Institute for In Vitro Sciences, Inc.

December 28, 2004

Advancing
Science &
A n i m a l
Welfare
Together

William Stokes, DVM
Director, NICEATM
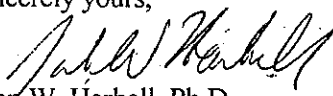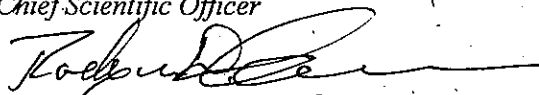NTP
Research Triangle Park, NC 27709

Dear Dr. Stokes:

This public comment is delivered in response to Federal Register Notice Volume 69, Number 212, Pages 64081-64082. It addresses the Background Review Document (BRD), "Current Status of In Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants: The Bovine Opacity and Permeability (BCOP) Test Method", November 1, 2004.

In this second public comment, we should like to address several points raised and/or omitted from the BRD. In some cases, these comments may be pertinent to all four BRDs (e.g., basis for interest in developing in vitro methods, predictive capacity of the Draize test, and examination of the USEPA categorization system relative to the Globally Harmonized System [GHS]). In this discussion, we will try to focus only on major issues that impact the overall validity of the analysis and conclusions drawn by the ICCVAM/ NICEATM. Our comments (enumerated as items 1-9 below) are arranged by the section of the BRD where the point is or should be addressed.

1. Definition of validation
2. Use of the BCOP assay by regulatory agencies
3. Use of the BCOP assay by industry
4. Use of initial degree and depth of injury as a predictor of overall degree and duration of injury
5. Motivation for developing new methods for predicting ocular irritation
6. Assessment of eye irritation potential in humans
7. Prediction of severe irritants under the USEPA classification scheme
8. Modes of action of chemicals in the eye
9. Using the BCOP assay across a range of chemistries

Sincerely yours,

John W. Harbell, Ph.D.
Chief Scientific Officer

Rodger D. Curren, Ph.D.
President

Institute for In Vitro Sciences, Inc.
21 Firstfield Road, Suite 220
Gaithersburg, MD 20878
USA

21 FIRSTFIELD ROAD
SUITE 220
GAITHERSBURG, MD 20878
TEL: 301.947.6523
FAX: 301.947.6538
WWW.IIVS.ORG

Preface

1) Page xxiv: Footnote 1. "**Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 1997, 2003).**" This definition of validation – which replaces the generally accepted word "relevance" with the word "accuracy"- is a significant change from the definition accepted by the stakeholders who supported the establishment of ICCVAM in 1997 and a change from the 1997 and 1999 (revised) ICCVAM guidelines. The 1997 document, "Validation and Regulatory Acceptance of Toxicological Test Methods", created as a result of a 1995 stakeholder workshop, defines validation as "a scientific process designed to characterize the operational characteristics, advantages, and limitations of a test method, and to demonstrate its reliability and relevance". Relevance is then subsequently described as a relatively encompassing term which includes both the mechanistic relationship of the model to the human or ecological target tissue as well as its ability to provide results equivalent to an original method. "The mechanistic relationship of the test endpoint to the toxic effect of concern should be established with a reasonable degree of rigor. In general, the closer the linkage between the effect measured and the toxicological effect of interest, the simpler the validation process will be." These general thoughts are carried through the 1999 (revised) General Guidelines for Submission to ICCVAM (page 3) reads, " Validation is a process designed to establish the operational characteristics of a proposed test method. These characteristics include the test method's reproducibility within and among laboratories, its relevance (i.e., the ability to measure or predict correctly), and its limitations." The use of reliability or reproducibility is consistent with general principles of validation but the direct substitution of accuracy for relevance would not be. In fact the Glossary of this 1999 document defines relevance in its more traditional sense as: "The extent to which the proposed test is related to the effect of interest and whether a test is meaningful and useful for a particular purpose. That is, the extent to which a test method will correctly predict or measure the biological effect of interest".

It is not until the 2003 ICCVAM document that the word accuracy has replaced the more correct term relevance, and then only in the Glossary, not in the text. ***Why is this "simple" change such a problem?*** Accuracy implies the need to directly predict a response from some reference test rather than produce a useful result for the evaluation of hazard. To my mind, this change is not just a small matter of semantics but a fundamental change in the philosophy of conducting and interpreting validation exercises. It elevates the reference test to an absolute standard and leads to the kind of overly simplistic almost complete reliance on sensitivity/specificity analysis that we see in all four BRDs. The mechanistic relationship of the proposed test method to the tissue toxicity being predicted, as well as the predictive capacity of the reference method (both for human effects and **for predicting itself** must be considered when determining the relevance. A detailed discussion of the mistakes that can be made when using sensitivity and specificity analyses alone can be found in a series of articles by Bruner et al. (2002)[1,2,3].

