2.0 THE BCOP TEST METHOD

2.1 BCOP Technical Summary

The following technical summary provides a synopsis of the performance analysis described in the BCOP BRD, which reviewed the available data and information for the test method.[6] The BRD describes the current validation status of the BCOP test method, including what is known about its reliability and accuracy, the scope of the substances tested, and a standardized protocol. The BRD may be obtained from the ICCVAM/NICEATM website (http://iccvam.niehs.nih.gov/).

2.1.1 Test Method Description

The BCOP test method is an organotypic model that provides short-term maintenance of normal physiological and biochemical function of the bovine cornea in an isolated system. 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 an ultraviolet/visible (UV/VIS) spectrophotometer, respectively. Both measurements are used to calculate an *In Vitro* Irritancy Score, which is used to assign an *in vitro* irritancy classification for prediction of the *in vivo* ocular irritation potential of a test substance. Although histopathological data could not be formally evaluated by ICCVAM, a histopathological assessment can be included on a case-by-case basis to discriminate borderline cases (i.e., substances that produce results that preclude assignment to a single category) or to identify ocular damage that does not produce opacity or permeability changes in the isolated cornea.[7] Histopathology also is used for chemical classes or formulations that are not well characterized in the BCOP assay, where the mode of action cannot be easily predicted, when delayed effects might be anticipated, or when a more complete characterization of damage is needed.

The BCOP test method protocols used in the various studies are similar, but not identical.[8] Variations in the publicly available BCOP protocols include different instrumentation to evaluate opacity, different decision criteria (i.e., prediction models) or *in vitro* classification systems, and differences in the use of positive controls, among other methodological variations. The essential principles of the test method protocol include isolating and culturing the bovine cornea, treating the isolated cornea with a test substance, collecting opacity and permeability data, and evaluating the data in relation to a prediction model. However, given the various uses and applications of the BCOP test method by different investigators and laboratories, and the evolution of the test method over time, a number of laboratory-specific differences have been noted regarding the conduct of the test method.

---

[6]Comparison of the performance analysis for BCOP to the other three *in vitro* test methods evaluated can be reviewed in **Section 6.0** and **Appendix B**.

[7]For the studies discussed here, histopathological endpoints were not evaluated or incorporated into the accuracy assessment.

[8]For additional information on this evaluation, please see the BCOP BRD (http://iccvam.niehs.nih.gov/methods/ocudocs/ocu_brd.htm#bcop).

2.1.2 Validation Database

A total of 158 substances in eight studies were used to evaluate BCOP test method accuracy. These substances represented a variety of chemical and product classes (ICCVAM 2006a). The chemical classes tested included alcohols, heterocyclic compounds, carboxylic acids, ketones, esters, inorganic salts, ethers, hydrocarbons, amines, and onium compounds. The product classes tested included solvents, surfactants, chemical/synthetic intermediates, drugs/pharmaceuticals/therapeutic agents, petroleum products, cleaners, personal care cleansers, hair shampoos, pesticides, plasticizers, reagents, bactericides, and insect repellents.

2.1.3 Test Method Accuracy

Based on all available data, the BCOP test method has an overall accuracy of 79% (113/143)[9] to 81% (119/147), when compared to *in vivo* rabbit eye test method data classified according to the EPA (1996), European Union (EU; 2001), or GHS (UN 2003) classification systems. Furthermore, the BCOP test method has an overall false positive rate of 19% (20/103) to 21% (22/103) and an overall false negative rate of 16% (7/43) to 25% (10/40), when compared to *in vivo* rabbit eye test method data classified according to the EPA (1996), EU (2001), or GHS (UN 2003) classification systems.

There were some notable trends in the performance of the BCOP test method among substances grouped according to chemical class and/or physicochemical properties (**Table 2-1**). The chemical classes of substances that were most consistently overpredicted (i.e., were false positives) by the BCOP test method, according to the GHS classification system, are alcohols (53%, 8/15) and ketones (40%, 4/10). With regard to physical form, liquids (26%, 18/68) appear more likely than solids (10%, 2/20) to be overpredicted by the BCOP test method.

