 2B substances (50%) were correctly identified, 29% (4/14) substances classified as GHS Category 2A were correctly identified, and 77% (17/22) GHS Category 1 substances were correctly identified.

The five participating laboratories were in 100% agreement on the ocular irritancy classification when distinguishing Not Labeled substances from all other classes of 92% (55/60) substances (**Table 7-1**).

All five participating laboratories agreed on the classification of 88% (15/17) substances that were correctly identified as GHS Category 1, 0% (0/4) substances correctly classified as GHS Category 2A, 50% (1/2) substances correctly classified as GHS Category 2B, and 50% (2/4) substances correctly classified as GHS Not Classified (**Table 7-2**).

The extent of agreement between testing laboratories was greatest for substances identified from *in vivo* rabbit eye data as corrosives or severe irritants when compared to any other combination of *in vivo* and *in vitro* results. Eighty-eight percent (15/17) of the accurately identified severe substances were shown to have 100% classification agreement among testing laboratories (**Table 7-2**).

There was 100% agreement on the 10 false positive substances among the five laboratories.

Gautheron et al. (1994): Of 34 substances classified by the GHS as Not Classified, 38% (13/34) were correctly identified, while 0% (0/2) GHS Category 2B substances were correctly identified, 33% (1/3) substances classified as GHS Category 2A was correctly identified, and 75% (6/8) GHS Category 1 substances were correctly identified.

The 11–12 participating laboratories were in 100% agreement on the ocular irritancy classification when distinguishing substances not labeled as irritants from all other classes of 65% (34/52) substances (**Table 7-1**).

All 11–12 participating laboratories agreed on the classification of 67% (4/6) substances that were correctly identified as GHS Category 1, 0% (0/1) substances correctly classified as GHS Category 2A, and 0% (0/13) substance correctly classified as GHS Not Classified (**Table 7-2**).

The extent of agreement between testing laboratories was greatest for substances identified from *in vivo* rabbit eye data as corrosives or severe irritants when compared to any other combination of *in vivo* and *in vitro* results: 67% (4/6) of the accurately identified severe substances were shown to have 100% classification agreement among testing laboratories) (**Table 7-2**).

Of the 21 false positive substances, 90% (19/21) were shown to have 100% agreement among the 11-12 laboratories.

Southee (1998): Of 3 substances classified by the GHS as Not Classified, 67% (2/3) were correctly identified, while 50% (1/2) GHS Category 2B substances was correctly identified, 67% (2/3) of substances classified as GHS Category 2A were correctly identified, and 57% (4/7) of GHS Category 1 substances were correctly identified.

The three participating laboratories were in 100% agreement on the ocular irritancy classification when distinguishing substances not labeled as irritants from all other classes of 88% (14/16) substances (**Table 7-1**).

All three participating laboratories agreed on the classification of 100% (4/4) substances that were correctly identified as GHS Category 1, 50% (1/2) substances correctly classified as GHS Category 2A, 100% (1/1) substance correctly classified as GHS Category 2B, and 100% (2/2) substances correctly classified as GHS Not Classified (**Table 7-2**).

Regarding the 1 false positive substance, there was 100% agreement among the three laboratories.

Table 7-1 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate/Mild Irritants, as Defined by the GHS Classification System,[1] by Study

Data Source	Classification (*In Vivo*/ *In Vitro*)	No. of Testing Labs	N	Substances with 100% Agreement Among Labs[2]	Substances with 91%-92% Agreement Among Labs	Substances with 82%-83% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 73%-75% Agreement Among Labs	Substances with 64%-67% Agreement Among Labs	Substances with 58%-60% Agreement Among Labs	Substances with ≤55% Agreement Among Labs
Balls et al. (1995)	+/+	5	40	38 (95%)	-	-	1 (3%)	-	-	-	1 (3%)
	+/-	5	0	-	-	-	-	-	-	-	-
	-/+	5	10	10 (100%)	-	-	-	-	-	-	-
	-/-	5	4	2 (50%)	-	-	1 (20%)	-	-	1 (20%)	-
	?/-	5	2	1 (50%)	-	-	-	-	-	-	1 (50%)
	?/+	5	4	4 (100%)	-	-	-	-	-	-	-
	Total		60	55 (92%)	-	-	2 (3%)	-	-	1 (2%)	2 (3%)
Gautheron et al. (1994)	+/+	11 12	13	11 (84%)	1 (8%)	-	-	1 (8%)	-	-	-
	+/-	11 12	0	-	-	-	-	-	-	-	-
	-/+	11 12	21	19 (90%)	-	-	-	2 (10%)	-	-	-
	-/-	11 12	13	-	-	1 (8%)	-	1 (8%)	2 (15%)	2 (15%)	7 (54%)
	?/-	11 12	1	-	-	-	-	-	1 (100%)	-	-
	?/+	11	4	4 (100%)	-	-	-	-	-	-	-
	Total		52	34 (65%)	1 (2%)	1 (2%)	-	4 (8%)	3 (6%)	2 (4%)	7 (13%)

