Draft Proposed ICCVAM Test Method Recommendations: Evaluation of the Validation Status of Alternative Ocular Safety Testing Methods and Approaches

April 1, 2009

The draft Background Review Documents supporting these draft recommendations are available at http://iccvam.niehs.nih.gov/methods/ocutox/PeerPanel09.htm. The draft Background Review Documents and the draft recommendations will be considered by an independent scientific peer review panel that will meet in public session on May 19-21, 2009 at the Consumer Product Safety Commission headquarters in Bethesda, MD. Public comments are welcome. More information is available in the Federal Register Notice of the meeting, available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/E9-7220.pdf. ICCVAM will finalize these recommendations after consideration of comments from the peer review panel, the public, and its scientific advisory committee.

These draft recommendations do not represent the official position of any Federal agency.

TABLE OF CONTENTS

1.0 Use of Topical Anesthetics, Systemic Analgesics, and Humane Endpoints in Ocular Toxicity Testing to Avoid or Minimize Pain and Distress

1.1 Draft Proposed ICCVAM Recommendations: Use of Topical Anesthetics Systemic Analgesics in Ocular Toxicity Testing to Avoid or Minimize Pain and Distress

ICCVAM proposes the following draft test method recommendations on the use of topical anesthetics and systemic analgesics to avoid or minimize pain and distress in acute eye irritation testing. ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document for this topic (available at http://iccvam.niehs.nih.gov/methods/ocutox/pretreat/BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, proposed modifications to the current standardized test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

Since 1984, the U.S. Consumer Product Safety Commission (CPSC) has recommended preapplication of tetracaine ophthalmic anesthetic for all rabbit eye toxicity studies. However, current EPA and OECD test guidelines for the rabbit eye test state that topical anesthetics can only be used if the user demonstrates that such pretreatments do not interfere with the results of the tests[1]. Therefore, they often are not used because a separate study would likely be necessary to provide such information.

The use of topical ophthalmic anesthetics and/or systemic analgesics during the conduct of the Draize rabbit eye irritation test was evaluated at a recent NICEATM/ICCVAM/ECVAM scientific symposium entitled “Minimizing Pain and Distress in Ocular Toxicity Testing”. While invited experts acknowledged that a single

[1] OECD TG 405 states, "The type, concentration, and dose of a local anesthetic should be carefully selected to ensure that differences in reaction to the test substance will not result from its use." Similarly, EPA (1998) states that, " The type and concentration of the local anesthetic should be carefully selected to ensure that no significant differences in reaction to the test substance will result from its use."

treatment with a topical anesthetic to anesthetize the surface of the cornea prior to the application of the test article to the eye could potentially cause slight physiologic changes, the consensus was that such alterations to the irritant response would be slight if any. Furthermore, the predominant view was that if there were any effects on the irritant response, it would tend to slightly increase the severity of the response. Therefore, the routine use of topical anesthetics was recommended, since the anesthetics at least avoid the discomfort experienced from installation of the test article on the eye, and temporarily avoid or minimize pain and distress that might result from immediate ocular damage. Experts also recommended that pretreatment with topical anesthetics combined with systemic analgesics should be routinely used to avoid pain, and that animals exhibiting clinical signs of pain or distress or with ocular lesions associated with painful conditions should continue to be treated with systemic analgesia.

A recent evaluation by NICEATM of the effects of pretreatment with tetracaine hydrochloride (0.5% w/v) on the ocular irritancy potential of 97 formulations indicate that such pretreatments had no statistically significant impact on the hazard classification severity category of observed ocular irritation. For a majority of the formulations tested, topical anesthetic pretreatment had no or minimal impact on:

- The hazard classification severity category of observed ocular irritation
- The variability in ocular irritation responses among animals treated with the same test article
- The number of days required for an ocular lesion to clear.

When a difference in ocular irritation response was observed in animals pretreated with topical anesthesia compared to animals that were not pretreated, the more severe response was more frequently observed in the pretreated animals. However, none of the observed differences were statistically significant. The observed differences occurred in both directions (increasing and decreasing the level of irritancy), which suggests that they are likely related to the inherent inter-individual biological variability of response rather than topical anesthetic pretreatment.

The draft proposed ICCVAM recommendations that follow were developed based on available data in conjunction with clinical experience and expert judgment.

Usefulness and Limitations

In order to avoid or minimize potential pain and distress caused by test article administration and initial injuries in the Draize rabbit eye test, ICCVAM proposes the routine use of a topical anesthetic (i.e., tetracaine or proparacaine, 1-2 drops of 0.5% w/v solution) and an opioid systemic analgesic (i.e., buprenorphine, 0.05 mg/kg) prior to instillation of a test substance, unless there is an adequate scientific rationale for not using these substances. Anti-inflammatory analgesics (e.g., nonsteroidal anti-inflammatory drugs) are not recommended because of their possible influence on study results due to demonstrated effects on the wound healing process. In addition, treatment with an opioid systemic analgesic (i.e., buprenorphine, 0.05 mg/kg, q 12 hr) should continue as long as a test animal displays clinical signs of more than momentary or slight pain or distress (e.g., blepharospasm, excessive lacrimation, pawing at the treated eye) or has ocular injuries expected to cause or be associated with pain or distress (e.g., opacity, iritis, conjunctival redness, chemosis scores ≥ 2). Users should also consider the humane endpoints detailed in **Section 1.2**, which could justify early termination of the study.