Alcohols (67%, 2/3) also were most often underpredicted (i.e., were false negatives) by the BCOP test method, according to the GHS classification system. With regard to physical form, solids (42%, 5/12) appear more likely than liquids (4%, 1/24) to be underpredicted by the BCOP test method. There was no definitive difference among the underpredicted substances for which pH information was available.

BCOP test method performance statistics also were evaluated when substances from the classes that gave the most discordant results were excluded (i.e., alcohols, ketones, solids). When using the GHS classification system, exclusion of alcohols and ketones individually resulted in small changes in the performance statistics. However, exclusion of solids from the data set caused a four-fold decrease in the false negative rate from 16% (7/43) to 4% (1/29). When both alcohols and ketones were excluded, the accuracy increased from 81% (119/147) to 88% (103/117) and the false positive rate decreased from 20% (21/104) to 12% (9/77). The largest changes were observed when all three discordant classes were excluded from the data set; accuracy increased to 92% (78/85), the false positive rate decreased to 12% (7/58), and the false negative rate decreased to 0% (0/27).

[9]The numbers in parentheses represent the data used to calculate the percentages noted.

Table 2-1 False Positive and False Negative Rates of the BCOP Test Method, by Chemical Class and Properties of Interest, for the GHS Classification System

Category	N[1]	False Positive Rate[2]		False Negative Rate[3]	
		%	No.[4]	%	No.
Overall	147	20	21/104	16	7/43
Chemical Class[5]					
Alcohol	18	53	8/15	67	2/3
Amine/Amidine	8	0	0/4	0	0/4
Carboxylic acid	15	38	3/8	14	1/7
Ester	12	12	1/8	0	0/4
Ether/Polyether	6	0	0/5	0	0/1
Heterocyclic	12	33	2/6	17	1/6
Hydrocarbon	12	8	1/12	-	0/0
Inorganic salt	5	0	0/3	0	0/2
Ketone	10	40	4/10	-	0/0
Onium compound	11	0	0/3	0	0/8
Properties of Interest					
Liquids	92	26	18/68	4	1/24
Solids	32	10	2/20	42	5/12
Pesticide	8	33	1/3	40	2/5
Surfactant – Total[6]	35	5	1/21	7	1/14
-nonionic	5	0	0/4	0	0/1
-anionic	3	0	0/2	100	1/1
-cationic	6	0	0/1	0	0/5
pH – Total[7]	28	-	-	21	5/24
- acidic (pH < 7.0)	11	-	-	18	2/11
- basic (pH > 7.0)	15	-	-	23	3/13
- equals 7	2	-	-	-	-
Category 1 Subgroup[8] **- Total**	38[10]	-	-	18	7/38
- 4 (CO=4 at any time)	20	-	-	15	3/20
- 3 (severity/persistence)	1	-	-	0	0/1
- 2 (severity)	4	-	-	25	1/4
- 2-4 combined[9]	25	-	-	16	4/25
- 1 (persistence)	13	-	-	23	3/13

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability; CO = corneal opacity; GHS = Globally Harmonized System (UN 2003).

[1]N = number of substances.

[2]False Positive Rate = the proportion of all negative substances that are falsely identified as positive *in vitro*.

[3]False Negative Rate = the proportion of all positive substances that are falsely identified as negative *in vitro*.

[4]Data used to calculate the percentage.

[5]Chemical classes included in this table are represented by at least five substances tested in the BCOP test method and assignments are based on the MeSH categories (www.nlm.nih.gov/mesh).

[6]Combines single chemicals labeled as surfactants along with surfactant-containing formulations.

[7]Total number of GHS Category 1 substances for which pH information was obtained.

[8]NICEATM-defined subgroups assigned based on the lesions that drove classification of a GHS Category 1 substance. 1: based on lesions that are persistent; 2: based on lesions that are severe (not including CO=4); 3: based on lesions that are severe (not including CO=4) and persistent; 4: CO=4 at any time.

[9]Subcategories 2 to 4 combined to allow for a direct comparison of GHS Category 1 substances classified *in vivo* based on some lesion severity component and those classified based on persistent lesions alone.

[10]The number of substances evaluated in the Category 1 subgroup analysis may be less than the total number of *in vivo* Category 1 substances evaluated since some substances could not be classified into the subgroups used in the evaluation.