Data Source	Classification (*In Vivo*/ *In Vitro*)	No. of Testing Labs	N	Substances with 100% Agreement Among Labs[2]	Substances with 91%-92% Agreement Among Labs	Substances with 82%-83% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 73%-75% Agreement Among Labs	Substances with 64%-67% Agreement Among Labs	Substances with 58%-60% Agreement Among Labs	Substances with ≤55% Agreement Among Labs
Southee (1998)	+/+	3	11	10 (91%)	-	-	-	-	-	-	1 (9%)
	+/-	3	1	-	-	-	-	-	-	-	1 (100%)
	-/+	3	1	1 (100%)	-	-	-	-	-	-	-
	-/-	3	2	2 (100%)	-	-	-	-	-	-	-
	?/-	3	0	-	-	-	-	-	-	-	-
	?/+	3	1	1 (100%)	-	-	-	-	-	-	-
	Total		16	14 (88%)	-	-	-	-	-	-	2 (12%)

Abbreviations: BCOP = bovine corneal opacity and permeability; GHS = Globally Harmonized System; N = number of substances.

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

[1] GHS classification system (UN 2007).

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

Table 7-2 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the GHS Classification System,[1] by Study

Data Source	*In Vivo* Classification	Classification (*In Vitro*)	Number of Substances	Number of Testing Labs	Substances with 100% Agreement Among Laboratories (%)	Substances with 70%–95% Agreement Among Laboratories (%)	Substances with 60%–69% Agreement Among Laboratories (%)	Substances with <60% Agreement Among Laboratories (%)
Balls et al. (1995)	NL (14)	Actual	4	5	2 (50%)	1 (25%)	1 (25%)	-
		Over	10	5	10 (100%)	-	-	-
	2B (4)	Under	0	5	-	-	-	-
		Actual	2	5	1 (50%)	1 (50%)	-	-
		Over	2	5	1 (50%)	-	1 (50%)	-
	2A (14)	Under	2	5	2 (100%)	-	-	-
		Actual	4	5	-	1 (25%)	1 (25%)	2 (50%)
		Over	8	5	2 (25%)	3 (38%)	3 (38%)	-
	1 (22)	Under	5	5	3 (60%)	1 (20%)	1 (20%)	-
		Actual	17	5	15 (88%)	1 (6%)	1 (6%)	-
Gautheron et al. (1994)	NL (34)	Actual	13	11		3 (23%)	2 (15%)	8 (62%)
		Over	21	11	19 (90%)	1 (5%)	1 (5%)	-
	2B (2)	Under	0	11	-	-	-	-
		Actual	0	11	-	-	-	-
		Over	2	11	1 (50%)	1 (50%)	-	-
	2A (3)	Under	0	11	-	-	-	-
		Actual	1	11	-	1 (100%)	-	-
		Over	2	11	1 (50%)	1 (50%)	-	-
	1 (8)	Under	2	11	1 (50%)	1 (50%)	-	-
		Actual	6	11	4 (67%)	1 (17%)	-	1 (17%)

Data Source	*In Vivo* Classification	Classification (*In Vitro*)	Number of Substances	Number of Testing Labs	Substances with 100% Agreement Among Laboratories (%)	Substances with 70%–95% Agreement Among Laboratories (%)	Substances with 60%–69% Agreement Among Laboratories (%)	Substances with <60% Agreement Among Laboratories (%)
Southee (1998)	NL (3)	Actual	2	3	2 (100%)	-	-	-
		Over	1	3	1 (100%)	-	-	-
	2B (2)	Under	0	3	-	-	-	-
		Actual	1	3	1 (100%)	-	-	-
		Over	1	3	1 (100%)	-	-	-
	2A (3)	Under	0	3	-	-	-	-
		Actual	2	3	1 (50%)	1 (50%)	-	-
		Over	1	3	-	-	-	1 (100%)
	1 (7)	Under	3	3	3 (100%)	-	-	-
		Actual	4	3	4 (100%)	-	-	-

Abbreviations: BCOP = bovine corneal opacity and permeability; GHS = Globally Harmonized System; NL = Not Labeled as Irritant; 2B = mild irritant; 2A = moderate irritant; 1 = severe irritant.

[1] GHS classification system (UN 2007).

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

Balls et al. (1995): Both of the substances classified by the EPA as Category IV (100%) were correctly identified, while 29% (6/21) EPA Category III substances were correctly identified; 29% (4/14) EPA Category II substances were correctly identified, and 74% (14/19) EPA Category I substances were correctly identified.

The five participating laboratories were in 100% agreement on the ocular irritancy classification when assessing substances not labeled as irritants from all other classes of 93% (56/60) substances (**Table 7-3**).

