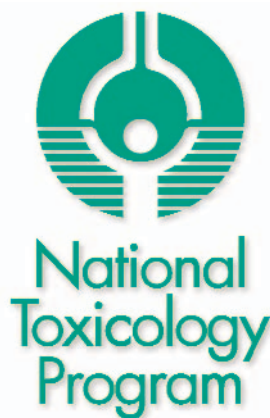


Alternative Tests Make the Grade

Toxicity testing is absolutely necessary for assessing the safety of substances in food, air, and water, in the workplace and at home. Although there are several tried-and-true toxicity assays, the search is always on for methods that can even better predict toxic effects. As scientific understanding of the effects of environmental toxicants grows, new tests are needed to evaluate previously unexamined end points and to take advantage of advances in biotechnology and the growing knowledge of how toxicants work at the molecular and cellular levels. Another issue is how to develop tests that can reliably and accurately assess toxicity using less time, money, and materials, and with greater regard for animal welfare. The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) was established in 1997 to address these needs through the development, validation, acceptance, and harmonization of new and revised toxicological test methods throughout the federal government.

ICCVAM is made up of representatives from the NIEHS and 13 other federal regulatory and research agencies. The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM, pronounced “nigh SEE tum”) was created in 1998 to support ICCVAM’s goals. NICEATM is housed on the NIEHS campus in Research Triangle Park, North Carolina. In collaboration with ICCVAM, center staff review all nominations for assays to be evaluated and decide whether there are sufficient data for an independent public scientific peer review to proceed. NICEATM also assists in arranging the peer review sessions and organizing the expert panels and workshops. The center prepares and publishes reports and information about each validated test including a detailed description of the methods and data used to substantiate its validity. This information, along with ICCVAM’s recommendations on the test’s use, is sent to research and regulatory agencies, which then decide whether and how the method will fit into their program activities.

The goal of ICCVAM is to facilitate the scientific validation and regulatory acceptance of new test methods that are more predictive of human health and environmental effects than current methods, and that will reduce, refine, and replace animal use where



scientifically possible. Reduction means employing methods that use fewer animals than standard historical models, or that obtain more types of information from each animal. Refinement refers to enhancing the animals’ well-being, for instance, by using more humane end points to end studies before the onset of significant pain and distress. Replacement can include using nonanimal systems or lower species (such as single-celled organisms) in place of higher species.

Alternative methods also include *in vitro* methods, such as cell cultures, and computer technologies that allow scientists to use existing animal data to build predictive models. Other alternative methods include transgenics, toxicogenomics, proteomics, high throughput technologies, molecular biomarkers, noninvasive imaging and labeling techniques, and tissue engineering—all of which take advantage of recent advances in science and technology.

Validation Projects to Date

Independent peer review panels are convened to review each test method. These panels are made up of technical experts from around the world who have no financial investment in the outcome of the review. Each panel must reach a scientific consensus on the extent to which the test method under review is useful for predicting human and/or environmental health effects. The panel must also identify any limitations of the test method.

In reviewing each method, the peer review panel considers two overall questions in addition to a series of detailed test-specific questions. First, has the method been evaluated sufficiently and is its performance satisfactory to support its proposed use? And second, has there been adequate consideration of animal welfare in terms of reduction, refinement, and replacement? To date, ICCVAM has completed reviews on two alternative test methods—the murine local lymph node assay (LLNA) and the Corrositex assay.



LLNA. The first peer review panel evaluated the LLNA in September 1998. Already used for over a decade to gauge the potential of chemicals to cause allergic contact dermatitis in a research setting, the assay was recommended for full-fledged endorsement as a viable stand-alone test method.

Traditionally, guinea pig assays are used to determine the potential of a chemical or

product to cause an allergic contact dermatitis response. Such assays involve the repeated application of a test substance to the guinea pigs’ skin and examination of the site over 4 weeks to see whether an allergic response has occurred or whether a delayed-sensitivity response can be induced with additional applications of the substance. While considered fairly reliable, these assays may produce false positive and false negative results. In addition, the results are somewhat subjective and require considerable experience and expertise to be accurately interpreted.

