National Institute of Environmental Health Sciences
Division of Translational Toxicology

Testing of Regulatory-Relevant Chemicals for Skin and Respiratory Sensitization Hazard

V.J. Johnson[1], E.N. Reinke[2], T. LaPratt[2], J. Strickland[2], K.T. To[2], O. Larne[3], H. Johansson[3], A. Forreryd[3], D.G. Allen[2*], N.C. Kleinstreuer[4], D. Germolec[4]

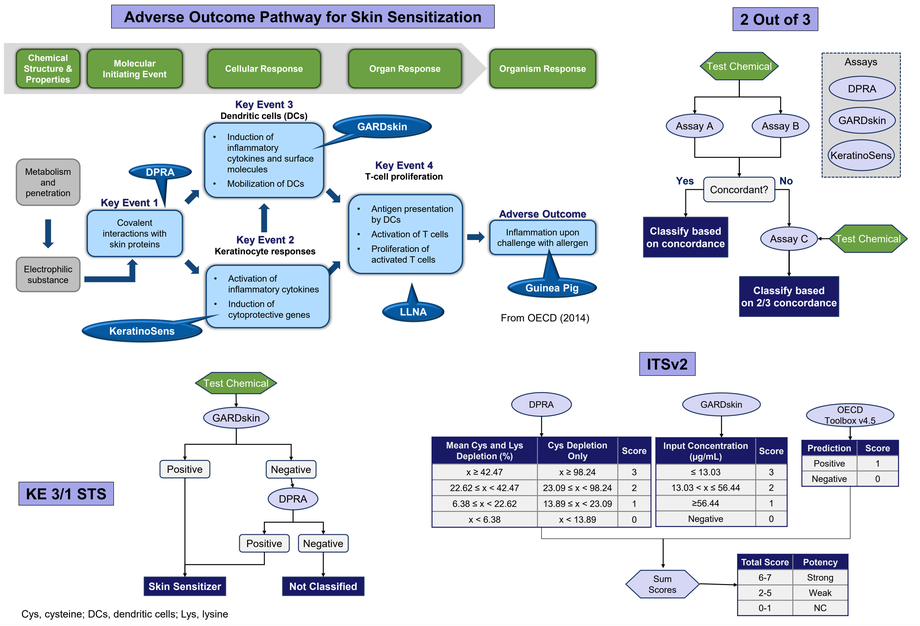
[1]Burleson Research Technologies, Inc., Morrisville, NC; [2]Inotiv, RTP, NC; [3]SenzaGen AB, Lund, Sweden; [4]NIH/NIEHS/DTT/STB/NICEATM, RTP, NC

Abstract 4189
Poster P738

Introduction

- Multiple U.S. federal agencies require the assessment of skin or respiratory sensitization potential for their chemical evaluation and management programs.
- Although several non-animal new approach methodologies (NAMs) have been internationally adopted as test guidelines, no single NAM is considered to be a complete replacement for animal tests for skin sensitization testing. To improve the performance, results from several NAMs can be combined in defined approaches (DAs) that inform on multiple key events in the adverse outcome pathway for skin sensitization (**Figure 1**).
- We compared the performance of the GARDskin assay to other NAMs using three different DAs: 2 out of 3 (2o3) (OECD 2023), Integrated Testing Strategy (ITSv2) (OECD 2023), and Key Event 3/1 Sequential Testing Strategy (KE 3/1 STS) (EPA 2018) (**Figure 1**). We tested 31 substances of interest to several U.S. federal agencies in the following skin sensitization NAMs for use in the DAs:
  - KE1: Direct peptide reactivity assay (DPRA; OECD 2022a)
  - KE 2: KeratinoSens™ assay, (KS; OECD 2022b)
  - KE 3: Human cell line activation test (h-CLAT; OECD 2022c)
  - KE 3: GARDskin (31 chemicals; OECD 2022c)
- GARDskin results were applied as a substitute for h-CLAT (Key Event 3) within the 2o3 and ITSv2 DAs.
  - Results were evaluated both for prediction of skin sensitization hazard (i.e., sensitizer vs. nonsensitizer) and potency classification according to United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS) categories.
  - Concordance and performance of the GARDskin results were compared to "classic" DAs and individual test methods, including murine local lymph node assay (LLNA) reference data.
- GARDair is a promising NAM to assess the potential for respiratory sensitization. Approximately 100 substances nominated by U.S Federal Agencies are being tested using GARDair. Results for chemicals tested thus far are provided and compared to expected respiratory hazard classifications based on occupational, animal, and NAM data.