---

[1] Bruner, L.H., Carr, G.J., Harbell, J.W., and Curren, R.D. (2002) An investigation of new toxicity test method performance in validation studies: 1. Toxicity test methods that have predictive capacity no greater than chance. **Human and Experimental Toxicology** 21:305-312.

We would ask the Peer Review Panel and other stakeholders to consider whether this change (accuracy for relevance) is consistent with the founding agreement among stakeholders and the scientific purpose of the ICCVAM process.

Executive Summary

2) Page xxxi, lines 53-54: "**The BCOP test method has not yet been considered by U.S. Federal agencies for regulatory use where submission of testing data is required.**" This statement is incorrect. In 2003, the Institute for In Vitro Sciences, Inc (IIVS) was given a study/facilities audit by the Office of Enforcement of the USFDA. which we were told was based on the submission of one or more BCOP studies to the USFDA. The laboratory and studies were found to be in compliance with the Good Laboratory Practices guidelines. In 2004, IIVS was given a study/facilities audit by the Office of Compliance of the USEPA based on the submission of one or more BCOP studies (submitted to the Office of Pesticide Programs). The laboratory and studies were found to be in compliance with the Good Laboratory Practices guidelines. At least one submission to the USEPA has been a matter of public record to the ICCVAM/NICEATM for over 18 months. Both of the audits of IIVS were "directed" in response to data submissions. It seems unlikely that either of these agencies would have used their compliance resources to perform audits were the data from these studies not to be used in some form of regulatory decision.

3) Page xxxi, lines 54-60: **Use of the BCOP assay by industry**. It is stated in the BRD that "Negative results and suspected false positive in vitro results proceed to standard in vivo testing …" Our experience has been that very few BCOP assay results are "confirmed" in the rabbit test outside of the pharmaceutical industry (please see the public comments on the BCOP BRD from N. Cuellar and J. Swanson, S.C. Johnson and Son, Inc.). We suspect (though can not absolutely confirm) that this is the case for most companies performing the test in house or with other contract research organizations. The important message here is that these companies have used the BCOP assay, some for over 10 years, and have established its usefulness in determining ocular irritation potential for their products before they are marketed.

Over the past 7 years, IIVS has used in excess of 20,000 bovine eyes (corneas) in support of product development and product safety evaluations for commercial clients. These studies have been performed on a wide range of chemical and formulation types for cosmetic, personal care, household products, agricultural chemicals, and pharmaceutical clients. The basic approach is summarized in the figure below:

---

[2] Bruner, L.H., Carr, G.J., Harbell, J.W., and Curren, R.D. (2002) An investigation of new toxicity test method performance in validation studies: 2. Comparison of three measures of toxicity test performance. **Human and Experimental Toxicology** 21:313-323.

[3] Bruner, L.H., Carr, G.J., Harbell, J.W., and Curren, R.D. (2002) An investigation of new toxicity test method performance in validation studies: 3. Sensitivity and specificity are not independent of prevalence or distribution of toxicity. **Human and Experimental Toxicology** 21:325-334.

Figure 1
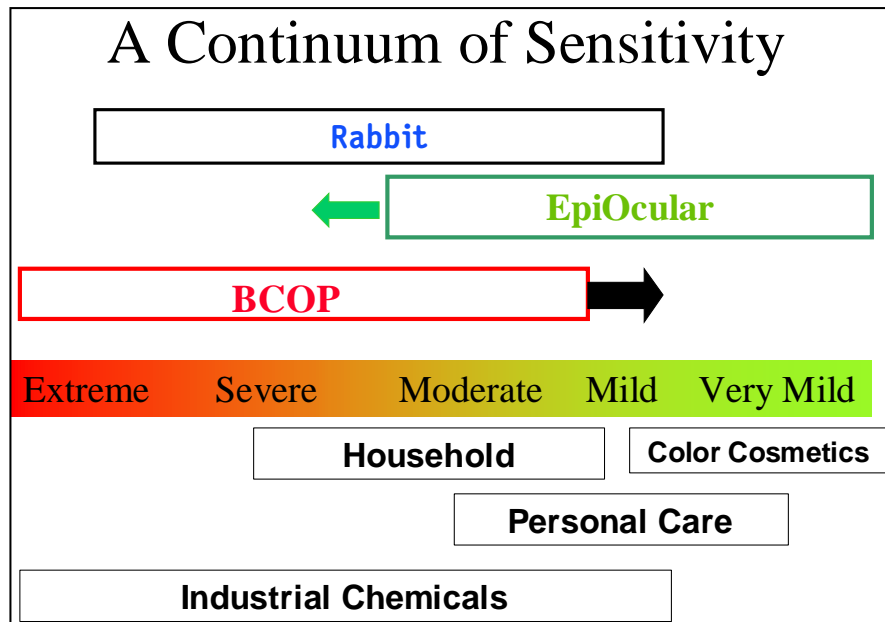
As the diagram suggests, the BCOP assay is used to resolve across a range of irritancy potential from the mild/moderate to extremely severe. While it is capable of showing a mild response (arrow), it does shows limited resolution in the very mild range. In that range, we would use the tissue construct assay (or similar) to determine the mildness for very mild products such as color cosmetics.

