National Institute of Environmental Health Sciences
Division of Translational Toxicology

BfR Bundesinstitut für Risikobewertung

Evaluation of Skin Sensitization Classification Rules to Reflect Human Potency and Support Weight-of-Evidence Assessments

U.S. FOOD & DRUG ADMINISTRATION

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

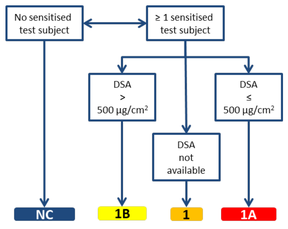
[1]Inotiv, RTP, NC; [2]RIFM, Woodcliff Lake, NJ; [3]NIH/NIEHS/DTT/NICEATM, RTP, NC; [4]CPSC, Rockville, MD; [5]FDA/CDER, Silver Spring, MD; [6]BfR, Berlin, Germany

Introduction

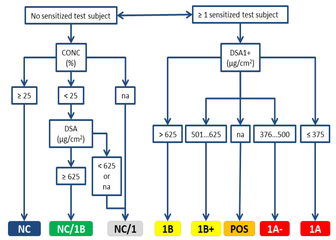
- Human reference data are needed to evaluate alternative test methods in the most human-relevant manner.
- To support the development of Guideline 497 on Defined Approaches for Skin Sensitization published by the Organisation for Economic Cooperation and Development (OECD 2021a), we collected historical human predictive patch test (HPPT) data used for the assessment of skin sensitization and evaluated data reliability.
- The 2255 HPPTs judged to be sufficiently reliable were assigned skin sensitization potency classifications according to the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (GHS; UN 2019) (Fig. 1a).
- We developed a modified approach to GHS classification (Fig. 1b) that incorporated uncertainty and potency.

Figure 1. GHS and Modified Decision Trees for Skin Sensitization Potential

a) Standard GHS Classification Approach



b) Modified GHS Classification Approach

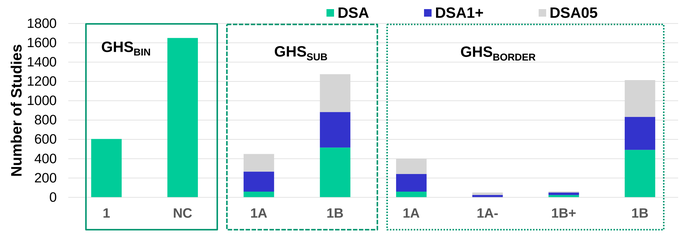


The modified approach developed for this project (b) incorporates uncertainty and potency. Two dose metrics were applied to this approach: DSA1+ and DSA05. DSA = dose per skin area.

GHS Classification of Human Predictive Patch Test Results

- 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 using dose per skin area (DSA) as the relevant dose metric where a:
  - Positive result at DSA ≤ 500 µg/cm² results in a classification as a 1A (strong) sensitizer.
  - Positive result at DSA > 500 µg/cm² typically indicates a 1B (moderate or weak) sensitizer, but 1A cannot be ruled out because a lower dose could produce a positive result.
- Chemicals that test negative result in a GHS designation of Not Classified (NC). However, in many cases the classification results based on the GHS are uncertain because:
  - Negative results at a concentration < 100% may not be unambiguously negative.
  - 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.
  - Negative test results at DSA ≥ 500 µg/cm² suggest no need for classification. Classification as 1A can certainly be ruled out, but 1B classification cannot because a higher test concentration might have produced a positive response.
- To resolve these uncertainties, we derived a borderline range around the 1A and 1B categories (Fig. 1b) and established a test concentration cut-off of at least 25% (the 99th percentile) 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.
  - Negatives tested at < 25% and with a DSA < 625 µg/ cm² were considered 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. In an effort to incorporate this measure into classification, we examined two additional dose metrics:
  - DSA1+, the hypothetical DSA producing exactly one sensitized test subject.
  - DSA05, the hypothetical DSA that sensitizes 5% of the test subjects.

