

***In Silico* Predictions of Skin Sensitization Using OECD QSAR Toolbox**

J Strickland¹, N Choksi¹, D Allen¹, W Casey²

¹ILS/NICEATM, RTP, NC, USA; ²NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

Abstract

Skin sensitization data are needed to develop precautionary labeling to protect workers and consumers from chemical exposures. To reduce or eliminate the use of animals for skin sensitization testing, a number of *in vitro* and *in silico* test methods have been proposed. NICEATM evaluated the utility of the OECD QSAR Toolbox for making read-across skin sensitization predictions using murine local lymph node assay (LLNA) outcomes as reference data. The Toolbox protocol identified analogs for 120 target substances (87 sensitizers and 33 nonsensitizers) using mechanism of protein binding and chemical structure schemes in the Toolbox. If protein binding alerts were not identified in a substance, auto-oxidation and skin metabolism products were predicted; a representative product with a protein binding alert was used in the evaluation. If neither parent nor products had protein binding alerts, the substance was classified as a nonsensitizer. For parent or products with protein binding alerts, *in vivo* skin sensitization data for analogs were used to predict the sensitization potential. Accuracy of the Toolbox protocol was 77% (92/120) with sensitivity = 77% (67/87) and specificity = 76% (25/33). Using only protein binding alerts in the parent compound to predict sensitization potential yielded accuracy = 69% (83/120), sensitivity = 66% (57/87), and specificity = 79% (26/33). Using only protein binding alerts in the parent or product to classify substances as sensitizers improved accuracy (82% [98/120]) and sensitivity (91% [79/87]) compared to the Toolbox protocol, but decreased specificity (58% [19/33]). Thus, potential skin sensitizers may be predicted with similar accuracy using either the Toolbox protocol or only protein binding alerts. Because the Toolbox protocol had a lower false positive rate, it will be evaluated as part of an integrated decision strategy for skin sensitization that includes *in vitro* data and physicochemical parameters.

Introduction

- Allergic contact dermatitis (ACD) is a skin reaction, characterized by localized redness, swelling, blistering, or itching (**Figure 1**), that can develop after repeated direct contact with a skin sensitizer. Skin sensitizers include commonly used substances such as neomycin and formaldehyde.
- National and international regulatory authorities require testing of pesticides, personal care products, and other products to assess their potential to cause ACD. The results of these tests are used to determine appropriate labeling to ensure safe use and handling.
- During the past 15 years, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) have evaluated a wide range of *in vivo*, *in vitro*, and *in silico* approaches to identify potential skin sensitizers.
- This poster evaluates the use of QSAR Toolbox v3.2, a freely available software package, for predicting skin sensitization hazard without using animals.

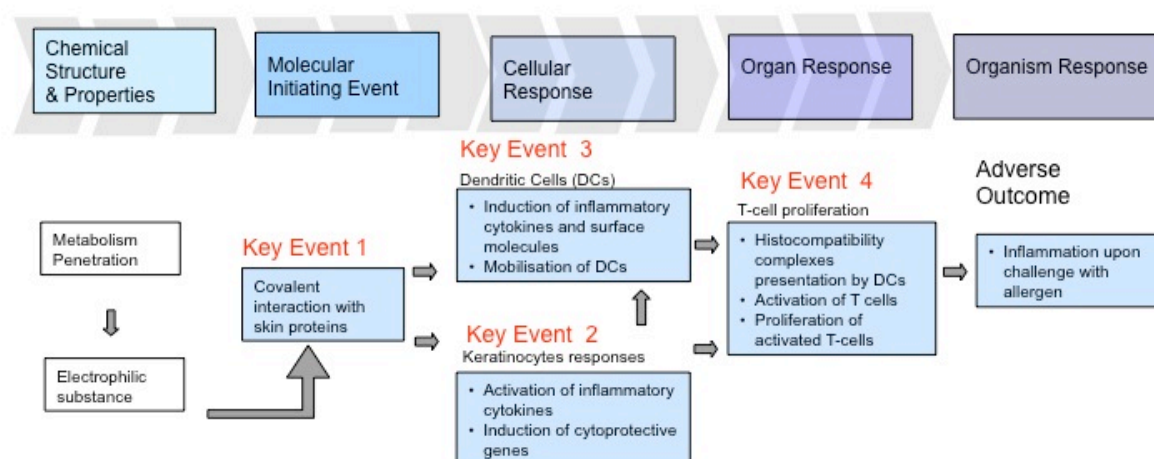
Figure 1 Skin Sensitization Reaction



The Skin Sensitization Adverse Outcome Pathway

- Although the development of skin sensitization is a complex process, a well-defined adverse outcome pathway (AOP) has been developed for substances that produce sensitization by covalently binding to proteins (OECD 2012) (**Figure 2**).

