2.0 ICCVAM Recommendations for the rLLNA Test Method

ICCVAM evaluated the validation status of the rLLNA test method as a reduction alternative to the traditional LLNA. The rLLNA should be used for the hazard identification of skin-sensitizing substances if dose-response information is not needed (e.g., for a compound presumed to be a strong sensitizer), provided there is adherence to all other LLNA protocol specifications as described in the updated ICCVAM-recommended LLNA test method protocol (available in **Appendix B** and at the NICEATM–ICCVAM website[16]). To further reduce animal use, the rLLNA should be used routinely as an initial test to determine allergic contact dermatitis (ACD) potential of chemicals and products before conducting the traditional LLNA. Negative substances can be classified as non-sensitizers, and positive substances can be classified as sensitizers.

Where dose-response information is required (e.g., for a compound presumed to be a weak or borderline sensitizer), positive substances must be tested in the traditional multidose LLNA. Accordingly, those substances for which dose-response information will be required and that are also suspected of having allergic contact dermatitis potential following consideration of all available information should be initially evaluated using the traditional LLNA.

2.1 ICCVAM Recommendations: Test Method Usefulness and Limitations

NICEATM and ICCVAM conducted a retrospective evaluation of rLLNA data to determine the test method's ability to distinguish between skin sensitizers and non-sensitizers. The performance assessment for the 465 unique substance and vehicle combinations evaluated in the study is provided in **Section 3.0**. Based on a review of the available data and comparison with the traditional LLNA, the scientific validity of the rLLNA has been adequately evaluated. ICCVAM concluded that, when conducted in accordance with the updated ICCVAM-recommended LLNA test method protocol specifications included in **Appendix B,** the rLLNA's performance is sufficient to distinguish between skin sensitizers and non-sensitizers when dose-response information is not required. This recommendation is based on its performance compared to that of the traditional LLNA. ICCVAM also concludes that use of the rLLNA can reduce by 40% the number of animals used for each test.

There is a small possibility of a false negative result (1.9% [6/318]) when compared to the traditional LLNA. This information should be considered when evaluating results from the rLLNA, and negative results should always be subjected to a weight-of-evidence evaluation of supplemental information (e.g., possibility of downturn in response at the high dose, test results with similar substances, peptide-binding activity, molecular weight, other testing data). If false negative results are suggested, confirmatory testing in the traditional LLNA or another accepted skin sensitization test method should be considered.

All of the testing limitations that apply to the traditional LLNA apply to the rLLNA also. For example, the rLLNA may not be suitable for use with certain types of test substances, such as nickel salts, mixtures, high-molecular weight compounds that cannot penetrate the stratum corneum, strong dermal irritants, or chemicals whose pharmacodynamic activity is to release dermal cytokines that cause local lymph node proliferation (e.g., certain pharmaceuticals such

[16] http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNAprotocol2008.pdf

as imiquimod [Gaspari 2007]). Additionally, the rLLNA may not be suitable for test substances that do not adhere for an acceptable period of time when applied to the dorsum of the ear.

Independent Peer Review Panel Conclusions and Recommendations

The Panel agreed that the available data support ICCVAM's draft recommendation that the rLLNA should be routinely recommended for hazard identification when dose-response information is not required. The Panel also agreed that to further reduce animal use the rLLNA should be routinely recommended as the initial test to identify sensitizers even if dose-response information *is* required, because negative results would not require additional testing. This is applicable in the occupational and public health setting in which obtaining hazard information is of critical importance. Subsequent traditional LLNA testing of substances that were positive in the rLLNA will provide dose-response information to assure detection of hazardous substances and allow potency estimates. The benefits of screening out the negatives, which do not require dose-response information, are clear; however, the animal welfare gains will depend on the proportion of test substances in any class that turn out to be non-sensitizers. The possible consequences of delays from another round of testing of those materials identified as sensitizers should also be considered.

The Panel agreed that the draft test method recommendations adequately addressed the low false negative rate by giving cautionary and weight-of-evidence consideration to the negative substances (and any possible false positive results). Furthermore, the Panel concluded that interspecies differences between the animal model and humans would probably make the false negative rate unimportant.

