Comparison of the DPRA with a Three-Test Battery for *In Vitro* Evaluation of Skin Sensitization

W Stokes[1], T Burns[2], J Strickland[2], L Rinckel[2] and D Allen[2]
[1]NICEATM/NTP/HHS, RTP, NC, USA; [2]ILS, Inc., RTP, NC, USA

Abstract

To protect people from allergic contact dermatitis (ACD), regulatory agencies require that the results from standardized animal tests be used for hazard labeling. Such labeling warns consumers and workers of the precautions necessary to avoid exposures to substances that may cause ACD. International legislation to ban animal testing of cosmetics has spurred efforts to develop *in vitro* replacements for ACD hazard tests that use animals. NICEATM retrospectively evaluated the performance of the direct protein reactivity assay (DPRA) against that of testing strategies using three *in vitro* assays: DPRA, the human cell line activation test (h-CLAT), and KeratinoSens. The murine local lymph node assay was used as the reference test for a set of 67 unique substances. The DPRA alone generated an accuracy of 85% (57/67), a false positive rate of 22% (5/23), and a false negative rate of 11% (5/44). Using the most prevalent result for each substance from all three assays yielded an accuracy of 82% (55/67), a false positive rate of 30% (7/23), and a false negative rate of 11% (5/44). A classification tree model was also evaluated for predicting the LLNA results. A structural reactivity assessment was used to divide the 67 substances into positive and negative groups, then a recursive partitioning routine was used to generate further branches based on the *in vitro* test results. This strategy did not improve the performance of the three *in vitro* tests relative to the DPRA (accuracy = 79% [53/67]). However, based on the classification tree results, an interim testing strategy that combines the DPRA and the LLNA was proposed. This strategy could potentially reduce animal use for skin sensitization testing by up to 77% compared to testing all substances in the LLNA. ILS staff supported by NIEHS contract N01-ES-35504.

Introduction

- Allergic contact dermatitis (ACD) is a significant public health problem, resulting in over 7 million outpatient visits annually, causing lost workdays and adversely affecting quality of life (Bureau of Labor Statistics 2010; Hutchings et al. 2001; Skoet et al. 2003).
- National and international regulatory authorities require testing of chemicals and products to identify potential contact allergens, which must be labeled with a hazard description and precautions necessary to avoid or minimize exposure (EPA 2003; ISO 2008; OECD 2010).
- Over 3700 contact allergens have been identified worldwide (Beltrani et al. 2006).
- To reduce animal use in testing, more than one *in silico* or *in vitro* assay is likely to be required for an accurate prediction of skin sensitization potential (Basketter and Kimber 2009).
  - Various integrated decision strategies that combine *in silico* and *in vitro* test methods have been proposed (Grindon et al. 2006).
  - Past approaches to develop integrated decision strategies have included test batteries (Bauch et al. 2011), classification tree models (Gerberick et al. 2007; Grindon et al. 2007) and Bayesian networks (Jaworska et al. 2011).
- This study compares the ACD hazard results obtained using three individual *in vitro* assays to those obtained using various combinations of these tests and consideration of structural alerts for reactivity.

Methods

- The following three *in vitro* test methods were evaluated:
  - Direct peptide reactivity assay (DPRA)
  - KeratinoSens™ assay
  - Human cell line activation test (h-CLAT)
- Each *in vitro* test monitors a different event in the adverse outcome pathway for the skin sensitization process **(Figure 1.)**.
- A database of 67 substances for which results from each of these assays are published was analyzed **(Table 1)**.
  - If multiple test method results existed for a substance, the most prevalent result was used for that substance.
  - A substance with an equal number of positive and negative results was deemed positive.
- The murine local lymph node assay (LLNA) was used as the reference test method for calculating performance statistics.
- Published results from a structural assessment of reactivity (Safford et al. 2011) were also considered for each of the 67 substances.
  - NICEATM conducted reactivity assessments for substances that were not included in Safford et al. (2011) according to the rules in that publication.
- Three test strategies were evaluated as predictors of the LLNA result **(Figure 2)**:
  1. Testing all substances independently in each of the three *in vitro* methods.
  2. Testing all substances in a three-test battery, consisting of DPRA, KeratinoSens assay, and h-CLAT.
     - The final outcome for each substance was determined using the most prevalent result from the three tests.
  3. Testing all substances using a decision tree model.
     - The final outcome for each substance was determined by the result of the terminal node.
- Strategies were developed to achieve the best accuracy and lowest false positive and false negative rates.