Figure 2 Adverse Outcome Pathway for Skin Sensitization



QSAR Toolbox

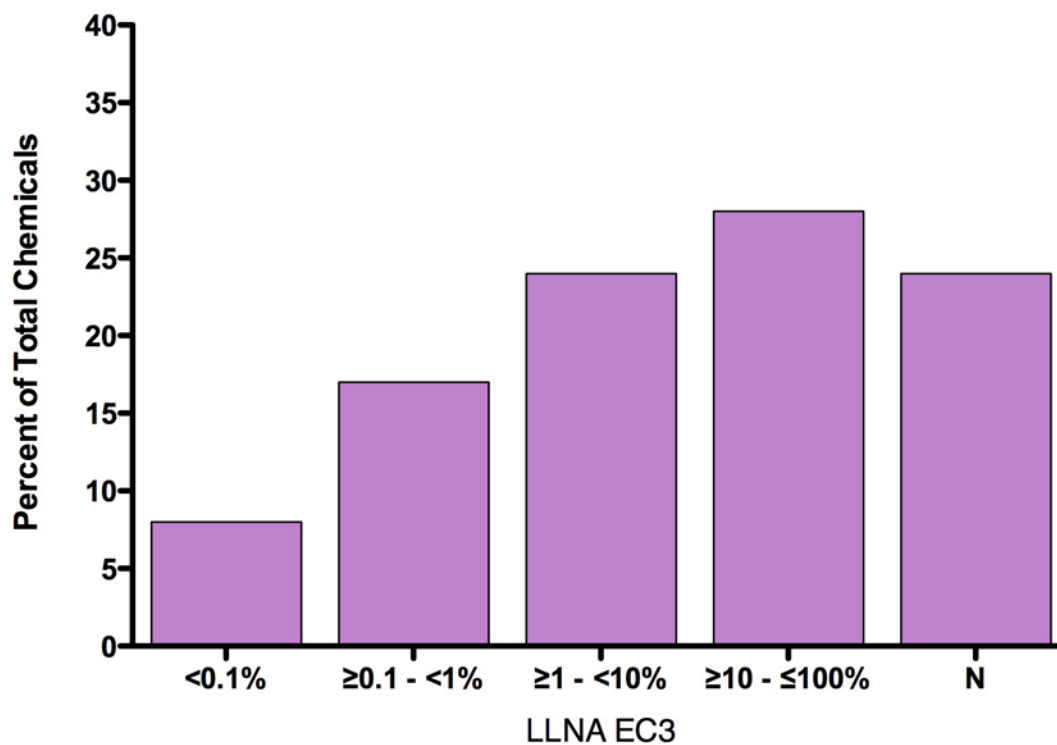
- The Organisation for Economic Co-operation and Development (OECD) developed the QSAR Toolbox software to make quantitative structure–activity relationship (QSAR) technology readily accessible, more transparent, and less costly (OECD 2014), thereby increasing regulatory acceptance of QSAR analyses.
- The QSAR Toolbox software is freely available at <http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm>
- QSAR Toolbox can be used to evaluate chemicals, including metals and the full range of organic functional groups and protein binding mechanisms, that are relevant to skin sensitization (ECHA 2014).
- QSAR Toolbox allows considerable user control over the prediction and output by including in its workflow the opportunity to
 - Identify relevant target chemical structural characteristics and potential mechanism or mode of action

- Identify other chemicals that have the same structural characteristics and/or mechanism or mode of action and existing data of interest (e.g., skin sensitization test results)
 - Use existing experimental data to fill any data gaps
- In this evaluation, QSAR Toolbox was used to predict skin sensitization via two methods:
 - **Method 1** assessed the ability of a substance to produce Key Event 1, the molecular initiating event in the AOP (**Figure 2**), by reporting the presence of protein binding alerts using two approaches:
 - **Approach 1** reported the presence of protein binding alerts only for the target substance.
 - **Approach 2** reported the presence of protein binding alerts for the target substance or its auto-oxidation or skin metabolism products
 - **Method 2** assessed the ability of a substance to produce Key Event 4 (a positive result in the murine local lymph node assay [LLNA]) and the adverse outcome in the AOP (a skin sensitization reaction in guinea pig and human tests) (**Figure 2**) by using a read-across prediction to classify substances as sensitizers or nonsensitizers

