

Reproducibility of allergy epicutaneous test techniques

S. P. Galant, M.D., and H. I. Maibach, M.D. *Orange, Calif.*

Three commonly used epicutaneous allergy skin test techniques, pinprick, chalazion knife, and Berkeley scarifier, were studied for size of response and reproducibility in the top, middle, and bottom regions of the back. All methods showed large variability, which was strongly influenced by the region of the back selected and the size of the wheal and flare reaction. Greater reproducibility was found with the larger skin test responses using the pinprick and scarifier techniques in the middle region of the back. The flare reaction was slightly less variable than the wheal response. Reproducibility and reaction size did not change significantly during the 4 week study period.

Since Blackley's¹ description of abrading the skin with a pollen to which he was clinically sensitive resulting in a wheal and flare response in 1873, various allergy skin test methods have been devised. Both epicutaneous and intradermal techniques have their advocates. The fact that there is no uniformity of method attests to the lack of a simple, reproducible technique that is universally acceptable.

Little data have been reported on the reproducibility* of epicutaneous and intradermal techniques. Reliability of a method is usually based on correlation with clinical history and previously performed skin tests.² While correlation with clinical history is important, it does not reflect the accuracy of the technique itself.

Although epicutaneous tests are thought to be a crude technique, several studies^{3, 4} have shown good reproducibility when certain conditions are met. The intradermal method, probably more reproducible,⁵ is less desirable for children because it causes more discomfort and is more dangerous.⁶ Furthermore, it does not correlate as well as epicutaneous methods with inhalation provocation testing in asthmatic children,⁷ and it appears to be clinically significant only when dilute solutions of allergen extracts are used.⁸

The back is thought to be the most sensitive⁹ and reproducible⁴ area of the body for skin testing, but the latter is based on only one study in which the Prausnitz-Küstner (P-K) reaction, not direct skin testing, was used. Only wheal responses were analyzed. The clinical impression, however, is that the back is not entirely uniform, and that certain regions are preferable for testing. Since the back permits more tests to be performed than the arms, it is of critical im-

Received for publication June 13, 1972.

Reprint requests to: S. P. Galant, M.D., Department of Pediatrics, University of California, Irvine, 101 City Drive South, Orange, Calif. 92668.

*Reproducibility is the inverse of variability.

portance to document the reproducibility of wheal and flare responses by the direct skin test method.

Three common epicutaneous techniques in clinical allergy use are the pinprick, scarifier, and chalazion knife methods.^{10, 11} Of these, only the pinprick^{3, 4} and chalazion knife¹² techniques have been carefully studied using quantitative criteria. Studies employing semiquantitative measurements (1+ to 4+) cannot be utilized since these parameters are inadequate for statistical analysis of comparative tests¹³; in addition, there are no universal standards for these designations.

This study directs itself to the following questions: (1) reproducibility of the 3 common epicutaneous techniques, (2) reproducibility of different regions of the back, and (3) size of reactions with the different techniques, and responses in different regions of the back.

METHODS

Subjects

Fifteen atopic Caucasian adults with known positive reactions to inhalant allergens were chosen.

Technique

The subject was tested with each technique in *duplicate* using 2 allergens to which he was known to be sensitive, a major and minor allergen,* a histamine diphosphate control (histamine base 1:1,000 w/v), and a phenol glycerin control (50 per cent glycerin-0.5 per cent phenol concentration) on the back. Allergens in 50 per cent glycerin were used in full concentration.†

Three different epicutaneous methods were used: the chalazion knife, pinprick, and Berkeley scarifier. All tests were performed by the same investigator (S.P.G.).

Sites on the back were equally divided with a marking pen into top, middle, and bottom regions. Test material and techniques were randomized as to site of testing each of the 4 weeks of the study. However, each of the tests in the duplicate pair was performed in the same region.

The wheal and flare responses were photographed with a fixed distance camera using Kodachrome II slide film 20 minutes after testing.¹⁴ After the film was developed, the lesions were magnified in a microfilm projector to simplify tracing on plain paper. This area was then computed with a compensating polar planimeter (K & E Model No. 4236).

The procedure was repeated for 4 weeks—one trial week and 3 further sessions to ensure testing of all techniques in the top, middle, and bottom areas of the back. The subjects were on no drugs during the test period.

Statistical analysis¹⁵

Since the first week did not differ significantly from the other weeks, it was included in the analysis.

1. Reproducibility of the duplicate skin tests was determined in 2 ways:

a. Analysis of variance was used to establish whether there was significant differences between techniques, regions of the back, subjects, and weeks.

b. The actual reproducibility was determined by calculating the coefficient of variation (standard deviation/mean).

*Major allergen (3+ to 4+ reaction) erythema diameter of 15 to 20 mm. usually with wheal and pseudopod formation; minor allergen (1+ to 2+ reaction) erythema diameter of 10 to 15 mm. usually without wheal formation.

†Extracts in weight/volume (w/v) included pollen 1:20 (w/v); house dust, Berkeley Biological 1:40 (w/v), Hollister-Stier 1:10 (w/v); molds, Hollister-Stier 1:100 (w/v).