All five participating laboratories agreed on the classification of 79% (11/14) substances that were correctly identified as EPA Category I, 0% (0/4) substances correctly classified as EPA Category II, 67% (4/6) substances correctly classified as EPA Category III, and 50% (1/2) substances correctly classified as EPA Category IV (**Table 7-4**).

When compared to any other combination of *in vivo* and *in vitro* results, the extent of agreement between testing laboratories was greatest for substances identified from *in vivo* rabbit eye data as corrosives or severe irritants. Of the accurately identified severe substances, 93% (13/14) were shown to have 80%–100% classification agreement among testing laboratories (**Table 7-4**).

Gautheron et al. (1994): Of 13 substances classified by the EPA as Category IV, 69% (9/13) were correctly identified, while 43% (9/21) EPA Category III substances were correctly identified, 25% (1/4 substances classified as EPA Category II was correctly identified, and 86% (6/7) EPA Category I substances were correctly identified.

The 11–12 participating laboratories were in 100% agreement on the ocular irritancy classification when assessing substances not labeled as irritants from all other classes of 65% (34/52) substances (**Table 7-3**).

All 11–12 participating laboratories agreed on the classification of 67% (4/6) substances that were correctly identified as EPA Category I, 0% (0/1) substances correctly classified as EPA Category II, 22% (2/9) substances correctly classified as EPA Category III, and 0% (0/9) substances correctly classified as EPA Category IV (**Table 7-4**).

All 4 false positive substances (100%) were shown to have 100% agreement among the 11–12 laboratories (**Table 7-4**).

Southee (1998): The one substance classified by the EPA as Category IV was correctly identified (100%), while 33% (2/6) EPA Category III substances were correctly identified, 50% (1/2) EPA Category II substances were correctly identified, and 50% (3/6) EPA Category I substances were correctly identified.

The three participating laboratories were in 100% agreement on the ocular irritancy classification when assessing substances not labeled as irritant from all other classes of 88% (14/16) substances (**Table 7-3**).

All three participating laboratories agreed on the classification of 100% (3/3) substances correctly identified as EPA Category I, 100% (1/1) substance correctly classified as EPA Category II, 100% (2/2) substances correctly classified as EPA Category III, and 100% (1/1) substance correctly classified as EPA Category IV (**Table 7-4**).

Table 7-3 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate/Mild Irritants, as Defined by the EPA Classification System,[1] by Study

Data Source	Classification (*In Vivo/In Vitro*)	No. of Testing Labs	N	Substances with 100% Agreement Among Labs[2]	Substances with 91%–92% Agreement Among Labs	Substances with 82%–83% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 73% Agreement Among Labs	Substances with 64%–67% Agreement Among Labs	Substances with 58%–60% Agreement Among Labs	Substances with ≤55% Agreement Among Labs
Balls et al. (1995)	+/+	5	50	48 (96%)	-	-	1 (2%)	-	-	-	1 (2%)
	+/-	5	2	1 (50%)	-	-	-	-	-	1 (25%)	-
	-/+	5	0	-	-	-	-	-	-	-	-
	-/-	5	2	1 (50%)	-	-	1 (50%)	-	-	-	-
	?/-	5	1	1 (100%)	-	-	-	-	-	-	-
	?/+	5	5	5 (100%)	-	-	-	-	-	-	-
	Total		60	56 (93%)	-	-	2 (3%)	-	-	1 (2%)	1 (2%)
Gautheron et al. (1994)	+/+	11 12	31	27 (87%)	-	1 (3%)	-	3 (10%)	-	-	-
	+/-	11 12	4	-	-	1 (25%)	-	-	-	-	3 (75%)
	-/+	11 12	4	4 (100%)	-	-	-	-	-	-	-
	-/-	11 12	9	-	-	-	-	2 (22%)	2 (22%)	3 (34%)	2 (22%)
	?/-	11 12	1	-	-	-	-	-	1 (100%)	-	-
	?/+	11	3	3 (100%)	-	-	-	-	-	-	-
	Total		52	34 (65%)	-	2 (4%)	-	5 (9%)	3 (6%)	3 (6%)	5 (10%)

continued

Table 7-3 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate/Mild Irritants, as Defined by the EPA Classification System,[1] by Study (continued)

Data Source	Classification (*In Vivo/In Vitro*)	No. of Testing Labs	N	Substances with 100% Agreement Among Labs[2]	Substances with 91%–92% Agreement Among Labs	Substances with 82%–83% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 73% Agreement Among Labs	Substances with 64%–67% Agreement Among Labs	Substances with 58%–60% Agreement Among Labs	Substances with ≤55% Agreement Among Labs
Southee (1998)	+/+	3	12	11 (92%)	-	-	-	-	-	-	1 (8%)
	+/-	3	2	1 (50%)	-	-	-	-	-	-	1 (50%)
	-/+	3	0	-	-	-	-	-	-	-	-
	-/-	3	1	1 (100%)	-	-	-	-	-	-	-
	?/-	3	0	-	-	-	-	-	-	-	-
	?/+	3	1	1 (100%)	-	-	-	-	-	-	-
	Total		16	14 (88%)	-	-	-	-	-	-	2 (12%)

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency; N = number of substances.

