LIMITATIONS IN CHARACTERIZATION OF ACUTE DERMAL IRRITATION WITH REDUCED NUMBER OF TEST ANIMALS

#1854

T.P. Kellner, J.R. Corlett, M.E. Hansen and P. Leung

Medical Toxicology Branch, Department of Pesticide Regulation., Cal/EPA, Sacramento, CA 95814

INTRODUCTION

Toxicology studies for pesticides registered in California are evaluated by the California Dept. of Pesticide Regulation (CDPR) according U.S. EPA's Health Effect Test Guidelines, which have been harmonized with those established by the Organization for Economic Co-operation and Development (OECD). As these guidelines have evolved, reduction or replacement of animal testing has been emphasized. Such was the case for the primary dermal irritation test (traditionally, six animals minimum); OECD guidelines now recommend using only three rabbits (OECD, 1981). Analyses to test the effect of reduced animal numbers have generally concluded that less than six rabbits would not be suitable for evaluating dermal irritation based solely on qualitative classification (*e.g.*, three-rabbit-derived classifications only agreed with the six-rabbit classifications for 70% of the chemicals tested; Derelanko, *et al*., 1993).

Current guidelines specify that erythema/eschar and edema are each scored on scale from 0 (no reaction) to 4 (maximum possible). The sum of erythema and edema scores at 1, 24, 48 and 72 hours is divided by (# test sites x 4 scoring intervals; normally 12) to obtain the Primary Dermal Irritation Index (PDII). These values, together with the nature and reversibility of the responses, are used to obtain a Toxicity Category (Tox. Cat.). Toxicity categories determine signal word, precautionary and first aid labeling for home and commercial-use products and restricted-entry intervals (REI) and personal protective equipment (PPE) for agricultural workers.

Classification System[2]

PDII	Classification	Toxicity Category
0-1.9	non– to slightly irritating	IV
2.0-5.0	moderately irritating	III
5.0-7.0	severely irritating	II
>7.0	corrosive	I

As stated in the U.S. EPA Health Effect Test Guidelines, the dermal irritation scores should be evaluated in conjunction with the nature and reversibility or otherwise of the responses observed. **The individual scores do not represent an absolute standard for the irritant properties of a material. They should be viewed as reference values which are only meaningful when supported by a full description and evaluation of the observations (U.S. EPA).** With only three animals per study, difficulties can arise in data interpretation and ultimately in assignment of Toxicity Category. For example, too much reliance can be placed on quantitative analysis (*i.e.,* PDII score) without adequate qualitative analysis (*e.g.*, demonstration of reversibility). When six or more animals are used, the reaction of a single animal that does not correlate with the remaining animals may be considered an isolated "outlier". When only three animals are tested, all animals must be considered when CDPR establishes a Toxicity Category.

REFERENCES

[1]Derelanko, M.J., Finegan, C.E., and Dunn, B.J. (1993). Reliability of Using Fewer Rabbits to Evaluate Dermal Irritation Potential of Industrial Chemicals. *Fundam. Appl. Toxicol.* 21, 159-163

[2]Draize, J., Woodward, G., and Calvery, O. (1944). Methods for the Study of Irritation and Toxicity of Substances Applied Topically to the Skin and Mucous Membrane. *J. Pharmacol. Exp. Ther.* 82, 377-390.

[3]U.S. Environmental Protection Agency. Health Effects Test Guidelines. OPPTS 870.2500 Acute Dermal Irritation. August, 1998.

[4]Organization for Economic Cooperation and Development, OECD Guideline for the Testing of Chemicals. Guideline 404: Acute Dermal Irritation/Corrosion. Adopted: April 24, 2002.

METHODS

Dermal irritation studies (summarized in Tables 1 - 4), submitted to CDPR to support product registrations, serve as case studies. The dorsal area and the trunk of rabbits were clipped free of hair and a single application of 0.5 ml of the test article was applied to a 6 cm^2 intact dose site and covered with a 1 in. square, 4-ply gauze pad. This area and the trunk were then wrapped with semi-occlusive tape for a 4-hour exposure. The pads were removed and the test sites were gently cleansed of residual test substance. Irritation was scored according the Draize scoring system, ranging from no reaction (0) to maximum (4)[1].

