35. Yamada K, Osamu A, Yoshimura T, Katayama K. Highly sensitive gas chromatographic-mass spectrometric method for the determination of platelet-activating factor in human blood. J Chromatogr 1988;433:243-7.
36. Prin L, Capron M, Tonnel AB, Bletry O, Capron A. Heterogeneity of human peripheral blood eosinophils: variability in cell density and cytotoxic ability in relation to the level and the origin of hypereosinophilia. Int Arch Allergy Appl Immunol 1983;72:336-46.
37. Winqvist I, Olofsson I, Persson A-M, Hallberg T. Altered density, metabolism, and surface receptors in eosinophilia. Immunology 1982;47:531-9.
38. Kajita T, Yui Y, Mita H, et al. Release of leukotriene C_4 from human eosinophils and its relation to cell density. Int Arch Allergy Appl Immunol 1985;78:406-10.
39. Frick WE, Sedgwick JB, Busse WW. Hypodense eosinophils in allergic rhinitis. J ALLERGY CLIN IMMUNOL 1988;82:119-25.
40. Sedgwick JB, Geiger KM, Busse WW. Superoxide generation by hypodense eosinophils from patients with asthma. Am Rev Respir Dis 1990;142:120-5.
41. Sedgwick JB, Calhoun WJ, Busse WW. Enhanced superoxide (O_2^-) generation by airway eosinophils (EOS) [Abstract]. J ALLERGY CLIN IMMUNOL 1991;87:240.
42. Henson PM. Activation and desensitization of platelets by PAF derived from IgE-sensitized basophils. J Exp Med 1976;143:937-44.

Precision of skin prick and puncture tests with nine methods

Pascal Demoly, MD,* Jean Bousquet, MD,*
Jean-Claude Manderscheid, MD, Sten Dreborg, MD,*****
Henriette Dhivert, MD,* and François-Bernard Michel, MD*
Montpellier, France, and Linkoping, Sweden

New devices for puncture tests have been proposed recently, but their precision by comparison to the prick test method is poorly known. Seven puncture tests (Allerprick, Morrow Brown standardized needle, Phazer, Pricker, Stallerpointe, Stallerkit, and Wyeth bifucated needle) were compared with the modified prick test performed with hypodermic or intradermal needles in eight carefully selected normal volunteers. Skin tests with histamine hydrochloride (10 mg/ml) were only performed when there was no factor that might interfere with their interpretation. The site of skin tests on the forearm was demonstrated not to significantly influence the reaction size. The coefficient of variation of the tests ranged from 8.4% to 21.7%. Modified skin prick tests are satisfactory since they are highly reproducible (coefficient of variation: 13.4% and 16.5%) and there is no subject effect. Phazet was found to be more reproducible without subject effect. Pricker is satisfactory since it has no subject effect and a reproducibility similar to that of modified prick tests. Other tests are less reproducible (Stallerkit or Morrow Brown) or vary between subjects (Allerkit, Stallerkit, Stallerpointe, and Wyeth Needle). (J ALLERGY CLIN IMMUNOL 1991;88:758-62.)

Key words: *Histamine, skin test, prick test, reproducibility, human*

SPTs are among the skin test methods from which to choose and, at least in Europe, they tend to be preferred to intradermal test.[1, 2] The modified SPT introduced by Pepys[3] is the current reference method, although the variability of this test has been demonstrated to be greater than that of the intradermal reaction.[4, 5] Investigators have therefore attempted during the past 10 years to decrease the variability of SPT by introducing puncture tests with various devices. The most popular instruments are the MB stan-

From the *Clinique des Maladies Respiratoires, Hopital l'Aiguelongue, Montpellier, and **Département d 'Informatique Médicale, Montpellier, France, and ***Department of Pediatrics, University Hospital, Linköping, Sweden.
Received for publication Nov. 27, 1990.
Revised June 13, 1991.
Accepted for publication June 13, 1991.
Reprint requests: Jean Bousquet, MD, Clinique des Maladies Respiratoires, Hopital l'Aiguelongue, 34059, Montpellier Cedex, France.
1/1/31693