TABLE I. Techniques

	Method					
	Wheal			Flare		
	Scarifier	Pinprick	Knife	Scarifier	Pinprick	Knife
Area (mean) (sq. cm.)	39.06	27.26	16.91	396.79	321.49	195.00
Variability (all regions)	0.83	0.81	1.22	0.66	0.64	0.94
Variability Region:						
Top	0.61	1.05	0.98	0.54	0.73	1.04
Middle	0.74	0.48	1.24	0.63	0.38	0.92
Bottom	0.68	0.52	1.01	0.51	0.63	0.79

Variability = Coefficient of variation (standard deviation/mean).

2. The comparison of area size for the different techniques, regions, subjects, and weeks utilized the analysis of variance method.

RESULTS

Data are presented for the major allergen *only* since this gave the best reproducibility, and the purpose of this paper is to compare different skin test techniques and areas of the back under the best circumstances. Skin test results from the glycerin-phenol control were insignificant and did not alter the following results.

Techniques

Size. All areas are magnified 40× (from projector and planimetry).

WHEAL. The scarifier gave the largest reaction (39.06 sq. cm.), while the knife gave the smallest reaction (16.91 sq. cm.); the difference is significant at the 1 per cent level ($p < 0.01$).

FLARE. Similar differences were noted, with the scarifier having the highest mean area (386.97 sq. cm.) and the knife the lowest (195.00 sq. cm.) ($p < 0.01$).

Reproducibility:

WHEAL. The variability (coefficient of variation) was large with each of the techniques, ranging from 0.81 (pinprick) to 1.22 (knife). It was smaller when the techniques were analyzed for the mean area within a region of the back, rather than for all regions considered together.

There was no statistically significant difference ($p > 0.05$) in reproducibility of any of the techniques, although pinprick and scarifier were considerably more reproducible than the knife (approximately 30 per cent).

FLARE. The results were similar to that found with the wheal response, although the reproducibility was better. The pinprick and scarifier techniques were the most reproducible (0.64 and 0.66, respectively), while the knife had the poorest result (0.94). The difference, while not statistically significant ($p > 0.05$), was again about 30 per cent (Table I).

Region of back

Size:

WHEAL. The middle of back gave the largest reactions (30.51 sq. cm.), and the bottom the smallest response (20.35 sq. cm.). The difference was significant at 1 per cent level ($p < 0.01$).

TABLE II. Region of back

	Region					
	Wheal			Flare		
	Top	Middle	Bottom	Top	Middle	Bottom
Area (mean) (sq. cm.)	28.38	30.51	24.35	306.56	337.29	269.44
Variability (all-techniques)	1.10	0.87	0.86	0.87	0.69	0.72
Variability Technique:						
Scarifier	0.61	0.74	0.68	0.54	0.63	0.51
Pinprick	1.05	0.48	0.52	0.73	0.38	0.63
Knife	0.98	1.24	1.01	1.04	0.92	0.79

Variability = Coefficient of variation (standard deviation/mean).

FLARE. The mid back region gave the largest flare (337.29 sq. cm.), and the bottom the smallest (269.44 sq. cm.) (p value of < 0.01).

Reproducibility:

WHEAL. There was no statistically significant difference ($p > 0.05$) between any of the regions of the back, but the middle (0.87) and bottom zones (0.86) were more reproducible than the top zone (1.10).

FLARE. Again no statistically significant differences ($p > 0.05$) were found, but the mid zone was most reproducible (0.69), the top least (0.87). The flare was more reproducible than the wheal (Table II).

Subject and week variability

There was a statistically significant difference between subjects as to size of wheal and flare response ($p < 0.01$), but there was no difference in the reproducibility of subjects' reactions. It is interesting to note that neither the size nor the reproducibility of the wheal and flare reaction varied significantly ($p > 0.05$) over the 4 week test period.

DISCUSSION

This study confirms the general impression that the epicutaneous techniques commonly employed have considerable variability. The variability was increased when a technique was analyzed over all regions of the back, compared to evaluation of the technique *within* one region. There was also a significant difference ($p < 0.01$) between skin test responses in different regions of the back. This suggests that the back is not uniformly responsive as previously reported.⁴ Stanworth, studying the PK wheal response in 7 subjects, found approximately 20 per cent variability using the pinprick technique on the back in 6. This occurred only with wheals 50 sq. mm. or greater and rose to 40 to 80 per cent with smaller reactions. One subject with a mean area of 22 sq. mm. had 85 per cent variability. Squire,³ studying skin tests on the arms, similarly found 20 per cent variability only with larger reactions, but in addition found this degree of reproducibility only when symmetric areas of the arms were compared.

The pinprick (and scarifier) produced the greatest reproducibility in our study; the overall variability for all regions was 81 per cent, while the best within a region was 48 per cent (mid back region). Although this is two- to

fourfold greater than Stanworth and Kuhns⁴ selected results, this includes many wheal reactions that were less than 50 sq. mm. Our results are not markedly different when all size responses are considered. (To determine actual skin test size in our study, areas are divided by 40, the magnification factor.)