Finally, the underpredicted substances were more likely to be classified *in vivo* (according to the GHS classification system) based on persistent lesions, rather than on severe lesions. However, three substances that caused severe lesions *in vivo* (corneal opacity=4) were false negatives in BCOP.

The performance statistics for the EPA and EU classification systems are similar to those discussed for the GHS classification system. Additional information on the performance characteristics of the BCOP test method for the EPA and EU classification systems can be obtained from **Section 6.0**, **Appendix B**, and the BCOP BRD.

2.1.4 Test Method Reliability (Inter- and Intra-Laboratory Reproducibility)

Quantitative BCOP test method data were available for replicate corneas within individual experiments and for replicate experiments within an individual laboratory for three studies. Therefore, an evaluation of the intralaboratory repeatability and reproducibility of the BCOP test method could be conducted. Intralaboratory repeatability of *In Vitro* Irritancy Scores was assessed by analyzing two studies (*In Vitro* Scores ≥55.1). For substances of varying irritancy in one study (three laboratories evaluated), the median coefficient of variation (CV) for *In Vitro* Irritancy Scores for replicate corneas (n=3) ranged from 11.8% to 14.2%. In a second study, mean and median CV values for *In Vitro* Irritancy Scores for replicate corneas (n=4) was 71% to 35%, respectively.

A CV analysis of intralaboratory data (*In Vitro* Irritancy Scores) from two studies indicated the following intralaboratory reproducibility of the BCOP test method. In one study, the between experiment (n=3) mean and median CV values for permeability values were 33.4% and 29.0%, respectively, for 25 surfactant-based personal care cleaning formulations. In the second study, the between experiment mean CV values of *In Vitro* Irritancy Scores for 16 substances tested two or more times in three laboratories ranged from 12.6% to 14.8%, while the median CV values ranged from 6.7% to 12.4%.

Additionally, comparable BCOP data were available for multiple laboratories within each of three comparative validation studies, which allowed for an evaluation of the interlaboratory reproducibility of the BCOP test method. For these studies, interlaboratory reproducibility was evaluated qualitatively based on the ocular irritancy classification assigned to each substance by each laboratory, and quantitatively using *In Vitro* Irritancy Scores. In the qualitative assessment of interlaboratory reproducibility of hazard classification category, 67% to 94% of the substances were classified the same by the participating laboratories. Substances with less than complete agreement in the testing laboratories include those representing such chemical classes as alcohols, ketones, and heterocyclic compounds, and such product classes as solvents, surfactants, chemical intermediates, and pesticides.

A quantitative evaluation of interlaboratory reproducibility also was conducted for these three studies by performing a CV analysis of *In Vitro* Irritancy Scores obtained for substances tested in multiple laboratories. In one study, the 17 substances predicted as severe in the BCOP assay had mean and median CV values of 36% and 17%, respectively, for results obtained in either 11 or 12 laboratories. In a second study, the 32 substances predicted as severe in the BCOP assay had mean and median CV values of 25% and 22%,

respectively, for results obtained in five laboratories. In a third study, the mean and median CV values for the *In Vitro* Irritancy Scores of the 16 tested substances were 32.4% and 22.8%, respectively, for results obtained in three laboratories.

Finally, the interlaboratory correlation between BCOP test method endpoint data generated by each laboratory was determined for 60 substances, as well as for various subsets of test substances (water-soluble, water-insoluble, surfactants, solids, solutions, and liquids). This analysis yielded a range of correlation coefficients for the subsets of test substances. Interlaboratory correlation coefficients for the *In Vitro* Irritancy Score generally spanned a range of 0.867 to 0.958 depending on the specific subsets of substances being evaluated.

2.2 ICCVAM Recommendations for the BCOP Test Method

2.2.1 Use of the BCOP Test Method

ICCVAM recognizes that the BCOP test method is not proposed as a stand alone replacement for the *in vivo* rabbit eye test method currently used for regulatory classification and labeling. ICCVAM concludes that there are sufficient data to support the use of the BCOP test method, in appropriate circumstances and with certain limitations, as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, UN GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach.[10]

The identified limitations for this test method are based on the false negative and false positive rates that are observed for certain chemical and physical classes. Based on the available database, the false negative rates for alcohols and solids range from 67% (2/3) to 100% (2/2) and 42% (5/12) to 50% (5/10), respectively, depending on the hazard classification system. Additionally, the false positive rates for alcohols, ketones, and solids range from 50% (7/14) to 56% (9/16), 40% (4/10), and 10% (2/20 to 2/21), respectively, depending on the hazard classification system. When substances within these chemical and physical classes are excluded from the database, the accuracy of BCOP across the EU, EPA, and GHS classification systems ranges from 87% (72/83) to 92% (78/85) and the false negative and false positive rates range from 0% (0/27) to 12% (3/26) and 12% (7/58) to 16% (9/56), respectively.