4) Page xxxii, lines 98-101: **Prediction of reversal/permanence of ocular effects**. In a series of seminal papers, Drs. Maurer and Jester (and their collaborators) have examined the relationship (measured both in vivo and in vitro) between the depth and degree of initial corneal injury and subsequent degree (tissue scores) and duration (days to clear) of the ocular irritation[4]. They have shown that an in vitro three dimensional model, even when cultured for a relatively short time period after exposure to a toxic agent, is capable of predicting the reversibility of ocular effects. Their work is discussed in more detail in the public comment of Harbell and Curren. It is surprising that their work is only mentioned in passing in the four BRDs since the relationship between the depth of initial injury and the subsequent degree and duration of irritation is pivotal to using the BCOP (and other ex vivo or tissue construct models) for the prediction of severe (and moderate and mild) irritation. These studies have examined a wide range of chemical "classes" and resulting modes of action on the eye[5,6]. Their focus has been on injury to the cornea, consistent with clinically significant injury in the human eye[7,8].

---

[4] Maurer, J.K., Parker, R.D., and Jester, J.V. (2002) Extent of initial corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. **Regulatory Toxicology and Pharmacology** 36:106-117.

[5] Jester, J.V., Li, H.F., Petroll, W.M., Parker, R.D., Cavanaugh, H.D., Carr, G.J., Smith, B., and Maurer, J.K. (1998) Area and depth of surfactant-induced corneal injury correlates with cell death. **Invest Ophthalmol Vis Sci** 39:922-936.

Section 1

5) Page 1-4, lines 83-85. **Basis for developing in vitro assays for ocular irritation**. The ICCVAM/NICEATM has stated that the motivation for developing in vitro test methods for the prediction of ocular irritation are limited to concerns about animal welfare and cost and time to conduct in vivo test [presumably for industry] as well as scientific interest in understanding injury at the tissue and molecular level. Certainly reduction in animal suffering is a noble goal but not the only one. In our experience, however, concerns about the lack of reproducibility of the rabbit ocular irritation test are frequently cited by industry as one reason to explore in vitro methods. These concerns have existed for decades, long before the current push for the development of in vitro methods. The work of Weil and Scala (1971)[9] highlighted these concerns for many users of in vivo ocular irritation data. In 1973, Marzulli and Ruggles[10] published their study of ocular irritation across 10 laboratories. This study examined 6 test chemicals and a concurrent positive control (70% isopropyl alcohol). The study was particularly powerful in that the positive control was tested concurrently with each of the test chemicals (one chemical per week). This design provided 60 trials of 6 rabbits each for the positive control (6 trials per laboratory). Unfortunately, the individual tissue scores were not reported. However, the number of animals, producing a positive tissue score (e.g., 1 or greater opacity) at 24 hours, was reported for each trial in each laboratory. Figure 2 shows the distribution of the number of rabbits in each study showing a positive corneal opacity at 24 hours after treatment with the positive control. Note that over a third of the trails produced no corneal opacity in any of the 6 rabbits while over a quarter of the trials produced opacity in 6 of 6 rabbits.

