NICEATM

National Toxicology Program Interagency Center for the Evaluation Of Alternative Toxicological Methods

ICCVAM

Interagency Coordinating Committee on the Validation of Alternative Methods

Preliminary Evaluation of the Under-Prediction Rate of *In Vivo* Dermal Corrosivity Test Methods
Part I: Introduction

William S. Stokes, D.V.M., D.A.C.L.A.M.

Scientific Advisory Committee on Alternative Toxiciological Methods
March 11, 2004
Bethesda, Maryland

National Toxicology Program

NIEHS

NATIONAL INSTITUTES OF HEALTH

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

Acknowledgements

ICCVAM Dermal Corrosivity and Irritation Working Group (DCIWG)

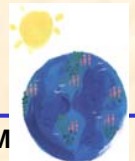
Marilyn Wind, CPSC
Susan Aitken, CPSC
Kailash Gupta, CPSC
Barnett Rattner, DOI
Steve Hwang, DOT
George Cushmac, DOT
Angela Auletta, EPA
Masih Hashim, EPA
Leonard Keifer, EPA
Marianne Lewis, EPA
John Redden, EPA
Mark Perry, EPA
Amy Rispin, EPA
Hari Mukhoty, EPA
Karen Hamernik, EPA (Co-Chair)
Debbie McCall, EPA
Leonard Schechtman, FDA
Robert Bronaugh, FDA
Jill Merrill, FDA
Donnie Lowther, FDA
Abby Jacobs, FDA (Co-Chair)
Doug Sharpnack, NIOSH
Vera Hudson, NLM
William Stokes, NIEHS
William Eastin, NIEHS
Buck Grissom, NIEHS
Rajendra Chhabra, NIEHS
Surender Ahir, OSHA

Acknowledgements

- Joe Haseman, NIEHS/NTP Biostatistician
- NICEATM Staff
  - Neepa Choksi, ILS Inc.
  - Ray Tice, ILS Inc.
- Data Sources
  - EPA: Richard Hill, Marianne Lewis
  - FDA: Donnie Lowther
  - InVitro International Inc.
  - European Centre for Ecotoxicology and Toxicology of Chemicals; Brussels, Belgium
    - § Reference Chemicals Data Bank

Background

- Draize rabbit skin test method
  - Used since the 1940's to identify dermal irritants and corrosives
- Current version used since 1981
  - 0.5 mL or 0.5 g of test substance applied to intact skin with a patch for 3 min, 1 hour, and/or 4 hours
  - Test substance removed at the end of exposure period
  - Observations at 24, 48, and 72 hours (optional at 1 and 4 hours)
  - Erythema and oedema scored at each observation
  - Observation for corrosive (irreversible) effects (e.g., eschar, necrosis)
- Animal welfare concerns led to development and validation of *in vitro* test methods over the past 15 years



In Vitro Dermal Irritation/Corrosion Test Methods

- ICCVAM evaluated and recommended four *in vitro* test methods for assessing dermal corrosivity
  - Corrositex®, EpiDerm™, EPISKIN™, Rat Skin TER assay
- Use in accordance with an internationally-harmonized tiered testing strategy (UN 2003)
  - Positive responses classified as corrosives
  - Negative responses followed by *in vivo* dermal irritation testing
    - Testing performed sequentially
    - Testing stops when an animal develops a corrosive lesion
- Recommendation based, in part, on the 12-17% false-negative rates of the *in vitro* assays
  - Corrosive products without proper hazard warning labels could result in irreversible permanent skin or eye injuries

UN. 2003. Globally Harmonized System of Classification and Labeling of Chemicals (GHS). [ST/SG/AC.10/30]. United Nations, New York and Geneva. Available: http://www.unece.org/trans/danger/publi/ghs/officialtext.html.

