Variability in the Rabbit Skin Irritation Assay

JP Rooney[1], NY Choksi[1], P Ceger[1], AB Daniel[1], JF Truax[1], D Allen[1], N Kleinstreuer[2]

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

Highlights

- Skin irritation testing has historically been conducted in rabbits. Results of these studies are the benchmark against which **new approach methodologies (NAMs)** are compared.
- Chemicals classified as **mild or moderate** irritants in one test were approximately **50% or less** likely to be replicated when tested again.
- **Variability** in the in vivo assay should be taken into **consideration** when **evaluating the performance** of NAMs.

Conclusions

- Chemicals classified as moderate irritants at least once are just as likely to be classified as mild irritants or non-irritants when tested again.
- Variability present in the in vivo assay should be taken into consideration when evaluating performance of NAMs.
- These analyses help provide much needed context both to assess “gold standard” reference test methods and to aid in setting expectations for NAM performance.

Conditional Probability Calculations

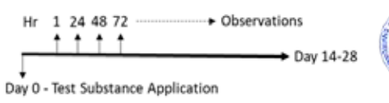
We used conditional probabilities, calculated iteratively for each category, to evaluate the reproducibility of the rabbit skin irritation test for identification of severe, moderate, mild, and non-irritants.

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

- Data subsets were created for each U.S Environmental Protection Agency (EPA) hazard classification category (I, II, III, IV**;** details at right) defined by the chemicals classified in the category by at least one test.
  - Frequency of classification for each category (Ci), given the total number of ESRs in that data subset (A), was determined.
- Probability of repeat testing was calculated for each category by dividing the frequency of each category by the frequency of all categories (total number of assays) in that data subset.
  - P = Ci/A

Background: EPA Skin Irritation Classification

In Vivo Testing Schematic



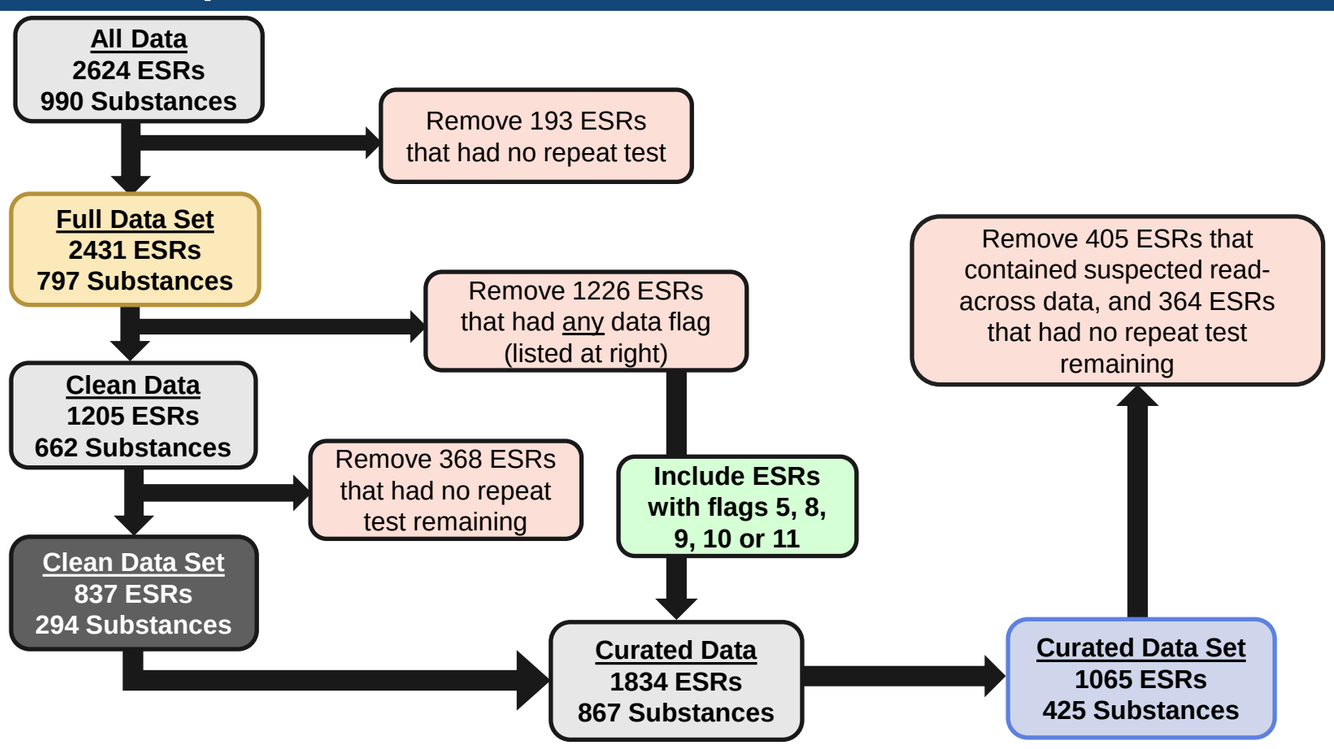
$$PDII = \frac{Sum\ erythema\ (all\ time\ points) + Sum\ edema\ (all\ time\ points)}{number\ of\ intervals \times number\ of\ animals}$$