Test Method Protocol

When required for ocular safety testing, the current Draize eye test protocol used for regulatory safety assessments of potential ocular hazards (EPA 1998, OECD 2002) should be conducted with the ICCVAM proposed modifications for the use of topical anesthetics and systemic analgesics. These modifications include the following procedures. Prior to instillation of a test substance, the animal is given a single dose of a systemic opioid analgesic (i.e., buprenorphine, 0.05 mg/kg SC, IM) and a topical anesthetic (i.e., tetracaine or proparacaine, 2 drops of 0.5% w/v solution). After test substance application, the animal is carefully observed for any clinical signs of pain and distress. Treatment with a systemic analgesic (i.e., buprenorphine, 0.05 mg/kg SC, IM, q 12 hr) should continue after instillation of the test substance if a test animal displays clinical signs of more than momentary or slight pain or distress (e.g., blepharospasm, excessive lacrimation, pawing at the treated eye) or ocular injuries expected to cause pain

or distress; in this case a regular treatment regimen (i.e., every 12 hr) should proceed until such signs or injuries are no longer present. While the choice of analgesic and its dosage should be made by the attending veterinarian because of the many variables associated with pain management, the recommended analgesic and associated dose (buprenorphine, 0.05 mg/kg) is based on its long history of successful veterinary use as an analgesic for moderate to severe pain in rabbits (Kohn et al. 2007[2]). Buprenorphine is also available in a transdermal patch that provides up to 4 days of controlled release drug, and this could be considered as an option to more frequent dosing.

Proposed Future Studies

Routine observation and recording of lesions and clinical signs is recommended during ocular irritation safety studies to evaluate efficacy in order to optimize analgesic dose and treatment schedule. Periodic review of these data should be performed to determine if adjustments are needed to improve the effectiveness of pre-treatment and post-treatment analgesia. Ideally, data should be collected during routine safety testing that could be analyzed periodically to determine the efficacy for specific types of lesions and clinical signs of pain and distress associated with ocular irritation/corrosivity testing.

ICCVAM encourages users to provide all data generated using the modified test method protocols to NICEATM to create a database that can be periodically evaluated to further characterize the usefulness and limitations of topical anesthetics and systemic analgesics for avoiding or minimizing pain and distress in ocular safety assessments.

[2] Kohn D, Martin E, Foley P, Morris T, Swindle M, Vogler G, Wixon S. 2007. Guidelines for the Assessment and Management of Pain in Rodents and Rabbits. J Am Assoc Lab Animal Sci. 46: 97-108.

1.2 Draft Proposed ICCVAM Recommendations on the Use of Humane Endpoints in Ocular Toxicity Testing

ICCVAM proposes the following draft test method recommendations on the use of humane endpoints to avoid or minimize pain and distress in ocular toxicity testing. ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document for this topic (available at: http://iccvam.niehs.nih.gov/methods/ocutox/pretreat/BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, proposed modifications to the current standardized test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

Public Health Service policy and U.S. Department of Agriculture (USDA) regulations on pain and distress in laboratory animals state that more than momentary or slight pain and distress:

- Should be limited to that which is unavoidable for the conduct of scientifically valuable research or testing
- Should be conducted with appropriate pain relief medication unless justified in writing by the principal investigator
- Should continue for only the necessary amount of time required to attain the scientific objectives of the study

These regulations also state that animals suffering severe or chronic pain or distress that cannot be relieved should be humanely killed after or, if appropriate, during the procedure, and finally, that Institutional Animal Care and Use Committees must ensure that the principal investigator complies with the requirements.

Participants at the 2005 symposium "Minimizing Pain and Distress in Ocular Toxicity Testing" also discussed early adverse responses predictive of ocular lesions associated with severe irritant or corrosive substances (GHS Category I [UN 2003], EU Category

R41 [EU 2001], or EPA Category I [EPA 1996]) that could be used routinely as humane endpoints to terminate a study. Among the invited participants were human and veterinary ophthalmologists and anesthesiologists, scientific experts in ocular hazard testing, research scientists, and industrial toxicologists. Subsequent to these discussions, the endpoints described below were recommended for routine use.

