Variability in the Rabbit Skin Irritation Assay

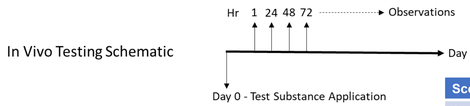
JP Rooney[1], N Choksi[1], R Rai[1], D Allen[1], N Kleinstreuer[2], W Casey[2]

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

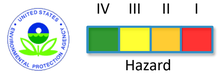
Background

- Characterizing the inherent variability in animal test methods, even when conducted according to accepted test guidelines, is critical to set appropriate expectations for performance and establish confidence in new approach methodologies (NAMs).
- Skin irritation testing has historically been conducted in vivo on rabbits, and results have been used as the benchmark against which NAMs have been compared.
- The Organisation for Economic Co-operation and Development's (OECD) Test Guideline 439 describes NAMs that are accepted as standalone replacement tests for in vivo skin irritation testing in certain jurisdictions.
- However, there are currently no NAMs that can identify chemicals as mild irritants and thus could serve as complete replacements where such classification and labeling is required (e.g., in the United States).
- Given the subjective scoring of erythema and edema responses in the in vivo rabbit test, the possibility exists for significant variability in results if the same chemical were to be evaluated in multiple independent tests.
- In this study, we used conditional probabilities to assess the reproducibility of hazard classifications resulting from the in vivo assay to provide an appropriate benchmark against which to evaluate NAMs.

EPA Skin Irritation Classification



$$PDII = \frac{Sum\ erythema\ (all\ time\ points) + Sum\ edema\ (all\ time\ points)}{number\ of\ intervals \times number\ of\ animals}$$



Score	Erythema	Edema
0	No erythema	No edema
1	Slight erythema	--
2	Well defined	Slight edema
3	Moderate-severe	Moderate
4	Severe	Severe

EPA	Category I	Category II	Category III	Category IV
PDII	Corrosive	>5.0	2.1-5.0	0-2.0
Signal Word	DANGER	WARNING	CAUTION	CAUTION
PPE Required	Coveralls worn over long-sleeved shirt and long pants	Coveralls worn over short-sleeved shirt and short pants	Long-sleeved shirt and long pants	Long-sleeved shirt and long pants
	Socks	Socks	Socks	Socks
	Chemical-resistant footwear	Chemical-resistant footwear	Shoes	Shoes
	Waterproof or chemical-resistant gloves	Waterproof or chemical-resistant gloves	Waterproof or chemical- resistant gloves	No minimum

- **Figure 1** (above) provides a general overview of the dermal irritation classification system used by the U.S. Environmental Protection Agency (EPA).
  - Test substance is applied to shaved, intact rabbit skin for at least 4 hours.
  - Erythema and edema formation are scored up to 72 hours after removal of the test substance using the Draize scoring system.
  - Primary dermal irritation index (PDII) is calculated as shown above using the scores from 1, 24, 48 and 72 hours after test substance removal.
  - PDIIs are used to classify test substances into categories II, III or IV.
  - Category I classifications are made based on direct evidence of corrosive properties.
  - Color coding indicates relative level of human hazard (i.e., red category is corrosive; green category is non-corrosive/minimal irritant).

Study Design: Data Collection and Curation

- A dataset of 3291 in vivo study records, representing 1071 unique chemicals, was obtained from the European Chemicals Agency database of registered substances (https://echa.europa.eu).
- Each chemical was tested at least twice in independent in vivo rabbit studies.
- Curation included identification of factors affecting test method reliability and methodological deviations/limitations.
- Study records were excluded from analysis if:
  - Skin status was abraded or abraded/intact.
  - Number of animals was less than 3 or unknown.
  - Test chemical dosing concentration was less than 90%.
  - Exposure duration was less than 4 hours (unless chemical is corrosive).
- PDIIs were calculated from the available data and used to classify chemicals according to the EPA skin irritation classification criteria.

Study Design: Conditional Probability Calculations

- Conditional probabilities, conducted iteratively for each category, were used to evaluate the reproducibility of the in vivo method for identification of severe, moderate, mild, and non-irritants.

$$P(T_2 = 1 | T_1 = 1) = \frac{P(T_2 = 1 \cap T_1 = 1)}{P(T_1 = 1)}$$

- Datasets were created for each category (I, II, III, IV) defined by the chemicals classified in the category by at least one test.
  - Frequency of classification for each category, given the total number of assays in that dataset, was determined.
    - Ci = frequency of each category.
    - A = Sum of total of number of assays in that dataset.
- Probability was calculated for each category by dividing the frequency of each category by the frequency of all categories (total number of assays) in that dataset.
  - P = Ci/A

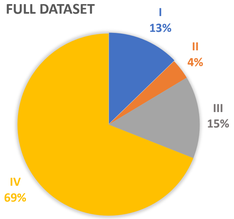
Example:

- In the curated dataset, 105 chemicals were classified as corrosive (Cat I) at least once.
- Those 105 chemicals had a total of 251 study reports (A = 251).
- 217 of those study reports had Cat I categorizations (Ci Cat I = 217).
  - Probability of receiving a Cat I result given a prior Cat I result:
    - P=Ci/A
    - 217/251 = 0.865
- Repeat for each category.