Tiered-Testing Strategy

Figure 1. Tiered testing and evaluation of dermal corrosion and irritation potential (see also the "Testing and evaluation strategy for eye irritation/corrosion")

Step	Parameter	Finding	Conclusion
5	Valid and accepted in vitro dermal corrosion test [e]	→ Positive response	→ Classify as corrosive [a]
	↓ Negative response or no data		
6	Valid and accepted in vitro dermal irritation test [f]	→ Positive response	→ Classify as irritant [a]
	↓ Negative response or no data		
7	*In vivo* dermal corrosion test (1 animal)	→ Corrosive response	→ Classify as corrosive [a]
	↓ Negative response		
8	*In vivo* dermal irritation test (3 animals total) [h]	→ Irritant response	→ Classify as irritant [a]
	↓ Negative response	→ No further testing	→ Classify as irritant [a]
9	When it is ethical to perform human patch testing [g]	→ Irritant response	→ Classify as irritant [a]
	Not as above	→ Non-irritant response	→ No further testing

a. Classify in the harmonised category, below.
b. Structure-activity and structure-property relationships are presented separately but would be conducted in parallel.
c. Measurement of pH alone may be adequate, but assessment of acid or alkali reserve is preferable; methods are needed to assess buffering capacity.
d. Pre-existing animal data should be carefully reviewed to determine if in vivo dermal corrosion/irritation testing is needed. As examples, testing may not be needed when a test material has not produced any dermal irritation in an acute dermal toxicity test at the limit dose, or produces very toxic effects in an acute dermal toxicity test. In the latter case, the material would be classified as being very hazardous by the dermal route for acute toxicity; it

From OECD 2001. OECD Series on Testing and Assessment, Number 33: Harmonised Integrated Classification System for Human Health and Environmental Hazards Of Chemical Substances and Mixtures. Document Number: ENV/JM/MONO(2001)6

In Vitro Dermal Irritation/ Corrosion Test Methods

- **Public Comments**
  - **Stated that ICCVAM should recommend, and agencies should adopt, these test methods as complete replacements for the rabbit assay**
    - **Basis was an assertion that the *in vitro* test methods are more accurate than the animal based methods**
    - **Cited low interlaboratory reproducibility of the rabbit assay per Weil and Scala study (1971)**
  - **Other claims that the rabbit assay has a 20-25% false negative rate and 20-25% false positive rate for corrosivity**
    - **No data provided to substantiate this claim**

Weil CS, Scala RA. 1971. Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. Toxicol. App. Pharmacol. 19:276-360.

Reliability of the *In Vivo* Rabbit Dermal Irritation Test

- Weil and Scala study (1971)
  - Evaluated the reproducibility of the Draize *in vivo* rabbit skin test method within and among 24 laboratories for 10 substances
- This study is the only formal evaluation of the reproducibility of the Draize *in vivo* rabbit skin test method
- Conclusions
  - Moderate intra-laboratory reproducibility
  - Low inter-laboratory reproducibility
  - Primary reasons for the low inter-laboratory reproducibility attributed to the subjective nature of the visual observations and variations in procedures among labs

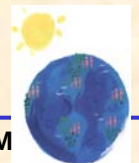
Weil CS, Scala RA. 1971. Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. Toxicol. App. Pharmacol. 19:276-360.

Limitations of the Weil and Scala Analysis

- The standard protocol used was different from the current Draize *in vivo* rabbit skin test method protocol in use since 1981
  - The Weil and Scala studies used a 24-hour exposure period versus the current maximum 4-hour exposure
  - Prolonged exposure likely responsible for corrosive lesions observed for several irritants
- Good Laboratory Practices (GLPs) had not yet been established and were not available for use
  - Impact unknown

NICEATM Study

- **Study Objective**
  - **Evaluate the likelihood of under-predicting a corrosive substance as a non-corrosive in the current rabbit dermal corrosivity test**
- **Data may assist in establishing an acceptable false-negative rate for corrosive effects for *in vitro* test methods proposed as complete replacements for the rabbit skin test**
  - **i.e., those tests where no *in vivo* confirmation would be performed:**
    - **negative *in vitro* corrosivity test**
    - **negative *in vitro* corrosivity test plus only any *in vitro* irritation test data**



NICEATM Study

- Data indicating corrosive severity was sought
  - UN Packing Group (PG) I: corrosive burns within 1 hour after a 3-minute exposure
  - UN PG II: corrosive burns within 14 days after a 1-hour exposure
  - UN PG III: corrosive burns within 14 days after a 4-hour exposure
  - Result: Limited Packing Group data available
- Corrosivity test data for humans sought
  - None found
  - Therefore, not possible to assess the false negative and/or false positive rates of the rabbit dermal corrosivity test method for humans
  - No reports found where human corrosive burns were reported for non-corrosive substances

