X. Report on the Independent Scientific Peer Review Panel on Alternative Ocular Safety Testing Methods

A. Presentations

Dr. Merrill, FDA, presented an introduction and overview of the proposed methods and approaches. She explained the public health importance of ocular safety testing and hazard labeling and said that 15% of all eye injuries are due to chemicals. The Draize Rabbit Eye Test, which involves instillation of 100 μL (liquids) or 100 mg (solids) of a

test substance into the lower conjunctival sac of one eye of a rabbit, is the *in vivo* test method currently accepted by US Federal and international regulatory agencies.

The Ocular Peer Review Panel, "the Panel," met on May 19 -21, 2009; their report will be available in July. ICCVAM plans to transmit recommendations to Federal agencies in December and request responses by June 2010.

The Panel evaluated:

- Routine use of topical anesthetics, systemic analgesics, and humane endpoints to avoid or minimize pain and distress during *in vivo* ocular irritation testing
- Validation status of four *in vitro* test methods for identifying mild/moderate ocular irritants and substances not labeled as irritants: BCOP, ICE, HET-CAM, and IRE
- Validation status of the *in vivo* low volume eye test (LVET)
- Validation status of the individual test methods and testing strategies to assess eye irritation potential of AMCPs, including use of the BCOP, Cytosensor Microphysiometer® (CM), and EpiOcular™ (EO) test methods

Dr. Merrill briefly reviewed the procedures for conducting the test methods, summarized the test method data, and then presented ICCVAM's draft proposed recommendations for their use and limitations. She then summarized the ICCVAM charges to the Panel and acknowledged ICCVAM and the ICCVAM Ocular Toxicity Working Group.

Dr. A. Wallace Hayes, Harvard School of Public Health and Peer Panel Chair, presented a summary of the Panel report. The Panel was composed of 22 members from six different countries and they came to complete consensus on all but one of the recommendations (see HET-CAM below). He acknowledged the support of NICEATM and in particular, the contract support staff. He detailed the ICCVAM charges to the Panel and summarized the Panel's recommendations:

- The Panel proposed an alternative preemptive pain management protocol that should be used for all *in vivo* rabbit eye irritation tests intended for regulatory safety testing, unless there is requirement for monitoring the pain response.
- The Panel concluded that, based on the available data and information, some humane endpoints recommended by ICCVAM are adequate to terminate a study.
- The Panel supported the ICCVAM draft recommendation that the available data and ICE test method performance do not support its use to identify substances from all hazard categories as defined by Globally Harmonized System (GHS), EPA, and EU classification systems.
- The Panel agreed with the ICCVAM draft recommendation that the available data and ICE test method performance do not support its use as a screening test to identify substances not labeled as irritants from all other hazard categories as defined by GHS, EPA, and EU classification systems.

