Estimation of the Underprediction Rates for the *In Vivo* Rabbit Dermal Irritation Assay

NY Choksi[1,2], JH Haseman[3], DG Allen[1,2], RR Tice[1,2], and WS Stokes[2]

[1]Integrated Laboratory Systems, Inc., Research Triangle Park, NC; [2]National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, NC; and [3]NIEHS, Research Triangle Park, NC

Original Abstract

Alternative *in vitro* test methods proposed to substitute or replace an *in vivo* test method should provide equivalent or improved protection of human or animal health in order to gain regulatory and general acceptance. The ICCVAM and NICEATM are collaborating with the ECVAM to conduct a validation study of three *in vitro* dermal irritation assays. To assess the acceptability of these *in vitro* assays, an effort was undertaken to estimate the false negative rate of the *in vivo* test as defined by its ability to consistently identify irritants, mild irritants, and non-irritants according to the Globally Harmonized Classification Scheme. Data for 187 substances was obtained from the ECETOC database for skin irritation and corrosion. The distribution of rabbits with mean erythema or oedema scores of <1.5, between 1.5 and 2.3, or >2.3 was determined for each of the substances classified as "negative", "mild irritant" or "irritant". Since the true classification of each substance is unknown, a simplifying assumption was made that the results classified as "negative" for substances tested once only. For multiple-tested substances, the classification obtained from a majority of the studies was used. The analysis indicated: (1) the likelihood of a mild irritant being under-classified as a non-irritant was <5% when based on all substances and <10% when based on multiple-tested substances, (2) the underclassification rate of irritants as non-irritants was <1%, and (3) the underclassification rate of irritants as mild irritants ranged from 9-30%, depending on whether all substances or only multiple-tested substances were considered. Additional *in vivo* irritation data for studies using currently accepted procedures was requested from US federal agencies and industry. Appropriate data received will be added to the database and the false negative analysis refined. This evaluation emphasizes the need for high quality *in vivo* dermal irritation data that can be used to assess the performance of proposed new alternative test methods.

Introduction

For almost 60 years, the Draize *in vivo* rabbit skin irritation/corrosivity assay has been used to predict the ability of test substances to induce skin irritation and/or corrosion in humans (Draize et al., 1944). However, to date, only one study has been conducted to assess the reliability of this test method. In 1971, Weil and Scala reported on a study that evaluated the reproducibility of the Draize *in vivo* rabbit skin test method within and among twenty-four laboratories for ten reference test substances. The resulting analysis indicated that there was moderate intra-laboratory reproducibility but low inter-laboratory reproducibility. Weil and Scala (1971) concluded that subjective classification of the skin response was the primary reason for the low reproducibility among the participating laboratories. Despite the limited number of substances tested and the use of a 24-hour exposure (compared to the currently accepted test method protocol of no greater than a 4-hour exposure), the results of this study have frequently been used to support the opinion that the *in vivo* rabbit skin irritation/corrosivity test is unreliable. Recently, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) initiated a study to estimate the underprediction rate of a positive response in the current version of the *in vivo* rabbit skin irritation test. The results of this analysis will be used to help establish the performance characteristics that alternative *in vitro* test methods would need to exhibit to replace the traditional *in vivo* rabbit skin test method.

Materials and Methods

Database
Data compiled for this analysis are from irritation studies using the *in vivo* rabbit skin test method recommended by the Organisation for Economic Co-operation and Development (OECD) in Test Guideline (TG) 404. The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) published the original data for these studies (ECETOC 1995). The test substances and the *in vivo* rabbit skin irritation data provided in the ECETOC publication met the following selection criteria:

Test Substances:
Most substances were single chemical entities that are commercially available at known high consistent purity and are expected to be stable on storage. Some commercial chemical mixtures, manufactured and supplied to a specification that ensures a consistent purity also are included in the database.