In Vivo Corrosivity Database

- Data compiled from corrosivity studies using the current exposures for the rabbit skin test protocols recommended by U.S. Federal agencies (e.g., FDA, EPA, CPSC, DOT) and OECD (TG 404)
- Data requested from Federal agencies and via an FR notice[1]
- Current database: 171 substances tested in 185 separate studies
- Several of the substances tested were commercial products with unknown formulations and chemical composition

Data Source	Total Studies	Substances
InVitro International (Bio-Technics)	117	103
ECETOC[2]	23	23
EPA (OPPTS)	26	26
FDA (CFSAN)	19	19
Total	185	171

[1]FR Notice (Vol. 68, No. 13, pp. 42067-42068, July 16, 2003). Available: http://iccvam.niehs.nih.gov/docs/FR/6842067.htm.
[2]ECETOC. 1995. Skin irritation and corrosion: reference chemicals data bank. Technical Report No. 66. Belgium.

In Vivo Corrosivity Database

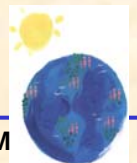
Distribution of Studies Provided By Each Data Source

Source	GLP Compliance	# Studies Provided	Studies per Substance				
			1	2	3	4	5
InVitro International	Yes	117	92	9	1	1	0
ECETOC	Yes	23	23	0	0	0	0
FDA	No	26	26	0	0	0	0
EPA	Unknown	19	19	0	0	0	0
Total		**185**	**160**	**9**	**1**	**1**	**0**

In Vivo Corrosivity Database

Number of Animals Used per Study

Source	Number of Animals Used per Study					
	1	2	3	4	5	6
InVitro International	53	50	11	0	3	0
ECETOC	3	1	10	4	0	5
FDA	0	0	6	0	0	20
EPA	0	0	2	0	0	17
Total	56	51	29	4	3	42



Limitations of the Database

- **Potency subcategories not known for most chemicals**
  - **Only corrosive versus non-corrosive**

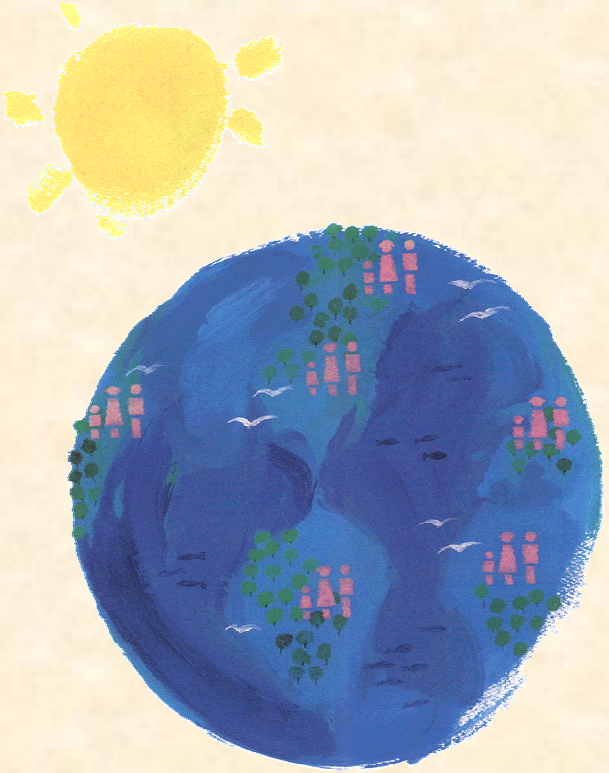
- **<u>Future Plans:</u>**
  - **Continue to seek high quality test data to add to the database**
  - **Currently collaborating with EPA OPPTS to obtain microfiche reports for 2400 commercially available chemicals with dermal test results from the EPA TSCATS database**
    - **Availability of individual animal data unknown**
    - **Distribution of corrosive chemicals unknown**

NICEATM

National Toxicology Program Interagency Center for the Evaluation Of Alternative Toxicological Methods

ICCVAM

Interagency Coordinating Committee on the Validation of Alternative Methods



Preliminary Evaluation of the Under-Prediction Rate of *In Vivo* Dermal Corrosivity Test Methods Part II: Data Analysis

Joseph Haseman, Ph.D.