- The Panel supported the ICCVAM draft recommendation that the available data and BCOP test method performance do not support its use to identify substances from all hazard categories as defined by GHS, EPA, and EU classification systems.
- The Panel agreed with the ICCVAM draft recommendation that the available data and BCOP test method performance support its use as a screening test to identify substances not labeled as irritants when results are used for EU or GHS hazard classifications.
- The Panel concluded that the BCOP test method cannot be used as a screening test to identify EPA Category IV substances.
- The Panel supported the ICCVAM draft recommendation that the available data and HET-CAM test method performance do not support its use to identify substances from all hazard categories as defined by GHS, EPA, and EU classification systems.
- The Panel (with one minority opinion) did not support the ICCVAM draft recommendation that the available data and HET-CAM test method performance support its use as a screening test to identify substances not labeled as irritants when results are used for EU or GHS hazard classifications.
- The Panel concluded that additional optimization and validation studies that include all four recommended endpoints are needed before definitive recommendations on the relevance and reliability of the IRE test method can be made.
- The Panel concluded that in the absence of all data, including the ECVAM BRD, they could not make definitive conclusions or recommendations on the validation status of the LVET.
- The Panel agreed with the ICCVAM draft recommendation that the CM test method can be used as a screening test to identify both ocular corrosive/severe irritants and substances not labeled as irritants in a tiered-testing strategy, as part of a weight-of-evidence approach, but this use is limited to surfactant chemicals and specific types of surfactant-containing formulations (e.g., cosmetics and personal care products).
- The Panel agreed with the ICCVAM draft recommendation that there were insufficient data to support use of the AMCPs testing strategy (i.e., using the BCOP, CM, and EO test methods) for classification of substances in all four ocular hazard categories.
- The Panel agreed with the ICCVAM draft recommendation that there were insufficient available data on which to base definitive recommendations on the proposed alternate testing strategy (i.e., using the BCOP and EpiOcular™ test methods) for classifying substances in all four ocular hazard categories.
- The Panel recognized that the use of histopathological evaluation as an additional endpoint does not improve the accuracy and predictability of the BCOP test method for the limited database of currently tested AMCPs; however, histopathological evaluation may prove to be a useful endpoint and as such, collection of ocular tissue and further efforts to optimize histopathological evaluation is strongly encouraged.

Dr. Levine said she saw nothing in the flow chart that required all three tests to be used at the same time. Dr. Hayes said that the concern of the Panel was that it would have been very helpful to know comparative results of compounds tested in all three tests to allow them to adequately evaluate the overall performance of the proposed testing strategy.

B. Public Comments

Dr. Rodger Curren, IIVS, asked the attendees to read the written comments he would be sending for posting on the Website. He addressed Dr. Levine's comment regarding materials not being tested in all three assays and said many antimicrobials were tested in each of the assay systems. Twenty-eight materials were fully evaluated in all the tests and there were no differences in results among the tests. He could understand if there were considerable differences in the chemistry of the materials, then testing in all three assays might be needed, but otherwise it was not. He suggested a way to strengthen the peer review process for additional studies going forward and to improve the efficiency of the reviews. He said there was no effective way for the proponents of an assay strategy or the developers of a new assay to interact with the Panel. Many questions arose in this and other reviews that could have been answered quickly by the writers of the BRD or the developers of the assay. The proponents of the assay were allowed to speak only to the methodologies and not to the interpretation. He was not proposing extended debate in the peer review process, but only some way to allow greater interaction with the Panel.

Dr. Kate Willett, PETA, expressed puzzlement that the Panel's evaluation involved such an enormous review when the original nomination was simply for the antimicrobial project.

Dr. Levine asked if EPA's specific charge to ICCVAM, regarding a review of the flow chart's use for making labeling decisions on AMCPs, was communicated to the Panel. Dr. Stokes said the charge was clearly communicated. He further added that bringing a peer review Panel together is very expensive and time-consuming process; therefore, NICEATM-ICCVAM wanted to take advantage of convening this international Panel of experts by having other related test methods reviewed. NICEATM-ICCVAM had other topics they wanted to review, so they consolidated them for one Panel at one meeting. It resulted in an aggressive agenda and the Panel was very thorough. They took their time to do a careful, comprehensive review that in the long-term would benefit the entire project.

C. SACATM Discussion

Dr. Freeman asked Dr. Levine about EPA's notice of proposed rule making for GHS adoption and whether EPA would adopt the GHS classification system. Dr. Levine said no decision would be made until a new Assistant Administrator is confirmed. Dr. Freeman said the Classification, Labeling and Packaging regulation in the EU system, which represents their acceptance of GHS, has been released. The EU system will merge with GHS in 2010. He said the GHS system represents the future and he was unsure what the United States is doing regarding the three scoring methods. He expressed confusion regarding the earlier conclusion for the use of BCOP to screen for

corrosives or severe irritants and the newer conclusion for its use to screen for substances not labeled as irritants using the GHS or EU system. He said it was at opposite ends of the spectrum and that if the United States were going to adopt GHS in the future, the EPA method should not matter. He expressed concern about classification of materials between those identified as severe irritants and non-irritants. Dr. Levine compared this classification to the issue with classifying skin irritation, where identification of the extremes is possible. She considered it a learning opportunity and suggested other agencies should also address this issue.