Overview of Decision Tree for the *In Silico* and *In Vitro* Test Battery

- Results from the structural reactivity assessment (Safford et al. 2011) and the three *in vitro* test methods were used to build a decision tree to predict the LLNA results **(Figure 3)**.
  - First, the structural reactivity assessment was used to divide the 67 substances into two groups (i.e., branches): reactive (positive) or nonreactive (negative).
  - Subsequently, branches were based on an analysis using the recursive partitioning routine in JMP® statistical software (SAS Institute, Cary, NC, USA) employing only the three *in vitro* tests to predict the LLNA results.
    - The *in vitro* tests with the highest likelihood-ratio chi-square (G^2) at each node were used to make the subsequent branches.
    - Dividing groups into branches continued until all three *in vitro* tests had been used.

Table 1. Results of the LLNA and Three *In Vitro* Test Methods for 67 Substances

Chemical Name	CASRN	Rxn Mechanistic Domain[1]	LLNA	DPRA	h-CLAT	KeratinoSens
Abietic acid	514-10-3	NR	+	+	-	+
4-Allylanisole	140-67-0	Pro-MA	+	+	+	+
Aniline	62-53-3	NR	+	-	+	+
Benzoquinone	106-51-4	MA	+	+	+	+
Benzoyl peroxide	94-36-0	SN2	+	+	-	-
Benzylidene acetone	122-57-6	MA	+	+	+	+
2,3-Butanedione	431-03-8	SB	+	+	+	+
Butyl glycidyl ether	2426-08-6	SN2	+	+	-	+
Cinnamic aldehyde	104-55-2	MA	+	+	+	+
Cinnamyl alcohol	104-54-1	MA	+	+	+	+
Citral	5392-40-5	SB	+	+	+	+
Cobalt (II) salts	7646-79-9	Special case[2]	+	+	+	+
1,2-Dibromo-2,4-dicyanobutane	35691-65-7	MA	+	+	+	+
Diethyl maleate	141-05-9	MA	+	+	+	+
2,4-Dinitrochlorobenzene	97-00-7	SnAr	+	+	+	+
Ethylenediamine	107-15-3	SB	+	+	+	+
Ethylene glycol dimethacrylate	97-90-5	MA	+	+	+	+
Eugenol	97-53-0	MA	+	+	+	+
Farnesal	502-67-0	SB	+	+	-	+
Formaldehyde	50-00-0	SB	+	+	+	+
Glutaraldehyde	111-30-8	SB	+	+	+	+
Hexadecyltrimethylammonium bromide	104302-76-3	NR	+	-	-	-
Hexyl cinnamic aldehyde	101-86-0	MA	+	-	+	+
Hydroxycitronellal	107-75-5	SB	+	+	+	+
2-Hydroxyethyl acrylate	818-61-1	MA	+	+	+	+
Imidazolidinyl urea	39236-46-9	Acyl	+	+	+	+
Isoeugenol	97-54-1	MA	+	+	+	+
Kathon	2682-20-4/ 26172-55-4	Special case[2]	+	+	+	+
2-Mercaptobenzothiazole	149-30-4	Acyl	+	+	+	+
Methyl methacrylate	80-62-6	MA	+	+	+	-
Methyl 2-nonynoate	111-80-8	MA	+	+	+	+
4-Nitrobenzyl bromide	100-11-8	SN2	+	+	+	+
Oxazolone	15646-46-5	Acyl	+	+	+	+
Phenylacetaldehyde	127-78-1	SB	+	+	+	+
Phenyl benzoate	93-99-2	Acyl	+	+	+	-
4-Phenylenediamine	106-50-3	MA	+	+	+	+
Phenylpropionaldehyde	93-53-8	SB	+	+	+	+
Phthalic anhydride	85-44-9	Acyl	+	+	+	-
Propyl gallate	121-79-9	Pro-MA	+	+	+	+
Pyridine	110-86-1	NR	+	-	+	-
Sodium lauryl sulfate[3]	151-21-3	NR	+	+	-	-
Tetramethylthiuram disulfide	137-26-8	Special case[2]	+	+	+	+
Undecylenic acid	112-38-9	NR	+	+	+	+
Xylene	95-47-6	NR[4]	+	-	-	-
p-Aminobenzoic acid	150-13-0	MA[4]	-	-	-	-
1-Butanol	71-36-3	NR	-	-	+	-
Chlorobenzene	108-90-7	NR	-	-	+	-
Diethyl phthalate	84-66-2	NR	-	-	+	-
Ethyl vanillin	121-32-4	NR	-	+	-	+
Fumaric acid	110-17-8	NR[4]	-	+	-	-
Glycerol	56-81-5	NR	-	-	-	-
Hexane	110-54-3	NR	-	-	-	+
4-Hydroxybenzoic acid	99-96-7	NR	-	-	+	-
Isopropanol	67-63-0	NR	-	-	-	-
Lactic acid	598-82-3	NR	-	-	-	-
4-Methoxyacetophenone	100-06-1	NR	-	-	+	+
6-Methylcoumarin	92-48-8	MA	-	-	+	+
Methyl salicylate	119-36-8	Acyl	-	-	-	-
Nickel (II) salts[5]	7718-54-9	Special case[2,4]	-	+	+	+
Octanoic acid	124-07-2	NR	-	+	+	-
Propylene glycol	57-55-6	NR	-	-	-	-
Propylparaben	94-13-3	NR	-	-	+	+
Resorcinol	108-46-3	Pro-MA	-	-	+	-
Salicylic acid	69-72-7	NR	-	-	+	-
Sulfanilamide	63-74-1	NR	-	-	-	-
Tartaric acid	87-69-4	NR	-	-	-	-
Vanillin	121-33-5	NR	-	+	-	+