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

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

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

Table 7-4 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System,[1] by Study

Data Source	*In Vivo* Classification (No.)[2]	Classification (*In Vitro*)	Number of Substances	Number of Testing Laboratories	Substances with 100% Agreement Among Laboratories (%)	Substances with 80%–92% Agreement Among Laboratories (%)	Substances with 61%–79% Agreement Among Laboratories (%)	Substances with 50%–60% Agreement Among Laboratories (%)	Substances with <50% Agreement Among Laboratories (%)
Balls et al. (1995)	IV (2)	Actual	2	5	1 (50%)	1 (50%)	-	-	-
		Over	0	5	-	-	-	-	-
	III (21)	Under	2	5	1 (50%)	-	-	1 (50%)	-
		Actual	6	5	4 (67%)	1 (17%)	-	1 (17%)	-
		Over	13	5	7 (54%)	2 (15%)	-	4 (31%)	-
	II (14)	Under	2	5	2 (100%)	-	-	-	-
		Actual	4	5	-	1 (25%)	-	1 (25%)	2 (50%)
		Over	6	5	3 (50%)	1 (17%)	-	2 (33%)	-
	I (19)	Under	5	5	3 (60%)	1 (20%)	-	1 (20%)	-
		Actual	14	5	11 (79%)	2 (14%)	-	1 (7%)	-
Gautheron et al. (1994)	IV (13)	Actual	9	11/12	-	-	3 (33%)-	5 (56%)	1 (11%)
		Over	4	11/12	4 (100%)	-	-	-	-
	III (21)	Under	2	11/12	-	-	-	-	2 (100%)
		Actual	9	11/12	2 (22%)	4 (44%)	3 (33%)	-	-
		Over	10	11/12	8 (80%)	2 (20%)	-	-	-
	II (4)	Under	0	11/12	-	-	-	-	-
		Actual	1	11/12	-	1 (100%)	-	-	-
		Over	3	11/12	-	1 (33%)	2 (67%)	-	-
	I (7)	Under	2	11/12	1 (50%)	1 (50%)	-	-	-
		Actual	6	11/12	4 (67%)	1 (17%)	1 (17%)	-	-

continued

Table 7-4 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EPA Classification System,[1] by Study (continued)

Data Source	*In Vivo* Classification (No.)[2]	Classification (*In Vitro*)	Number of Substances	Number of Testing Laboratories	Substances with 100% Agreement Among Laboratories (%)	Substances with 80%–92% Agreement Among Laboratories (%)	Substances with 61%–79% Agreement Among Laboratories (%)	Substances with 50%–60% Agreement Among Laboratories (%)	Substances with <50% Agreement Among Laboratories (%)
Southee (1998)	IV (1)	Actual	1	3	1 (100%)	-	-	-	-
		Over	0	3	-	-	-	-	-
	III (6)	Under	1	3	1 (100%)	-	-	-	-
		Actual	2	3	2 (100%)	-	-	-	-
		Over	3	3	3 (100%)	-	-	-	-
	II (2)	Under	0	3	-	-	-	-	-
		Actual	1	5	1 (100%)	-	-	-	-
		Over	1	5	1 (100%)	-	-	-	-
	I (6)	Under	3	5	3 (100%)	-	-	-	-
		Actual	3	5	3 (100%)	-	-	-	-

Abbreviations: BCOP = bovine corneal opacity and permeability; EPA = U.S. Environmental Protection Agency; IV = Not Labeled as Irritant; III = mild irritant; II = moderate irritant; I = severe irritant.

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

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

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

Balls et al. (1995): Of 16 substances classified by the EU as Not Labeled, 25% (4/16) were correctly identified, while 47% (7/15) EU R36 substances were correctly identified, and 74% (14/19) EU R41 substances were correctly identified.

The five participating laboratories were in 100% agreement on the ocular irritancy classification when assessing substances not labeled as irritant from all other classes of 93% (56/60) substances (**Table 7-5**).

All five participating laboratories agreed on the classification of 86% (12/14) substances that were correctly identified as EU R41, 29% (2/7) substances correctly classified as EU R36, and 50% (2/4) substances correctly classified as EU Not Labeled (**Table 7-6**).

When compared to any other combination of *in vivo* and *in vitro* results, the extent of agreement between testing laboratories was greatest for substances identified from *in vivo* rabbit eye data as corrosives or severe irritants. All (100%) of the accurately identified severe substances were shown to have 95%–100% classification agreement among testing laboratories (**Table 7-6**).