Database

- Of the 120 substances included in this evaluation, 73% (87/120) were classified as sensitizers by the LLNA and 27% (33/120) were classified as nonsensitizers. The distribution of LLNA potency for the evaluated substances is shown in **Figure 3**.

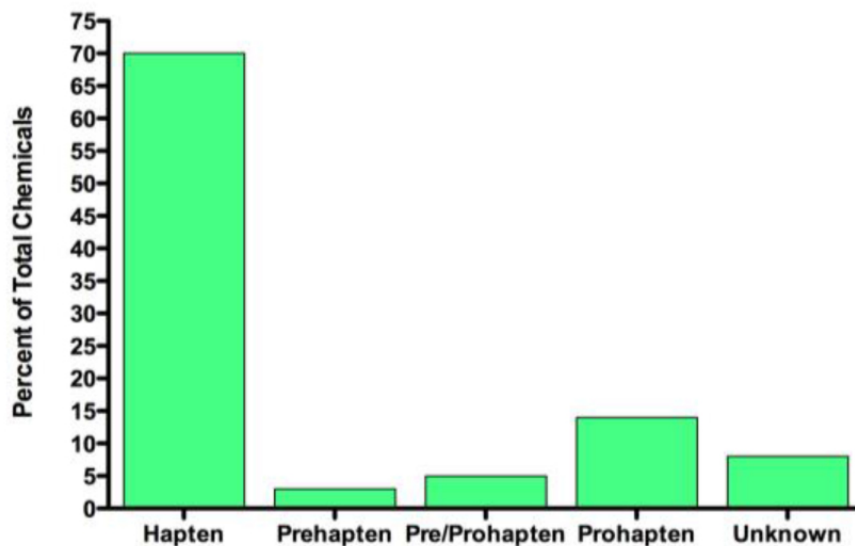
Figure 3 **Distribution of LLNA EC3 Values for 120 Substances**



Abbreviations: EC3 = estimated effective concentration that produces a stimulation index of 3, the threshold for a positive response in the LLNA; LLNA = murine local lymph node assay; N = nonsensitizers.

- Haptens are substances with the potential to induce skin sensitization without modification or metabolism. Prehaptens require oxidation in order to elicit a skin sensitization reaction and prohaptens require metabolism. The distribution of prehaptens and prohaptens among the 87 sensitizers in the 120-substance database is shown in **Figure 4**.

Figure 4 **Distribution of Prehaptens and Prohaptens for 87 LLNA Sensitizers**



Abbreviation: LLNA = murine local lymph node assay.

- The distribution of substances among product categories is shown in **Table 1**.

Table 1 Distribution of 120 Substances in the Database Among Product Categories

Product Category	Percentage of Substances ^a	Number of Substances
Manufacturing	49%	59
Food additive	37%	44
Pharmaceutical	28%	33
Intermediate in chemical synthesis	25%	30
Fragrance agent	20%	24
Pesticide (other) ^b	18%	21
Personal care product	17%	20
Cosmetic	16%	19
Pesticide (antimicrobial) ^b	15%	18
Solvent	7%	8
Household product	6%	7
Plastic	2%	2
Rubber	1%	1
Antioxidant	1%	1