In Vivo Rabbit Skin Irritation Data:
- Data generated since 1981 in studies carried out according to OECD TG 404 (OECD 1981, 1992) and following principles of Good Laboratory Practices
- Data was obtained in tests that:
  - normally used at least three albino rabbits tested at the same time
  - involved application of 0.5 mL or 0.5 g of the chemical to intact skin for not more than 4 hours of exposure
  - observations were made at least 24, 48, and 72 hours after patch removal or over sufficient duration to enable reversibility/irreversibility to be assessed
  - individual erythema and oedema scores (using Draize scoring scale) for each observation in each tested rabbit
  - test substances were tested undiluted except for those test substances where higher concentrations were expected to cause severe effects
- Data for test substances that were generated in multi-patch studies (i.e., more than one chemical was tested on the same rabbit at the same time) were identified separately

Relevant data were extracted and entered into Excel spreadsheets. Information compiled for each entry included test substance name or unique identifier, number of rabbits tested, and the number of animals that exhibited irritation reactions to the test substance. The database includes 164 different substances tested in 197 experiments; 23 substances were tested multiple times. The list of test substances used in this analysis is shown in **Table 2**.

Statistical Analysis
Two separate approaches were used to evaluate the underprediction rate of the *in vivo* irritation test method. In Approach 1, all of the data in the database (164 test substances evaluated in 197 experiments) were evaluated. Each test substance was weighted equally towards the total contribution to the estimate of the underlying distribution of animal responses and overall irritancy classifications. In Approach 2, only data from the 23 substances evaluated multiple times were evaluated. For all but four substances, the multiple studies were conducted in the same laboratory.

Irritancy Classification Decision Rules
To assess irritancy severity, the approach specified in the United Nations Global Harmonization System (GHS) (UN 2003) was used. In this classification system, results for erythema and oedema are averaged independently over three days (24, 48, and 72 hours post-application of the chemical) for each tested animal. As per the GHS system, for experiments that utilize three animals to assess irritancy, a chemical is classified as:
1. an irritant if two or more animals have an average erythema or average oedema score greater than 2.3
2. a mild irritant if two or more animals have an average erythema or average oedema score that ranges between 1.5 and 2.3
3. a non irritant if no more than one animal has an average erythema or average oedema score greater than 1.5.

For studies that utilized three animals and a one animal had a score less than 1.5, one animal had a score between 1.5 and 2.3, and one animal had a score greater than 2.3, the test substance was classified as a mild irritant.

For studies that utilized more than three animals, classification was based on a random sample of three of the animals tested. For example, if a study used 4 animals and 2 of the animals exhibited oedema scores >2.3 and the other 2 animals exhibited oedema scores between 1.5 and 2.3, then the substance would be classified half the time as an irritant and half the time as a mild irritant. For substances tested multiple times, if there was disagreement in irritation classification among the different experiments, the majority classification was used (i.e. the classification most often assigned). If no single classification was prevalent, the more severe classification was used.

Classification of Potential Outcomes Based on Score
In order to estimate the classification and underclassification rate of a test substance when three animals are tested, all possible experimental outcomes when three animals are tested need to be identified, overall irritancy classification assigned, and the contribution of the probability of each outcome to the overall classification and underclassification rate calculated. **Table 1** provides all 10 possible outcomes that may occur when three animals are tested to determine if a test substance is an irritant. Based on the decision rules noted above, each of these possible outcomes is assigned an irritancy classification (negative, mild irritant, or irritant). The final column in **Table 1** provides the probability used to calculate the contribution of each outcome to the overall underclassification rate.

Table 1. Table of Possible Experiment Classification Outcomes

Erythema or Oedema Score			Classification	Probability Calculation*
<1.5	1.5-2.3	>2.3		
3	0	0	Negative	$(P_N)^3$
2	1	0	Negative	$3P_N{}^2P_M$
2	0	1	Negative	$3P_N{}^2P_I$
1	1	1	Mild Irritant	$6P_NP_MP_I$
1	2	0	Mild Irritant	$3P_NP_M{}^2$
0	3	0	Mild Irritant	$(P_M)^3$
0	2	1	Mild Irritant	$3P_M{}^2P_I$
1	0	2	Irritant	$3P_N{}^2P_I$
0	1	2	Irritant	$3P_MP_I{}^2$
0	0	3	Irritant	$(P_I)^3$

*P_N represents the probability that a treated animal will exhibit an erythema or oedema score that is less than 1.5, P_M represents the probability that a treated animal will exhibit an erythema or oedema score that is between 1.5 and 2.3, P_I represents the probability that a treated animal will exhibit an erythema or oedema score that is greater than 2.3.