2.2 ICCVAM Recommendations: Test Method Protocol

ICCVAM recommends basing the protocol for rLLNA testing on the updated ICCVAM-recommended LLNA protocol, which addresses the rLLNA procedure (**Appendix B)**. The only difference between the traditional LLNA and the rLLNA test methods is that the middle- and low-dose groups are omitted in the rLLNA. On the basis of Panel comments, ICCVAM updated the traditional LLNA test method protocol to provide guidance on identifying the appropriate maximum dose for testing. In the rLLNA, in addition to the concurrent vehicle and positive-control groups, each test substance is tested at only one dose level (the high dose), whereas in the traditional LLNA each test substance is tested at a minimum of three dose levels. The test substance concentration should be the highest soluble concentration that does not induce overt systemic toxicity and/or excessive local irritation. Any other approach, such as one based on a pre-established threshold dose level, is inappropriate. For example, Kimber et al. (2006) proposed a 10% threshold concentration at which all negative results would be considered valid. However, 51 (16% [51/315]) of the test substances evaluated were non-sensitizers at concentrations of at least 10%[17] but were sensitizers at higher concentrations.

In the traditional LLNA test method protocol, a stimulation index (SI) is calculated as the ratio of the mean incorporation of ^{3}H-thymidine or ^{125}I-iododeoxyuridine by the auricular lymph nodes of the treated animals and that of the vehicle control animals. In the rLLNA, as in the traditional LLNA, the threshold for classifying a substance as a skin sensitizer is an SI $\geq$ 3.

[17] An initial dose was tested at 10% or greater and resulted in a stimulation index (SI) < 3, while a subsequent higher dose resulted in an SI $\geq$ 3.

In the updated LLNA test method protocol (**Appendix B**), ICCVAM recommends collecting individual animal data in order to allow identification and exclusion of outlier values that could result in false negative or false positive results. This is especially important to help avoid false negative results for weaker sensitizers (i.e., substances that induce an SI just above 3). The U.S. Environmental Protection Agency (EPA) Health Effects Test Guideline 870.2600 (EPA 2003) also requires the collection of individual animal data for the assessment of interanimal variability and a statistical comparison of test- and control-group measurements. While the Organisation for Economic Co-operation and Development (OECD) Test Guideline (TG) 429 (OECD 2002) allows for both the collection of individual animal measurements and the pooling of the lymph nodes for each treatment group, the latter eliminates any measure of interanimal variability and/or identification of outlier values, as well as statistical identification of a positive/negative response.

OECD TG 429 requires that each dose group consist of at least four animals if pooled animal data are collected and a minimum of five animals if individual animal data are collected (OECD 2002). To determine if the required number of animals for individual animal data collection could be the same as the required number for pooled data without diminishing accuracy, NICEATM evaluated data from 83 LLNA studies (275 dose groups) from six different laboratories **(Appendix C)**. This is important because most animal-use regulations require that the minimum number of animals be used in studies, which currently results in many countries collecting only pooled data because doing so requires fewer animals. This evaluation indicated that a reduction in the sample size from five to four animals per group is unlikely to have a significant impact on the results of an LLNA study; therefore, the ICCVAM-recommended LLNA test method protocol (**Appendix B**) was revised to require a minimum of four animals per dose group.

The updated ICCVAM-recommended LLNA test method protocol (**Appendix B**) also recommends that each test include a concurrent positive-control substance. Use of a positive-control substance can ensure that all protocol procedures are conducted properly and that all aspects of the test system work properly such that they produce a positive response. However, similar to OECD TG 429 (OECD 2002), the updated ICCVAM-recommended test method protocol states that testing of the positive-control substance at intervals of no more than six months may be considered in laboratories that conduct the LLNA at least once per month and that have a history and a documented proficiency for obtaining consistent results with positive controls.

