Questions:

I. The approach by ICCVAM to validate the LLNA for the prediction of strong and weak skin sensitizers poses a methodological challenge. The reason is that the possibility of misclassification in humans of a substance's potency may negatively influence the outcome of the validation; i.e., it is possible that available HRIPT and HMT data may lead to a false human skin sensitization potency categorization. It is often difficult to correctly interpret the total dose used in the human tests due to insufficient documentation of total area dosed or possible prior patient exposure history.

- In their analysis, Schneider and Akkan (2004) used the chemicals included in the 1999 ICCVAM validation as a starting point for a literature search to identify skin sensitizers for which quantitative human data on induction doses were available expressed as dose per unit area (ug/cm^2). They were able to identify and assess 46 substances. They were not able to identify more substances as "relevant uncertainties are related to limitations in the human data, which mostly come from older studies. First, the reporting of size of the skin area to which the test substance has been applied and of the volume of test solution used is often insufficient. In some cases, skin area and test solution volume could be deduced from information given on types of patches and application systems used. Moreover, in human HRIPT and HMT studies observed incidences for sensitization reactions depend on the concentrations applied during both the induction and elicitation phase. Often, but not in all cases, the same concentration was applied for both phases. Otherwise, the overall outcome of the test may have been influenced by different elicitation concentrations, a factor not considered in the regression analysis."

  In the evaluation performed by ICCVAM in 2008, 76 substances with quantitative human data among them 16 with negative LLNA results have been included. With respect to the points raised by Schneider and Akkan, it is important that it is described why it was possible in the current analysis to include more substances with both positive human and LLNA data (n=60) than Schneider and Akkan (n=46). Therefore, detailed information on ICCVAM's assessment of human dose per unit area is needed and the possibility of misclassification arising from such approach needs to be described. This is important with respect to the assessment of the rate of putative misclassification of strong/weak skin sensitizers using the human data in order to interpret the outcome of the validation study.

NICEATM response:
There are a total of 112 substances included in the ICCVAM Potency BRD. These include 81 substances with LLNA EC3 values coupled with either NOEL and/or LOELs from HMT and/or HRIPT studies, 10 substances classified as non-sensitizers in humans and in the LLNA, 16 substances with either NOELs and/or LOELs from HMT and/or HRIPT studies that are classified as negative in the LLNA, and 5 human non-sensitizers classified as positive in the LLNA. Among the 81 substances referenced above, 55 are included in the Schneider and Akkan (2004) analysis. These authors state that:

- **The database was limited to "*Compounds for which response on experimental sensitization has been tested in both predictive human tests and the local lymph node assay*."**
- **For the human data, *"the lowest effective concentration applied during the induction phase of the study for each chemical was converted to a dose per unit area ($\mu g/cm^2$) using the information on substance concentration, application volume, and area of application given in the publication. Using the information on sensitization incidence given in the publications from this value a dose per skin area leading to a sensitization incidence of 5% (DSA05) was derived by linear interpolation. This low but existent effect level was assumed to be comparable to the EC3 effect level in the LLNA*."**
- ***Where multiple results in both human and LLNA data (including multiple results for one ranking level in the LLNA) were present for a particular chemical, an arithmetic mean was calculated. Negative results were not considered in the calculation of the mean values. The mean values include comparable results with different vehicles except strikingly discordant results from tests with varying vehicles.***

The ICCVAM Potency BRD includes data on an additional 32 substances obtained from references more recent than 2004. These include:

- **Seven substances from Lalko and Api (2006) who state:**
  - ***"For comparison to the LLNA results, human non-diagnostic patch test data for the tested essential oils were gathered from both published and unpublished literature sources. Priority was given to searching the two historically most relevant study types—the HMT and the HRIPT."***
  - ***"In most cases the only data available are HMTs conducted in 25 subjects at a single concentration in petrolatum at a dose that was related to the reported use in consumer products at the time the study was conducted—typically 10x the maximum reported use level. Nevertheless, it was instructive to compare the calculated EC3 values to the highest reported concentration tested in humans that did not result in sensitization reactions—the Maximum Tested No Observed Effect Level (MT-NOEL).***
- **19 substances from Api (2007)**
  - **Includes NOELs, MT-NOELs, and LOELs derived from either HMT or HRIPT studies obtained from the RIFM historical**

database of fragrance ingredients that have exhibited dermal sensitization potential.