Calculation
To estimate the underprediction rate of the *in vivo* rabbit skin irritation test method, the distribution of animals within each irritancy class (i.e., irritant, mild irritant, negative) with scores less than 1.5, between 1.5 and 2.3, and greater than 2.3 was calculated. Using this distribution and the possible experimental outcomes provided in Table 1, the individual probabilities are calculated for each possible outcome for a specific irritancy classification. For test substances that are identified as irritants (according to the GHS classification system), the probabilities of the possible experimental outcomes that would lead to an irritant classification (last three rows in Table 1) are determined using the distribution of animal scores. The individual probabilities are then totaled to provide an overall prediction or underprediction rate.

Assumptions
The following assumptions were used in the statistical analysis:
1. All studies in the database are independent.
2. The *in vivo* rabbit skin irritation test method protocol used (except for the number of animals) for all substances followed OECD TG 404.
3. Only irritating substances induced a positive response (i.e., there were no false positive responses).
4. For each irritating substance tested once, the observed positive response is accurate.
5. For test substances tested multiple times and result in the same classification each time, the classification is "correct".
6. The distribution of substances included in this evaluation, in terms of dermal irritation, is a true representation of the total population of irritating substances.

Table 2

List of Test Substances Used in Analysis

Test #	Test Lab*	Test Substance	Tested	Score <1.5	Score 1.5-2.3	Score >2.3	Irritancy Classification
1	A	1,1,1-trichloroethane	3	0	0	3	Irritant
1	A	1,13-tetradecadiene	4	2	2	0	Mild Irritant
1	A	1,3-dibromopropane	3	0	3	0	Mild Irritant
1	B	1,4-cineole (eucalyptol)	3	3	0	0	Negative
1	C	1,5-hexadiene	3	3	0	0	Negative
1	A	1,6-bromohexane	3	3	0	0	Negative
1	A	1,9-decadiene	3	0	3	0	Mild Irritant
1	D	10-undecenoic acid	4	0	4	0	Mild Irritant
1	A	1-bromo-2-chloroethane	3	1	2	0	Mild Irritant
1	A	1-bromo-4-chlorobutane	3	3	0	0	Negative
1	A	1-bromo-4-fluorobenzene	3	3	0	0	Negative
1	A	1-bromohexane	3	0	3	0	Mild Irritant
1	A	1-bromopentane	3	0	1	2	Irritant
1	E	1-decanol	4	0	2	2	Irritant
1	E	1-formyl-1-methyl-4(4-methyl-3-pentene-1-yl)-3-cyclohexene	4	0	4	0	Mild Irritant
1	F	2,3-dichloropropionitrile	3	0	3	0	Mild Irritant
1	D	2,4-decadienal	4	0	1	3	Irritant
1	D	2,4-dimethyl-3-cyclohexen-1-carboxaldehyde	4	0	4	0	Mild Irritant