Of the 12 false positive substances, 100% (12/12) were shown to have 100% agreement among the 5 laboratories (**Table 7-6**).

Gautheron et al. (1994): Of 36 substances classified by the EU as Not Labeled, 36% (13/36) were correctly identified, while 50% (2/4) EU R36 substances were correctly identified, and 75% (6/8) EU R41 substances were correctly identified.

The 11–12 participating laboratories were in 100% agreement on the ocular irritancy classification when assessing non labeled substances from all other classes of 65% (34/52) substances (**Table 7-5**).

All 11–12 participating laboratories agreed on the classification of 67% (4/6) substances that were correctly identified as EU R41, 0% (0/2) substances correctly classified as EU R36, and 54% (7/13) substances correctly classified as EU Not Labeled (**Table 7-6**).

Of the 23 false positive substances, 91% (21/23) were shown to have 100% agreement among the 11-12 laboratories (**Table 7-6**).

Southee (1998): Of the 4 substances classified by the EU as Not Labeled, 50% (2/4) were correctly identified, while 50% (2/4) EU R36 substances were correctly identified, and 67% (4/6) EU R41 substances were correctly identified.

The three participating laboratories were in 100% agreement on the ocular irritancy classification when assessing substances not labeled as irritant from all other classes of 88% (14/16) substances (**Table 7-5**).

All three participating laboratories agreed on the classification of 100% (4/4) substances correctly identified as EU R41, 100% (3/3) substances correctly classified as EU R36, and 100% (2/2) substances correctly classified as EU Not Labeled (**Table 7-6**).

Of the 2 false positive substances, 50% (1/2) was shown to have 100% agreement among the three laboratories (**Table 7-6**).

Table 7-5 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate Irritants, as Defined by the EU Classification System,[1] by Study

Data Source	Classification (*In Vivo/In Vitro*)	No. of Testing Labs	N	Substances with 100% Agreement Among Labs[2]	Substances with 91%–92% Agreement Among Labs	Substances with 82%–83% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 73% Agreement Among Labs	Substances with 64%–67% Agreement Among Labs	Substances with 58%–60% Agreement Among Labs	Substances with ≤55% Agreement Among Labs
Balls et al. (1995)	+/+	5	34	32 (94%)	-	-	1 (8%)	-	-	-	1 (8%)
	+/-	5	0	-	-	-	-	-	-	-	-
	-/+	5	12	12 (100%)	-	-	-	-	-	-	-
	-/-	5	4	2 (50%)	-	-	1 (25%)	-	-	1 (25%)	-
	?/-	5	1	1 (100%)	-	-	-	-	-	-	-
	?/+	5	9	9 (100%)	-	-	-	-	-	-	-
	Total		60	56 (93%)	-	-	2 (3%)	-	-	1 (2%)	1 (2%)
Gautheron et al. (1994)	+/+	11 12	12	10 (83%)	1 (8%)	-	-	1 (8%)	-	-	-
	+/-	11 12	0	-	-	-	-	-	-	-	-
	-/+	11 12	23	21 (91%)	-	-	-	2 (9%)	-	-	-
	-/-	11 12	13	-	-	1 (17%)	-	2 (16%)	2 (16%)	3 (23%)	5 (38%)
	?/-	11 12	1	-	-	-	-	-	1 (100%)	-	-
	?/+	11	3	3 (100%)	-	-	-	-	-	-	-
	Total		52	34 (65%)	1 (2%)	1 (2%)	-	5 (10%)	3 (6%)	3 (6%)	5 (10%)

continued

Table 7-5 Reliability of the BCOP Test Method in Predicting Not Labeled Ocular Substances or Corrosives/Severe/Moderate Irritants, as Defined by the EU Classification System,[1] by Study (continued)

Data Source	Classification (*In Vivo/In Vitro*)	No. of Testing Labs	N	Substances with 100% Agreement Among Labs[2]	Substances with 91%–92% Agreement Among Labs	Substances with 82%–83% Agreement Among Labs	Substances with 80% Agreement Among Labs	Substances with 73% Agreement Among Labs	Substances with 64%–67% Agreement Among Labs	Substances with 58%–60% Agreement Among Labs	Substances with ≤55% Agreement Among Labs
Southee (1998)	+/+	3	10	9 (90%)	-	-	-	-	-	-	1 (10%)
	+/-	3	1	-	-	-	-	-	-	-	1 (100%)
	-/+	3	2	2 (100%)	-	-	-	-	-	-	-
	-/-	3	2	2 (100%)	-	-	-	-	-	-	-
	?/-	3	0	-	-	-	-	-	-	-	-
	?/+	-	1	1 (100%)	-	-	-	-	-	-	-
	Total		16	14 (88%)	-	-	-	-	-	-	2 (12%)

Abbreviations: BCOP = bovine corneal opacity and permeability; EU = European Union; N = number of substances.