Abbreviations used

ANOVA:	Analysis of variance
CV:	Coefficient of variation
DF:	Degree of freedom
SPT:	Skin prick test
A:	Allerprick
MB:	Morrow Brown standardized needle
MPT 1:	Modified prick test with hypodermic needle
MPT 2:	Modified prick test with intradermal needle
P:	Østerballe standardized needle (Pricker)
PH:	Uncoated Phazer
SP:	Stallerpointe
SK:	Stallerkit
W:	Wyeth bifurcated smallpox vaccination needle

dardized needle,[6] the P precision needle,[7] the SP,[8] the A,[9] the Ph,[10-12] the Greer "pen,"[13] the Multi-Test,[14, 15] and the W.[16] Opinions concerning these so-called standardized methods vary according to the skill, experience, and the preference of the investigator, as well as the aims of the use of the skin tests. Although some variability may be accepted for the clinical diagnosis of allergy, skin tests must be highly reproducible for standardization purposes or for the assessment of antiallergic treatments. Some studies have been conducted after a few methods, and their variability has been demonstrated to range from 10% to 30% when the mean wheal area was used,[6, 9, 13, 17-23] although Basomba et al.[21] observed a much greater variability. However, (1) except in the study of Adinoff et al.,[13] only one to three different techniques were examined, (2) most studies compared skin test methods rather than examining their true reproducibility, and (3) the interpatient variability of skin tests was seldom studied.

Since newly "standardized" skin puncture devices are now very popular, it was important to assess their performance in a single study. The reproducibility of nine different methods of SPT or puncture tests was examined in normal individuals with histamine. The reproducibility of these methods was examined both within and between patients.

MATERIAL AND METHODS

Subjects

Eight normal volunteers, ranging in age from 22 to 32 years, were enrolled in the study. None of the volunteers was atopic according to the criteria defined by Pepys.[24] None of these subjects was suffering from any disease that might affect the performance of skin tests.

Histamine reagent

Histamine hydrochloride (Pharmacia Diagnostics AB, Uppsala, Sweden) was used. The revised Nordic Guidelines on Standardization of Allergenic Extracts[25] recommends the use of histamine hydrochloride at a strength of 10 mg/ml (54.3 mmol/L) for standardization purposes, and as other investigators, we used this concentration.[26-28]

Methods of skin test

The modified prick test was performed according to the method of Pepys.[3] A small drop of each test extract was placed on the forearm, and a needle was passed through the drop and inserted into the epidermal surface at a low angle with a bevel up facing away from the surface. The needle tip was then gently lifted to elevate a small portion of the epidermis without inducing bleeding. The test solution was gently wiped away 1 to 2 minutes later. Two different needles were used since Pepys[3] suggests a hypodermic needle (⅝ inch, 25-gauge, MPT 1), whereas other investigators prefer an intradermal needle (22-gauge, 1¼ inch, MPT 2).

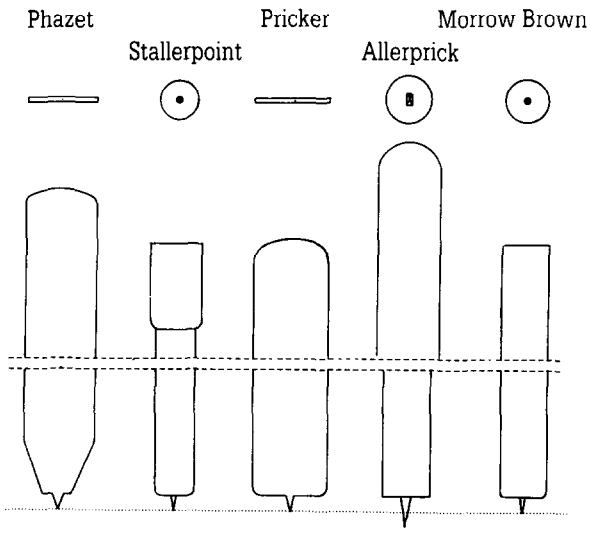
All other devices were inserted perpendicular to the skin through the test extract drop. The solution was wiped away 1 to 2 minutes after the test. The following devices were used (Fig. 1): uncoated PH (Pharmacia Diagnostics, Uppsala, Sweden),[10-12] SP (Stallergènes Laboratories, Fresnes, France),[8] P (Dome-Hollister-Stier, Paris, France),[7] W,[16] A (Hall Allergenen Laboratories, Haarlem, The Netherlands),[9] MB,[6] and SK (Stallergènes Laboratories). This device was based on the Multi Test (Lincoln Diagnostics, Miami, Fla.)[14, 15] but used single 1 mm long plastic pins.

