# Evaluation of Skin Sensitization Classification Rules to Reflect Human Potency

M. Herzler[1], J. Abedini[2*], K.T. To[2], D.G. Allen[2], A.M. Api[3], D. Germolec[4], J. Gordon[5], N. Kleinstreuer[4], H.-S. Ko[6], J. Matheson[5], H.-J. Thierse[1], J. Truax[2], J.T. Vanselow[1], J. Strickland[2]

[1]BfR, Berlin, Germany; [2]Inotiv, RTP, NC; [3]RIFM, Woodcliff Lake, NJ; [4]NIH/NIEHS/DTT/NICEATM, RTP, NC; [5]U.S. CPSC, Rockville, MD; [6]FDA/CDER, Silver Spring, MD

Abstract Number: 3177
Poster Number: P292

## Introduction

- Human reference data are needed to evaluate alternative test methods in the most human-relevant manner.
- To support the development of Guideline 497 on Defined Approaches for Skin Sensitization published by the Organisation for Economic Co-operation and Development (OECD; OECD 2021), we collected historical human predictive patch test (HPPT) data used for the assessment of skin sensitization.
- Data from 2255 HPPTs, representing 1366 different substances, were judged to be sufficiently reliable and used to assign skin sensitization potency classifications according to the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS; UN 2019) (Fig. 1a).
- Approaches currently used to assign skin sensitizers to GHS potency subcategories consider only the dose inducing the skin sensitization response and not the frequency of induced sensitization in human subjects. Variations in conduct of assays may also introduce uncertainty into otherwise valid data.
- To address these limitations, we developed a modified approach to GHS classification (Fig. 1b) that incorporates a frequency metric into potency classification and also addresses uncertainty in assay results.

## Figure 1. Standard and Modified GHS Classification Decision Trees

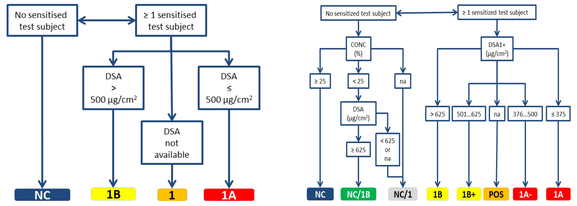


Figure 1 (a) represents the standard GHS classification approach. The modified approach we developed, shown in (b), incorporates sensitization incidence as well as ambiguous/borderline cases. Two dose metrics were applied to this approach: DSA1+ or DSA05 (not shown). Derivation of the dose metrics DSA1+ and DSA05 is explained below.
DSA = dose per skin area.

## GHS Classification of Human Predictive Patch Test Results

- The dose metric for assessing potency of a skin sensitizer in humans is the dose per skin area (DSA), the amount in micrograms of chemical per cm² (µg/cm²) of skin area required to induce an allergic reaction.
- In the standard GHS classification system (Fig. 1a), a substance is classified as a skin sensitizer (Category 1) if at least one subject is sensitized.
  - A positive result at DSA ≤ 500 µg/cm² results in a classification as a 1A (strong) sensitizer.
  - A positive result at DSA > 500 µg/cm² typically indicates a 1B (other) sensitizer, but 1A cannot be ruled out because a lower dose could produce a positive result.
- Chemicals that test negative are assigned a GHS designation of Not Classified (NC). However, in many cases this is uncertain because:
  - Negative results at a concentration < 100% may not be unambiguously negative.
  - Negative test results at DSA ≥ 500 µg/cm² suggest no need for classification. Classification as 1A can be ruled out, but 1B classification cannot because a higher test concentration might have produced a positive response.
  - Negative results at DSA < 500 µg/cm² suggest that a classification for skin sensitization hazard might not be needed. However, classification as 1A or 1B sensitizer cannot be ruled out with certainty because the concentration tested was not high enough to exclude these possibilities.
- To resolve these uncertainties, we derived a borderline range of 375 to 625 µg/cm² (± 25% around the 500 µg/cm² cut-off) (Fig. 1b) and established a test concentration cut-off of at least 25% (the 99$^{th}$ percentile of the top concentrations of negative tests) to classify negative tests as NC. Under this proposed modification:
  - Chemicals testing negative at concentrations < 25% with DSA ≥ 625 µg/cm² were classified as NC/1B, an outcome that, while ambiguous, enables exclusion of a strong skin sensitization potential.
  - Chemicals testing negative at concentrations < 25% with DSA < 625 µg/cm² were classified as NC/1, an ambiguous classification that provides no information on the skin sensitization potential.
- GHS classification does not account for the number of sensitized individuals contributing to a positive result, thereby ignoring an important measure of potency. To incorporate this measure into classification, we examined two additional dose metrics:
  - DSA1+, the hypothetical DSA producing one sensitized test subject.
  - DSA05, the hypothetical DSA that sensitizes 5% of the test subjects.