TABLE 1: Individual Skin Irritation Scores - Erythema/Edema

Animal No.	Sex	Time After Patch Removal						
		30-60 min	24 hrs	48 hrs	72 hrs	Day 7	Day 10	Day 14
1	M	3/2	3/2	3/2	3/2	2/2[1]	2/1[2]	1/1[2]
2	M	3/2	3/2	3/2	2/1	2/1[1]	2/1[2]	1/1[2]
3	M	3/2	3/2	3/2	2/1	2/1[1]	2/1[2]	1/0[2]
Total		9/6	9/6	9/6	7/4	6/4	6/3	3/2
Mean		3.0/2.0	3.0/2.0	3.0/2.0	2.3/1.3	2.0/1.3	2.0/1.0	1.0/0.7

[1]Hyperkeratosis
[2]Desquamation

PDII = 4.7

TABLE 1. The registrant and US EPA reviewer considered these data reflective of Toxicity Category III; **Rationale**: 1. PDII of 4.7; 2. Animals showed recovery from irritation on days 7, 10 and 14. In contrast, the CDPR reviewer assigned Toxicity Category II for three reasons:

- With only three animals tested, additional focus was given to the most extreme irritation response noted in the study. One animal showed score 2 erythema and edema up to day 7, with score 1 being noted for these signs at study termination.
- Reversibility of irritation was not fully demonstrated. At day 14, all animals showed erythema and desquamation and 2 of 3 showed persistent edema. This indicated a degree and persistence of irritation beyond that of the usual "moderate" Tox. Cat. III dermal irritant.
- The PDII of 4.7 is close to the threshold for Category II (5 to 7).

Discussion: If the observations had been extended beyond 14 days, the CDPR review may have resulted in Tox. Cat. III (*i.e.,* possible reversibility demonstrated). The lack of data beyond 14 days, the near-threshold PDII score and the persistence of significant irritation prompted the Toxicity Category II classification by CDPR.

TABLE 2: Individual Skin Irritation Scores - Erythema/Edema

Animal No.	Sex	Time After Patch Removal						
		30-60 min	24 hrs	48 hrs	72 hrs	Day 7	Day 10	Day 14
1	M	1/0	2/0	2/0	2/0	2/1	1/0	0/0
2	M	1/0	2/1	2/1	2/1	2/1	2/0	2/0
3	M	1/0	1/0	0/0	-	-	-	-
Total		3/0	5/1	4/1	4/1	4/2	3/0	2/0
Mean		1.0/0.0	1.7/0.3	1.3/0.3	1.3/0.3	1.3/0.7	1.0/0.0	0.7/0.0

PDII = 1.6

TABLE 2. After evaluation, CDPR determined that the primary dermal irritation study was unacceptable due to equivocal data. One rabbit (#2) had moderate erythema (score 2) up to the final day of the study (day 14) in contrast to another rabbit (#3), which showed complete clearing by 48 hours. The duration of dermal irritation in the third rabbit (#1) was intermediate to these extremes, with score 2 erythema reported up to day 7, score 1 at day 10 and clearing by day 14. Assignment of a Toxicity Category based on this study alone would be very speculative**.** A rebuttal by the registrant was based on a guidance document from the US EPA (Rejection Rate Analysis "Conduct of Acute Studies" dated Sept. 1997). This includes the following: "Observations should be continued until irritation has subsided (no scores other than "1" for dermal erythema and/or edema) or 14 days had been reached. If there is still irritation at 14 days the study may be ended. The exception to this rule would be ending a study early when it is evident that it will be placed into Toxicity Category I".

Discussion: Additional irritation data on file with CDPR using a similar test article were cited (data not shown). In a study in rats, there were indications of dermal irritation in three females (remaining seven animals had no reported dermal abnormalities). One of these females had a lesion on the back (right side), appearing on day 2 and clearing by day 7. Another had a lesion on the back (left side) on day 2, with clearing by day 9. A third female had redness on the back on day 2 and clearing on day 3. This rat also had redness on day 4, with clearing by the next day. These limited data indicated that dermal irritation was slight to moderate for the test article and that clearing of irritation occurred by day 9 after exposure. This was consistent with the rabbit that showed score 1 erythema up day 10 and complete clearing by day 14 (*i.e.,* the "intermediate" response). The combined data were reflective of Toxicity Category III.

TABLE 3: Individual Skin Irritation Scores - Erythema/Edema

Animal No.	Sex	Time After Patch Removal						
		30-60 min	24 hrs	48 hrs	72 hrs	Day 7	Day 10	Day 14
1	M	2/3	3/3	3/3	3/3	3/2	3/2	2/1
2	M	2/3	3/3	3/2	3/2	2/2	2/2	1/0
3	M	2/3	3/3	3/2	3/2	2/2	2/2	2/0
Total		6/9	9/9	9/7	9/7	7/6	7/6	5/1
Mean		2/3	3/3	3/2.3	3/2.3	2.3/2	2.3/2	1.7/0.3

PDII = 5.4

TABLE 3. CDPR considered these data clearly reflective of Toxicity Category II using methods of analyses prescribed in the original Draize publications, with the dermal irritation hazard based on the quantitative severity (PDII of 5.4) as well as qualitative considerations (*e.g.,* persistence of significant dermal irritation through day 14). However, the registrant was specifically instructed by the Reg. Div. Of US EPA / OPP to label this product as a Toxicity Category III dermal irritant.