---

[6] Maurer, J.K., Molai, A., Parker, R.D., Li, L., Carr, G.J., Petroll, M.W., Cavanagh, D.H., and Jester, J.V. (2001) Pathology of ocular irritation with bleaching agents in the rabbit low-volume eye test. **Toxicological Pathology** 29(3):308-319.
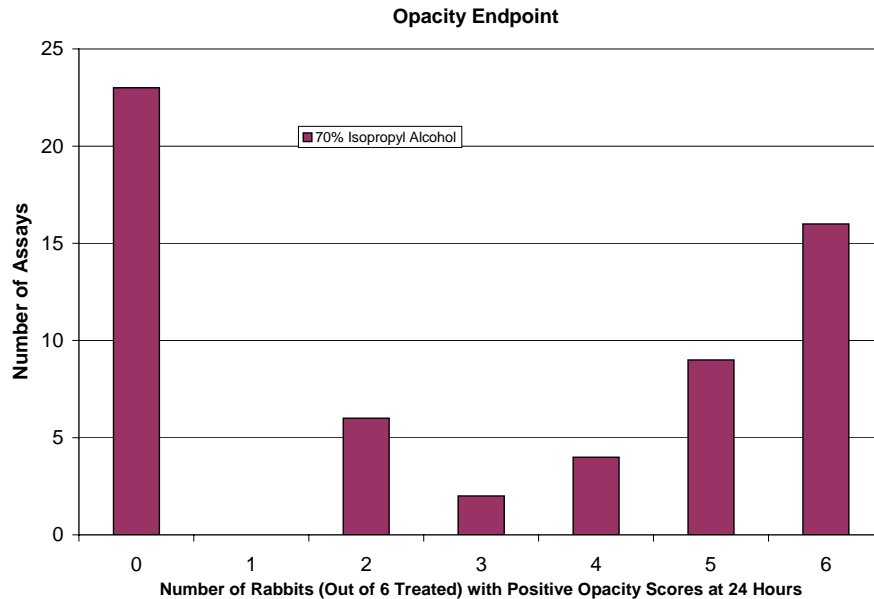
[7] Nussenblatt, R.B., Bron, A., Chambers, W., McCulley, J.P., Pericoi, M., Ubels, J.L., and Edelhauser, H.F. (1998) Ophthalmologic perspectives on eye irritation testing. **J. Toxicol – Cut. & Ocular Toxicol**. 17(2&3):103-109.

[8] Maurer, J.K., McCulley, J.P., Edelhauser, H.F., and Nussenblatt, R.B. (1998) A proposed new classification scheme for chemical injury to the human eye. The Association for Research in Vision and Ophthalmology Annual Meeting.

[9] Wiel, C.S. and Scala, R.A. (1971) Study of intra- and interlaboratory variability in the results of the rabbit eye and skin irritation tests. **Toxicology and Applied Pharmacology** 19:276-360.

[10] Marzulli, F.N. and Ruggles, D.I. (1973) Rabbit eye irritation test: collaborative study. Journal of the AOAC 56(4):905-914.

Figure 2

Opacity Endpoint

Number of Rabbits (Out of 6 Treated) with Positive Opacity Scores at 24 Hours

In their conclusion, these authors stated "The experimental finding of this present collaborative study are in substantial agreement with those of Weil and Scala. Our interpretation of these findings, however, would differ from theirs. In both studies, laboratories were able, in most cases, to distinguish an eye irritant from a nonirritant, if this is all they are asked to do, and if all 4 criteria (change in cornea and iris and conjunctival redness and chemosis) were used for judging eye irritation in a simple pass-fail procedure. (Compare these findings with results cited in Table 60 of the Weil-Scala report.) The test is inadequate if only 1 single parameter (rather than all 4) is used to make this judgment. Furthermore, collaborative results indicate that additional study to identify and eliminate the sources of variability is necessary before reproducible results with regard to comparison of degrees of irritancy can be obtained."

A number of researchers have sought to improve the consistency of the evaluation in vivo. Many have focused on measurements of damage such as corneal swelling. The work of Kennah et al (1989)[11] is a good example of such efforts.

Industrial toxicologists are often faced with the need to understand the potential action of a chemical or formulation on the eye with considerable precision. Some products must be extremely mild. Others, while potentially somewhat irritating, must not exceed the "industry norm" for that given class of products. Thus, a degree of precision is needed to effectively evaluate eye irritation potential (mildness) within a product development/product safety framework. To this end, developers and users of the in vitro methods have expended considerable effort to understand the consistency of the assay. The in vitro assays (such as

---

[11] Kennah II, H.E., Hignet, S., Laux, P.E., Dorko, J.D., and Barrow. C.S.(1989) An objective procedure for quantitating eye irritation based upon changes in corneal thickness. Fundamental and Applied Toxicology 12:258-268.

the BCOP) have been designed to address some of the potential sources of variation observed in the in vivo test. In particular, the assay was developed to precisely control the exposure of the test material to the target tissue and to score the responses of the target tissue in an as objective a manner as possible. Furthermore, concurrent testing of positive and negative control substances provides a measure of the performance of the assay with each trial.

6) Page 1-18, lines 431-432. "**It is proposed that the current animal test provides a suitable assessment of eye irritation potential in humans.**" The ICCVAM/NICEATM has expressed this opinion/conclusion this section and lines 297-300 where a similar statement appears. Within the scientific community, there are those who would disagree with this conclusion[12] and thus the inclusion, in the BRD, of the data supporting the ICCVAM/NICEATM conclusion would be helpful to all readers.

Section 4

7) Predictive capacity of the reference test. By necessity, Draize data for this analysis were taken from studies performed in many laboratories over an extended period of time. To produce a relevant analysis of the predictive capacity of the in vitro tests, one first needs to understand the predictive capacity (in this case, precision) of the Draize test over time, within laboratories and between laboratories. Such data are not presented in the BRDs and so the analysis of the relevance of the in vitro tests to predicting the in vivo rabbit result reflects the "worst case" for prediction since major metrics of the reproducibility of the Draize results are not presented.