- **Six substances from Basketter et al. (2005) which states:**
  - ***"HRIPT data were obtained from the published literature and RIFM-FEMA database. For each chemical, a maximal no observed effect level (NOEL) was determined by examination of all sources. In the absence of positive data (where the NOEL was the maximal concentration tested), this has been highlighted."***

The remaining 25 substances obtained from references prior to Schneider and Akkan (2004) are:

- **15 substances identified in five different published articles as nonsensitizers in humans based on clinical experience.**
- **Seven substances from Griem et al. (2003), which reports that they:**
  - ***"Identified known human sensitizing chemicals for which both an EC3 value from LLNA and a NOEL and/or LOEL from HRIPT or HMT were available."***
  - ***"In some cases, LOELs were extrapolated from studies in humans where only one dose was tested based on sensitization rates (i.e., using a divisor of 3 for sensitization rate of 10-25%, or divisor of 10 for sensitization rate of 25-50%)."***
- **Three substances from Gerberick et al. (2001) which reports that:**
  - ***"A review of the limited, but nevertheless valuable, published literature on non-diagnostic human repeat patch testing, including both the human maximization test (HMT) and the human repeat insult patch test (HRIPT) was conducted. To help rank order the contact allergens tested in humans, a no-effect level (NOEL) for each chemical was determined. For comparison with LLNA EC3 values, NOELs were expressed as a function of dose per unit area of skin mg/cm^2 calculated from the concentration tested, patch size, and application volume. However, in some instances in which a true NOEL was not defined, we were limited to using either the lowest effect concentration (lowest effect level [LOEL]), or the highest concentration tested that did not give a response in an HRIPT or HMT procedure.***
  - ***These data, along with expert judgment based extensive clinical experience of ACD (e.g., clinical diagnostic patch test data), were used to classify the compounds as strong, moderate, weak, extremely weak, or nonsensitizers."***

- Should the HMT and HRIPT data be treated as equivalent?

- Is a correction factor/uncertainty factor/safety factor of 10 the most appropriate for the extrapolation of LOAEL values to NOAEL values? Schneider & Akkan (2004) used arithmetic means for human and LLNA data except when there were discordant results with varying vehicles. The authors interpolated linearly from the LOEL to a dose corresponding to an estimated sensitization incidence of 5%

("DSA_{05}"). Griem et al (2003) used LOAELs, which were divided by an arbitary factor in cases of high observed incidences.

- ICCVAM analyzed 250 ug/cm^2 and 500 ug/cm^2 as the cut-off values for a stronger sensitizer. Has the reverse analysis been performed where the LLNA (e.g., at EC_3 1% or 2%) and the GP data have been set as the standard and an optimal human cut-off calculated (does it vary between the LLNA and the GP data)?

NICEATM response: Appendix D of the draft ICCVAM BRD details the performance characteristics for use of LLNA EC3 values to predict the proposed categories of human and guinea pig sensitization potency. Each table includes EC3 cutoffs on (or about) 1.0% and 2.0% (see Table 1). In this evaluation, the availability of a LOEL and/or NOEL was used as the criteria for classifying a substance as a human sensitizer, and LOEL values were divided by 10.

Table 1 Performance Characteristics for LLNA EC3 Values of Approximately 1% and 2%

Comparison	EC3 cutoff	Correct Classification	Over-classification	Under-classification
EC3 vs. 250 μg/mL	1.02%	62%	0%	69%
	2.00%	69%	0%	56%
EC3 vs. 250 μg/mL*	1.02%	55%	10%	77%
	2.00%	59%	10%	67%
EC3 vs. 500 μg/mL	1.02%	53%	0%	74%
	2.00%	60%	0%	50%
EC3 vs. 500 μg/mL*	1.02%	47%	11%	79%
	2.00%	53%	11%	71%
EC3 vs. GP	0.90%	74%	18%	32%
	2.15%	79%	18%	23%
EC3 vs. GP*	0.90%	55%	47%	40%
	2.15%	58%	47%	31%
EC3 vs. GP WOE	1.05%	69%	21%	43%
	1.95%	73%	28%	26%
EC3 vs. GP WOE*	1.05%	52%	46%	52%
	1.95%	54%	49%	37%

*Includes false negative and false positive substances in the database
EC3 = Estimated concentration needed in the LLNA to induce an SI=3; GP = guinea pig; WOE = weight of evidence classification

Since the ultimate goal in sensitization testing is the prediction of human sensitization potency, optimizing human threshold levels based on LLNA data was not considered useful and therefore not conducted.

II. Once criteria are determined for acceptability and use of human data, questions arise about the data from LLNA studies:

- Can the LLNA protocols be narrowed, e.g., by selection of solvents or choice of other test parameters to improve correlation coefficients? Is it meaningful to combine results for different solvents?
- For repeat LLNA studies for a chemical substance, which EC_3 value should be selected? Should the geometric mean or the most conservative value be used?

III. How representative of sensitizers may the selection of chemicals with human data be? Does the set of chemicals analyzed by ICCVAM emphasize strong sensitizers?

NICEATM response:
The draft ICCVAM BRD uses two different proposed human thresholds to delineate between strong and weak sensitizers for the 97 substances included in the performance evaluation. In this evaluation, the availability of a LOEL and/or NOEL was used as the criteria for classifying a substance as a human sensitizer, and LOEL values were divided by 10.