Scientific Advisory Committee on Alternative Toxiciological Methods
March 11, 2004
Bethesda, Maryland

Definition of False Negative Rate

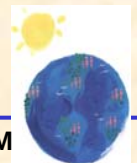
- The false negative rate of a corrosivity test is defined as the probability that a corrosive substance will not produce a positive response when subjected to the test
  - i.e., will produce a response that "falsely" identifies the substance as non-corrosive in the rabbit model

The false negative rate depends upon …

- The specific corrosivity test used
- The responsiveness of the animals to the corrosivity test (mean response rate and variability in response)

Assumptions

- Corrosivity is based on a three animal test, in which one or more positive responses indicates that a test substance is corrosive
- The distribution of test substances in the database, in terms of corrosivity (i.e., the proportion of responding rabbits), is representative of the “real world” of corrosive substances



Basis of False Negative Analysis

	Approach 1: Based on Studies	Approach 1: Based on Test Substances	Approach 2: Average Response Rate
All Data Used			
1 Animal Tested Excluded			
1 and 2 Animal Tested Excluded			

Approach 1 - Specific distribution of response rates
Approach 2 - Average response rates

Approach 1: Distribution of Substances in Database

Positive Response Rate	Based on Studies			Based on Test Substances		
	All Data	1 Animal Tested Excluded	1 and 2 Animal Tested Excluded	All Data	1 Animal Tested Excluded	1 and 2 Animal Tested Excluded
100%	134	78	36	121	71	42
83.3%	4	4	4	4	4	4
75%	0	0	0	3	3	3
71.4%	0	0	0	1	1	1
66.7%	14	14	14	14	14	14
60%	1	1	1	1	1	1
58.3%	0	0	0	0	0	0
50%	15	15	6	10	10	6
40%	1	1	1	1	1	1
33.3%	10	10	10	10	10	10
16.7%	6	6	6	6	6	6
Total Studies	**185**	**129**	**78**	**171**	**121**	**88**

Negative Response Likelihood Calculations

Probability of a Positive Response	Likelihood of 0/3 Negative Response
100%	0.0000
83.3%	0.0046
75%	0.0156
71.4%	0.0234
66.7%	0.0370
60%	0.0640
58.3%	0.0723
50%	0.1250
40%	0.2160
33.3%	0.2963
16.7%	0.5787

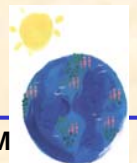
Example Calculation for Determining the Likelihood for a 50% Positive Response Rate:
$(1-0.5)^3=0.5^3=0.1250$

False Negative Calculation Example

Based on Studies Using All the Available Data

Probability of a Positive Response	Frequency	Likelihood of 0/3 Negative Response	Contribution to False Negative Rate
100%	134	0.0000	0.0000
83.3%	4	0.0046	0.0184
75%	0	0.0156	0.0000
71.4%	0	0.0234	0.0000
66.7%	14	0.0370	0.5180
60%	1	0.0640	0.0640
58.3%	0	0.0723	0.0000
50%	15	0.1250	1.8750
40%	1	0.2160	0.2160
33.3%	10	0.2963	2.9630
16.7%	6	0.5787	3.4722
	185		**9.1266**

Estimated False Negative Rate:
(9.1266/185)*100=4.9%



Approach 2: Distribution of Animals With a Corrosive Response in Database

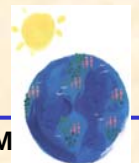
	Proportion of Animals With a Corrosive Response
All Data Used	412/528 (78.0%)
1 Animal Tested Excluded	356/472 (75.4%)
1 and 2 Animal Tested Excluded	263/370 (71.1%)

Sample False Negative Rate Calculations Using Approach 2

All Available Data Used

Positive response rate: 412/528 or 0.780

False Negative Rate Calculation:
$(1-0.780)^3 = (0.22)^3 = 0.0106$ or 1.1%



Estimated False Negative Rates

	Approach 1: Based on Studies	Approach 1: Based on Test Substances	Approach 2: Average Response Rate
All Data Used	4.9%	5.0%	1.1%
1 Animal Tested Excluded	7.1%	7.1%	1.5%
1 and 2 Animal Tested Excluded	10.3%	9.2%	2.4%

Conclusions

- Within the limits of the assumptions, the false negative rates ranged from 1.1% to 10.3%
- The false negative rate most likely to be representative of this group of corrosive substances is from 5 to 7%
- Additional data will allow us to refine these false negative rates, but it is unlikely to significantly change these estimates