Usefulness and Limitations

ICCVAM recognizes that current ocular testing guidelines include guidance that allow for certain types of severe ocular injuries, or evidence of severe pain and distress, to be used as criteria for study termination for humane reasons (OECD 2000, OECD 2002, EPA 1998). In addition there is international guidance on general humane endpoints that can be used as the basis for ending an experiment (OECD, 2002). ICCVAM recommends that the following ocular lesions, which are considered to be predictive of a severe irritant or corrosive response and are not expected to fully reverse by the end of the 21-day post-treatment observation period, should be considered and used as humane endpoints to terminate studies early where determined appropriate:

- Endpoints currently accepted for study termination (OECD, 2000):
  - Draize corneal opacity score of 4 that persists for 48 hr
  - Corneal perforation or significant corneal ulceration including staphyloma
  - Blood in the anterior chamber of the eye
  - Absence of light reflex that persists for 72 hr
  - Ulceration of the conjunctival membrane
  - Necrosis of the conjunctiva or nictitating membrane
  - Sloughing
- Vascularization of the corneal surface (i.e., pannus)
- Greater than 75% of the limbus destroyed
- Area of fluorescein staining not diminishing over time based on daily assessment
- Lack of re-epithelialization five days after application of the test substance

- Extent of depth of injury to the cornea (routinely using slit-lamp and fluorescein staining) where corneal ulceration extends beyond superficial layers of the stroma or the depth of injury increases over time

Given the many years of clinical experience represented by the Symposium participants, ICCVAM considers that consideration and use of the recommended humane endpoints where determined appropriate can aid in further minimizing the duration and severity of pain and distress for animals used in ocular toxicity testing. However, while these endpoints are recommended for consideration as additional humane endpoints, a minority view expressed by some members of the ICCVAM Ocular Toxicity Working Group is that some of the recommended endpoints should not automatically be used as a basis to terminate a study (i.e. pannus, fluorescein staining).

Test Method Protocol

Ocular safety assessment studies should be conducted using the ICCVAM recommended modifications to the current Draize eye test protocol for regulatory safety assessments of potential ocular hazards (EPA 1998, OECD 2002). These include incorporation of the recommended humane endpoints and the following language.

As described in EPA (1998) and OECD (2002), eyes should be examined at 24, 48, and 72 hours after treatment with a test substance. Evaluations can be facilitated by use of a hand slit-lamp or other appropriate ophthalmologic device. After recording observations at 24 hr post-treatment, the eyes can be examined with the aid of fluorescein at each observation time point. Accordingly, one drop of sodium fluorescein U.S.P (or equivalent) is dropped directly onto the corneal surface. After flushing out excess fluorescein with sodium chloride solution (or equivalent) injured areas of the cornea appear yellow. Digital photographs during all fluorescein staining observations may add clarity toward accurately evaluating changes in the extent or depth of staining corresponding to a lesion that is not likely to reverse

Proposed Future Studies

ICCVAM encourages users to provide to NICEATM all data that are generated using these modifications so NICEATM can create a database that can be periodically

237 evaluated to further characterize the usefulness and limitations of using the proposed
238 humane endpoints to avoid or minimize pain and distress in ocular safety assessments.
239
240

2.0 Draft Proposed ICCVAM Recommendations The Low Volume Eye Test (LVET)

ICCVAM proposes the following draft test method recommendations on the low volume eye test (LVET). ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document for this test method (available at: http://iccvam.niehs.nih.gov/methods/ocutox/antimicro/LVET-BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, standardized test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

The review of the validity of the LVET was undertaken because LVET data is used to support the validity of one if the *in vitro* test methods proposed in the *in vitro* testing strategy for antimicrobial cleaning products. The accuracy of the LVET was compared to the Draize test and to available human data and experience.

The LVET data, as well as the comparative traditional Draize rabbit data with which to evaluate the accuracy of the LVET, are only available for limited types and numbers of substances (i.e., surfactant-containing personal and household cleaning products). The available comparative LVET and human (clinical studies and accidental exposures) data proposed to support its accuracy are largely with substances that are mild irritants or nonirritating (which also are predominantly surfactant containing cosmetic and personal care product formulations). Ethical considerations have limited the types of substances that can be tested in human clinical studies. As a result, LVET comparisons to human clinical study data are based on tests with mild irritants or substances not labeled as irritants. Such data provide little assurance to the regulatory agencies charged with protecting public health that the LVET can provide adequate protection from substances that may cause moderate or severe ocular injuries in humans.

Accidental exposures are not generally considered to be a reliable source of the true ocular hazard potential since such exposures are likely immediately followed by flushing

the eyes with large volumes of water, and may not represent the most severe lesion that might be produced by such an exposure. Such accidental exposures as human reference data do not allow definitive quantitative measures of amount and time of exposure.

Thus while the LVET is proposed as more likely to approximate the volume of a substance that could enter the human eye experimentally, there is limited data to indicate whether it can accurately identify the ocular hazard of substances known to cause moderate, severe, or permanent human ocular injuries. In contrast, there are no documented instances where a substance with a hazard category determined in the Draize eye test produced a more severe hazard category response in humans following accidental exposures or ethical human studies.