Abbreviations: + = positive, - = negative; Acyl = acylating agent; DPRA = direct protein reactivity assay; h-CLAT = human cell line activation test; LLNA = murine local lymph node assay; MA = Michael addition: NR = non-reactive; Rxn = reaction; SB = Shiff base former; SN2 = SN2 electrophile; SNAr = SNAr electrophile

[1] Reaction mechanisms from Safford et al. (2011).
[2] Special cases are (1) substances that can be classified into one of the reaction domains but require further comment or (2) reactive chemicals that do not fit any of the domains.
[3] Sodium lauryl sulfate is a known false positive in the LLNA, compared to human results.
[4] Reaction mechanisms assigned by NICEATM according to criteria in Safford et al. (2011).
[5] Nickel is a known false negative in the LLNA, compared to human results.

Figure 1. Relationship of Test Methods to Key Events in Skin Sensitization

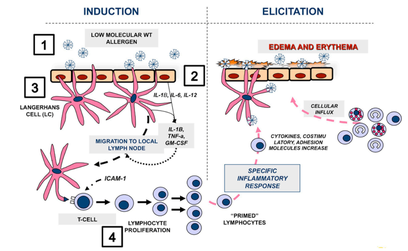


Illustration by D. Sailstad

1. Haptenation: attachment of allergen to skin protein
   - DPRA. HPLC monitors depletion of a nucleophile-containing synthetic peptides (cysteine and lysine).
   - KeratinoSens. A luminometer monitors induction of the Keap1-Nrf2-ARE regulatory pathway, in a cell line based on the human HaCaT keratinocyte cell line (Boukamp et al. 1988) containing a luciferin/luciferase reporter construct with a single copy of the ARE-element of the human AKR1C2 gene**.** The Keap1-Nrf2-ARE regulatory pathway appears to be induced by many skin sensitizers.
2. Epidermal inflammation: release of pro-inflammatory signals by epidermal keratinocytes
3. Dendritic cell (DC) activation and maturation
   - h-CLAT. Flow cytometry monitors induction of two protein markers (CD54 and CD86) on the surface of a human monocytic leukemia cell line.
4. T-cell proliferation: clonal expansion of hapten-peptide specific T-cells
   - LLNA. Scintillation counting monitors the content of ^{3}H-thymidine in mouse lymph node cells as an indicator of lymphocyte proliferation.