1	B	2,4-dimethyltetrahydrobenzaldehyde	4	0	4	0	Mild Irritant
1	G	2,4-dinitromethylaniline	3	3	0	0	Negative
1	F	2,4-hexadienal	4	0	0	4	Irritant
1	I	2,4-xyldine	3	3	0	0	Negative
1	A	2,5-methylene-6-propyl-3-cyclohexen-carbaldehyde	4	1	3	0	Mild Irritant
1	J	2,6-dimethyl-2,4,6-octatriene	4	0	4	0	Mild Irritant
1	D	2,6-dimethyl-4-heptanol	3	3	0	0	Negative
1	A	2-bromobutane	3	3	0	0	Negative
1	A	2-bromopropane	3	3	0	0	Negative
1	A	2-ethoxyethyl methacrylate	3	1	2	0	Mild Irritant
1	D	2-ethylhexanal	4	0	3	1	Mild Irritant
1	A	2-ethylhexylcocoate	3	3	0	0	Negative
1	L	2-ethylhexylpalmitate	3	3	0	0	Negative
1	M	2-fuorotoluene	3	3	0	0	Negative
1	F	2-methoxyethyl acrylate	3	0	0	3	Irritant
1	N	2-methyl butyric acid	4	0	2	2	Irritant
1	E	2-methyl-4-phenyl-2-butanol	4	0	4	0	Mild Irritant
1	E	2-phenylpropionaldehyde	4	0	4	0	Mild Irritant
1	A	2-tertiarybutyl phenol	6	0	0	6	Irritant
1	A	3,3'-thiodipropionic acid	4	4	0	0	Negative
1	E	3,7-dimethyl-2,6-nonadien-1-a	4	0	1	3	Irritant
1	A	3-chloro-4-fluoronitrobenzene	6	3	3	0	Mild Irritant
1	F	3-diethylaminopropionitrile	3	3	0	0	Negative
1	F	3-mercapto-1-propanol	6	0	5	1	Negative
1	A	3-methylbutyraldehyde	3	0	3	0	Mild Irritant
1	A	4-(methylthio)benzaldehyde	3	0	0	3	Irritant
1	A	4,4-methylene bis(2,6-ditertiary butyl phenol)	3	3	0	0	Negative
1	A	4-amino-1,2,4-triazole	4	4	0	0	Negative
1	B	4-tricyclo-decylidene-8-butanal	4	0	3	1	Mild Irritant
1	K	55/45 caprylic/capric acid	3	0	0	3	Irritant
1	K	60/40 caprylic/capric acid	3	0	0	3	Irritant
1	O	65/35 caprylic/capric acid	3	0	0	3	Irritant
1	K	65/35 caprylic/capric acid	3	0	0	3	Irritant
1	B	6-butyl-2,4-dimethyldihydropyrane	4	1	3	0	Mild Irritant
1	Unk	70/30 oleine/capric acid	3	0	2	1	Mild Irritant
1	Unk	80/20 oleine/capric acid	3	0	0	3	Irritant
1	Unk	90/10 oleine/capric acid	3	0	0	3	Irritant
1	E	acetaldehyde	4	0	0	4	Irritant
1	H	allyl heptanoate	4	1	3	0	Mild Irritant
1	E	allyl phenoxyacetate	4	0	4	0	Mild Irritant
1	H	alpha-ionol	4	4	0	0	Negative
1	E	benzyl acetone	4	2	2	0	Mild Irritant
1	H	beta-ionol	4	4	0	0	Negative
1	K	caprylic acid	3	0	0	3	Irritant
1	D	cinnamaldehyde	4	0	4	0	Mild Irritant
1	P	cinnamon leaf oil	4	0	2	2	Irritant
1	D	cis-cyclooctene	4	0	2	2	Irritant
1	E	cis-jasmone	4	4	0	0	Negative
1	A	citrathal	4	0	4	0	Mild Irritant
1	B	clove leaf oil	4	0	3	1	Mild Irritant
1	D	diacetyl	4	4	0	0	Negative
1	A	dichloromethane	4	0	4	0	Mild Irritant
1	A	dimethyl disulphide	6	1	5	0	Mild Irritant
1	A	di-n-propyl disulphide	3	0	3	0	Mild Irritant
1	Unk	erucamide	4	4	0	0	Negative
1	D	ethyl tiglate	4	0	4	0	Mild Irritant
1	A	ethyl trimethyl acetate	6	6	0	0	Negative