Usefulness and Limitations

A review of available data regarding the usefulness and limitations of the LVET (see ICCVAM Background Review Document available at (http://iccvam.niehs.nih.gov/methods/ocutox/antimicro/LVET-BRD.pdf) determined that:

- LVET under-predicts severe irritants compared to the Draize;
- There are insufficient data to evaluate the extent of under-prediction relative to known human severe ocular irritants
- There is an inconsistent relationship between LVET and Draize results (i.e., time-to-clear) for substances with available human data.

Accordingly, ICCVAM proposes that the LVET has not been adequately validated and does not have adequate demonstrated performance (sensitivity and specificity) to serve as an acceptable reference test method against which to determine the validity of *in vitro* alternative test methods for hazard classification and labeling purposes.

Test Method Protocol

Any future validation studies conducted to further evaluate the usefulness and limitations of the LVET should use the LVET protocol as originally developed by Griffith et al. (1980). The LVET differs from the Draize rabbit eye test by applying 10 µL instead of 100 µL volume of the test substance, and applying the test substance directly on the

cornea instead of in the conjunctival sac. Scoring of corneal, iridal, and conjunctival lesions in the LVET is identical to that of the Draize rabbit eye test (EPA 1998, OECD 2002). In addition, due to the increased potential for pain from administering the test article directly onto the corneal surface, routine pre-treatment with topical anesthetics and systemic analgesics is recommended unless there is an adequate scientific rationale for withholding such pretreatments.

Proposed Future Studies

If an organization or sponsor desires to more adequately characterize the usefulness and limitations of the LVET, ICCVAM recommends that a comprehensive set of reference substances be tested and compared to Draize eye test results and human responses, where available. This reference list should be representative of the many types of substances that are evaluated for their ocular toxicity potential and include substances that are known to cause moderate, severe, and corrosive responses in humans.

3.0 *In Vitro* Testing Strategies for Ocular Hazard Classification of Antimicrobial Cleaning Products

ICCVAM proposes the following draft test method recommendations on in vitro testing strategies for ocular hazard classification of antimicrobial cleaning products. ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document and Summary Review Document for this topic (available at: http://iccvam.niehs.nih.gov/methods/ocutox/antimicro/BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, proposed test method protocols, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

The AMCP BRD included data for 228 substances tested in one or two of the three in vitro test methods proposed for use in the testing strategy. However, none of the substances had been tested in all three *in vitro* test methods. Therefore, there are no data available for the proposed substances with which to characterize the actual performance of a testing strategy that includes BCOP, CM, and EO. Of the 228 substances, 28 are EPA registered anti-microbial cleaning products, with eight additional materials being in-use dilutions of EPA registered antimicrobial concentrates.

In addition, the test method protocol used to generate the *in vivo* reference data varied among the 228 substances included in the validation database. Most of the substances tested in the BCOP (85% [58/68]) were tested in the traditional Draize rabbit eye test protocol (i.e., EPA 1998; OECD 2002). Approximately half (54% [29/54]) of the substances tested in EO were tested in the Draize rabbit eye test, while the remaining substances (46% [25/54]) were tested in the low volume eye test (LVET). All 105 of the substances tested in CM were tested in the LVET. The LVET is a modification to the rabbit eye test that involves application of 10 µL of the test substance directly to the corneal surface instead of 100 µL of the test substance applied into the conjunctival sac. As noted in **Section 2.0**, the draft OTWG position is that the LVET predictivity for the Draize test and the lack of LVET data for substances that are known to cause moderate

and severe irritation and ocular corrosion makes it inadequate to serve as a reference test method to support the validity of *in vitro* test methods. For this reason, the CM and some EO data for which only LVET data exists, were not considered adequate to support the proposed testing strategy.

However, additional data on 53 surfactant and surfactant-containing formulations were provided in a BRD prepared by ECVAM where there was data from the traditional Draize rabbit test available to assess the accuracy of the CM test method. These substances were not claimed as AMCPs, but they were surfactant-containing formulations with similar composition to many AMCPs. The database of 53 water-soluble surfactants tested in CM includes 21 surfactant chemicals and 32 surfactant-containing formulations tested across seven different laboratories. Based on the performance of CM using these 53 substances, ICCVAM has proposed[1] that the CM test method can be used as a screening test to identify water-soluble surfactant chemicals and certain types of surfactant-containing formulations (e.g., cosmetics and personal care product formulations, but not pesticide formulations) as either EPA Category I, GHS Category 1, or EU Category R41; or as EPA Category IV, GHS Not Labeled, EU Not Classified in a tiered-testing strategy, as part of a weight-of-evidence approach. A substance that is not classified into one of these two categories would need to be tested in another test method that is capable of correctly identifying possible *in vitro* false positives. Positives would also need to be additionally tested with methods that can correctly identify severe, moderate, and mild ocular irritants (for more detail, see ICCVAM Draft Proposed Recommendations on Cell Function-Based Assays for Identifying All Categories of Ocular Hazard). Analyses performed to identify the ocular hazard potential of these non-AMCP test substances based on Draize reference data suggest that the CM test method could be useful in a testing strategy.