Dr. Barile, a lead discussant, agreed with Dr. Freeman. He understood from the 2006 review that both BCOP and ICE were approved for identifying corrosives and severe irritants, but in the 2009 conclusions, only BCOP was approved for the classification of corrosives and severe irritants. Dr. Stokes clarified that BCOP and ICE are still recommended for identifying corrosives and severe irritants. In the 2009 review, the recommendations for the use of ICE have not changed; ICE was not recommended for the identification of all ocular hazard categories as defined by the EPA, EU, and GHS classification systems. In addition, ICE was not recommended as a screening test to identify substances not labeled as irritants from all other hazard categories as defined by the GHS, EPA, and EU classification systems. Dr. Stokes emphasized that one of the reasons NICEATM-ICCVAM is using the term "not labeled as irritant," is that under the EU and GHS classification systems, even if a material is consider not labeled as an irritant, it can still cause a considerable amount of irritation. For example, substances not classified as irritants in the GHS or EU scheme are EPA Category III or higher. Category III substances cause lesions that persist for more than urs, but clear by seven days. Dr. Stokes also noted that the IRE was not recommended because there are not enough data using all four endpoints, as in the current ICCVAM-recommended protocol. HET-CAM was proposed by ICCVAM to identify non-labeled surfactants and surfactant-containing compounds. The Panel disagreed with the ICCVAM recommendations because they considered the number of substances in the intermediate irritancy categories (i.e., mild and/or moderate irritants) to be insufficient.

Dr. Barile asked about use of the CM in ocular testing and the status of the testing, given that the machine is no longer available. Dr. Stokes said a new version of the CM is being developed that will measure additional endpoints. The new machine will need to meet or exceed the performance for the existing CM. Dr. Barile said little information had been presented on the CM as to what it tested and he asked why mouse fibroblasts were used. He suggested a more extensive review of CM by ICCVAM and more background information. Dr. Levine said the EPA has a policy of not recommending a brand or product; guidelines are based on performance standards. Dr. Barile also asked about use of the 2006 BRD database. Dr. Merrill said data from the AMCP submission were added to the BCOP database from 2006, but the available database for ICE had not changed since 2006.

Dr. Fox, a lead discussant, said he agreed with the report but had some comments on the science. The CM is an antiquated tool that is not sophisticated enough for use in ocular methods; there are better tools available. The methodology should be validated, not the instrument. He said in the original review, the BCOP was found acceptable to

detect corrosives, but has been upgraded to detect non-labeled materials. Dr. Stokes said more data had been added from the AMCP submission. BCOP was originally evaluated for its accuracy in classifying substances as either severe or non-severe, with irreversible or reversible effects, respectively. Accuracy for identifying moderate, mild, and non-labeled categories was not performed in the original review. Dr. Freeman said there was some dissension on the BCOP conclusions in 2006. Dr. Fox said that the local anesthetics recommended for use are esters, which have short half-lives; he asked why amides, which are longer acting, were not chosen for use. A disadvantage of local anesthetics is that they create tear breakup time and allow the compound increased access to the eye. He said a topical ophthalmic amide anesthetic might be a better option for pain control in the Draize test.