A tiered-testing strategy for ocular irritation/corrosion (e.g., as described in the Globally Harmonized System of Classification and Labelling of Chemicals; UN 2003) allows for the use of validated and accepted *in vitro* methods prior to the use of animals for ocular safety testing. In a tiered-testing strategy, when a positive result is obtained in an appropriately validated *in vitro* test, a test substance may be classified as an ocular hazard without testing in rabbits. A substance that tests negative in the *in vitro* ocular toxicity test would need to be tested in the *in vivo* ocular test to identify possible *in vitro* false negatives and to identify moderate and mild ocular irritants. As is appropriate for any test system, there is the opportunity for confirmatory testing if false positive results are suggested based on a weight-

[10]The recommendations are based on the performance results for BCOP without the use of histopathology for decision making purposes.

of-evidence evaluation of supplemental information (e.g., pH, structure-activity relationships, other testing data). Using *in vitro* data in a tiered-testing strategy with a weight-of-evidence decision process to classify substances as ocular corrosives or severe irritants will avoid the potential pain and distress that might be experienced by rabbits who otherwise would have been administered these test substances. A tiered-testing strategy may not be applicable to purposes other than regulatory classification and labeling.

Users should be aware that BCOP's performance characteristics could be revised as additional data become available. For example, the current validation database did not allow for adequate evaluation of all chemical or product classes (e.g., formulations). Additional data may allow for further evaluation of this, as well as other, chemical and product classes. Therefore, prior to initiation of BCOP studies, investigators are encouraged to consult the ICCVAM/NICEATM website (see http://iccvam.niehs.nih.gov/methods/eyeirrit.htm) to review the most current validation database, overall performance characteristics, and chemical and physical class performance characteristics. Evaluation of the most current information will allow users to determine the appropriateness of this test method for evaluating substances that are within a specific chemical, physical, or product classes.

2.2.2 BCOP Test Method Protocol

ICCVAM recommends that when testing is conducted, the BCOP test method protocol should be based on the BCOP standardized test method protocol provided in **Appendix D**. This will facilitate collection of consistent data and expand the current validation database. Exceptions and/or changes to the proposed standardized test method protocol should be accompanied by a scientific rationale. Users should be aware that the test method protocol could be revised based on future optimization and/or validation studies. ICCVAM, therefore, recommends that test method users consult the ICCVAM/NICEATM website to ensure use of the most current recommended test method protocol (http://iccvam.niehs.nih.gov/methods/eyeirrit.htm).

2.2.3 Optimization of the Current BCOP Test Method Protocol

The current ICCVAM recommendations are focused on the use of the BCOP test method as a screening test for ocular corrosives and severe irritants (see **Section 2.2.1**). For that use, the current test method protocol should be sufficient. To further the use of this test method and to evaluate the use of the BCOP test method as a potential replacement for the *in vivo* rabbit eye test method or for the identification of mild and moderate ocular irritants (e.g., EPA Category II, III, and IV; GHS Category 2; EU R36), ICCVAM recommends additional studies be considered and undertaken to decrease the false positive rate of this test method.

A histopathological evaluation of the corneal tissue, using a standardized scoring scheme, should be conducted. Such data will allow for the development of standardized decision criteria and a more comprehensive evaluation of the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.

Studies should be conducted to evaluate the impact of using a corneal holder that maintains normal corneal curvature (e.g., the corneal mounting system designed by Ubels et al. 2002) on accuracy and/or reliability of the BCOP test method.

ICCVAM also recommends that an evaluation be conducted on the effect of modifying various test method protocol components (e.g., duration of test substance exposure) on the accuracy and/or reliability of the BCOP test method.