The chalazion knife resulted in the poorest reproducibility and yet it was the technique most familiar to one of the authors (S.P.G.). The response with this technique was significantly smaller than the pinprick and scarifier methods. In a previous study¹² we reported a 30 per cent reproducibility of the flare with the chalazion knife on the back compared to 94 per cent in this study. One explanation may be the much larger response elicited in the previous study (mean area 1,190 sq. mm. compared to 485 sq. mm.). That this is not the entire explanation can be seen by comparing pinprick and scarifier techniques that have comparable variability, but moderate differences in area size.

Contrary to several reports,^{16, 17} the flare was slightly *less* variable than the wheal, perhaps reflecting size difference. Histamine release following allergen challenge in the skin elicits the triple response of Lewis consisting of both wheal and flare reactions. Each should be considered in the complete evaluation of the allergy skin test, since the flare can add a measure of accuracy to this bioassay.¹⁸

It is interesting to note that there was no significant difference in reaction size or reproducibility over the 4 week period of study, confirming the data reported by Stanworth and Kuhns⁴ and Gottlieb, Stupniker, and Askovitz.¹⁶ Thus, although skin test reactions are thought to vary with time, they do not appear to change significantly over several weeks and perhaps up to 6 to 7 months.¹⁶

It is concluded that while epicutaneous techniques are variable, this variability is increased by the relative nonuniformity of the back, and the failure to elicit an adequate-sized response, perhaps on the basis of not penetrating the epidermis sufficiently. The latter consideration may be overlooked by those performing epicutaneous testing and thus should be emphasized as a major factor in reducing variability of this technique.

Future studies should include investigation of the distribution of mast cells in the skin to determine if a pattern exists that would assist in selecting optimal sites for skin testing. Until that time, the mid back region would appear to be best suited for that purpose.

REFERENCES

- 1 Blackley, C. H.: Experimental researches on the causes and nature of *Catarrhus aestivus* (hay fever and asthma), London, 1873, Bailliere, Tindall & Cox, Ltd.
- 2 Fineman, A. H.: Studies in hypersensitiveness. XXII. A comparative study of intradermal, scratch, and conjunctival tests in determining the degree of pollen sensitivity, *J. Immunol.* 11: 465, 1926.
- 3 Squire, J. R.: The relationship between horse dandruff and horse serum antigens in asthma, *Clin. Sci.* 9: 127, 1950.
- 4 Stanworth, D. R., and Kuhns, W. J.: Quantitative studies on the assay of human skin sensitizing antibodies (reagins). I. An examination of factors affecting the accuracy of the Prausnitz-Kustner (P-K) test, *Immunology* 8: 323, 1965.
- 5 Becker, E. L., and Rapport, B. Z.: Quantitative studies in skin testing. II. The form of the dose-response curve utilizing a quantitative response, *J. ALLERGY* 19: 317, 1948.

- 6 VanArsdel, P. P., Jr., and Sherman, W. B.: The risk of inducing constitutional reactions in allergic patients, *J. ALLERGY* **28**: 251, 1957.
- 7 Bronsky, E. A., and Ellis, E. F.: Inhalation bronchial challenge testing in asthmatic children, *Pediatr. Clin. North Am.* **16**: 85, 1969.
- 8 Linblad, J. H., and Farr, R. S.: The incidence of positive intradermal reactions and the demonstration of skin sensitizing antibody to extracts of ragweed and dust in humans without history of rhinitis or asthma, *J. ALLERGY* **32**: 392, 1961.
- 9 Alexander, H. L., and McConnell, F. S.: The variability of skin reactions in allergy, *J. ALLERGY* **2**: 23, 1930.
- 10 Sheldon, J. M.: A manual of clinical allergy Philadelphia, 1967, W. B. Saunders Co.
- 11 Kaufman, H. S.: Allergy scratch testing with a uniform technique: Demonstration of a new device, *Ann. Allergy* **24**: 26, 1966.
- 12 Galant, S. P., Zippin, C., Bullock, J., and Crisp, J.: Allergy skin test. I. Antihistamine inhibition, *Ann. Allergy* **30**: 53, 1972.
- 13 Pabst, M. R., Boatner, C. H., and Efron, B. G.: Studies with antigens: Analysis of comparative skin reactions by means of the chi square statistic, *New Orleans Med. & Surg. J.* **93**: 142, 1940.
- 14 Maibach, H., and Kligman, A.: The micrometer syringe for quantitative skin testing, *Arch. Dermatol.* **87**: 743, 1963.
- 15 Snedecor, G. W., and Cochran, W. G.: Statistical methods ed. 6, Ames, Iowa, 1967, Iowa State University Press.
- 16 Gottleib, P. M., Stupniker, S., and Askovitz, S. T.: The reproducibility of intradermal skin tests: A controlled study, *Ann. Allergy* **18**: 949, 1960.
- 17 Joubert, L., Gaut, Z., and Abrams, W. B.: A new approach to the use of histamine skin tests in the study of antihistamine drugs, *Clin. Pharmacol. Ther.* **10**: 250, 1969.
- 18 Dannenberg, T. B., and Feinberg, S. M.: The development of tolerance to antihistamines, *J. ALLERGY* **22**: 330, 1951.