Users should be aware that the decision to include a positive control only periodically instead of concurrently could affect the adequacy and acceptability of negative study results generated without a concurrent positive control. For example, if a false negative result is obtained in the periodic positive-control test, all negative test-substance results obtained since the last acceptable periodic positive-control test and the unacceptable periodic positive-control test could be questioned. In order to demonstrate that the prior negative test-substance results are acceptable, a laboratory could be expected to repeat all negative tests, which would require additional expense and increased animal use.

Independent Peer Review Panel Conclusions and Recommendations

The Panel agreed with ICCVAM's draft test method recommendations and recommended adherence to the ICCVAM-recommended LLNA protocol (with modifications omitting the

middle- and low-dose groups) for future rLLNA testing. The Panel also advised collecting individual animal data for future studies because it would allow an estimate of interanimal variability and conducting a statistical analysis to determine if the test substance is significantly different from the control substance.

The Panel agreed that the current recommendation to select a maximum applied dose for the rLLNA based on the absence of overt systemic toxicity and/or excessive local irritation is appropriate. The Panel also agreed that the data did not support establishment of a uniform concentration threshold for the maximum concentration to be tested. Thus, it seemed justifiable that preliminary experimentation (as would be typically performed during a dose range-finding study) should be conducted for vehicle selection, test substance solubility, and stability in the vehicle.

2.3 ICCVAM Recommendations: Future Studies

ICCVAM recommends additional studies to further characterize and potentially improve the usefulness and applicability of the rLLNA for identifying potential skin sensitizers. For instance, to improve the predictive performance of the rLLNA compared to the traditional LLNA, ICCVAM recommends investigating the basis for abnormal dose responses for six substances that would have resulted in false negative results using the rLLNA rather than the traditional LLNA. This information should help identify ways to improve the accuracy of the rLLNA compared to the traditional LLNA.

Efforts should also be made to identify data from guinea pigs and humans for substances like these that exhibit abnormal dose responses in the traditional LLNA. Information from post-marketing and/or occupational exposures should be collected and assessed.

ICCVAM recommends that all future LLNA studies should collect and analyze individual animal data. This will allow detection of outliers and avoidance of false negative results that can occur from pooling data that include one or more abnormally low values. Existing LLNA studies using data pooled from all animals in a dose group, such as four of the six false negative rLLNA results in this evaluation, should be evaluated further with data obtained from individual animals within each dose group to determine if data pooling may have led to false negative outcomes.

ICCVAM also recommends that users identify opportunities to use fewer animals per dose group without compromising test method accuracy. Thus, laboratories conducting the LLNA should collect and analyze data from individual animals. The updated ICCVAM-recommended LLNA test method protocol includes statistical procedures necessary for such determinations (**Appendix B**). This includes evaluating the laboratory's historical positive-control database to determine if the number of animals in the concurrent positive-control group can be reduced.

Independent Peer Review Panel Conclusions and Recommendations

The Panel indicated that, though limited in scope, the available data supported ICCVAM's draft test method recommendations for additional studies. The Panel agreed that attempts should be made to investigate if maximum solubility was achieved (e.g., use of chemical-specific methods to document solubility). For hazard assessment, it was troublesome that there were so many vehicle choices, because the vehicle could have a significant effect on whether (and how much) a test substance penetrated the skin barrier. Observed vehicle effects may relate to dermal

penetration as well as to immunomodulation. The Panel considered it desirable to follow the hierarchy of vehicles recommended in the ICCVAM-recommended LLNA protocol. The Panel suggested that it might be informative to test both known mild and severe sensitizers concurrently in all recommended vehicles to evaluate whether a specific vehicle choice(s) might influence the results.

2.4 ICCVAM Recommendations: Performance Standards

ICCVAM developed performance standards for the traditional LLNA, which may in turn be applied to the rLLNA.[18] These test method performance standards are proposed to evaluate modified LLNA test methods that are functionally and mechanistically similar to the traditional LLNA. Thus, modified rLLNA test method protocols that adhere to the LLNA performance standards would be considered acceptable for hazard identification purposes.

[18] Available at http://iccvam.niehs.nih.gov/methods/immunotox/PerfStds/llna-ps.htm