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

US-EPA	Category I	Category II	Category III	Category IV
PDII	Corrosive	>5.0	2.1 - 5.0	0 - 2.0
Signal Word (Color)	DANGER (Red)	WARNING (Orange)	CAUTION (Yellow)	CAUTION (Green)
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

Abbreviations:
ESR = Endpoint Study Report
PDII – Primary Dermal Irritation Index
NAMs = New Approach Methodologies
EPA = U.S. Environmental Protection Agency

Data Preparation Workflow

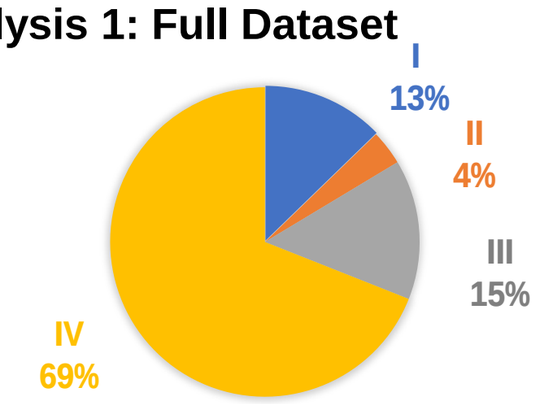


Data Flags: Studies were flagged for methodological deviations from study guidelines.

Flag	Description
1	Only tested on abraded skin
2	Scores reported from combined intact and abraded skin
3	Concentration tested was less than 90%
4	Exposure duration <4 hours and substance was not corrosive
5	Exposure duration >4 hours
6	Tested on unknown number of animals
7	Tested on <3 animals and substance was not corrosive
8	Scoring at incomplete time points (24 hr, 48 hr, 72 hr)
9	Parameter reported other than erythema/edema/ PDII
10	Reported and calculated PDII do not match
11	Reported scores are a range

Dataset Summary and Conditional Probabilities

Analysis 1: Full Dataset



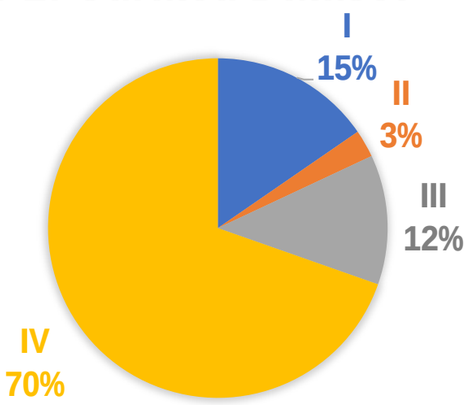
Conditional Probabilities

Prior result	I	II	III	IV	N
I	76.0%	8.0%	8.5%	7.5%	313
II	12.0%	**28.1%**	35.3%	24.6%	89
III	5.8%	5.0%	**43.5%**	45.7%	357
IV	2.2%	1.9%	11.6%	84.4%	1672

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

- Includes all study results with sufficient data to generate a PDII.
- No exclusion of studies for methodological concern (see Flags table above).
- **Results based on 2431 total study reports including 797 individual chemicals.**

Analysis 2: Curated Dataset



Conditional Probabilities

Prior result	I	II	III	IV	N
I	86.3%	4.2%	7.1%	2.5%	207
II	14.1%	**44.9%**	20.5%	20.5%	35
III	6.9%	5.2%	**53.6%**	34.3%	133
IV	0.9%	2.0%	9.1%	88.0%	690

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

- The full dataset was curated to exclude studies with methodological deviations/limitations**.**
- **Results based on 1065 total study reports including 425 individual chemicals.**

More Information

This project was funded with federal funds from the National Institute of Environmental Health Sciences, National Institutes of Health, under Contract No. HHSN273201500010C.

Data available at: https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION

Publication: **https://doi.org/10.1016/j.yrtph.2021.104920**