Figure 2. Testing Strategies Considered

Strategy 1: Individual Tests

Test substances in DPRA, h-CLAT or KeratinoSens alone

Strategy 2: Test Battery

Test substances in three-test battery (DPRA, h-CLAT & KeratinoSens)

Strategy 3: Decision Tree Model

Assess reactivity of substances per Safford et al. 2011 → *In vitro* tests with the highest likelihood ratio (G^2) at each node determine the subsequent branches → Continue branches until all *in vitro* tests have been used

Table 2. Performance Characteristics of *In Vitro* Test Methods and the 3-test Battery With the LLNA As the Reference Test

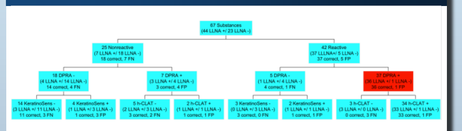
Model	Accuracy	False Positive Rate	False Negative Rate	Sensitivity	Specificity
DPRA[1]	85% (57/67)	22% (5/23)	11% (5/44)	89% (39/44)	78% (18/23)
h-CLAT[2]	73% (49/67)	48% (11/23)	16% (11/23)	84% (37/44)	52% (12/23)
KeratinoSens[3]	78% (52/67)	30% (7/23)	18% (8/44)	82% (36/44)	70% (16/23)
3-test Battery	82% (55/67)	30% (7/23)	11% (5/44)	89% (39/44)	69% (16/23)

Abbreviations: DPRA = direct protein reactivity assay, h-CLAT = human cell line activation test
[1] DPRA data was taken from Bauch et al. (2011), Gerberick et al. (2004), Gerberick et al. (2007), Natsch and Gfeller (2008) and Ryan et al. (2011).
[2] h-CLAT data was taken from Bauch et al. (2011) and Nukuda et al. (2010).
[3] KeratinoSens data was taken from Bauch et al. (2011) and Emter et al. (2010).

Results

- Among the three individual tests, the DPRA (accuracy = 85% [57/67]) had the highest accuracy, followed by KeratinoSens (accuracy = 78% [52/67]) and h-CLAT (accuracy = 73% [49/67]).
- The DPRA alone also had a higher accuracy than the test battery composed of all three assays (accuracy = 82% [55/67]) **(Table 2)**.
  - The accuracy of the test battery was lower than the DPRA alone because the battery resulted in two additional false positive outcomes.
- A decision tree strategy with an initial structural assessment of reactivity (accuracy = 79% [53/67]) did not improve the performance of the three *in vitro* tests relative to the DPRA **(Table 3)**.
  - The accuracy of the decision tree was improved if the reactive DPRA-positive substances were considered as sensitizers and not tested in the h-CLAT (accuracy = 84% [56/67]).

Figure 3. Decision Tree for Classification of LLNA Sensitizers and Nonsensitizers



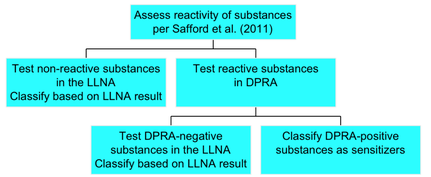
Abbreviations: + = positive; - = negative; DPRA = direct peptide reactivity assay; FN = false negative; FP = false positive; LLNA = murine local lymph node assay; h-CLAT = human cell line activation test.

Table 3. Performance of the Decision Tree Models

Model	Accuracy	False Positive Rate	False Negative Rate	Sensitivity	Specificity
Full decision tree	79% (53/67)	26% (6/23)	18% (8/44)	82% (36/44)	74% (17/23)
Decision tree without h-CLAT for testing the reactive DPRA positives	84% (56/67)	26% (6/23)	11% (5/44)	89% (39/44)	74% (17/23)
Decision tree without h-CLAT or KeratinoSens (reactivity assessment [1] and DPRA only)	85% (57/67)	22% (5/23)	11% (5/44)	89% (39/44)	78% (18/23)

[1] The reactivity assessment was per Safford et al. (2011).

Figure 4. Proposed Interim Testing Strategy Using DPRA and the LLNA to Reduce the Number of Animals Used for Testing



Proposed Interim Testing Strategy

- Until more accurate integrated decision strategies are developed, the following interim integrated testing strategy is proposed. Based on the database evaluated, this strategy minimizes animal use, avoids false negatives, and maximizes testing efficiency **(Figure 4)**.
  1. Assess structural reactivity per Safford et al. (2011).
  2. Test all reactive substances in DPRA.
  3. Classify reactive DPRA-positive substances as sensitizers (positive predictivity = 97% [35/36]).
  4. Test all nonreactive and DPRA-negative substances in the reduced LLNA (rLLNA). This is necessary to detect substances that would otherwise be classified as false negatives if only evaluated using the reactivity and DPRA tests. (Note that the rLLNA uses 40% fewer mice than the LLNA [ICCVAM 2009]).