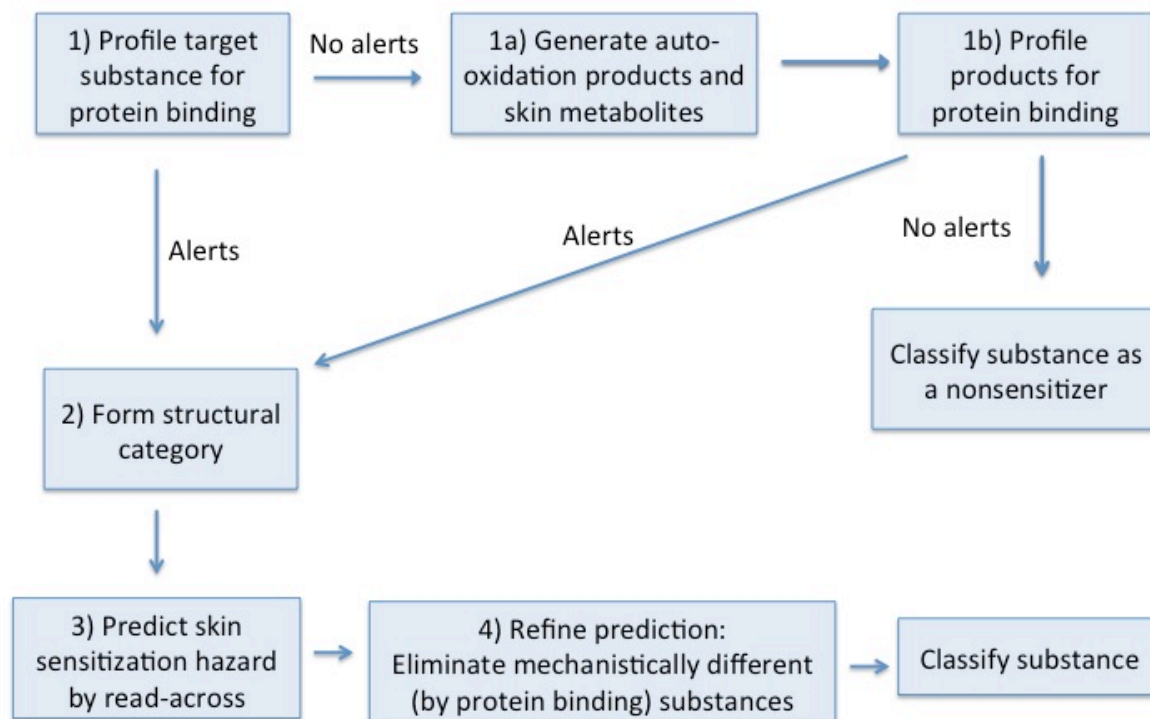
^a Total of all percentages exceeds 100% because most substances were associated with more than one product category.

^b Twenty-six of the 39 pesticides (antimicrobials plus other pesticides) are currently registered with the U.S. Environmental Protection Agency.

Protocol

- Training materials for QSAR Toolbox v3.2 were used to develop a read-across protocol to predict the skin sensitization hazard of 120 substances.

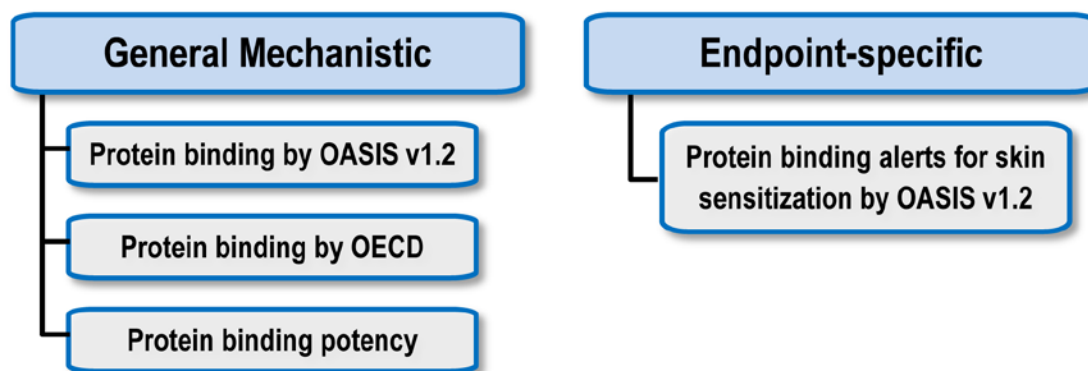
Figure 5 **QSAR Toolbox Protocol for Skin Sensitization Predictions**