## Evaluation of Substances with Multiple Discordant Tests

- After classification of each of the 2255 HPPTs using Fig. 1b, substances with discordant tests were classified by combining the multiple results using three weight-of-evidence (WoE) approaches:
  - WoE score: average of individual test data scores (Fig. 2)
  - Median-like location parameter (MLLP) (adapted from Hoffmann et al. 2018) (Table 1)
  - Median sensitization potency estimate (MSPE) (Table 2), a slightly modified version of the MLLP.
- Substances were classified using three different modes based on GHS categories:
  - $GHS_{BIN}$: substance classified in a binary manner as a sensitizer or not classified.
  - $GHS_{SUB}$: substance assigned to one of three classes: 1A sensitizer, 1B sensitizer, or not classified.
  - $GHS_{BORDER}$: substance assigned to one of five classes: the three classes used in $GHS_{SUB}$ with different criteria (except NC); "1" (sensitizer, but subclassification not possible); and "NC/1B" (substance may or may not be a sensitizer, but 1A can be ruled out).

## Figure 2. WoE* Score for Classifying Substances with Multiple Discordant Tests

**Individual Test Data**

| Extrapolated Classification (Figure 1) | Score |
|---|---|
| 1A | 2 |
| 1A- | 1.75 |
| POS | 1.5 |
| 1B+ | 1.25 |
| 1B | 1 |
| NC/1 | NA |
| NC/1B | 0.5 |
| NC | 0 |

**Calculate Weight-of-Evidence Score**

**Combined Chemical Classification**

| WoE Score | Classification Mode | | |
|---|---|---|---|
| | $GHS_{BIN}$ | $GHS_{SUB}$ | $GHS_{BORDER}$ |
| 1.76-2 | 1 | 1A* | 1A |
| 1.51-1.75 | | | 1** |
| 1.50 | | NA | |
| 1.26-1.49 | | 1B* | |
| 0.76-1.25 | | | 1B |
| 0.26-0.75 | NA | NA | NC/1B |
| 0-0.25 | NC | NC | NC |

*WoE score is calculated as the mean of the scores for the individual test data.
**NC/1 tests were excluded from the combined chemical classification.
NA = not applicable, not assigned

## Median-like Location Parameter Approach

- Hoffmann et al. (2018) described a "median-like location parameter" (MLLP) approach to establish a representative value for describing skin sensitizer potency of substances with multiple test results.
- We applied the MLLP approach to the HPPT data set as summarized in Table 1.
  - Test results with NC/1 or NC/1B outcomes were included as negatives if tests with a positive outcome (1A, 1B, or 1) were available and the concentration applied in the ambiguous tests was at least equal to the median DSA1+ of the positive tests.
  - The substance was considered a sensitizer under $GHS_{BIN}$ if the majority of the unambiguous tests were positive; the $GHS_{BIN}$ classification was NC if the majority of the unambiguous tests were negative. If the number of positives and negatives was equal, the overall reference classification was 1.
  - For GHS subcategorization ($GHS_{SUB}$), ambiguous test results for subclassification (1) were excluded (in addition to the NC/1 and NC/1B tests with concentrations that were too low).
  - Classification was performed using the approach shown in Table 1 after calculating the MLLP of the remaining tests.