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

[1] EU classification system (EU 2001).

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

Table 7-6 Interlaboratory Variability of the BCOP Test Method in Predicting Ocular Irritant Classes Compared to the *In Vivo* Rabbit Eye Test Method, as Defined by the EU Classification System,[1] by Study

Data Source	*In Vivo* Classification (No.)	Classification (*In Vitro*)	Number of Substances	Number of Testing Laboratories	Substances with 100% Agreement Among Laboratories (%)	Substances with 76%-95% Agreement Among Laboratories (%)	Substances with 50%-75% Agreement Among Laboratories (%)
Balls et al. (1995)	NL (16)	Actual	4	5	2 (50%)	1 (25%)	1 (25%)
		Over	12	5	12 (100%)	-	-
	R36 (15)	Under	0	5	-	-	-
		Actual	7	5	2 (29%)	2 (29%)	3 (42%)
		Over	8	5	3 (38%)	2 (24%)	3 (38%)
	R41 (19)	Under	5	5	3 (60%)	1 (20%)	1 (20%)
		Actual	14	5	12 (86%)	2 (14%)	-
Gautheron et al. (1994)	NL (36)	Actual	13	11/12	7 (54%)	2 (15%)	4 (31%)
		Over	23	11/12	21 (91%)	-	2 (9%)
	R36 (4)	Under	0	11/12	-	-	-
		Actual	2	11/12	-	1 (50%)	1 (50%)
		Over	2	11/12	1 (50%)	1 (50%)	-
	R41 (8)	Under	2	11/12	1 (50%)	1 (50%)	-
		Actual	6	11/12	4 (67%)	1 (17%)	1 (17%)
Southee (1998)	NL (4)	Actual	2	3	2 (100%)	-	-
		Over	2	3	1 (50%)	1 (50%)	-
	R36 (5)	Under	1	3	-	-	1 (100%)
		Actual	3	3	3 (100%)	-	-
		Over	1	3	1 (100%)	-	-
	R41 (6)	Under	2	3	2 (100%)	-	-
		Actual	4	3	4 (100%)	-	-

Abbreviations: BCOP = bovine corneal opacity and permeability; EU = European Union; NL = Not Labeled as Irritant; R36 = moderate/mild irritant; R41 = severe irritant.
[1] EU classification system (EU 2001).

8.0 Bovine Corneal Opacity and Permeability Test Method Data Quality

8.1 Adherence to National and International GLP Guidelines

The original evaluation of BCOP test method data quality is detailed in the previous BCOP BRD (ICCVAM 2006a). As indicated in Section 8.0 of the AMCP BRD (2008) submission, it could not be determined whether all of the *in vitro* data contained in the AMCP BRD were generated under full GLP compliance. Where possible, that information is contained in the spreadsheets that form the database from which the AMCP BRD was generated. All of the new *in vitro* data that were generated during the course of constructing the current ICCVAM 2010 BRD were conducted with full GLP compliance.

9.0 Reports in the Peer-Reviewed Literature

NICEATM located among the peer-reviewed literature a total of four BCOP studies published since the previous evaluation of the BCOP method for identification of ocular corrosives and severe irritants (ICCVAM 2006a) that contained BCOP data (Cater and Harbell 2006, 2008; Debbasch et al. 2005; Van Goethem et al. 2006). The four publications contained BCOP test method analyses; however, none of these publications included raw data and therefore were not added to the database.

In Debbasch et al. (2005), 12 makeup removers were tested both in the BCOP and in a clinical in-use test under ophthalmological control after their application to the external eyelid. The undiluted test product (750 μL) was pipetted onto the corneas and exposure conducted for 4 hours. Corneal opacity was determined using an adapted spectrophotometer and barrier disruption by fluorescein update using OD_{490} mm. *In vitro* scores were classified according to Gautheron et al. (1994) and Harbell and Curren (1998). However, no *in vivo* rabbit eye data were reported, and these data have not been obtained. For this reason, Debbasch et al. (2005) was not included in the BCOP performance analyses detailed in this BRD.

In Cater and Harbell (2006), surfactant-based "rinse-off" personal care formulations were tested in the BCOP test method using slight modifications of the BCOP protocol reported by Sina et al. (1995). Corneas were exposed to the test substances (750 μL) for 10, 30, or 60 minutes either undiluted or diluted in deionized water. Corneas were evaluated for opacity, fluorescein uptake, and histological alterations. No *in vivo* rabbit reference data were reported, and thus this study was not included in the BCOP performance analyses detailed in this BRD.

Van Goethem et al. (2006) tested 20 substances in the BCOP test method (7 compounds classified as GHS Not Classified and 13 GHS Category 1). These results were published in Vanparys et al. (1993) and Gautheron et al. (1994), which were included in the previous BCOP BRD (ICCVAM 2006a).