The testing laboratory defined a score of 3 for erythema as "moderate". If any severe irritation was noted, then they would give it score 4. This was despite the testing laboratory's published grading scale, in which a score of 3 for erythema was considered "moderate to severe" and the **study report describing the test article as "severely irritating".**

Discussion: Evaluation Scientists at US EPA stated that they looked primarily at the 72-hour score and saw "moderate" (score 3) redness and "moderate" (score 2 and 3) edema and considered this reflective of Toxicity Category III. CDPR requested further clarification; according to US EPA guidance, the primary focus is the dermal reaction at 72 hours. According to their system, score 3 erythema and edema at 72 hours would be considered "moderate" irritation and unless the lesion gets worse thereafter, the resulting assignment is Toxicity Category III. Obviously, US EPA and CDPR dermal irritation evaluation methodology is not harmonized.

TABLE 4: Individual Skin Irritation Scores - Erythema/Edema

Animal No.	Sex	Time After Patch Removal						
		4.5 hrs	24 hrs	48 hrs	72 hrs	Day 7	Day 10	Day 14
1	F	2/2	2/2	1/0	0/0	-	-	-
2	F	2/2	2/1	1/0	0/0	-	-	-
3	F	2/2	2/1	1/0	0/0	-	-	-
Total		6/6	6/4	3/0	0/0	-	-	-
Mean		2.0/2.0	2.0/1.3	1.0/0.0	0.0/0.0	-	-	-

PDII = 2.08

TABLE 4. The CDPR reviewer assigned Toxicity Category III, based mainly on the PDII score of 2.08, while the US EPA, the study author and the registrant concluded that Tox. Cat. IV was appropriate, based on the resolution of irritation at 72 hours. This was despite the fact that the US EPA reviewer and the study author described the formulation qualitatively as "moderately irritating", a characterization normally reserved for a Tox. Cat. III dermal irritant, not a Tox. Cat. IV "non- or slight irritant".

Discussion: The registrant stated the following: "None of the animals at any point in the test exhibited a score indicative of a Tox. Cat. III hazard (3 or higher)." This indicated confusion between an individual irritation score and an overall Toxicity Category. The registrant went on to concede the following: "The resulting PDII is very slightly higher than 2, which according to OECD guidelines would indeed suggest a Tox. Cat. III classification". These conflicting viewpoints were reminiscent of the previous example where differing methods of analysis and small sample size have led to different Toxicity Category determinations.

CONCLUSION

Strategies that may be employed to reduce the uncertainty associated with decreased group size in studies of this type include the following:

- In the case of persistent irritation, observations should be continued beyond 14 days to demonstrate possible reversibility of skin damage. As illustrated in Table #1, this additional data could have made the difference between assignment of Toxicity Category II and III by CDPR.
- With borderline or equivocal data (Table #2), refer to dermal irritation results for similar formulations and dermal irritation data from acute dermal toxicity and dermal sensitization studies (*i.e.,* range-finding data to obtain the highest non-irritating concentration or HNIC) that accompany the dermal irritation study for standard product registrations.
- With the use of fewer animals, standardization and harmonization of the classification of dermal irritants between registrants, regulators and test laboratories is critical, as illustrated by Table # 3 and # 4. If the use of the quantitative PDII scores in conjunction with qualitative analysis (*i.e.,* the norm at CDPR) is being supplanted by a 72-hour "decision-point" strategy (as favored by US EPA), then this new approach needs to be standardized and documented.
- As demonstrated in Table # 4, PDII differences as little as 0.1 obtained from fewer animals (i.e., with 2.1 receiving a moderate classification and 2.0 classified as mild) can result in disagreement with regard to appropriate Toxicity Category. The inability of the three-rabbit study to reliably distinguish between negligible and non-irritants should be clearly understood. Where the need to detect negligible irritants is critical (*e.g.,* disinfectant wipes or dermal applied insect repellents), the use of additional rabbits should be considered (Derelanko, *et al.,* 1993).
- Thorough wiping of excess test article with an "appropriate" solvent after exposure can reduce variability of irritation results (data not shown). Residual test material can intensify dermal irritation in primary dermal irritation studies and lead to false positive data in dermal sensitization studies.