Figure 1. Adverse Outcome Pathway for Skin Sensitization and Defined Approaches for Hazard and/or Potency Classification



Cys, cysteine; DCs, dendritic cells; Lys, lysine

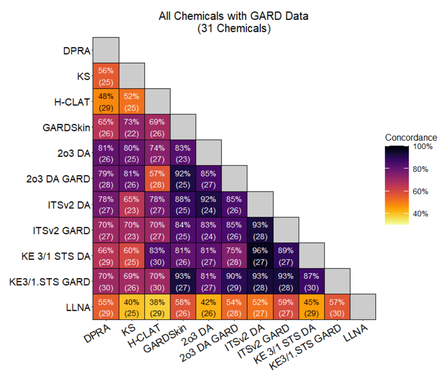
Chemical Selection and Testing Approach

- For skin sensitization testing, substances were nominated by: National Institute of Environmental Health Sciences, U.S. Environmental Protection Agency (EPA), U.S. Food and Drug Administration (FDA), and the U.S. Consumer Product Safety Commission (CPSC).
  - 185 nominated, 181 tested, 174 with in vivo results (172 LLNA and 2 human).
  - A subset of 31 substances was nominated for additional testing in the GARDskin.
- Substances were nominated for respiratory sensitization testing by: National Cancer Institute, CPSC, EPA, FDA, National Institute for Occupational Safety and Health, and U.S. Air Force; 94 of 101 nominated substances will be tested.
- Testing was carried out according the relevant OECD guidelines or testing protocol as supplied by method developers.
- GARDskin test results were compared to existing reference data, including LLNA, human predictive patch test, DPRA, KeratinoSens, h-CLAT, and the DAs (**Figures 2 and 3**).
- GARDair test results (**Figure 4**) were compared to existing occupational human reference data and results from the LLNA, DPRA, KeratinoSens, and h-CLAT, when available, due to a lack of applicable reference data for respiratory sensitization.

Results: GARDskin Concordance with Individual Methods and DAs for Skin Sensitization

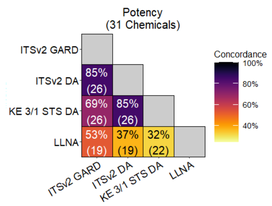
- In the figures below, the heatmaps include a percent concordance calculated based on total number of chemicals that shared a prediction (positive or negative) over the total number of chemicals shared between the assays or DAs.
- Numbers in parentheses indicate the numbers of chemicals predicted by both comparators.
- Tables show performance of GARDskin both as a standalone assay or as a drop-in replacement in the DAs for hazard and GHS potency (ITSv2 only).
- Bold italics indicate "classic" DA/method against reference LLNA data (**Figure 3**).

Figure 2. GARDskin Concordance and Performance of the DAs



Reference	Sensitivity	Specificity	Balanced Accuracy	False Positive Rate	False Negative Rate	Number of Chemicals Predicted
h-CLAT	83%	38%	61%	33%	17%	26
LLNA	53%	67%	60%	33%	47%	26
2o3	93%	63%	78%	38%	7%	23
ITSv2	100%	57%	79%	43%	0%	25
KE3/1 STS	100%	38%	69%	63%	0%	26
hCLAT (LLNA)	*60%*	*33%*	*55%*	*67%*	*24%*	*30*
2o3 (LLNA)	*57%*	*25%*	*41%*	*75%*	*43%*	*26*
ITSv2 (LLNA)	*80%*	*17%*	*52%*	*37%*	*11%*	*27*
KE3/1 STS (LLNA)	*80%*	*7%*	*44%*	*93%*	*20%*	*29*

Figure 3. GARDskin GHS Potency Classification



Reference	Accuracy	Under-predicted	Over-predicted	Number of Chemicals Predicted
LLNA	53%	42%	5%	19
ITS	85%	15%	0%	26
KE3/1 STS	69%	19%	12%	26
ITSv2 (LLNA)	*37%*	*42%*	*21%*	*22*

Figure 4. GARDair Initial Results

GARDair Concordance with Expected Hazard Classification for Respiratory Sensitization

GARDair Classification \ Expected Classification*	Sensitizer	Non-Sensitizer
Sensitizer	81.8% (9/11)	42.9%[b] (3/7)
Non-Sensitizer	18.2%[a] (2/11)	57.1% (4/7)

Total chemicals scheduled for GARDair testing: 94
Chemicals tested using GARDair to date: 18

*Concordance of classification is based on comparison to existing occupational human reference data, LLNA, and DPRA, KeratinoSens, and h-CLAT assessments.