An alternative testing strategy, which would include only BCOP and EO, was also evaluated using two approaches: 1) test in BCOP first and then in EO, or 2) test in EO first and then BCOP. For the first approach, the BCOP was evaluated for its ability to

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

identify substances as either Category I or II. All substances that were classified as Category I or II in BCOP (n=15) were removed from the database and the remaining 13 substances were evaluated based on EO results for identifying Category III or IV substances. The reverse was done for the second approach; the EO was evaluated for its ability to identify substances as either Category III or IV and all substances that were classified as Category III or IV in EO (n=13) were removed from the database and the remaining 15 substances were evaluated based on BCOP results for identifying Category I or II substances. Regardless of which approach was used, the performance of the proposed BCOP/EO testing strategy was the same. The BCOP/EO testing strategy correctly classifies 79% (22/28) of the substances, which includes identifying 100% (14/14) of the Category I substances, 100% (4/4) of the Category III substances, and 44% (4/9) of the Category IV substances. The one Category II substance in the database was underclassified as a Category III. None of the irritant categories (i.e., Category I, II, or III) were underclassified as Category IV substances.

Usefulness and Limitations

Given the limitations of the available database for three *in vitro* test methods (the cytosensor microphysiometer [CM], the EpiOcular™ [EO], and the bovine corneal opacity and permeability [BCOP] test methods), there are currently insufficient data with which to adequately demonstrate that an *in vitro* testing strategy using these BCOP, CM, and EO can identify all four required EPA hazard categories for ocular irritation/corrosion.

None of the 228 AMCPs included in the validation database have been tested in all three *in vitro* methods. There are a limited number of AMCPs (n = 28) that have been tested in both BCOP and EO. However, of these, there is only one EPA Category II substance and only four EPA Category III substances (based on Draize eye test results). Therefore, although the performance of a testing strategy using BCOP and EO appears to be useful for identifying Category I substances using BCOP and Category IV substances using EO, there are insufficient data with which to adequately demonstrate that this strategy can identify all four required EPA hazard categories for ocular irritation/corrosion.

Therefore, definitive recommendations on the usefulness and limitations of an *in vitro* testing strategy cannot be made at this time.

Test Method Protocols

The detailed test method protocols appended to the AMCP BRD submission use a variety of endpoints to predict ocular irritation potential. While they have not been demonstrated to be adequately validated for use in a testing strategy for AMCPs, decision criteria have been developed to correspond to the four different categories of ocular irritation defined by the EPA hazard classification system (i.e., EPA Categories I-IV). ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of an *in vitro* testing strategy.

Proposed Future Studies

Given the limitations in the validation database, a reference list of AMCPs (for which high quality Draize eye test data are available) should be tested prospectively in each of the proposed test methods (BCOP, Cytosensor, and EpiOcular) to allow for a more complete evaluation of the usefulness and limitations of an *in vitro* testing strategy.

Industry stakeholders are encouraged to provide strategies and approaches that are currently used for corporate decisions on product safety in an integrated decision strategy, including the various types of data and information and the respective qualitative and quantitative decision criteria.

4.0 *In Vitro* Alternative Test Methods for Identifying Ocular Hazard Categories

ICCVAM previously evaluated the validation status of the BCOP, HET-CAM, ICE, and IRE test methods for their ability to identify ocular corrosives and severe irritants, and considered BCOP and ICE to have sufficient performance to substantiate their use for regulatory hazard classification testing of some types of substances. The IRE and HET-CAM assays lacked sufficient performance and/or sufficient data to substantiate their use for regulatory hazard classification. ICCVAM subsequently recommended that the BCOP and ICE should be used in a tiered-testing strategy, where positive substances can be classified as ocular corrosives or severe irritants without the need for animal testing. ICCVAM is now reviewing the validation status of these *in vitro* test methods for identifying nonsevere ocular irritants (i.e., those that induce reversible ocular damage) and substances not labeled as irritants.

4.1 The ICE Test Method

ICCVAM proposes the following draft test method recommendations on the ICE test method. ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document for this topic (available at: http://iccvam.niehs.nih.gov/methods/ocutox/mildmod/ICE-BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, a proposed test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

The test method recommendations described herein are based upon two analyses of ICE test method performance:

- The overall correct classifications for ICE test method ranged from 59% (83/141) to 77% (118/153), depending on the hazard classification system evaluated when using the entire database; and 64% (49/77) to 80% (66/82) depending on the hazard classification system evaluated when discordant classes are removed.

- Overall accuracy for identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) from all other categories ranged from 78% (110/141) to 85% (130/153) depending on the hazard classification system used. False positive and false negative rates ranged from approximately 11% (10/93) to 34% (27/79) and 6% (4/62) to 22% (13/60), respectively whether or not discordant classes were included in the evaluation. The lowest false negative rate (6% [4/62]) was noted for the GHS system, followed by 14% (11/81) for the EPA system, and 22% (13/60) for the EU system. However, among these false negatives, at least one substance was classified as an ocular corrosive/severe irritant based on Draize data (n = 1 each for the EPA and GHS systems, and n = 6 for the EU system). Considering the public health impact of misclassifying a corrosive substance as Not Labeled, these false negative results cannot be minimized.