In Cater and Harbell (2008), the BCOP test method was used on four commercial and one unregistered body wash developed for children or as mild bath products. The purpose was to determine if the BCOP test method could be used as a prediction model for relative ranking of human eye responses under conditions of a standard human eye sting test to surfactant-based formulations. Test articles were prepared as 25% solutions in deionized water; 750 μL was applied to the corneas for a 30 minute exposure. Following exposure, opacity and fluorescein uptake were determined *in vitro*, but no *in vivo* rabbit eye data were reported.

10.0 Animal Welfare Considerations (Reduction, Refinement, and Replacement)

10.1 How the BCOP Test Method Will Reduce, Refine, or Replace Animal Use

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

Reducing animal use through improved science and experimental design

Refining experimental procedures such that animal suffering is minimized

Replacing animal models with nonanimal procedures (e.g., *in vitro* technologies) where possible (Russell and Burch 1992)

The BCOP test method refines animal use. Because these animals are being humanely killed for nonlaboratory purposes, the testing procedure inflicts no additional pain or distress on animals. Substances that are identified as corrosive or severe irritants *in vitro* are excluded from *in vivo* testing. Furthermore, the ability to identify mild and moderate ocular irritants would eliminate the need for *in vivo* testing, thus sparing rabbits from the pain associated with these types of substances.

The BCOP test method can also reduce animal use because the test method utilizes animal species routinely raised as a food source in large numbers and thereby replaces laboratory animals. Additionally, with the ability to identify ocular corrosives and severe ocular irritants as well as mild and moderate ocular irritants from the *in vitro* method, the animals that would have been used in the *in vivo* rabbit eye test would be spared.

10.2 Requirement for the Use of Animals

Although cattle are required as a source of corneas for this *in vitro* test method, only cattle humanely killed for food or other nonlaboratory purposes are used as eye donors (i.e., no live animals are used in this test method).

11.0 Practical Considerations

Practical considerations for the BCOP method are detailed in the previous BCOP BRD (ICCVAM 2006a).

12.0 References

AMCP Background Review Document. 2008. Submission to ICCVAM: An *in vitro* approach of EPA toxicity labeling of anti-microbial cleaning products and annexes A–G (http://iccvam.niehs.nih.gov/methods/ocutox/antimicro/BRD.pdf).

Bailey PT, Freeman JJ, Phillips RD, Merrill JC. 2004. Validation of the BCOP assay as a predictor of ocular irritation of various petrochemical products. Poster presentation at the Society of Toxicology 2004 meeting (Baltimore, MD).

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.

Berdasco N, Gilbert K, Lacher J, Mattsson J. 1996. Low rate of severe injury from dermal and ocular irritation tests and the validity of using fewer animals. J Am Coll Toxicol 15:177–193.

Calabrese EJ. 1983. Dermatotoxicity: predictive models. In: Principles of Animal Extrapolation. New York:Wiley Interscience.

Carpenter CP, Smyth HF. 1946. Chemical burns of the rabbit cornea. Am J Opthalmol 29:60–73.

Cater KC and Harbell JW. 2006. Prediction of eye irritation potential of surfactant-based rinse-off personal care formulations by the bovine corneal opacity and permeability (BCOP) assay. Cutan Ocul Toxicol 25:217–233.

Cater KC and Harbell JW. 2008. Comparison of *in vitro* eye irritation potential by bovine corneal opacity and permeability (BCOP) assay to erythema scores in human eye sting test of surfactant-based formulations. Cutan Ocul Toxicol 27:77–85.

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

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

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

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

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

EEC. 1991. Commission Directive 91/325/EEC of 1 March 1991 adapting to technical progress for the twelfth time Council Directive 67/548/EEC on the approximation of the laws, Regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Communities L180:52–56.

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

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

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

EPA. 2003c. Good Laboratory Practice Standards. 40 CFR 160.

EU. 2001. Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Communities L225:1–333.

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

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

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

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

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

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

ICCVAM. 2006a. Background Review Document: Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method. NIH Publication No. 06-4512. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

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

ICCVAM. 2009. Independent Scientific Peer Review Panel Report: Evaluation of the Validation Status of Alternative Ocular Safety Testing Methods and Approaches. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

ICCVAM Authorization Act of 2000. 2000. Public Law 106-545.

Kay JH, Calandra JC. 1962. Interpretation of eye irritation tests. J Soc Cosmet Chem 13:281–289.

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

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

OECD. 2005. OECD Series on Testing and Assessment Number 34. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Paris:Organisation for Economic Co-operation and Development. Available: http://appli1.oecd.org/olis/2005doc.nsf/linkto/env-jm-mono(2005)14.

Russell WMS, Burch RL. 1992. The principles of humane experimental technique. 14th ed. Herts, England:Universities Federation for Animal Welfare.