Three distinct labeling systems are evaluated in the BRDs. Each uses the same in vivo data but may come to rather different conclusions as to what is "severe". For simplicity, let us consider the Globally Harmonized System (GHS) and the USEPA systems (6-animal and 3-animal test formats). For this purpose, we shall use the results of the Draize data from the Cosmetic, Toiletries, and Fragrance Association (CTFA) Phase III evaluation of surfactant-containing formulations. This data set is selected because it is public (see Appendix H). The data were developed in a highly respected laboratory and the study used a random block design so that each animal (of the 6 treated with each formulation) was scored without reference to the other animals treated with the test article. This data set of 25 formulations is also interesting because each of us has personal experience with formulations of these types (soaps, shampoos etc.). Table 1 shows 25 formulations arranged in increasing order of irritation based on the 6-rabbit responses. Using the USEPA system for a 6-rabbit test, the last 10 formulations (in yellow) were labeled as Category 1 eye irritants because at least one of the rabbits (of the 6) showed a severe response (did not clear in 21 days)[13]. The column "# uncleared" shows the number of animals that did not show recovery by day 21 and the column "Mean Days to Clear" indicates the mean number of days required for the remaining animals to show recovery.

---

[12] Ibid references 7 and 8

[13] Seabaugh, V.M. and Vocci, F.J. (1988) Standard Evaluation Procedure: Eye Irritation Studies, EPA-540/09-88-105.

The categorization of these in vivo data can now be compared between the traditional USEPA system and the GHS system. The GHS system uses 3 rabbits per test article and so 20 unique combinations of 3 rabbits can be generated from the 6-rabbit test. This analysis allows one to examine how consistent the categorization would be between the 3 and 6 rabbit tests. The GHS criteria were then used to evaluate each of the 20 combinations and assign a GHS category. For a set of 6-rabbits where only one animal does not clear, 10 of the combinations included that animal and the irritation category would be a GHS 1. However, 10 of the combinations do not include that severe animal. The GHS categories, for these combinations, would be determined by the tissue scores that were relatively mild. Where 2 animals do not clear, 16 of the combinations would be categorized as GHS 1 and where 3 animals do not clear, 19 of the combinations would be categorized as GHS 1. Four or more uncleared (or tissue scores of "severe") animals are required to make all 20 combinations category 1. For the 10 severe irritants in this data set, there are 200 combinations of 3 animals each. The distribution of the 200 combinations is shown in the lower portion of the table. In this analysis, 161 combinations (80.5%) were GHS category 1, 18 combinations (9%) were GHS category 2a, 7 combinations (3.5%) were GHS category 2b, and 14 combinations (7%) were GHS category NI (non-irritating). In this data set, 19.5% of the 3-rabbit combinations were one to three categories below the severe (category 1) rating. Table 2 shows the same kind of analysis using the USEPA 3-rabbit system. Note that the total fraction of combinations below the severe rating was the same but the distribution in the other three categories was slightly different. Thus, the use of different categorization systems can produce different profiles of severe and nonsevere "responses" for a given data set.

This analysis is not intended to denigrate either the Draize test or the categorization schemes of the regulatory agencies. It is intended to provide a strong note of caution against taking a single in vivo result (irrespective of the number of animals tested) as the one and only possible response category that might be generated.

Table 1. Globally Harmonized System Categorization

Name	Material	GHS 1	GHS 2a	GHS 2b	GHS NI	% Under predicted	# uncleared	Mean Days to Clear
Shampoo 5	HZD*				20	na	0	3.0
Shampoo 8	HZG*				20	na	0	3.5
Eye Makeup re.	HZH				20	na	0	0.0
Mild Shampoo	HZJ				20	na	0	0.0
Shampoo 3	HZM*				20	na	0	2.3
Shampoo 6	HZN*				20	na	0	1.0
Baby Shampoo 1	HZP				20	na	0	2.8
Cleaning Gel	HZQ				20	na	0	0.0
Polishing Scrub	HZT				20	na	0	0.0
Facial Cleaner	HZZ				20	na	0	4.2
Liquid Soap 1	HZB*			4	16	na	0	4.3
Hand Soap	HZU*			4	16	na	0	2.8
Shampoo 4	HZV*			4	16	na	0	5.0
Shampoo 1	HZC*			10	10	na	0	0.5
Liquid Soap 2	HZW*			16	4	na	0	6.3
Gel Cleaner	HZE	10		0	10	50%	1	2.4
Facial Cleaner Foam	HZR*	10		6	4	50%	1	4.6
Shampoo 7	HZA	16	4			20%	2	12.3
Baby Shampoo 2	HZF	16	4			20%	2	10.5
Shampoo 2	HZX	16	4			20%	2	12.3
Shampoo Anti-Dandruff	HZY	16	4			20%	2	12.3
Skin Cleaner	HZI	19	1			5%	3	9.3
Shower Gel	HZS	19	1			5%	3	7.0
Foam Bath	HZL	19		1		5%	3	9.3
Bubble bath	HZK	20				0%	5	7.0
Total number	200	161	18	7	14			
of combinations								
Fraction		80.5%	9.0%	3.5%	7.0%	19.5%		
Predicted		concordant	1 under	2 under	3 under	Total under		