-
- In the first step of the protocol, QSAR Toolbox assesses the likelihood that the target substance will bind to protein using four structural feature recognition (“alert”) systems for protein binding (**Figure 5, Step 1**).
 - QSAR Toolbox recognizes structural features that have been identified as binding to skin proteins using four alert systems, termed “profilers.” Three of the profilers are general mechanistic profilers and one is an endpoint-specific (i.e., skin sensitization) mechanistic profiler (**Figure 6**).
 - **Protein binding by OASIS v1.2** identifies whether a substance contains any of 101 structural features (alerts) in 11 mechanistic domains that are responsible for the interaction with proteins.
 - **Protein binding by OECD** determines whether a substance binds to proteins using a scheme that has five overarching mechanistic domains related to structural alerts grouped by the presence of a common reactivity site into mechanistic alerts. This profiler has 102 categories of protein binding that includes 16 mechanistic

- alerts (e.g., ring-opening acylation) that cover 52 structural alerts (e.g., alpha-lactams).
- **Protein binding potency** assesses the likelihood that the target substance will covalently bind with the thiol group of glutathione. The profiler contains 49 Michael addition and 46 bimolecular nucleophilic substitution (SN2) categories. The 95 structural alerts are separated into five potency categories of reactivity: extremely, highly, moderately, slightly, and suspect.
 - **Protein binding alerts for skin sensitization by OASIS v1.2** is the endpoint-specific profiler that evaluates whether a substance binds to proteins. Compared to the protein binding by OASIS v1.2 profiler (discussed in the first bullet above), this profiler accounts for the inhibition of protein binding in the skin, produced by electronic and steric factors, for substances with protein binding alerts. It uses 100 structural alerts that have been categorized into 11 mechanistic domains. Each mechanistic domain has more than two mechanistic alerts.
- The protocol for **Method 1, Approach 1** ends here.
- Substances with protein binding alerts were classified as sensitizers.
 - Substances without protein binding alerts were classified as nonsensitizers.

Figure 6 **QSAR Toolbox Protein Binding Profilers**



- If QSAR Toolbox identified no protein binding alerts for the target substance, a further assessment was conducted to determine if the substance might be a pre- or prohaptan. The autoxidation and skin metabolism modules were used to generate oxidation/metabolism products (**Figure 5, Step 1a**). These products were then profiled as above for protein binding alerts. If none were identified, the target substance was classified to be a nonsensitizer (**Figure 5, Step 1b**).
- The protocol for **Method 1, Approach 2** ends here.

- Substances with products that had protein binding alerts were classified as sensitizers
 - Substances that had products without protein binding alerts were classified as nonsensitizers.
- If protein binding alerts were identified in parent substances or products, QSAR Toolbox formed a category of similar substances with *in vivo* skin sensitization data based on structural similarity (**Figure 5, Step 2**).
 - A read-across algorithm was then applied to the *in vivo* skin sensitization data (mouse, guinea pig, or human) of the category members to predict the skin sensitization hazard of the target substance (**Figure 5, Step 3**). The read-across algorithm uses the skin sensitization outcome that appears most often for the five nearest neighbors, based on log K_{ow} , to predict the skin sensitization hazard of the target substance.
 - Substances with dissimilar mechanisms of protein binding (compared with the target substance) were then eliminated to refine the skin sensitization hazard prediction (**Figure 5, Step 4**).
 - The protocol for **Method 2** ends here with the read-across prediction.
 - All skin sensitization predictions were evaluated for concordance with LLNA outcomes.

Results

- The performance statistics for the three methods are shown in **Table 2**.

Table 2 Performance Statistics with Respect to LLNA Outcomes

Method	Accuracy	Sensitivity	False Negative Rate	Specificity	False Positive Rate
Method 1, Approach 1^a	69% (83/120)	66% (57/87)	35% (30/87)	79% (26/33)	21% (14/33)
Method 1, Approach 2^b	82% (98/120)	91% (79/87)	9% (8/87)	58% (19/33)	42% (14/33)
Method 2^c	77% (92/120)	77% (67/87)	23% (20/87)	76% (25/33)	24% (8/33)

Abbreviation: LLNA = murine local lymph node assay.

^a **Method 1, Approach 1** used only the presence of protein binding alerts in the target substance to predict sensitization potential.

^b **Method 1, Approach 2** used the presence of protein binding alerts in either the target compound or auto-oxidation or skin metabolism products to predict sensitization potential.

^c **Method 2** used *in vivo* skin sensitization data in a read-across algorithm. Read-across predictions for five substances (4-nitrobenzyl bromide, ethylene glycol dimethacrylate, farnesal, glyoxal, and phenylacetaldehyde) were unreliable because the log K_{ow} of each target substance was outside the range of the nearest neighbors. However, the predictions were concordant with LLNA outcomes.