While the traditional assays measure the allergic reaction itself, the LLNA measures the lymphocyte proliferation response, a necessary and inevitable biological precursor to sensitization. In the LLNA, the test substance is applied to the ears of 4–5 young adult mice for 3 consecutive days. The animals are rested for 2 days and then euthanized, and their lymph nodes are removed and examined. If the test substance spurs an immune response, there will be a rapid proliferation of lymph cells after the exposure. This measurable increase in lymph cells can be used to characterize the sensitization potential of the test substance.

The peer review panel evaluated data on 209 chemicals. Of these, both LLNA and guinea pig data were available for 126 chemicals, and both LLNA and human data were available for 74 chemicals. From the data submitted, the panel concluded that the accuracy (the proportion of correct outcomes) of the LLNA was about 86% when compared to data from all guinea pig tests and about 72% when compared to human data (guinea pig tests have an accuracy of about 73% compared to human data). In terms of accuracy, sensitivity (the proportion of all test substances that are correctly classified as positive), specificity (the proportion of all test substances that are correctly classified as negative), positive and negative predictivity (the proportion of positive and negative test substances that are correctly identified by the assay as such), and comparability to human data, the LLNA performs at least as well as traditional guinea pig assays and in some cases was a better predictor of a human allergic response. The LLNA was therefore determined to be a viable alternative to traditional guinea pig assays for identifying strong to moderate chemical sensitizing agents and predicting the risk of human allergic contact dermatitis.

The LLNA offers many advantages over traditional guinea pig assays. It allows scientists to measure changes in allergic response over several concentrations. Because the method evaluates the induction phase rather than the elicitation phase of the response, the mice used in the LLNA don’t suffer the

discomfort of the secondary allergic skin reaction. William Stokes, associate director for animal and alternative resources in the NIEHS Environmental Toxicology Program and ICCVAM cochair, says, "The incorporation of an earlier mechanistic end point in the LLNA avoids virtually all pain and distress that can occur in the traditional guinea pig test. This test system is an excellent example of an ideal humane end point." The LLNA can also be performed in a much shorter period of time, at a lower cost, and with fewer animals than the traditional assays.

The main weakness identified was the assay's proclivity toward false negative results with some weak sensitizing agents and with metals, and false positive results with some strong irritants. Despite these limitations, the peer review panel unanimously recommended the LLNA as a stand-alone alternative for contact sensitization hazard assessment. However, the panel did suggest some minor standardization changes to the protocol. In addition, because formal audited reports indicating adherence to good laboratory practices were not prepared for many of the validation studies examined (the report notes this was because the data were primarily intended for publication), the panel suggested that retrospective data audits be conducted on at least three of the validation studies conducted by the sponsors.



Corrositex.

In January 1999, an ICCVAM peer review panel evaluated the Corrositex assay, manufactured by

InVitro International of Irvine, California, which is used to determine whether a chemical will cause irreversible damage (corrosion) to skin. Even before ICCVAM was established, the U.S. Department of Transportation had accepted Corrositex as an assay for certain specific chemical classes.

In the past, corrosivity of chemicals and chemical mixtures was determined by the ability of the test substance to visibly damage skin tissue—traditionally, that of a rabbit—at the site of contact within 4 hours of exposure. In the Corrositex method, a glass vial is filled with a chemical detection system consisting of water and pH indicator dyes, and overlaid with a collagen matrix membrane. If the sample chemical penetrates the membrane, the fluid will change color from yellow to reddish pink.

Before testing with Corrositex, all test chemicals are prescreened by putting a small amount of the test material directly in the detection fluid. If the material does not shift the pH of the fluid to less than 4.5 or greater than 8.5 and thereby change the color of the detection fluid, it does not qualify for testing.

The panel reviewed Corrositex data from tests of 163 different materials for which there were corresponding *in vivo* rabbit corrosivity data. The panel confirmed what the Department of Transportation had already decided for acids, bases, and acid derivatives—that the test is useful as a stand-alone assay. The panel also concluded that Corrositex is useful as part of a tiered assessment strategy for testing those and other chemical and product classes.

For acids, bases, and acid derivatives, Corrositex had an overall sensitivity of 90%, an overall specificity of 70%, and an overall accuracy of 87%. The main drawback the panel noted was the proportion of test chemicals that did not qualify for use with the test (about 18%). Further review of the data indicated, however, that the qualification step eliminates primarily chemicals that are noncorrosive, although some nonqualifying chemicals may actually be corrosive.