Figure 2. Skin Sensitization Classification Results



- In our data set of 2255 HPPTs, applying the standard GHS approach (GHS_{BIN}) to binary hazard categorization (Fig. 1a) using the DSA classified 605 substances as Category 1 positives and 1650 substances as NC (Fig. 2, far left). Fig. 2 also shows the classification of positive results into GHS subcategories 1A and 1B (GHS_{SUB}) and into the modified GHS classifications to account for uncertainty (GHS_{BORDER}) as shown in Fig. 1b using the DSA, DSA1+, and DSA05 dose metrics.
- Only 575 of the 605 positive HPPTs had sufficient information for subcategorization into GHS_{SUB} and GHS_{BORDER} classifications.
- When these approaches were applied, the dose metrics diverged.
- Using the GHS_{SUB} classification approach:
  - The DSA metric classified 59 substances as 1A and 516 substances as 1B.
  - The DSA1+ metric classified 208 substances as 1A and 367 substances as 1B.
  - The DSA05 metric classified 182 substances as 1A and 392 substances as 1B.
  - Thus, the DSA classified the most sensitizers into the less potent group, 1B.
- Using the GHS_{BORDER} classification approach:
  - The DSA metric only included 24 substances in the borderline 1A- and 1B+ categories, compared to 51 for the DSA1+ and 36 for the DSA05.
- The DSA1+ dose metric was selected to support the development of OECD Guideline 497.

Table 1. Reproducibility of Substance Classifications With at Least Two Test Results

Classification Mode	Number of Test Results	Number of Substances	Reproducibility	
			Mean	SD
GHS_{BIN}	≥ 2	69	98.1	7.4
	≥ 3	40	96.8	9.6
	≥ 4	30	96.8	9.4
	≥ 5	25	98.2	8.0
GHS_{SUB}	≥ 2	67	83.2	19.7
	≥ 3	38	77.0	17.9
	≥ 4	28	77.2	18.8
	≥ 5	23	77.6	19.4

- The reproducibility of DSA1+ classification results was calculated using the modified GHS classification algorithm (Fig. 1b) as the fraction of all unambiguous HPPT results for a given chemical that correctly predicted the overall call determined based on a weight-of-evidence approach (OECD 2021b).
  - For GHS_{BIN}, Category 1 or NC, tests resulting in classifications of NC/1B were excluded.
  - For GHS_{SUB}, which includes 1A, 1B, and NC, tests resulting in a Category 1 classification were excluded. Studies resulting in classifications of NC/1B were excluded if the overall classification for a chemical was 1B or NC. However, if the overall call was 1A, studies resulting in classifications of NC or 1B were counted as contradictory.
- As shown in Table 1, reproducibility was 97-98% for binary classification, GHS_{BIN}, and 77-83% for classification into subcategories, GHS_{SUB}.

Weight-of-Evidence Evaluation

- The DSA1+ was used to classify substances for human skin sensitization hazard and potency using the modified GHS approach (Fig. 1b) as reference data for development of the OECD defined approach project referred to in the Introduction. Murine local lymph node assay (LLNA) data were also collected and evaluated as reference data.
- Of the 200 OECD reference chemicals, 91 substances had GHS_{BIN} classifications for both HPPT and LLNA data, and 87 substances had GHS_{SUB} classifications.
  - HPPT and LLNA GHS_{BIN} classifications were concordant for 95% (86/91) of the substances and GHS_{SUB} classifications were concordant for 86% (75/87).
  - In 2/5 (GHS_{BIN}) and 9/15 (GHS_{SUB}) discordant cases, the overall classification decision was made based on the LLNA reference classification, while in all but one of the remaining cases, it was made based on the HPPT data. The remaining case was decided by expert judgment.

Summary

- We collected a large dataset of historical HPPT studies from the scientific literature to use as reference data for development of OECD Guideline 497.
- We developed a new approach based on GHS categories for hazard and potency classification of these tests. The modified approach considers the number of sensitized subjects as a measure of potency and accounts for uncertain/borderline results.
- To classify the HPPT studies, we used the DSA1+ dose metric, which is the DSA estimated to sensitize a single subject.
- Using this approach for classifying chemicals with multiple results resulted in a reproducibility of 97-98% for binary classification and 77-83% for subcategory classification.
- 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 (i.e., the test was less likely to be negative because the dose was too low) (OECD 2021b).

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

OECD. 2021a. Guideline No. 497.
OECD. 2021b. Series on Testing and Assessment No. 336, Annex 4.
UN. 2019. Globally Harmonized System of Classification and Labelling of Chemicals.

Acknowledgements

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.

* Affiliation was Inotiv-RTP during this project. Current affiliation is RTI International.