DRAFT ICCVAM TEST METHOD RECOMMENDATIONS

Assessment of the Use of the LLNA to Address Relative Skin Sensitization Potency

January 2008

1.0 Draft Recommendations: Test Method Uses and Limitations

Background: ICCVAM is currently evaluating the validation status of the LLNA for the classification of skin sensitization potency. The information included in this BRD is based on a retrospective review of LLNA, guinea pig, and human data derived from a database of over 500 substances, 170 of which have comparative LLNA, guinea pig, and/or human data. Among these 170 substances, there are 112 substances with comparative human data (97 sensitizers, 15 non-sensitizers), 105 substances with comparative guinea pig data (52 sensitizers, 53 non-sensitizers), and 47 substances with comparative human and guinea pig data (34 sensitizers, 13 non-sensitizers). The performance of the LLNA, using LLNA EC3 values, was evaluated against the potency classification assigned based on either human or guinea pig data.

Draft Recommendations:

Using LLNA Data to Predict Human Potency Classification Categories (i.e., Strong vs. Weak[1]*):* Based on this analysis, there is a significant positive correlation ($p < 0.0001$) between EC3 values and human threshold values used to distinguish strong from weak sensitizers. An accuracy analyses was conducted using a range of LLNA EC3 values versus human threshold concentrations using either ≤ 250 □g/cm^2 or ≤ 500 □g/cm^2 as the threshold for discriminating between strong and weak human sensitizers. However, this correlation is not very strong, as evidenced by R^2=0.405. The data suggest that there is only a small difference in accuracy when the human threshold concentration for distinguishing weaker sensitizers from strong sensitizers is ≤ 250 □g/cm^2 or ≤ 500 □g/cm^2. The LLNA EC3 threshold values that provide optimal (albeit modest) classification when compared to either of the two proposed human threshold values (≤ 250 □g/cm^2 and ≤ 500 □g/cm^2) were 6.8% and 8.1%, respectively. Using these two EC3 values, the correct classification rate was 74% (60/81) and 70% (57/81) for 250 and 500 μg/cm^2, respectively, while the over- and under-classification rates ranged from 28% (10/36) to 31% (9/29) and 24% (11/45) to 29% (15/52), respectively.

[1] Although the skin sensitization categories proposed by the GHS (United Nations Globally Harmonized System for the Labelling and Classification of Chemicals) are Category 1: Strong Sensitizer and Category 2: Sensitizer, to avoid confusion in this document these are instead referred to as "strong" and "weak" sensitizers.

When substances incorrectly identified as false positive or false negative in the LLNA (when compared to highest accuracy for human data) are included in these analyses, as well as those classified as non-sensitizers in both the LLNA and in humans, the optimal EC3 value was 9.4% for either human threshold concentration. Using all 112 substances with both LLNA and human data, the correct classification rate was 62% (70/112) and 60% (67/112) for 250 and 500 $\mu g/cm^2$, respectively, while the over- and under-classification rates ranged from 26% (13/50) to 33% (5/15) and 21% (10/47) to 33% (14/43), respectively.

Using LLNA Data to Predict Guinea Pig Potency Classification Categories (i.e., Strong vs. Weak): Based on this limited dataset, there does not appear to be a significant association between LLNA EC3 values and sensitization potency based on guinea pig data. In one analysis, which focused only on substances classified as sensitizers in both the LLNA and in guinea pigs, overclassification means that weak sensitizers are missclassified as strong while underclassification means that strong sensitizers are missclassified as weak. Using the optimal EC3 value of 2.0%, the correct classification rate was 73% (38/52), while the over- and under-classification rates were 28% and 26%, respectively. In a second analysis, which included substances classified as sensitizers in both the LLNA and in humans as well as substances classified in the LLNA as false positives and false negatives compared to the human, and substances classified as non-sensitizers in both the LLNA and in humans. In this analysis, overclassification means that nonsensitizers are misclassified as weak or strong sensitizers and weak sensitizers are missclassified as strong while underclassification means that strong sensitizers are missclassified as weak or nonsensitizers and weak sensitizers are misclassified as nonsensitizers. Using the optimal EC3 value of 3.6%, the correct classification rate was 57% (60/105), while the over- and under-classification rates ranged from 25% (8/32) to 61% (30/49) and 9% (3/32) to 17% (4/24), respectively.

Considered together, these data indicate that although there is a significant positive correlation between LLNA EC3 values and human sensitization threshold doses, this correlation is not strong. Therefore, the LLNA should not be considered as stand-alone test method for predicting sensitization potency, but must instead be used as part of a weight-of-evidence evaluation to discriminate between strong and weak sensitizers.

2.0 Draft Recommendations: Test Method Protocol for the LLNA

The ICCVAM recommended LLNA protocol, which is based on recommendations from an independent expert peer review panel evaluation of the LLNA (ICCVAM 1999), can be found on the ICCVAM-NICEATM website at http://iccvam.niehs.nih.gov/methods/immunotox/llnadocs/LLNAProt.pdf (ICCVAM 2001). The LLNA procedure is also described in the EPA Health Effects Test Guidelines (EPA 2003) and OECD TG 429 (OECD 2002). Although not included in these protocols, a description of how to calculate an EC3 is included in the draft ICCVAM LLNA performance standards (http://iccvam.niehs.nih.gov/methods/immunotox/PerfStds/LLNAPerfStd07Jan08FD.pdf)

3.0 Draft Recommendations: Future Studies

To further evaluate the usefulness and limitations of the LLNA for potency determinations, efforts should be made to identify additional human data and human experience and guinea pig data for substances with comparative LLNA data. Emphasis should be placed on identifying substances that are classified as strong sensitizers based on a threshold concentration between 250 $\mu g/cm^2$ to 500 $\mu g/cm^2$ to more adequately evaluate the usefulness and limitations of choosing one of these two reference thresholds that have been proposed.