Tests

The entire study was performed by the same well-trained investigator between 9 AM and 12 PM hours to avoid circadian variations.[29] None of the patients was taking any drug that might interfere with skin test.[30] Skin tests were performed on the volar surface of the forearm in octuplicate with each technique, the test sites being placed 3 cm apart, as suggested by Pepys.[3] This distance might not be sufficient, because Terho et al.[31] observed that histamine may cause false positive reactions when tests are placed 3 to 4 cm apart. We attempted to confirm this finding, but with other investigators,[32, 33] we were unable to find any effect of histamine at this close distance. Since there might be differences in skin reactivity between the *medial* and the ulnar sides of the forearm,[34, 35] the techniques were performed in random order. The wheal reaction was recorded after 10 minutes by means of the cellophane tape technique, and mean diameters were averaged. Since it is impossible to perform nine octuplicate tests in the same patient the same day, we performed two testing sessions 1 week apart. To avoid possible differences, the techniques were done in random order, and for each session, the reference test, that is, MPT 1, was repeated.

Statistical analysis

The reproducibility (CV = SD/arithmetic mean × 100) was calculated for each octuplicate series and for each device by averaging the means of the octuplicates of the



point length (mm)	1	1	1	2	1
inclusive angle (°)	45	25	25	25	25
needle length (cm)	4	3	3	10	3
needle width (mm)	4.5	2.8	5	3	3
material	metal	plastic	metal	plastic	plastic

FIG. 1. Description of the standardized SPT needles.

eight subjects. Low values of the CV represent a high reproducibility and vice versa.

Statistical analysis was performed by use of ANOVA and nonparametric tests with BMDP Statistical Software, Inc. (University of California-Los Angeles). Since some studies have indicated differences in wheal sizes according to the site of skin testing on the forearm, we tested this possibility by use of one-way ANOVA. The effects of the skin tests, subjects, and of their interaction were analyzed by two-way ANOVA with the following formula:

$$E^2_{ijk} = (X_{ijk} - M_{ij})^2$$

where the subject is j, skin test is i, point is k, X equals the value, and M equals arithmetic mean for the 8 points.

The comparison of means of E^2 between subjects for each test or between tests for each subject was done with the Kruskal-Wallis test. Finally, the relative difference between skin tests was studied by the Kruskal-Wallis test with the following formula: Relative difference between skin tests = E^2/M^2.

For the comparison of the reproducibility between the tests, we selected a reference test, that is MPT 1, and tested the statistical differences by ANOVA between all other tests and the reference test. When F was >3.50, the reproducibility was considered to be significantly different from the reference test.

RESULTS

Effects of test sites

There was no significant difference between test sites on the forearm (DF, 7; F, 0.67).

Effects of testing sessions

There was no significant difference in MPT 1 reproducibility performed during two sessions (DF, 7; F, 0.55). Moreover, there was no difference between mean sizes of the tests.

Mean wheal sizes induced by different techniques

The mean wheal sizes are presented in Table I that were induced by the different methods used. P induced bleeding in all test sites, and metallic devices elicited the largest wheals, whereas the plastic devices with a small pin (SP and MB) or MPT never induced bleeding and elicited the smallest wheals.