Table 2. USEPA 3-Rabbit Categorization

Name	Material	EPA 1	EPA 2	EPA 3	EPA 4	% Under Predicted	# uncleared	Mean Days to Clear
Shampoo 5	HZD*			20		na	0	3.0
Shampoo 8	HZG*			20		na	0	3.5
Eye Makeup re.	HZH				20	na	0	0.0
Mild Shampoo	HZJ				20	na	0	0.0
Shampoo 3	HZM*			10	10	na	0	2.3
Shampoo 6	HZN*			20		na	0	1.0
Baby Shampoo 1	HZP			19	1	na	0	2.8
Cleaning Gel	HZQ			20		na	0	0.0
Polishing Scrub	HZT				20	na	0	0.0
Facial Cleaner	HZZ				20	na	0	4.2
Liquid Soap 1	HZB*			20		na	0	4.3
Hand Soap	HZU*			20		na	0	2.8
Shampoo 4	HZV*			20		na	0	5.0
Shampoo 1	HZC*			20		na	0	0.5
Liquid Soap 2	HZW*			20		na	0	6.3
Gel Cleaner	HZE	10		10		50%	1	2.4
Facial Cleaner Foam	HZR*	10		10		50%	1	4.6
Shampoo 7	HZA	16	4			20%	2	12.3
Baby Shampoo 2	HZF	16	4			20%	2	10.5
Shampoo 2	HZX	16	4			20%	2	12.3
Shampoo Anti-Dandruff	HZY	16	4			20%	2	12.3
Skin Cleaner	HZI	19	1			5%	3	9.3
Shower Gel	HZS	19	1			5%	3	7.0
Foam Bath	HZL	19		1		5%	3	9.3
Bubble bath	HZK	20				0%	5	7.0
Total number	200	161	18	21	0			
of combinations								
Fraction		80.5%	9.0%	10.5%	0.0%	19.5%		
Predicted		concordant	1 under	2 under	3 under	Total under		

<u>Section 5 and 6</u>

8) Suggested addition of thoughts on chemical exposure and modes of action in the eye:
A complementary approach to determining the relevancy of a new test is to ask whether that new test can model the essential components of the reference test. In the case of ocular irritancy tests, one of the first components might be to have a relevant target tissue. The second might be to be able to model the test material exposure to the target tissue. Modeling the exposure that a rabbit cornea might receive in a Draize test is by no means trivial. Figure 3 shows a rather simplified diagram of the factors (variables) that are present in any such exposure.