## Table 1. MLLP Approach to Classification of Each Substance

| Median-like Location Parameter (based on DSA1+ in µg/cm²) | Classification Mode | | |
|---|---|---|---|
| | $GHS_{BIN}$ | $GHS_{SUB}$ | $GHS_{BORDER}$ |
| ≤ 375 | 1 | 1A | 1A |
| 375 < MLLP ≤ 500 | | | 1 |
| 500 < MLLP ≤ 625 | | 1B | |
| > 625 | | | 1B |
| NC/1B | NA | NA | NC/1B |
| NC | NC | NC | NC |

NA = not applicable, not assigned

## Median Sensitization Potency Estimate Approach

- The median sensitization potency estimate (MSPE) approach was developed due to concern that the MLLP approach was insufficiently conservative in some cases. The MSPE approach is summarized in Table 2.
  - NC/1 test results were excluded, but positive test results with no DSA1+ values were included in the median calculation.
  - The MSPE was calculated by sorting all values from low to high potency in the following order:
    NC → NC/1B → Numerical DSA1+ values > 500 µg/cm² in descending order → POS → Numerical concentration values ≤ 2.0% in descending order.
  - When only positive tests were available and the number of 1A results equaled that of 1B, the MSPE was "POS".

## Table 2. MSPE Approach to Classification of Each Substance

| Median Sensitization Potency Estimate (based on DSA1+ in µg/cm²) | Classification Mode | | |
|---|---|---|---|
| | $GHS_{BIN}$ | $GHS_{SUB}$ | $GHS_{BORDER}$ |
| ≤ 375 | 1 | 1A | 1A |
| 375 < MSPE ≤ 500 | | | 1 |
| POS | | NA | |
| 500 < MSPE ≤ 625 | | 1B | |
| > 625 | | | 1B |
| NC/1B | NA | NA | NC/1B |
| NC | NC | NC | NC |

NA = not applicable

## Table 3. Comparison of Classification Approaches for DSA1+ and DSA05

| | DSA1+ | | | DSA05 | | |
|---|---|---|---|---|---|---|
| | $GHS_{BIN}$ | $GHS_{SUB}$ | $GHS_{BORDER}$ | $GHS_{BIN}$ | $GHS_{SUB}$ | $GHS_{BORDER}$ |
| Total available outcomes* | 286 | 273 | 1309 | 287 | 276 | 1309 |
| Outcomes based on more than one approach | 271 | 256 | 289 | 277 | 264 | 289 |
| Approach outcomes identical | 271 (100%) | 249 (97.3%) | 196 (67.8%) | 277 (100%) | 260 (98.5%) | 199 (68.9%) |
| Outcomes not identical, but consensus classification possible | 0 | 0 | 91 (31.5%) | 0 | 0 | 88 (30.4%) |
| Outcomes not identical - decided by expert judgment | 0 | 7 (2.7%) | 2 (0.7%) | 0 | 4 (1.5%) | 2 (0.7%) |

* "Outcomes" refers to classification outcomes of the WoE score, MLLP, and MSPE approaches for combining multiple discordant tests for individual substances.

## Table 4. Comparison of Classifications Using DSA1+ and DSA05

$GHS_{BIN}$

| | | DSA05 | | | Total |
|---|---|---|---|---|---|
| | | 1 | NC | NA | |
| DSA1+ | 1 | 234 | 0 | 0 | 234 |
| | NC | 0 | 52 | 0 | 52 |
| | NA | 1 | 0 | 1079 | 1080 |
| Total | | 235 | 52 | 1079 | 1366 |

$GHS_{SUB}$

| | | DSA05 | | | | Total |
|---|---|---|---|---|---|---|
| | | 1A | 1B | NC | NA | |
| DSA1+ | 1A | 55 | 9 | 0 | 0 | 64 |
| | 1B | 7 | 150 | 0 | 0 | 157 |
| | NC | 0 | 0 | 52 | 0 | 52 |
| | NA | 0 | 3 | 0 | 1090 | 1093 |
| Total | | 62 | 162 | 52 | 1090 | 1366 |

Table shows number of substances. NA = not available; NC = not classified
- Dose metrics DSA1+ and DSA05 classified an identical number of substances in the sensitizer (1) and nonsensitizer (NC) classes for $GHS_{BIN}$. DSA1+ classified more substances in subcategory 1A and DSA05 classified more substances in subcategory 1B.