Conclusions

- The proposed interim integrated testing strategy provides time and cost efficient accurate hazard classification while significantly reducing animal use.
  - No false negatives were produced
  - The only false positive was a nickel salt; however, nickel is a well-known sensitizer in humans.
  - This strategy could potentially reduce animal use for skin sensitization testing by up to 77% compared to testing all substances in the LLNA.
    - 55% (37/67) of the substances are classified without animal testing.
    - 45% of the substances are classified using the rLLNA.
- An example of the animal savings by using this strategy follows:
  - Testing all 67 substances using the LLNA uses:
    67 substances x 20 animals/substance = 1540 animals
  - Using the proposed interim strategy, 30 substances would be tested in the rLLNA:
    30 substances x 12 animals/substance = 360 animals
    1540 – 360 =1180 animals saved by using the interim testing strategy
- DPRA testing alone was more accurate than testing in KeratinoSens or h-CLAT, and more accurate than a 3-test battery consisting of DPRA, KeratinoSens, and h-CLAT.
- A decision tree approach incorporating an initial reactivity assessment and incorporating DPRA, KeratinoSens, and h-CLAT did not provide better performance than DPRA alone.
- Future improved integrated testing and decision strategies that incorporate methods that also identify sensitizing substances that are pro-haptens and pre-haptens can be expected to further reduce the number of substances requiring animal testing for accurate hazard classification.

References

Basketter D, Kimber I. 2009. J Appl Toxicol 29(6): 545-550.

Bauch C, Eltze T, Fabian E, Kolle S, Pachel C, Ramirez T, Wiench B, Wruck C, van Ravenzwaay B, Landsiedel R. 2011. Poster presented at 8th World Congress on Alternatives and Animal Use in the Life Sciences, August 21-25, Montreal, Canada.

Beltrani VS, Bernstein IL, Cohen DE, Fornacier L. Ann Allergy Asthma Immunol 97 (Suppl 2).

Bureau of Labor Statistics. 2010. Injuries, Illnesses, and Fatalities homepage. Available: http://www.bls.gov/IIF [accessed 15 February 2011].

Emter R, Ellis G, Natsch A. 2010. Toxicol Appl Pharmacol. 245: 281-290.

EPA. 2003. Health Effects Test Guidelines: OPPTS 870.2600 - Skin Sensitization. EPA 712–C–03–197 . Washington, DC:U.S. Environmental Protection Agency. Available: http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series870.htm

Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP. 2004. Toxicol Sci 81(2): 332-343.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP. 2007. Toxicol Sci 97(2): 417-427.

Grindon C, Combes R, Cronin MTD, Roberts DW, Garrod JF. 2006. Altern Lab Anim 34(4): 407-427.

Grindon C, Combes R, Cronin MTD, Roberts DW, Garrod JF. 2007. Altern Lab Anim 35(6): 683-697.

Jaworska J, Harol A, Kern PS, Gerberick GF. 2011. ALTEX 28(3): 211-225.

Hutchings CV, Shum KW, Gawkrodger DJ. 2001. Contact Dermatitis. 45: 17-20.

ICCVAM. 2009. NIH Publication Number 09-6439.

ISO Vol. ISO/DIS 10993-10, 2008.

Natsch A, Gfeller H. 2008. Toxicol Sci 106(2): 464-478.

Nukada Y, Ashikaga T, Sakaguchi H, Sono S, Mugita N, Hirota M, et al. 2011. Contact Dermatitis 65(6): 343-353.

OECD. 2010. Test No. 429. OECD Publishing, Paris.

Ryan C, Jaworska J, Kern P, Chaney JG, Gerberick GF. 2011. Poster presented at Society of Toxicology Annual Meeting, March 6–10, Washington, DC.

Safford RJ, Aptula AO, Gilmour N. 2011. Regul Toxicol Pharmacol 60: 218-224.

Skoet R, Zachariae R, Agner T. 2003. Br J Dermatol. 149, 452-456.

Acknowledgements

The Intramural Research Program Division of the National Toxicology Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS, Inc., under NIEHS contract N01-ES 35504.

This poster reflects the views of the authors. The views expressed above have not been reviewed or approved by any U.S. Federal agency and do not necessarily represent the official positions of any U.S. Federal agency.

Because the poster was written as part of the official duties of the authors, it can be freely copied.

The authors thank Mr. Frank Stack of ILS, Inc., for helpful discussions regarding this poster.

Posters and abstracts presented by NICEATM–ICCVAM at SOT 2012 are available on the NICEATM–ICCVAM website at http://iccvam.niehs.nih.gov/.