Misclassified Substances for Method 2 Read-across

- The QSAR Toolbox read-across protocol (**Method 2**) misclassified 28 substances.
- **Table 3** lists characteristics of the 20 false negatives (LLNA sensitizers that were misclassified as nonsensitizers).
 - Nine of the false negatives were weak LLNA sensitizers ($EC3 > 10\%$), two were moderate LLNA sensitizers ($1\% < EC3 < 10\%$), and nine were strong sensitizers ($EC3 < 1\%$).
 - Four false negatives were prohaptens and one was a pre/prohaptent. Of these five substances, three were strong LLNA sensitizers.
 - 60% (12/20) of the false negatives had no protein binding alerts for the parent substances, and 40% (8/20) had no protein binding alerts for either the parent or an auto-oxidation product or skin metabolite.
 - Human data were available for 16 of the false negatives; 11 substances are human skin sensitizers. Thus, five substances are false positive in the LLNA (i.e., the read-across protocol may be more relevant to human sensitization).

Table 3 Twenty Substances with False Negative QSAR Toolbox Read-across Predictions (Method 2)

Chemical Name	CASRN	Molecular Weight (g/mol)	Pre/Pro-hapten?	Parent Protein Binding Alert	Any Protein Binding Alert	Geometric Mean EC3 (%)	LLNA Result ^a	Toolbox Read-across Result	Human Result ^b
Pyridine	110-86-1	79.10	Pro	No	Yes	72	POS	NEG	NEG
Nonanoic acid	112-05-0	158.24	N	No	No	35	POS	NEG	NA
Undecylenic acid	112-38-9	184.28	N	No	No	19.4	POS	NEG	NA
Lauryl gallate	1166-52-5	338.44	Pre/Pro	No	Yes	0.3	POS	NEG	POS
Propyl gallate	121-79-9	212.20	N	No	Yes	0.32	POS	NEG	POS
Benzylidene acetone	122-57-6	146.19	N	Yes	Yes	3.7	POS	NEG	POS
Xylene	1330-20-7	106.08	N	No	No	95.8	POS	NEG	NEG
Oxalic acid	144-62-7	90.05	N	No	No	0.10	POS	NEG	NA
Sodium lauryl sulfate	151-21-3	288.38	N	No	No	3.8	POS	NEG	NEG
2,4,6-Trinitrobenzenesulfonic acid	2508-19-2	293.17	N	No	No	0.26	POS	NEG	NA
Methylisothiazolinone	2682-20-4	115.15	N	Yes	Yes	0.90	POS	NEG	POS
1,2-Dibromo-2,4-dicyanobutane	35691-65-7	265.93	Pro	Yes	Yes	0.9	POS	NEG	POS
Imidazolidinyl urea	39236-46-9	388.29	N	Yes	Yes	24	POS	NEG	POS
Abietic acid	514-10-3	302.46	Pro	No	Yes	15	POS	NEG	POS
Penicillin G	61-33-6	334.39	N	Yes	Yes	17.4	POS	NEG	POS
Salicylic acid	69-72-7	138.12	N	No	No	12.2	POS	NEG	NEG

Chemical Name	CASRN	Molecular Weight (g/mol)	Pre/Pro-hapten?	Parent Protein Binding Alert	Any Protein Binding Alert	Geometric Mean EC3 (%)	LLNA Result ^a	Toolbox Read-across Result	Human Result ^b
Benzalkonium chloride	8001-54-5	339.26	N	No	No	0.1	POS	NEG	NEG
2-Hydroxyethyl acrylate	818-61-1	116.12	N	Yes	Yes	0.37	POS	NEG	POS
Coumarin	91-64-5	146.14	N	Yes	Yes	29.6	POS	NEG	POS
Benzoyl peroxide	94-36-0	242.23	Pro	Yes	Yes	0.01	POS	NEG	POS

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; EC3 = estimated effective concentration that produces a stimulation index of 3, the threshold for a positive response in the LLNA; LLNA = murine local lymph node assay; NA = not available; N = hapten; NEG = negative; POS = positive; Pro = prohaptent; Pre/Pro = pre/prohaptent.

^a From the NICEATM LLNA database at <http://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/immunotoxicity/nonanimal/index.html>

^b From ICCVAM (2011) or Basketter et al. (2014).