Figure 3.
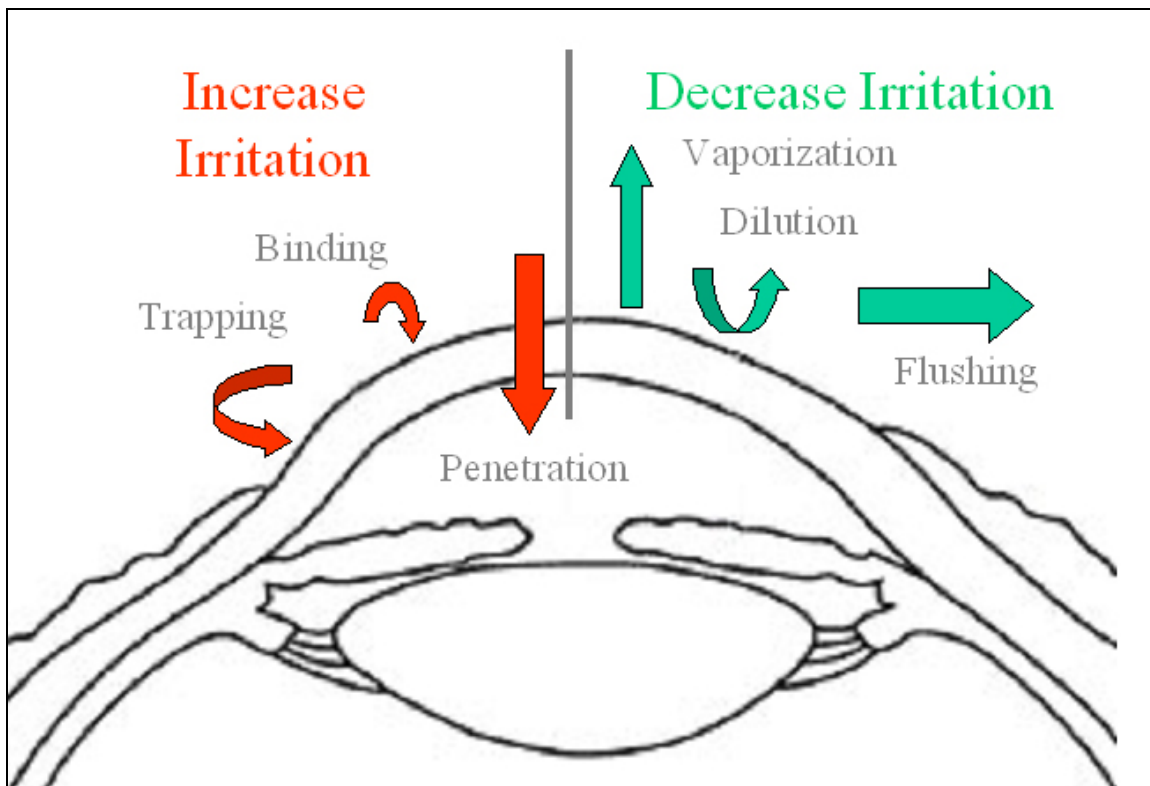
The rabbit eye is treated with the test article as an open system except that the initial installation is placed in the lower conjunctival sac. The consistency of these variables, from rabbit to rabbit, is not measured in the standard assay. Even the trapping of solid test articles, under the lower lid, is not controlled so that the effective exposure is not known. In the development of the BCOP assay, a closed dosing system was selected so that exact control over the exposure could be achieved. However, this control meant that the time of exposure might not be the same for all categories of test materials. The basic exposure periods for the assay of 10 minutes for liquids and 240 minutes for solids (20% suspension) were empirically derived (Sina, personal communication). However, we have seen that certain classes of liquids, notably volatile organic solvents, produce more damage in the 10-minute exposure than they seem to in the Draize test. This observation was made in the study of Balls et al (1995)[14] reported in the BRD and has lead to several studies on a more appropriate exposure time for such solvents. The results of one such study are provided in Appendix H-1[15]. Volatile organic solvents may perhaps also be difficult to assess in vivo as the data of Marzulli and Ruggles have suggested[16].

---

[14] Balls, M., Botham, P.A., Bruner, L.H., and Spielman, H. (1995) The EC/HO international validation study on alternative methods to the Draize eye irritation test. **Toxicology In Vitro** 9(6):871-929.

[15] Cuellar, N., Lloyd, P.H., Swanson, J.E., Merrill, J.C., Mun, G., Harbell, J.H., and Bonnette, K.L. (2004) Phase Two: Evaluating the eye irritancy of solvents in a simple fragrance mixture with the bovine corneal opacity and permeability (BCOP) assay. The Toxicologist 78: abstract 1306.

[16] Ibid reference 10

Another component of the analysis is to examine the ways in which chemicals might act on the cornea (in vivo and in vitro). Again, the work of Drs. Maurer and Jester sheds considerable light on ocular irritation at the cellular level. While potentially not all inclusive, we might break modes of action on the corneal (and other tissues of the eye) into four basic categories with some examples of the types of chemicals that might act through one or more of these modes of action:

- Membrane lysis
  - Surface active agents
  - Organic solvents
- Protein Coagulation/Denaturation
  - Acids and certain solvents
- Saponification
  - Alkali (often progressive)
- Alkylation, Oxidative Damage
  - Reactive materials such as bleaches and peroxides

While the first three modes of action on the cornea have been shown to act rapidly and progressively through the tissue (anterior to posterior)[17], action of reactive chemistries is often delayed both in vivo and in vitro[18]. Furthermore, the keratocytes within the stroma may be selectively impacted and their degeneration leads to specific inflammatory changes.

In our experience, membrane lysis, protein coagulation/denaturation, and saponification lead to changes in opacity, permeability or both in the BCOP assay. Depth of injury (as assessed histologically) generally parallels the increases in opacity and permeability. However, certain reactive chemicals, particularly peroxides, act on the keratocytes without inducting a commensurate increase in either opacity or permeability in the BCOP. Several of the chemicals evaluated in the Balls et al (1995) (EC/HO study) showed a similar pattern of action and thus under prediction from these measures. Curren et al (2000)[19] reported that the addition of histology to the original endpoints of the assay was able to identify tissue changes suggestive of more severe irritation potential. These tissue changes were not observed in the corneas treated with the milder test articles.