Dataset Summary and Conditional Probabilities

Figure 2: Summary of Complete and Curated Datasets

FULL DATASET



EPA Category	Number of Study Reports	Percent
I	312	12.8
II	89	3.6
III	358	14.6
IV	1689	69.0

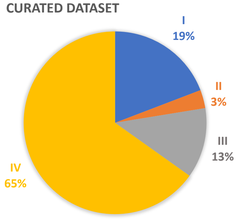
- Includes all study results with sufficient data to generate a PDII.
- No exclusion of studies for methodological concern (e.g. 24-hour exposure).
- **2448 total study reports**
- **806 individual chemicals**

Conditional Probabilities

Prior result	I	II	III	IV
I	75.9%	8.0%	8.5%	7.5%
II	12.0%	28.1%	35.3%	24.6%
III	5.8%	5.0%	43.5%	45.7%
IV	2.2%	1.9%	11.6%	84.5%

Table 1 describes the conditional probabilities for the **full** dataset for receiving the same categorical classification when a chemical is tested multiple times.

CURATED DATASET



EPA Category	Number of Study Reports	Percent
I	207	19.2
II	35	3.2
III	134	12.4
IV	703	65.2

- The full dataset was curated to exclude studies with methodological deviations/limitations and suspected read-across data
- **1079 total study reports**
- **432 individual chemicals**

Conditional Probabilities

Prior result	I	II	III	IV
I	86.3%	4.2%	7.1%	2.5%
II	14.1%	44.9%	20.5%	20.5%
III	6.8%	5.2%	53.6%	34.4%
IV	0.9%	2.0%	9.0%	88.1%

Table 2 describes the conditional probabilities for the **curated** dataset for receiving the same categorical classification when a chemical is tested multiple times.

Subscribe to the NICEATM News Email List

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."

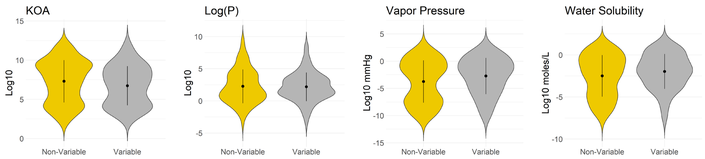
Acknowledgements

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.

Comparison of Physical-chemical Properties

Figure 3: Physical-chemical Properties Do Not Explain Variability



- The physical-chemical properties of each chemical were predicted using quantitative structure-activity relationship models. Groups of chemicals were compared using the Chemical Characterization tool in the NTP Integrated Chemical Environment (https://ice.ntp.niehs.nih.gov/).
- **Fig. 3** shows representative plots of predicted octanol-air partition coefficient (KOA), octanol-water partition coefficient (LogP), vapor pressure and water solubility for each group of chemicals.
- In these examples, chemicals were grouped as follows:
  - Non-variable: chemicals consistently classified in a single hazard category (yellow).
  - Variable: chemicals classified in more than 1 hazard category (grey).
- No physical-chemical property examined clearly distinguishes chemicals that are classified in multiple categories from those that are consistently classified in a single category.

Proposed Binary Approach

Table 3: Conditional Probabilities with Binary Approach Using Curated Dataset

EPA	Cat I	Cat II	Cat III	Cat IV
PDII	Cor.	>5.0	2.1-5.0	0-2.0
Binary	Irritant		Non-Irritant	

The four classifications are reduced to two by combining categories I and II (irritant), and III and IV (non-irritant), based largely on PPE requirements (**Fig. 1**).

Prior Result	Irritant (Cat I / II)	Non-Irritant (Cat III / IV)
Irritant (Cat I / II)	81.5%	18.5%
Non-Irritant (Cat III / IV)	5.0%	95.0%

EPA Category	Number of Study Reports	Percent
Irritant	242	22.4
Non-Irritant	837	77.6

Summary and Future Directions

- Chemicals classified as moderate irritants at least once are just as likely to be classified as mild irritants or non-irritants when tested again **(Tables 1 and 2).**
- Reducing the categorization scheme to a binary outcome improves reproducibility (**Table 3**).
- Variability present in the in vivo assay should be taken into consideration when evaluating the performance of NAMs.
- These analyses help provide much needed context not only to assess "gold standard" reference test methods, but also to aid in setting expectations for NAM performance.

References

Bell SM. et al. 2017 EHP DOI 10.1289/EHP1759 (https://ice.ntp.niehs.nih.gov/)

EPA. 1998a. Health Effects Test Guidelines, OPPTS 870.2500, Acute Dermal Irritation, EPA 712-C-98-196. Available:https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/epa/epa_870_2500.pdf

Luechtefeld et al. 2016. ALTEX 33(2) doi: 10.14573/altex.1510053

Mansouri K. et al. 2018. J Cheminform https://doi.org/10.1186/s13321-018-0263-1

OECD. 2019. Test No. 439: In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264242845-en.

US EPA. 2016. OPP Label Review Manual. Chapter 10: Worker Protection Label. Revised Feb 2016. Available: https://www.epa.gov/sites/production/files/2016-02/documents/chap-10-feb-2016.pdf