Reproducibility of skin tests

The results of the reproducibility of the nine techniques of skin tests are presented in Table I. The CVs range from 8.4% to 21.7%. The CVs of MPT 1 and MPT 2 were not very different. Two devices presented a lower CV than MPT (PH and SP), three other devices had a similar CV (P, A, and W), whereas the two remaining devices had a high CV (MB and SK). SK was the only test method that had a reproducibility significantly different from the reference test.

The overall difference between skin tests was significant (DF, 8; F, 3.05; $p < 0.0023$) as well as the interaction between skin tests and subjects (DF, 56; F, 1.88; $p < 0.0002$) but differences between subjects

TABLE I. Results of skin tests

Skin test	Mean size (mm)	CV%	Rank in relative difference	Subject interaction
PH	5.31 ± 0.6	8.4 ± 1.7	1	No
SP	4.05 ± 1.0	8.9 ± 2.8	2	Yes
A	4.63 ± 0.7	13.3 ± 6.1	3	Yes
MPT 2	3.75 ± 0.7	16.5 ± 5.2	4	No
W	4.36 ± 0.8	14.7 ± 3.9	5	Yes
P	4.97 ± 0.5	12.3 ± 3.7	6	No
MPT 1	3.55 ± 0.7	13.4 ± 5.6	7	No
SK	4.21 ± 0.6	21.7 ± 14.9	8	Yes
MB	3.82 ± 0.3	20.8 ± 6.8	9	No

Results listed in mean ± SD.

were not significantly different (DF, 7; F, 1.85; not significant). By Kruskal-Wallis test, results of some skin test devices were significantly different according to the subjects tested: A, $p < 0.0067$; W, $p < 0.0054$; SP, $p < 0.006$; and SK, $p < 0.0012$; whereas no significant subject effect was observed for the other skin test methods.

The relative difference between skin tests demonstrated a rank order of reproducibility (Table I) slightly different from the results obtained by CVs. However, the same patterns were observed. PH and SP are the most reproducible tests, A, MPT 1, MPT 2, W, and P are in the same range, and SK and MB are the less reproducible tests.

DISCUSSION

The techniques used in this study are simple and widely used. Normal volunteers were carefully selected, and skin tests were only performed when there was no factor that might interfere with their interpretation. The site of skin tests on the forearm was demonstrated not to influence, significantly, the reaction size. It was also found that when MPT 1 was repeated at weekly intervals, there was no difference in reproducibility; therefore, comparisons of reproducibility between different devices can be performed with two testing sessions when they are performed carefully. Although all tests might have been done in one session on the back, we preferred to use the forearm, since there is more variability in the back and since most doctors use the forearm in their clinic. The MPT appears to be satisfactory since it is highly reproducible and there is no subject effect. PH was found to be more reproducible without subject effect. P is satisfactory since it has no subject effect and a reproducibility similar to that of MPT. Other tests are not as reproducible (SK and MB) or vary between subjects (SP, A, W, and SK).

In Europe, many, if not most, common allergens, are standardized, and the reference material is either histamine, 54.3 mmol/L, or codeine phosphate, 9%, representing a similar standardization method. Thus, most allergen-induced SPTs are ranging between 4 to 6 mm, the reason we have deliberately chosen the concentration of histamine 54.3 mmol/L.

The mean size of skin tests was larger with metallic devices (P, PH, and W) than with plastic devices (SP and MB) except for A that had a longer pin. Most metallic devices induced bleeding, revealing that they penetrate deeper into the skin. MPT performed with two needles induced wheals of similar sizes. The mean size of wheals elicited by MPT and plastic devices were the smallest. These findings should be taken into consideration when the positivity of skin tests is based on their wheal size only.

The results observed in this study demonstrate that prick or puncture tests performed by a well-trained investigator are highly reproducible with most techniques used, since CV ranged from 8.4% to 21.7%.