1	F	ethylthioethyl methacrylate	3	0	3	0	Mild Irritant
1	F	ethyltriglycol methacrylate	3	0	3	0	Mild Irritant
1	E	eugenol	4	4	0	0	Negative
1	M	fluorobenzene	3	3	0	0	Negative
1	H	geranyl dihydrolinalool	4	2	2	0	Mild Irritant
1	E	geranyl linalool	4	4	0	0	Negative
1	K	glycerol tri-isostearate	3	3	0	0	Negative
1	E	glycerol triundecanoate	3	3	0	0	Negative
1	E	glycolbromoacetate	3	0	0	3	Irritant
1	D	guaiacol (o-methoxyphenol)	4	1	3	0	Mild Irritant
1	J	heptanal	4	1	3	0	Mild Irritant
1	N	heptyl butyrate	4	1	3	0	Mild Irritant
1	F	hydrogenated tallow amine	3	1	0	2	Irritant
1	E	hydrogenated tallow propylene diamine	6	6	0	0	Negative
1	D	isobutyraldehyde	4	4	0	0	Negative
1	Unk	isolongifolene ketone	4	4	0	0	Negative
1	K	isopropanol	3	2	1	0	Negative
1	K	isopropylisostearate	3	3	0	0	Negative
1	K	isopropylmyristate	3	3	0	0	Negative
1	K	isopropylpalmitate	3	2	1	0	Negative
1	K	isostearic acid	3	0	1	2	Irritant
1	E	lauric acid	4	3	1	0	Negative
1	E	linalool oxide	4	4	0	0	Negative
1	B	litsea cubeba oil	4	0	2	2	Irritant
1	N	m-chloronitrobenzene	4	4	0	0	Negative
1	A	methyl 2-methylbutyrate	4	4	0	0	Negative
1	A	methyl caproate	3	0	3	0	Mild Irritant
1	A	methyl laurate	3	3	0	0	Negative
1	A	methyl lavender ketone	4	0	3	1	Mild Irritant
1	A	methyl linoleate	3	1	2	0	Mild Irritant
1	A	methyl palmitate	3	3	0	0	Negative
1	A	methyl stearate	3	3	0	0	Negative
1	E	methyl trimethyl acetate	3	3	0	0	Negative
1	S	N,N-dimethylbenzylamine	3	0	0	3	Irritant
1	N	n-butyl propionate	4	4	0	0	Negative
1	F	n-decylidene methyl anthranilate	4	4	0	0	Negative
1	A	nonanal	4	0	4	0	Mild Irritant
1	P	oleyl propylene diamine dioleate	3	0	2	1	Mild Irritant
1	A	parsley herb oil	4	0	4	0	Mild Irritant
1	E	perilla oil	4	0	4	0	Mild Irritant
1	E	phenethyl bromide	3	3	0	0	Negative
1	B	pimenta leaf (allspice) oil	4	0	0	4	Irritant
1	Q	p-isopropylphenylacetaldehyde	4	4	0	0	Negative
1	B	p-mentha-1,8-dien-7-ol	4	0	2	2	Irritant
1	D	potassium hydroxide	3	0	0	3	Irritant
1	K	p-tert-butyl dihydrocinnamaldehyde	4	2	2	0	Mild Irritant
1	A	p-tolyl alcohol (p-methyl phenol)	4	0	0	4	Irritant
1	K	salicylaldehyde	4	1	3	0	Mild Irritant
1	A	soap from 20/80 coconut palm	3	3	0	0	Negative
1	A	soap from 20/80 coconut/tallow	3	3	0	0	Negative
1	Unk	sodium bicarbonate	3	3	0	0	Negative
1	A	sodium bisulphite	3	3	0	0	Negative
1	Unk	sodium carbonate	3	1	2	0	Mild Irritant
1	Unk	sodium chloride	3	3	0	0	Negative
1	R	sodium lauryl sulphate 20% aq	3	0	0	3	Irritant
1	A	sodium lauryl sulphate 50% aq	3	0	0	3	Irritant
1	K	sodium metasilicate	3	0	0	3	Irritant
1	A	sodium metasilicate	3	1	1	1	Mild Irritant
1	P	stearyl alcohol	3	1	2	0	Mild Irritant
