Institute for In Vitro Sciences, Inc.

25 August 2005

Dr. Raymond Tice
NICEATM, NIEHS
P.O. Box 12233
MD EC-17
Research Triangle Park, NC 27709

Dear Dr. Tice,

On 25 July 2005 NICEATM published a revised analysis of four in vitro test methods proposed for detecting ocular corrosives and severe irritants. The Institute for In Vitro Sciences, Inc. appreciates the opportunity to comment on this new report. Our comments and observations are attached.

Sincerely,

Rodger D. Curren, Ph.D.
President
Institute for In Vitro Sciences, Inc.

John W. Harbell, Ph.D.
Vice President/Treasurer
Institute for In Vitro Sciences, Inc.

Comments from the Institute for In Vitro Sciences, Inc. on "Revised Analysis and Proposed Reference Substances for In Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants"

It is encouraging that additional data has been received dealing with the identification of ocular corrosives and severe irritants and that a reanalysis of the performance of the in vitro methods using these new data has been conducted. We will, for the most part, confine our comments to information concerning the Bovine Cornea Opacity and Permeability (BCOP) assay, although a number of our general comments will apply to the analysis of the other assays as well.

Overall Method of Analysis

The basis of the analysis of the BCOP method, as well as the other three methods, is a calculation of the "accuracy" of the method. This "accuracy" consists of comparing the classification results determined by the in vitro assay with the classification determined by the animal method, i.e. asking the question of whether the in vitro assay correctly identifies the classification of a specific chemical or not. **The assumption here is that the animal results have correctly determined the hazard classification of that chemical**. We believe that this last statement is not necessarily true, considering that even the EPA, GHS and EU classification systems can categorize the hazard differently given the same set of animal results (often due to the different weighting they give to the results from a single animal).

Of greater concern to us is the significant probability that a retest of an individual chemical in a rabbit test on a second day, in a second laboratory, or with a different set of technicians will result in a different classification for that chemical. NICEATM's own calculations (Draft Report: Interim Analysis of the Estimated Potential Underclassification Rates of the Current Rabbit Test for Detecting Ocular Corrosives and Severe Irritants. 6 January 2005), based only on the least problematic parameter – reproducibility between animals within an assay, estimates an underclassification rate of up to 16% (for solids). An average underclassification rate would of course depend on the spectrum of chemicals being tested, but we might assume it to be 12 – 13%.

The important point is that the underclassification rate stated above only considers one aspect of variability. It does not incorporate the day-to-day variability that would likely be higher than variability within an assay, nor does it incorporate laboratory-to-laboratory variability where differences in environmental conditions, animal strains, and technician training etc. may have a profound effect on reproducibility.

Thus we, and many others, believe that the "classifications" given to chemicals in this analysis are hardly absolute. However the accuracy analysis that is conducted implies that they are. We feel that whenever the results of such analyses are presented that they should be accompanied by information indicating – to the extent possible – what the expected underclassification rate of the animal test itself would be. The in vitro test cannot be expected to perform any better than this because it is limited by the standard itself.

We are aware that it is difficult to obtain current data about between laboratory reproducibility for individual chemicals so that a more exact estimate of the underclassification rate of the animal method can be obtained. However just because it cannot be accurately obtained doesn't mean it should be ignored. At the very least the estimates of underclassification based only on intraassay reproducibility should be expressed as minimum estimates and that the other sources of variability which we know exists would very likely move this rate higher. **Only when the performance of the animal test itself is well estimated can a reasonable comparison to the in vitro test be made.**

Removal of some data from the general BCOP analysis

Some of the data submitted by Casterton, et al. have been removed from the general analysis because they use different endpoint measurements than that of the primary prediction model being tested. While the Casterton data provide useful information about the performance of the BCOP methodology overall, we agree that these data should be excluded from the general calculations.

Evaluation of solids

Of the six solids that were underpredicted, we have evaluated two of these materials using histopathology of the corneas as an additional endpoint and found that more damage had occurred to the corneas than was estimated by opacity and permeability measurements alone. We continue to believe that when testing new chemistries that histological analysis and a longer post-exposure time point needs to be included in the protocol.

We intend to analyze more of the underpredicted solids in our laboratory using histopathogy as an additional endpoint and will report that information when it is obtained.

Correlation of BCOP (or other in vitro method) results with the EPA's hazard classifications

Since the response of a single animal, even if it appears to be a clear outlier with respect to the response of other test animals, can drive the classification of a chemical to a higher category, it is unrealistic to expect any well controlled in vitro system to match that classification. We believe the in vitro results will give a

much better reflection of the real hazard of that material than will a classification based on the response of a single high responding rabbit.

In addition, since the dosing parameters of the rabbit test (and subsequent rabbit behavior) does not reflect what would be expected to occur in a human accidental exposure, it is very questionable to expect that variability in rabbit response would reflect the range of human responses.

NICEATM has begun to address this problem in their draft report on underclassification rates (see "Overall methods of analysis" above), by analyzing three subgroups of strong, moderate and weak responders within each classification level. We believe it would be important to extend this analysis even further than is currently being done to each of the four in vitro test methods under consideration.

<u>Comment on the criteria for acceptance of hazard classification information for test chemicals</u>

We are aware that some test chemicals have not been included in the performance evaluation of the Isolated Chicken Eye (ICE) method (and perhaps others), because of a lack of sufficient eye injury data. The chemicals to which we are referring are those which, because they were corrosive in the ICE test and therefore could not ethically be tested in the rabbit eye, were instead tested on the skin and found to be skin corrosives. We believe that such data adds significant information to understanding the performance of the ICE method and therefore should be included in an additional analysis unless quantitative information comes forward to indicate a more secure link between skin corrosively results and eye corrosively results.

We again thank the NICEATM for the opportunity to comment on this reanalysis document, and look forward to the final report on the program that hopefully will document the acceptability of several in vitro methods for the classification of ocular corrosive and severe irritants.