[a]These discordant results were for chemicals that are known to react with water changing their reactivity for biomolecules.

[b]These discordant results were all for chemicals known to be skin, eye, and/or respiratory irritants.

Results

- For skin sensitization hazard classification, concordance against LLNA was higher for DAs that incorporated GARDskin than those that did not (**Figure 2**).
  - The highest hazard concordance of GARDskin was noted for comparisons against ITSv2 DA (88%) whereas the lowest hazard concordance for all methods was seen for comparisons against the LLNA (42–59%).
  - For potency classification, the highest concordance was observed between the two ITSv2 options (classic vs. GARDskin, 85%) (**Figure 3**).
  - The classic ITSv2 had higher concordance with the KE3/1 STS than ITSv2 with GARDskin. Overall concordance for any of the DAs with LLNA was low, with ITSv2 with the GARDskin method being the highest at 69%.
  - LLNA data have been shown to be less predictive of human responses than AOP-based DAs; however, human data were not available for these chemicals
- Classifications of chemicals for respiratory sensitization hazard using the GARDair assay were compared against expected hazard classifications based on existing occupational human reference data and results from LLNA, DPRA, KeratinoSens, and h-CLAT, when available. The GARDair assay had an overall concordance of 72.2% with expected classifications (**Figure 4**).
  - Two expected sensitizers were classified as nonsensitizers in the GARDair assay. These substances are known to react with water resulting in alteration or elimination of reactivity for biomolecules. Because the GARDair is an aqueous assay, such substances are more likely to be misclassified.
  - Four expected nonsensitizers were classified as sensitizers by GARDair but are known irritants to the skin, eyes, and/or respiratory tract. Irritants can give false positives in some sensitization NAMs.

.

Conclusions

- GARDskin tended to overpredict sensitization hazard when compared to reference data, but overall performed well against LLNA reference data for this substance set.
- In vitro testing and DAs provide a useful alternative to animal testing for skin sensitization hazard and potency classification of substances relevant to a wide range of federal agency programs.
- Preliminary results indicate that hazard classification using the GARDair assay showed promising concordance with hazard classifications based on existing occupational human reference data and results from LLNA, DPRA, KeratinoSens, and h-CLAT.
  - Consistent with GARDskin, the preliminary GARDair tended to overpredict sensitization for chemicals expected to be negative. However, definitive testing for respiratory allergy has not been conducted for those chemicals.

References

EPA 2018. Interim Science Policy (KE 3/1 STS). https://www.regulations.gov/document/EPA-HQ-OPP-2016-0093-0090.

OECD 2014. Guidance Document No. 168. https://doi.org/10.1787/9789264221444-en.

OECD 2022a. Test No. 442C: In Chemico Skin Sensitisation (DPRA). https://doi.org/10.1787/9789264229709-en.

OECD 2022b. Test No. 442D: In Vitro Skin Sensitisation (KS). https://doi.org/10.1787/9789264229822-en.

OECD 2022c. Test No. 442E: In Vitro Skin Sensitisation (h-CLAT). https://doi.org/10.1787/9789264264359-en.

OECD 2023. Guideline No. 497: Defined Approaches on Skin Sensitization (2o3, ITSv2). https://doi.org/10.1787/b92879a4-en.

Acknowledgments

The Intramural Research Program of NIEHS supported this poster. Technical support was provided by Inotiv under NIEHS contract HHSN273201500010C and by Burleson Research Technologies, Inc., under NIEHS contract HHSN273201400017C.

*D.G. Allen's current affiliation is with the International Collaboration on Cosmetics Safety, New York, NY.

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.

To get announcements of NICEATM activities, visit the NIH mailing list page for NICEATM News at https://list.nih.gov/cgi-bin/wa.exe?SUBED1=niceatm-l&A=1 and click "Subscribe."