1	B	tea tree oil	4	0	3	1	Mild Irritant
1	P	thyme oil, red	4	0	0	4	Irritant
1	N	tetrahydrogeranial	4	0	4	0	Mild Irritant
1	A	thyme oil, red	4	0	0	4	Irritant
1	F	tonalid	3	3	0	0	Negative
1	A	trichloroethylene	3	0	0	3	Irritant
1	A	undecylenate acid	3	1	2	0	Mild Irritant
2	E	alpha terpinyl acetate	4	0	3	0	Mild Irritant
2	E	alpha terpinyl acetate	4	0	1	0	Irritant
3	E	alpha terpinyl acetate	4	1	3	0	Mild Irritant
2	E	alphahexyl cinnamic aldehyde	3	0	2	1	Mild Irritant
3	E	alphahexyl cinnamic aldehyde	4	1	3	0	Mild Irritant
1	E	alphahexyl cinnamic aldehyde	4	4	0	0	Negative
2	E	alphaterpineol	4	0	4	0	Mild Irritant
1	E	alphaterpineol	3	0	3	0	Mild Irritant
3	E	alphaterpineol	4	4	0	0	Negative
1	N	benzyl acetate	3	2	1	0	Mild Irritant
2	E	benzyl acetate	3	3	0	0	Negative
1	N	benzyl alcohol	4	3	0	0	Negative
2	N	benzyl alcohol	4	4	0	0	Negative
1	N	benzyl benzoate	4	2	2	0	Mild Irritant
2	E	benzyl benzoate	3	3	0	0	Negative
1	N	benzyl salicylate	3	0	3	0	Mild Irritant
2	E	benzyl salicylate	4	0	4	0	Mild Irritant
1	E	cinnamyl alcohol	4	0	4	0	Negative
2	E	cinnamyl alcohol	6	6	0	0	Negative
1	E	cyclamen aldehyde	3	0	0	3	Irritant
3	E	cyclamen aldehyde	4	0	0	4	Irritant
2	N	cyclamen aldehyde	3	0	0	3	Irritant
4	N	cyclamen aldehyde	4	0	0	4	Irritant
1	E	di-citronellol	3	0	2	1	Mild Irritant
2	E	di-citronellol	4	0	4	0	Mild Irritant
1	E	di-citronellol	3	0	3	0	Mild Irritant
2	E	diethyl phthalate	3	3	0	0	Negative
1	A	diethyl phthalate	4	4	0	0	Negative
2	N	dihydromercenol	4	4	0	0	Negative
1	E	dihydromercenol	4	4	0	0	Negative
3	N	dihydromercenol	3	0	0	3	Irritant
2	N	dimethylbenzylcarbinylacetate	3	0	0	3	Irritant
1	N	dimethylbenzylcarbinylacetate	6	0	5	1	Mild Irritant
2	N	dipropylene glycol	4	4	0	0	Negative
1	N	dipropylene glycol	4	4	0	0	Negative
2	E	d-limonene	4	0	4	0	Mild Irritant
1	E	d-limonene	4	0	4	0	Mild Irritant
2	E	geraniol	4	2	2	0	Mild Irritant
1	E	geraniol	4	0	4	0	Mild Irritant
2	E	geraniol	4	0	3	1	Mild Irritant
1	N	hexyl salicylate	3	0	3	0	Mild Irritant
2	E	hexyl salicylate	4	0	4	0	Mild Irritant
4	E	hexyl salicylate	3	0	3	0	Mild Irritant
1	E	hydroxycitronellal	3	3	0	0	Negative
2	E	hydroxycitronellal	4	4	0	0	Negative
3	N	isobornyl acetate	4	4	0	0	Negative
1	E	isobornyl acetate	3	0	3	0	Mild Irritant
2	E	lilestralis/lilial	3	0	2	1	Mild Irritant
1	E	lilestralis/lilial	4	1	2	0	Irritant
2	N	linalol	4	4	0	0	Negative
1	E	linalol	3	3	0	0	Negative
2	N	linalol	4	4	0	0	Negative
1	E	linalol	4	4	0	0	Negative
2	E	linalyl acetate	3	0	1	2	Irritant
1	E	linalyl acetate	3	0	3	0	Mild Irritant
1	E	phenylethyl alcohol	3	3	0	0	Negative
2	N	phenylethyl alcohol	3	2	1	0	Negative