- **Table 4** shows the eight LLNA nonsensitizers that were misclassified as sensitizers (false positives) by the read-across protocol (**Method 2**).
 - Only 38% (3/8) of the false positives had protein binding alerts for the parent substances, but 100% (8/8) had protein binding alerts for the parent or an auto-oxidation product or skin metabolite.
 - Two of the false positives had discordant results in multiple LLNA tests, although the majorities were negative.
 - Aniline had 4/9 positive tests.
 - Streptomycin had 1/5 positive tests.
 - Human data were available for 6/8 false positives. Of these, 3 substances are human skin sensitizers.

Table 4 **Eight Substances with False Positive Toolbox Read-across Predictions (Method 2)**

Chemical Name	CASRN	Molecular Weight (g/mol)	Parent Protein Binding Alert	Any Protein Binding Alert	LLNA Result ^a	Toolbox Read-across Result	Human Result ^b
4-Methoxyacetophenone	100-06-1	150.17	No	Yes	NEG	POS	NEG
Ethyl vanillin	121-32-4	166.18	No	Yes	NEG	POS	NEG
Vanillin	121-33-5	152.15	No	Yes	NEG	POS	NEG
3-Phenoxypropiononitrile	3055-86-5	147.18	No	Yes	NEG	POS	NA
Streptomycin sulfate	3810-74-0	581.57	Yes	Yes	NEG	POS	POS
Aniline	62-53-3	93.13	No	Yes	NEG	POS	POS
Saccharin	81-07-2	183.18	Yes	Yes	NEG	POS	NA
2-Hydroxypropyl methacrylate	923-26-2	144.17	Yes	Yes	NEG	POS	POS

Abbreviations: CASRN = Chemical Abstracts Service Registry Number; LLNA = murine local lymph node assay; NA = not available; NEG = negative; POS = positive.

^a From the NICEATM LLNA database at <http://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/immunotoxicity/nonanimal/index.html>

^b From ICCVAM (2011) or Basketter et al. (2014).

Conclusions

- For the database used in this evaluation, using protein binding alerts to predict LLNA outcome yielded the best performance when alerts for both the target substance and its auto-oxidation products or skin metabolites were considered.
 - When using alerts for only the target substances (**Method 1, Approach 1**), accuracy was 69%, sensitivity was 66%, and specificity was 79%.
 - When using alerts for target substances or auto-oxidation products or metabolites (**Method 1, Approach 2**), accuracy increased to 82% and sensitivity increased to 91%, but specificity decreased to 58%.
- Using the read-across protocol to predict LLNA outcome (**Method 2**) provided a more balanced evaluation of positives and negatives.
 - With this approach, accuracy was 77%, sensitivity was 77% and specificity was 76%.
- The results from the read-across protocol were used in an evaluation of an integrated decision strategy for skin sensitization that includes *in vitro* data and physicochemical parameters (see **Matheson et al. Abstract 421, Poster Board 108** in this session)

References

Basketter DA, Alépée N, Ashikaga T, Barroso J, Gilmour N. et al. 2014. Categorization of chemicals according to their relative human skin sensitizing potency. *Dermatitis* 25:11-21.

ECHA. 2014. Illustrative examples with the OECD QSAR Toolbox workflow. Part 1: Introductory note. [Internet]. Helsinki, Finland:European Chemical Agency. Available: http://echa.europa.eu/documents/10162/21655633/illustrative_example_qsar_part1_en.pdf

ICCVAM. 2011. ICCVAM Test Method Evaluation Report: Usefulness and Limitations of the Murine Local Lymph Node Assay for Potency Categorization of Chemicals Causing Allergic Contact Dermatitis in Humans. NIH Publication No. 11-7709. Research Triangle Park, NC: National Institute of Environmental Health Sciences.

OECD. 2012. OECD Series on Testing and Assessment No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Assessment. Paris:OECD Publishing.

OECD. 2014. The OECD QSAR Toolbox [Internet]. OECD Publishing. Available: <http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm>.

Acknowledgements

The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS under NIEHS contract HHSN27320140003C.

The views expressed above do not necessarily represent the official positions of any Federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied.



A summary of NICEATM activities at the 2015 SOT Annual Meeting is available on the National Toxicology Program website at <http://ntp.niehs.nih.gov/go/742110>.