The available validation database for the ICE test method has remained unchanged since the original ICCVAM evaluation (ICCVAM 2006). Therefore, the original ICCVAM recommendation for the use of the ICE test method to identify substances as ocular corrosives/severe irritants remains unchanged (i.e., *that there are sufficient data to support the use of the ICE test method, in appropriate circumstances and with certain limitations, as a screening test to identify substances as ocular corrosives and severe irritants [i.e., EPA Category I, UN GHS Category 1, EU R41] in a tiered-testing strategy, as part of a weight-of-evidence approach*.)

Usefulness and Limitations

The ICE test method has been previously recommended for identification of ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1) in appropriate circumstances and with certain limitations. Based on an evaluation of available data and corresponding performance (sensitivity and specificity), ICCVAM proposes that the ICE test method not be recommended to identify all categories of ocular hazard classification as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001; UN 2003). Furthermore, the ICE test method is not recommended as a screening test to identify substances as not labeled as irritants (i.e., EPA Category IV, EU

Not Labeled, GHS Not Classified) from all other hazard categories (i.e., EPA Category I, II, or III; EU R41 or R36; GHS Category 1, 2A, or 2B) as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001; UN 2003).

Test Method Protocol

An ICCVAM recommended test method protocol for the ICE test method is included in ICCVAM (2006). This same protocol should be used for all future ICE studies with the modification of including decision criteria for all categories of ocular irritation as described in the current ICE BRD. ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of the ICE test method for the identification of all ocular hazard categories.

Proposed Future Studies

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

- Additional optimization studies/evaluations should be conducted in an attempt to improve the correct classification of mild and moderate ocular irritants and substances not labeled as irritants. After optimization, additional studies to further assess the reliability and accuracy of the test method are recommended.

- ICCVAM recommends that a histopathological evaluation of the corneal tissue, using standardized procedures, be included when the ICE test method is conducted. Such data will allow for development of decision criteria and future assessments on the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.

4.2 The BCOP Test Method

ICCVAM proposes the following draft test method recommendations on the BCOP test method. ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document for this topic (available at: http://iccvam.niehs.nih.gov/methods/ocutox/mildmod/BCOP-BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, a proposed test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

The test method recommendations described herein are based upon two analyses of BCOP test method performance:

- Overall correct classifications that ranged from 49% (91/187) to 54% (101/186), depending on the hazard classification system evaluated when using the entire database; and 47% (31/66) to 54% (35/65) depending on the hazard classification system evaluated when discordant classes are removed. Using alternative decision criteria for the identification of corrosive/severe ocular irritants (i.e., IVIS ≥ 75 as the cutoff to define such substances [used in the AMCP submission protocol] instead of IVIS ≥55.1 as the cutoff to define such substances [as per the ICCVAM recommended BCOP protocol]) does not does not improve test method performance.

- Overall accuracy for identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) from all other categories ranged from 64% (76/118) to 83% (154/186) depending on the hazard classification system used. While false positive rates were high (53% [24/45] to 70% [63/90] depending on the hazard classification system used), the false negative rates were low (6% [8/141] for EPA the system, and 0% [0/54 or 0/97] for the EU and GHS systems, respectively). Among the eight false negatives for the EPA system, 100% (8/8) were EPA Category III substances based on Draize

data. For 38% (3/8) of these substances, the categorization was based on at least one rabbit with a corneal opacity score of one that was not resolved until day three of the study. Another substance was categorized based on all six rabbits with a conjunctival redness score of three that was not resolved until day seven of the study. Considering the severity and number of ocular lesions noted *in vivo*, these false negative results cannot be minimized as they present a significant risk to the user that could be exposed to these types of materials.

In the original ICCVAM evaluation of BCOP, which was based on 145 substances, overall accuracy, false positive, and false negative rates were 79% (113/143) to 81% (119/147), 19% (20/103) to 21% (22/103), 16% (7/43) to 25% (10/40) depending on the hazard classification system evaluation (i.e., EPA, EU, or GHS). Based on the current BCOP validation database, which has increased to 211 substances, overall accuracy, false positive, and false negative rates are 77% (91/118) to 79% (147/186), 24% (20/85 to 29/123), 15% (10/65) to 21% (7/33). Based on these similar performance statistics, the original ICCVAM recommendation for the use of the BCOP test method to identify substances as ocular corrosives/severe irritants remains unchanged (i.e., *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*.)

Usefulness and Limitations

The BCOP test method has been previously recommended for identification of ocular corrosives and severe irritants (i.e., EPA Category I, EU R41, GHS Category 1) in appropriate circumstances and with certain limitations. Based on an evaluation of available data and corresponding performance (sensitivity and specificity), ICCVAM proposes that the BCOP test method is not recommended to identify substances from all hazard categories as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001; UN 2003). The BCOP test method can be used as a screening test to identify substances as not labeled as irritants (i.e., EU Not Labeled, GHS Not Classified), from all other hazard categories (i.e., EU R41 or R36; GHS Category 1, 2A, or 2B) when

results are to be used for EU or GHS hazard classifications. Because of the significant lesions associated with 50% (4/8) of the EPA Category III substances that were false negative in BCOP (i.e., identified as Category IV), the BCOP cannot be recommended as a screening test to identify EPA Category IV substances.