9) Using the BCOP assay across a range of chemistries

Many users of the BCOP assay are evaluating formulations where the basic chemistry of the ingredients is well known. This allows the tailoring of the protocol to match the expected modes of action and physical form and possible exposures (either rabbit or human). Selection of the exposure and post-exposure times are selected according to the needs of the client, the expected action(s) on the cornea and time course of those actions (e.g., where the lesions are manifested in 2 to 4 hours). The "standard" protocol outlined in Appendix A is an example of this approach. Where reactive chemistries are included in the formulation, the

---

[17] Ibid reference 4

[18] Ibid reference 6

[19] Curren, R., Evans, M., Raabe, H., Ruppalt, R., Harbell, J. (2000) Correlation of histopathology, opacity, and permeability of bovine corneas exposed in vitro to known ocular irritants. **Veterinary Pathology** 37(5):557.

post-exposure incubation time is selected to allow the damage (if any) to be manifested in the tissue and histology is used to determine the presence or absence of any lesion that might not be reflected in opacity or permeability measurements. An example of the use of a long post-exposure protocol is provided in Appendix H-3[20]. In this study, 4 and 20/24 hour post exposure incubations were employed to evaluate this reactive chemical (percarbonate). A similar kind of approach was used in the evaluation of model bleach-containing cleaning formulations[21].

In Table 12-8, the ICCVAM/NICEATM propose over 25 chemical classes for their reference set of chemicals/formulations. Rather than focus on the organic chemist's view of the world, it might be more helpful to consider what these various chemicals do to tissues of the eye.

It is important to have an approach for assessing unknown chemicals (chemistries). Curren et al (2000) used the experience from the EC/HO study to propose such a testing strategy. Test chemicals should be tested neat and at 20% (if solids) using unbuffered dosing solutions. The protocol should also include a shorter (2 to 4 hours) and longer (16-20 hours) post-exposure incubation and histopathological evaluation. This approach is intended to cast the widest net to address possible modes of action and resulting lesions. This strategy is based on the observation that those materials acting to lyse membranes, denature proteins or other macromolecules, or saponify lipids will be detected through the opacity and permeability endpoints. Those materials that act on specific cells to produce delayed cell death will be detected through the histological evaluation. In both cases, the histological evaluation provides a direct/confirmatory measure of depth of injury. Certain lesions are characteristic of severe irritants. These include necrosis or pyknosis of the keratocytes in the deep stroma (below mid depth) and loss of functional endothelium across the majority of the cornea. The loss of functional endothelium is reflected in the presence of collagen matrix vacuolization directly above Descemet's Membrane.

Conclusions

We have presented the following points for the ICCVAM/NICEATM to consider in future drafts of the BRDs.

- The "relevance" of a test method's results should be understood through the validation process, not just the "accuracy" of the method. This is especially true when the estimate of accuracy is based on a comparison to a benchmark test method where the reproducibility of the benchmark test method is not known.
- To the best of our knowledge, the BCOP BRD has erred in stating that results from this test have "not yet been considered by US Federal agencies for regulatory use where submission of testing data is required."

---

[20] Gran, B.P., Swanson, J.E., Merrill, J.C., and Harbell, J.W. (2003) Evaluating the irritancy potential of sodium percarbonate: a case study using the bovine corneal opacity and permeability (BCOP) assay. **The Toxicologist** 72:220.

[21] Swanson, J.E., White, B.T., Gran, B.P., Merrill, J.C., and Harbell, J.W. (2003) Evaluating oxidizing/reactive cleaning products in the bovine corneal opacity and permeability (BCOP) assay. **The Toxicologist** 72:220-221.

- To the best of our knowledge, the majority of companies, utilizing the BCOP assay to predict severe injury, do not retest negatives or suspected false positives in vivo. Therefore the statement in the BRD to the contrary should be modified
- There is evidence that initial depth of injury in a three dimensional model can give insight as to the potential for reversibility or lack of reversibility of the injury. Thus the fact that the bovine corneas are generally held no longer than 24 hours after exposure should not be portrayed in as negative a fashion as it currently is in the BRD.
- To the best of our knowledge, the lack of reproducibility of the animal test has been one of the driving factors behind developing in vitro methods. This thought should be presented in the BRD.
- If there are data to support the contention that the current animal test is a satisfactory predictor of the human ocular response, these data should be presented.
- The classification of severe irritants can vary considerably depending on the exact classification rules (EPA, GHS, etc) applied. A major reason that they vary is due to the fact that not all animals respond the same to the irritants. This influence of the variability of the animal test on final classification should be presented.
- Different chemicals cause injury by different modes of action; therefore this fact should be accounted for in the exposure parameters of the in vitro model.
- For completely unknown chemicals, a broad screening protocol, which includes histological examination of the corneas, should be employed.