Dr. Karen Brown, a lead discussant, said the use of anesthetics for the Draize test was overdue. She said it should be a requirement unless there is justification for non-use. Systemic anesthesia should be used as well as topical anesthetics. She agreed with the Panel's recommendations, but asked for more information on the two AMCP testing strategies saying more work should be done in that area and it should move forward quickly. Individual tests were done with the BCOP and EO and it appeared they could differentiate severe from moderate and mild AMCPs. She asked how companies could be encouraged to generate more data for the AMCPs, similar to GlaxoSmithKline doing more research on the IRE. Dr. Stokes said the Ocular Toxicology WG's recommendation was to encourage industry to generate more data. Accordingly, the EPA just issued a proposal for a pilot project to encourage industry to generate data that would utilize the methods in the strategy. Dr. Levine said the EPA is proposing an eighteen-month pilot. Companies will provide both *in vitro* data and Draize data on similar products. The project will collect incident information on products that have been on the market for eight to ten years without labeling. The EPA will then make labeling decisions and evaluate how it is working. Dr. Karen Brown said the sequence of tests looked very promising.

Dr. Hansen, a lead discussant, concurred with the previous comments and said it was encouraging and long overdue that ICCVAM was moving toward requiring topical anesthetics and systemic analgesics.

Mr. Wnorowski, a lead discussant, said his company had developed some of the data several years ago on the anesthetics. His company has been successfully using anesthetic pretreatments for all its studies. He supported the other models moving forward and being accepted for regulatory purposes.

Dr. Freeman concurred with discussion on the use of anesthetics in the Draize test. Dr. Meyer asked how much is enough with respect to ICCVAM, and would the regulatory agencies accept a partial solution in the identification of classes II, III, and IV. Most of the pain and distress occurs with class I chemicals. She suggested moving forward rather than continuing to address the low rates of performance for the other classifications. Dr. Ehrich asked Dr. Hayes about the Panel's specific recommendation for the use of the analgesic buprenorphine. Dr. Hayes said this was based on strong recommendations from the veterinary anesthesiologist and ophthalmologists on the

Panel, based on their clinical experience. He said the important concept was to use a systemic analgesic first followed by a topical anesthetic prior to test substance application, and then to continue treatment with systemic analgesics as long as necessary.

Dr. Charles said harmonization is needed for assessing the performance criteria for the assays from a drug development perspective. Once there is harmonization, there is a need for guidance and strategy. He suggested assessing the other methods in a similar fashion to the AMCPs, and categorizing the test article based on a multiple assay strategy as opposed to doing more work on each individual assay. Dr. Stokes said an ECVAM-sponsored workshop suggested a top-down, bottom-up approach using a three-category system. An *in vitro* test or battery of tests would be needed that could identify all substances that could cause irreversible effects (i.e., all category I substances, with a high degree of certainty). All other categories would involve reversible damage and not cause permanent effects. Another test or battery would then be used only to identify substances that do not cause significant irritation (i.e., non-labeled substances). It would not require a high degree of sensitivity, but would identify most substances in this category without significant over-labeling. All other substances would be classified as mild or moderate. Further testing could be done to differentiate mild and moderate substances yielding a lower hazard warning for mild substances. This top-down, bottom-up approach is being pursued for both dermal and ocular irritation. Dr. Stokes explained that for ocular testing the methods are not available for the top or bottom for the level of performance needed. Not enough data are currently available to support a completely non-animal approach. Dr. Freeman said it was debatable because the BCOP could identify the highs and lows using GHS. Dr. Stokes said there are significant restrictions on categories of substances for which BCOP can be used, as some chemical classes and physical properties result in significant false negative results, which would not be acceptable in a top-down decision model.

Dr. Nicolaysen asked why the Panel recommended the dose of 0.01 mg/kg buprenorphine, which is lower than the 0.05 mg/kg used clinically. Dr. Hayes said the dose was based on clinical experience. Dr. Nicolaysen said there should be better evidence to use the lower dose. Dr. Marilyn Brown expressed some concern about the handling-stress induced in animals with the administration of the both analgesics and anesthetics, but with differing dosing schedules. Dr. Corcoran said there did not appear to be a consensus standard of care. He thought the recommendations to be overly proscriptive and suggested establishing an expectation for care with the goal of relieving pain with an antinociceptive and an anesthetic at appropriate doses and dosage schedules.