Test Method Protocol

An ICCVAM recommended test method protocol for the BCOP test method is included in ICCVAM (2006). This same protocol should be used for all future BCOP studies with the modification of including decision criteria for all categories of ocular irritation as described in the current BCOP BRD. ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of the BCOP test method for the identification of all ocular hazard categories

Proposed Future Studies

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 and III; GHS Category 2A and 2B; EU R36), ICCVAM recommends additional studies be considered and undertaken.

- Additional optimization studies/evaluations should be conducted in an attempt to improve the correct classification of mild and moderate ocular irritants and substances not labeled as irritants. After optimization, additional studies to further assess the reliability and accuracy of the test method are recommended.

- ICCVAM recommends that a histopathological evaluation of the corneal tissue, using standardized procedures, be included when the BCOP test method is conducted. Such data will allow for development of decision criteria and future assessments on the usefulness of this endpoint for classifying and labeling substances, especially those that may otherwise produce borderline or false negative results.

4.3 The HET-CAM Test Method

ICCVAM proposes the following draft test method recommendations on the HET-CAM test method. ICCVAM developed the draft recommendations after considering available relevant data, information, and analyses, which are provided in the draft Background Review Document for this topic (available at: http://iccvam.niehs.nih.gov/methods/ocutox/mildmod/HETCAM-BRD.pdf). This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, a proposed test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

HET-CAM performance analyses compared to the Draize rabbit eye test were performed for each classification system (i.e., GHS, EPA, EU) each of the six HET-CAM protocols (i.e., IS [A], IS [B], Q-Score, S-Score, IS, and ITC protocols). With the exception of the IS(A) and IS(B) protocols, all analysis methods had at least one *in vivo* moderate or severe irritant substance classified *in vitro* as not labeled as an irritant (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified). The IS(B) overclassified over 90% (39/42) of the Not Classified (GHS) substances. Therefore, more extensive analyses of HET-CAM were restricted to the IS(A) protocol.

The test method recommendations described herein are is based upon two analyses of ICE test method performance:

- Overall correct classifications that ranged from 40% (23/58) to 41% (24/59), depending on the hazard classification system evaluated when using the entire database; and 62% (5/8) to 78% (7/9) depending on the hazard classification system evaluated when discordant classes are removed.
- Overall accuracy for identification of substances not labeled as irritants (i.e., EPA Category IV, EU Not Labeled, GHS Not Classified) from all other categories ranged from 58% (36/58) to 60% (47/60) depending on the hazard classification system used. False positive and false negative rates ranged from

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

The available validation database for the HET-CAM test method has remained unchanged since the original ICCVAM evaluation (ICCVAM 2006). Therefore, the original ICCVAM recommendation for the use of the HET-CAM test method to identify substances as ocular corrosives/severe irritants remains unchanged (i.e., *Based on these rates, the use of these analyses methods and decision criteria for screening and identifying ocular corrosives and severe irritants [i.e., EPA Category I, GHS Category 1, EU R41] in a tiered-testing strategy, as part of a weight-of-evidence approach, is not recommended.*)

Usefulness and Limitations

Based on an evaluation of available data and corresponding performance (sensitivity and specificity), ICCVAM proposes that the HET-CAM test method is not recommended to identify substances from all hazard categories as defined by the GHS, EPA, and EU classification systems (EPA 1996; EU 2001; UN 2003). However, based on an analysis of 60 compounds (25 surfactant based formulations, 18 oil/water emulsions and 17 individual substances), the HET-CAM IS(A) test method can be used as a screening test to identify substances as not labeled as irritants (i.e., EU Not Labeled, GHS Not

Classified), from all other hazard categories (i.e., EU R41 or R36; GHS Category 1, 2A, or 2B) when results are to be used for EU or GHS hazard classifications. However, based on the limited database for HET-CAM IS(A), this recommended use is limited to cosmetic and personal care formulations that are oil/water emulsions or surfactant containing formulations. Furthermore, while the limited database also indicates that HET-CAM could identify substances labeled as EPA Category IV, the database does not include substances that are actually regulated by EPA (e.g., pesticide formulations, antimicrobial cleaning products). For this reason, additional testing of such products in HET-CAM may be necessary before definitive recommendations can be made on its usefulness for identifying Category IV substances.

Test Method Protocol

An ICCVAM recommended test method protocol for the HET-CAM test method is included in ICCVAM (2006). This same protocol should be used for all future HET-CAM studies with the modification of including decision criteria for all categories of ocular irritation as described in the current HET-CAM BRD. ICCVAM encourages users to provide all data that are generated from future studies, as they could be used to further characterize the usefulness and limitations of the HET-CAM test method for the identification of all ocular hazard categories.