*A: Aldrich; B: Quest; C: Fluka; D: IFF; E: Givaudan-Roure; F: Elf Atochem; G: Hoechst AG; H: Bedoukian; I: Bayer AG; J: BBA; K: Unichema International; L: DS Industries; M: Reidel; N: Firmenich; O: Procter & Gamble; P: Prodarom; Q: MallindKrodt; R: Fisher Scientific; S: BRI; Unk: Unknown

Results: Approach 1

All Test Substances in Database Are Given Equal Weight

Table 3. Distribution of Animal Scores for Each Irritancy Class

Estimated Probability of ...	True Classification of Test Substance		
	Negative	Mild Irritant	Irritant
An animal scoring <1.5	95.70%	14.20%	0.70%
An animal scoring between 1.5 and 2.3	3.90%	81.50%	19.20%
An animal scoring >2.3	0.40%	4.20%	80.10%

Table 4. Estimated Probabilities of Classification

Our Classification of Test Substance	True Classification of Test Substance		
	Negative	Mild Irritant	Irritant
Negative	99.46%	5.50%	0.01%
Mild Irritant	0.54%	94.00%	10.30%
Irritant	<0.01%	0.50%	89.69%

Example Calculation of Estimated Probability of Negative Classification

For test substances that are non irritants (Negative Test Substances), there are three potential animal outcomes:

Erythema or Oedema			Classification	Probability Calculation
<1.5	1.5-2.3	>2.3		
3	0	0	Negative	$(P_N)^3$
2	1	0	Negative	$3P_N{}^2P_M$
2	0	1	Negative	$3P_N{}^2P_I$

Using the distributions calculated in **Table 3**, the likelihood of an animal test correctly classifying a test substance as Negative is:

$$(P_N)^3 + 3P_N{}^2P_M + 3P_N{}^2P_I = (0.957)^3 + [3(0.957)^2(0.039)] + [3(0.957)^2(0.004)] = 0.9946 \text{ or } \mathbf{99.46\%}$$

Results: Approach 2

Only Test Substances Tested Multiple Times Used

Table 5. Distribution of Animal Scores for Each Irritancy Class

Estimated Probability of ...	True Classification of Test Substance		
	Negative	Mild Irritant	Irritant
An animal scoring <1.5	91.70%	11.60%	0%
An animal scoring between 1.5 and 2.3	8.30%	79.50%	42.40%
An animal scoring >2.3	0%	8.90%	57.60%

Table 6. Estimated Probabilities of Classification

Our Classification of Test Substance	True Classification of Test Substance		
	Negative	Mild Irritant	Irritant
Negative	98.00%	3.70%	0%
Mild Irritant	2.00%	94.00%	38.70%
Irritant	0%	2.00%	61.30%

Table 7

Estimated Underprediction Rates of *In Vivo* Dermal Irritation Test Method

Outcome	Approach 1	Approach 2	Average
Underprediction of Irritant as Mild Irritant	10.30%	38.70%	24.50%
Underprediction of an Irritant as Negative	0.01%	0%	<0.01%
Underprediction of a Mild Irritant as a Negative	5.50%	3.70%	4.60%
Underprediction of an Irritant and Mild Irritant as a Negative	5.51%	3.70%	4.60%

Conclusions and Discussion

- Within the limits this evaluation, the underprediction rate of:
  - an irritant as a mild irritant ranged from 10.30% to 38.70%,
  - an irritant as a non irritant ranged from 0% to 0.01%, and
  - a mild irritant as a non irritant ranged from 5.50% to 3.70%.

- Based on these data, the likelihood that an irritant would be misclassified as a non-irritant is less than 0.01%.

- Approach 2, which is based on largely within lab reproducibility, is considered the best approach for assessing the underprediction rate for the traditional *in vivo* rabbit skin irritation test method. However, the database for substances tested multiple times is limited in number and the distribution of substances is weighted toward mild irritants, which would be expected to have the greatest lack of reproducibility. Therefore, the calculated values in this approach are likely an overestimation of the underprediction rate.

Acknowledgements

References

Draize JH, Woodward G, Calvery HO. 1944. Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J. Pharmacol. Exp. Therap. 82:377-389.

ECETOC. 1995. Skin irritation and corrosion: reference chemicals data bank. Technical Report No. 66. Belgium.

OECD. 1981. OECD Test Guideline 404 Skin Irritation/Corrosion. OECD, Paris.

OECD. 1992. OECD Test Guideline 404 Skin Irritation/Corrosion (1992 Revision). OECD, Paris.

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.

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.

More information on ICCVAM and NICEATM can be accessed at:
http://iccvam.niehs.nih.gov