- **When 250 μg/cm^2 is used as the threshold, there are 47 strong and 50 weak human sensitizers.**
- **When 500 μg/cm^2 is used as the threshold, there are 54 strong and 43 weak human sensitizers.**
- **See Table 6-3 of the draft ICCVAM BRD**

IV. What are the differences between the validation approach used by Basketter, Gerberick, and Kimber (BRD Appendix A) with the approach taken by ICCVAM.

NICEATM response:
The approach used by Basketter et al. focuses specifically on the correlation of LLNA EC3 values and human threshold values. The authors used HRIPTs obtained from Griem et al. (2003), Schneider and Akkan (2004), and Basketter et al. (2005). They do not select specific human threshold values against which to calculate over- or under-classification rates for LLNA EC3 versus human threshold values and therefore did not include substances that are considered non-sensitizers nor those that would be considered false negative or false positive based on LLNA results.

The ICCVAM evaluation uses NOELs and LOELs (DSA_{05} values reported by Schneider and Akkan were considered LOELs for the purposes of the ICCVAM evaluation) obtained from either HMT or HRIPT studies. The values used were those provided in the published and unpublished reports detailed under the response to question I. LOELs were divided by a safety factor of 10 (i.e., LOEL/10). In this evaluation, the availability of a LOEL and/or NOEL was used as the criteria for classifying a substance as a human sensitizer.

Two approaches were used to evaluate the ability of the LLNA to predict sensitization potency in humans.

In the first approach, for each substance classified as a sensitizer in both the LLNA and in humans, the LLNA EC3 concentration (expressed in µg/cm2 and not as a percent) was correlated against the human threshold response (i.e., either the NOEL or LOEL/10, expressed in µg/cm^2). This approach mimics that of the Basketter et al. submission described above.

In the second approach, using the same set of 81 sensitizers used in the first approach, the human sensitizers were classified into strong or weak based on using either of two proposed decision criteria (strong sensitizers <250 or <500 µg/cm^2). Next, the optimal EC3 value that maximized obtaining the correct skin sensitization calls for strong and weak sensitizers (using one or the other proposed decision criterion) was pragmatically determined and the correct classification rate as well as the over- and under-classification rates calculated.

In a variant of the second approach, substances that were classified in the LLNA as false positives (i.e., sensitizers in the LLNA but non-sensitizers in humans), false negatives (i.e., non-sensitizer in the LLNA but sensitizers in human tests), and non-sensitizers in both the LLNA and in human tests were included, the optimal EC3 values were re-calculated, and then the correct classification rate as well as the over- and under-classification rates re-calculated for each sensitization category (strong sensitizer, weak sensitizer, non-sensitizer).

In these analyses, for substances that had more than one EC3 or human threshold value, two methods for arriving at a single EC3 or threshold value were used. First, the most potent (i.e., the lowest) LLNA EC3 or human threshold concentration was used. Second, the geometric mean of all LLNA EC3 or human threshold concentrations was used. In the latter case, the HMT and the HRIPT were not classified as repeat tests for the same substance (i.e., geometric means were calculated only for repeat HMT or repeat HRIPT).

V. With regard to Table 6-2, please compare and contrast the approaches taken by the various investigators represented. That is, analyze the possible sources of variability in the various approaches.

VI. Note that ICCVAM presents the variability among EC_3 values for repeat LLNA tests. Can the panel estimate variability for human data points.

VII. When weighing evidence in human or animal data, what are the critical parameters to be considered?

References:

Api AM. 2007. The use of human data when conducting dermal sensitization quantitative risk assessments for fragrance ingredients. Research Institute for Fragrance Materials, Inc. Unpublished submission to NICEATM.

Basketter DA, Clapp C, Jefferies D, Safford RJ, Ryan CA, Gerberick GF, Dearman RJ, Kimber I. 2005. Predictive identification of human skin sensitisation thresholds. Contact Dermatitis: 53: 260 – 267.

Gerberick GF, Robinson MK, Ryan CA, Dearman RJ, Kimber I, Basketter DA, Wright Z, Marks JG. 2001. Contact allergenic potency: correlation of human and local lymph node assay data. American Journal of Contact Dermatitis 12: 156-161.

Griem P, Goebel C, Scheffler H. 2003. Proposal for a risk assessment methodology for skin sensitization based on sensitization potency data. Regulatory Toxicology and Pharmacology 38(3): 269-290.

Lalko J, Api AM. 2006. Investigation of the dermal sensitization potential of various essential oils in the local lymph node assay. Food and Chemical Toxicology 44(5):739-746.

Schneider K, Akkan Z. 2004. Quantitative relationship between the local lymph node assay and human skin sensitization assays. Regulatory Toxicology and Pharmacology 39: 245-255.