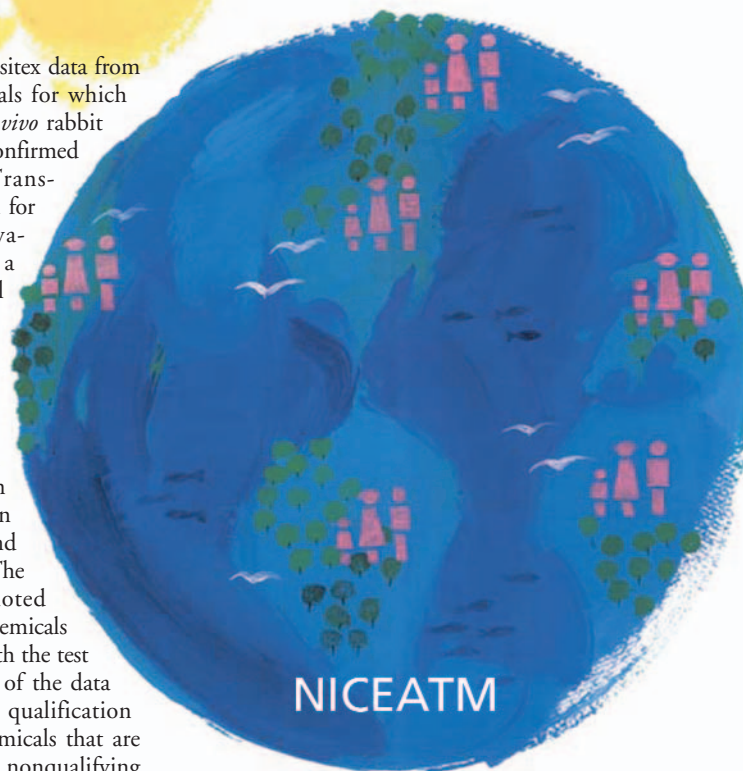
When used as a stand-alone assay for identifying corrosives, Corrositex can replace animal use altogether; when used as part of a tiered approach, it reduces and refines the need for animal use by providing a basis for decisions on whether further *in vivo* tests are needed. When follow-up *in vivo* tests are necessary, they may use fewer animals. Corrositex also has the advantages of being less expensive than the rabbit skin test, displaying results more quickly, and requiring no special equipment, facilities, or training to use.

The panel recommended several minor changes to the current test method protocol in order to provide more detailed instructions with the test kit and to address issues of variability in testing conditions.

What's Next

A workshop is planned for this May to assess the validation status of the Frog Embryo Teratogenesis Assay—*Xenopus* (FETAX), a method used to evaluate the developmental toxicity potential of chemicals. FETAX uses early-stage embryos of the South African clawed frog (*Xenopus laevis*) to measure the effects of chemicals on mortality, malformation, and growth inhibition. In July, a peer review panel will review the validation of a revised version of the up-and-down procedure, a reduction alternative for the traditional LD50 test. And this fall there will be a workshop to discuss currently available *in vitro* methods for predicting acute toxicity.

In another project, ICCVAM and NICEATM are working with the U.S. Environmental Protection Agency on the val-



idation of screening and testing methods for endocrine disruptors. The center is also in the process of gathering background documentation on studies evaluating the use of transgenic mice to replace standard mice in carcinogenicity studies. These transgenic mice are genetically engineered to be more susceptible to developing cancer in response to chemical carcinogens, allowing for studies that can be performed in less time using fewer animals.

Based on experience gained through reviewing the LLNA and Corrositex, in October 1999 ICCVAM revised the guidelines for stakeholders wishing to submit assays for peer review. The new guidelines delineate the type of data the committee will need to evaluate a new or revised test method, and include specifications for organizing the supporting information so that the peer review panel can more quickly evaluate the validity of a test method. Says Stokes, "We encourage those involved in developing and validating new test methods to use these guidelines during the planning stages. This will increase the likelihood that data needed to characterize the usefulness and limitations of the test method will be generated. It will also increase the likelihood that there will be adequate information for agencies to make decisions on the acceptability of the method." The new guidance was opened for public comment with a 2 December 1999 *Federal Register* notice. Depending on the comments received, the guidance may be revised further and reissued later this year. —Susan M. Booker