Proposed Future Studies

ICCVAM recommends that additional studies should be conducted to further optimize the HET-CAM prediction models and the decision criteria that would be used to identify ocular corrosives and severe irritants (EPA Category I, EU R41, GHS Category 1), as well as moderate (EPA Category II, EU R36, GHS Category 2A) and mild irritants (EPA Category III, GHS Category 2B), as defined by the EPA, GHS, or EU classification systems. Such studies could potentially improve the usefulness of the HET-CAM test method for identifying these types of substances.

4.4 The IRE Test Method

ICCVAM proposes the following draft test method recommendations on the IRE test method. ICCVAM developed the draft recommendations after considering available relevant data and information. This section provides a brief summary of the background and rationale for the draft proposed recommendations, followed by the specific draft recommendations on proposed usefulness and limitations, a proposed test method protocol, and proposed future studies and activities.

Background and Rationale for the Draft Proposed ICCVAM Recommendations

Currently, there is no widely accepted, standardized IRE test method for detecting ocular irritants. Evaluation of the IRE test method for its usefulness as a partial or full replacement for the Draize rabbit eye test has been confounded by the lack of a standardized protocol. As an indication of the diversity among IRE protocols used, consider the following list of endpoints evaluated among published IRE studies:

- CEC (1991): Corneal opacity, corneal swelling, and fluorescein retention (1 and 4 hours)
- Balls et al. (1995): Corneal opacity and corneal swelling (1 and 4 hours)
- Gettings et al. (1996): Mean extent of corneal swelling across time (1 to 4 hours)
- Guerriero et al. (2004): Maximal corneal opacity (opacity x area), maximal corneal swelling, fluorescein penetration (intensity x area) and assessment of epithelial integrity (0.5, 1, 2, 3, and 4 hours)

Although initially developed by Burton et al. (1981) for the assessment of severe eye irritants using a relatively small set of eleven test substances, the IRE test method has been modified for use in the assessment of either selective types of irritants (e.g., severe irritants) or for specific classes of chemical substances or products (e.g., surfactant-containing chemicals, cosmetic and hair care products) (Gettings et al. 1966; Chamberlain et al. 1997; Cooper et al. 2001; Jones et al. 2001). In other studies, protocols were geared to evaluate a wider range of chemical classes over the entire range of irritancy for test method assessment or validation purposes (Price and Andrews 1985; Koeter and Prinsen 1985; CEC 1991; Balls et al. 1995; Gettings et al. 1996) or for

711 interlaboratory trials (Whittle et al. 1992). Guerriero et al. (2004) modified the original
712 IRE test method protocol to refine assessment of pharmaceutical worker safety by using
713 decision criteria designed to identify severe eye irritants using a chemical database of 30
714 pharmaceutical ingredients, chemical intermediates, and raw materials and an additional
715 14 reference chemicals from ECETOC (1998).

716 The available validation database for the IRE test method has remained unchanged since
717 the original ICCVAM evaluation (ICCVAM 2006). Therefore, the original ICCVAM
718 recommendation for the use of the IRE test method to identify substances as ocular
719 corrosives/severe irritants remains unchanged (i.e., *the use of the IRE test method for*
720 *screening and identifying ocular corrosives and severe irritants [i.e., EPA Category I,*
721 *GHS Category 1, EU R41] in a tiered-testing strategy, as part of a weight-of-evidence*
722 *approach, is not recommended. There also are insufficient data using all four*
723 *recommended IRE endpoints (corneal opacity, fluorescein penetration, corneal swelling,*
724 *and observations of significant effect on corneal epithelium) to assess test method*
725 *accuracy and reliability when all these endpoints are evaluated in a single study.*

726 ***Usefulness and Limitations***

727 There are insufficient data using all four recommended IRE endpoints (corneal opacity,
728 fluorescein penetration, corneal swelling, and observations of significant effect on
729 corneal epithelium) to assess test method accuracy and reliability when all these
730 endpoints are evaluated in a single study. Furthermore, among the studies that included
731 each endpoint, decision criteria are focused on distinguishing ocular corrosives and
732 severe irritants from all other ocular hazard categories (i.e., moderate and mild irritants
733 and substances not labeled as irritants), and do not specify decision criteria for each
734 ocular hazard category. For these reasons, an adequate evaluation of the IRE test method
735 for its ability to identify all ocular hazard categories is not feasible at this time.

736 ***Test Method Protocol***

737 An ICCVAM recommended test method protocol for the ICE test method is included in
738 ICCVAM (2006). This same protocol should be used for all future ICE studies with the
739 modification of including decision criteria for all categories of ocular irritation as
740 described in the current ICE BRD. ICCVAM encourages users to provide all data that are

generated from future studies, as they could be used to further characterize the usefulness and limitations of the IRE test method for the identification of all ocular hazard categories.

Proposed Future Studies

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

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