## Table 5. Reproducibility of Test Classifications

| | Number of test results | No. of substances | | Reproducibility (%) Mean (SD) | |
|---|---|---|---|---|---|
| | | DSA1+ | DSA05 | DSA1+ | DSA05 |
| $GHS_{BIN}$ | > 1 | 109 | 110 | 90.1 (18.8) | 89.9 (18.8) |
| | > 2 | 59 | 60 | 90.3 (18.0) | 89.9 (17.8) |
| | > 3 | 42 | 42 | 90.3 (17.1) | 90.3 (17.1) |
| | > 4 | 28 | 28 | 92.6 (14.4) | 92.6 (14.4) |
| $GHS_{SUB}$ | > 1 | 90 | 93 | 85.2 (20.1) | 85.8 (21.5) |
| | > 2 | 49 | 55 | 83.1 (18.6) | 80.6 (22.8) |
| | > 3 | 37 | 37 | 80.3 (19.0) | 79.3 (24.0) |
| | > 4 | 26 | 26 | 78.7 (20.0) | 80.1 (21.3) |

Table shows reproducibility results for classifications of substances with at least two test results relevant to binary ($GHS_{BIN}$) or subcategory ($GHS_{SUB}$) classifications.

## Concordance with Local Lymph Node Assay

- To further explore the utility of our proposed classification approach we applied it to classification of the reference chemicals identified in OECD Guideline 497 (OECD 2021).
  - Classifications derived from HPPT data using DSA1+ and DSA05 dose metrics in the modified GHS approach (Fig. 1b) were compared to classifications based on murine local lymph node assay (LLNA) results.
- Of the 196 OECD reference chemicals, 55 substances had $GHS_{BIN}$ classifications for HPPT (using both DSA1+ and DSA05) and LLNA data, and 46 substances had $GHS_{SUB}$ classifications.
  - The concordance of HPPT classifications with LLNA classifications was similar when based on DSA1+ or DSA05.
    - For binary classifications, concordance of both HPPT DSA1+ and HPPT DSA05 with LLNA was 83% (44/55).
    - For subcategory classification, the concordance of HPPT DSA1+ classifications with LLNA classifications was 61% (28/46), while the concordance of HPPT DSA05 was 63% (29/46).

## Summary

- We collected a large data set of historical HPPT studies from the scientific literature to use as reference data for development of OECD Guideline 497.
- We developed a new approach for hazard and potency classification of these tests based on GHS categories. The modified approach accounts for uncertain or borderline results and considers the number of sensitized subjects as a measure of potency using DSA1+ and DSA05 dose metrics.
- Both DSA1+ and DSA05 provided reproducible results when used with three different WoE approaches for combining multiple discordant results for single substances.
- Use of borderline ranges around the 1A/1B cutoff value identified ambiguous subclassifications.
- A test concentration cut-off of 25% was used to define the minimum concentration at which a negative test result would be accepted to provide more certainty for negative results.
- Substance classifications based on HPPT results were consistent with LLNA classifications.

## Conclusion

- We conclude that using a modified GHS approach to classifying HPPT data provided good reproducibility and concordance with animal reference data while considering potency and uncertainty.
- DSA1+ or DSA05 may be a more relevant dose descriptor for potency determination.

## References

Hoffmann et al. Crit Rev Toxicol 48(5):344-358. https://doi.org/10.1080/10408444.2018.1429385
OECD. 2021. Guideline No. 497. https://doi.org/10.1787/b92879a4-en
UN. 2019. Globally Harmonized System of Classification and Labelling of Chemicals. https://unece.org/transport/standards/transport/dangerous-goods/ghs-rev9-2021

## Acknowledgments

Learn more about ICE at SOT 2023:

**Multiple demonstrations daily at the NIEHS booth.**

**Visit ICE**
**https://ice.ntp.niehs.nih.gov**

This project was funded in whole or in part with federal funds from the National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN273201500010C.

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.

JAA was an Inotiv employee during this project. Her current affiliation is RTI International.

**Subscribe to NICEATM News email list https://ntp.niehs.nih.gov/go/niceatm**