Sina JF, Galer DM, Sussman RG, Gautheron PD, Sargent EV, Leong B, et al. 1995. A collaborative evaluation of seven alternatives to the Draize eye irritation test using pharmaceutical intermediates. Fundam Appl Toxicol 26:20–31.

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

Swanson JE, Harbell JW. 2000. Evaluating the eye irritancy potential of ethanolic test materials with the bovine corneal opacity and permeability assay. The Toxicologist 54:188–189.

Swanson JE, Lake LK, Donnelly TA, Harbell JW, Huggins J. 1995. Prediction of ocular irritancy of full-strength cleaners and strippers by tissue equivalent and bovine corneal assays. J Toxicol Cutaneous Ocul Toxicol 14:179–195.

United Nations (UN) (2007). Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Second revised edition, UN New York and Geneva, 2007. Available at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html.

Van Goethem F, Adriaens E, Alepee N, Straube F, De Wever B, Cappadoro M, et al. 2006. Prevalidation of a new *in vitro* reconstituted human cornea model to assess the eye irritating potential of chemicals. Toxicol in Vitro 20:1–17.

Vanparys P, Deknudt G, Sysmans M, Teuns G, Coussement W, Van Cauteren H. 1993. Evaluation of the bovine corneal opacity-permeability assay as an in vitro alternative to the Draize eye irritation test. Toxicol In Vitro 7:471–476.

13.0 Glossary[2]

Accuracy:[3] (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of "relevance." The term is often used interchangeably with *concordance* (see also *two-by-two table*). Accuracy is highly dependent on the prevalence of positives in the population being examined.

Assay:[3] The experimental system used. Often used interchangeably with *test* and *test method*.

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

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

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

Blepharitis: Inflammation of the eyelids.

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

Chemosis: A form of eye irritation in which the membranes that line the eyelids and surface of the eye (*conjunctiva*) 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:[3] 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.

[2] The definitions in this Glossary are restricted to their uses with respect to the Draize rabbit eye test method and the BCOP test method.

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

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

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

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

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

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

False negative:[3] A substance incorrectly identified as negative by a test method.

False negative rate:[3] 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:[3] A substance incorrectly identified as positive by a test method.

False positive rate:[3] The proportion of all negative substances that are falsely identified by a test method as positive (see *two-by-two table*). It is one indicator of test method accuracy.

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

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

Good Laboratory Practices (GLP):[3] Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the Organization for Economic Cooperation and Development and Japanese authorities that describe record keeping and quality assurance procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.

Hazard:[3] The potential for an adverse health or ecological effect. A hazard potential results only if an exposure occurs that leads to the possibility of an adverse effect being manifested.

Interlaboratory reproducibility:[3] 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:[3] 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:[3] The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.

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

***In vitro* irritancy score:** An empirically derived formula used in the BCOP assay 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 = mean opacity value + (15 x mean permeability value).

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

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

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

Negative predictivity:[3] The proportion of correct negative responses among substances testing negative by a test method (see *two-by-two table*). It is one indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

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

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

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.

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

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.

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

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

Performance:[3] The accuracy and reliability characteristics of a test method (see *accuracy*, *reliability*).

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

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

Positive predictivity:[3] 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:[3] The proportion of positives in the population of substances tested (see *two-by-two table*).

Protocol:[3] The precise, step-by-step description of a test method, including a listing of all necessary reagents, criteria, and procedures for evaluation of the test data.

Quality assurance:[3] 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:[3] A new or modified test method that reduces the number of animals required.

Reference test method:[3] 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:[3] A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals, or enhances animal well-being.

Relevance:[3] 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:[3] 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:[3] 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:[3] The consistency of individual test results obtained in a single laboratory (intralaboratory reproducibility) or in different laboratories (interlaboratory reproducibility) using the same protocol and test substances (see intra- and inter-laboratory reproducibility).

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

Sensitivity:[3] The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy (see *two-by-two table*).

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

Severe irritant: (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye that is not reversible within 21 days of application or causes serious physical

decay of vision. (b) Substances that are classified as EPA Category I, GHS Category 1, 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:[3] 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:[2] The experimental system used; used interchangeably with *test method* and *assay*.

Test method:[3] A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with *test* and *assay*. See also *validated test method* and *reference test*.

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

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

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

Transferability:[3] The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Two-by-two table:[3] The two-by-two table can be used for calculating accuracy (concordance) ([a+d]/[a+b+c+d]), negative predictivity (d/[c+d]), positive predictivity (a/[a+b]), prevalence ([a+c]/[a+b+c+d]), sensitivity (a/[a+c]), specificity (d/[b+d]), false positive rate (b/[b+d]), and false negative rate (c/[a+c]).

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

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

Validated test method:[3] 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:[3] The process by which the reliability and relevance of a procedure are established for a specific purpose.

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

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

This page intentionally left blank