The newly developed techniques usually compare favorably with the MPT proposed by Pepys,[3] and the results of this study confirm previous data.* Since all the tested techniques have a CV <30%, they can be used in clinical practice, the choice being made depending on the preference of the clinician and their relative cost. However, the skill of the investigators is of importance, and we did not test in the present study this parameter; thus, results might have been different if investigators of different skills had been performing the study. This study demonstrates also that prick or puncture tests can be used for standardization purposes or drug trials as already proposed,[2, 27] but for this purpose, PH and both MPT are the optimal methods, since they have a low CV and are not variable between subjects.

*References 3, 5, 6, 10, 13, 17, 18, 20, and 21.

REFERENCES

1. Nelson HS. Diagnostic procedures in allergy. I. Skin testing. Ann Allergy 1983;51:411-7.
2. Dreborg S, Backman A, Basomba A, Bousquet J, Dieges P, Malling H-J. Skin tests used in type I allergy testing. Position paper of the European Academy of Allergology and Clinical Immunology. Allergy 1989;44(suppl 10):1-59.
3. Pepys J. Skin testing. Br J Hosp Med 1975;14:412-7.
4. Malling HJ. Diagnosis and immunotherapy of mould allergy. II. Reproducibility and relationship between skin sensitivity estimated by end-point titration and histamine equivalent reaction using skin prick test and intradermal test. Allergy 1985;40:354-62.
5. Voorhorst R, Van Krieken H. Atopic skin test reevaluated. II. Variability in results of skin testing donor in octuplicate. Ann Allergy 1973;31:195-204.
6. Brown HM, Su S, Thantrey N. Prick testing for allergens standardized by using a precision needle. Clin Allergy 1981;11:95-8.
7. Østerballe O, Weeke B. A new lancet for skin prick testing. Allergy 1979;34:209-12.
8. Bousquet J, Djoukhadar F, Hewitt B, Guérin B, Michel FB. Shelf life of a pollen and a mite standardized extract. Clin Allergy 1985;15:29-35.
9. Nierop G, Voorhost R, Temmerman-van de Vijver RL. Atopic skin test re-evaluated. X. Comparison of a perfected technique of intracutaneous skin testing and the prick test performed with a modified Morrow Brown needle. Ann Allergy 1981;46: 105-9.
10. Belin L, Dreborg S, Einarsson R, et al. PHAZET™—a new type of skin prick test. Allergy 1985;40(suppl 4):60-3.
11. Østerballe O, Nielsen JP. Allergen-coated lancets (Phazet) for skin prick testing in children. Allergy 1989;44:356-62.
12. Chanal I, Horst M, Segalen C, Dreborg S, Michel FB, Bousquet J. Comparison between modified skin prick test with standardized allergen extracts and Phazet. J ALLERGY CLIN IMMUNOL 1988;82:878-87.
13. Adinoff AD, Rosloniec DM, McCall LL, Nelson HS. A comparison of six epicutaneous devices in the performance of immediate hypersensitivity skin testing. J ALLERGY CLIN IMMUNOL 1989;84:168-74.
14. Kniker WT, Anderson CT, Roumiantzeff M. The Multi-Test system: a standardized approach to evaluation of delayed-type hypersensitivity and cell-mediated immunity. Ann Allergy 1979;43:73-9.
15. Ménardo JL, Bousquet J, Michel FB. Comparison of three prick test methods with the intradermal test and with the RAST in the diagnosis of mite allergy. Ann Allergy 1982;48:235-9.
16. Santilli J Jr, Potsus RL, Goodfriend L, Marsh DG. Skin reactivity to purified pollen allergens in highly ragweed-sensitive individuals. J ALLERGY CLIN IMMUNOL 1980;65:400-8.
17. Taudorf E, Malling HJ, Laursen LC, Lanner A, Weeke B. Reproducibility of histamine skin prick test: inter- and intravariation using histamine dihydrochloride 1, 5, and 10 mg/ml. Allergy 1985;40:344-9.
18. Malling HJ. Reproducibility of skin sensitivity using a quantitative skin prick test. Allergy 1985;40:400-4.
19. Dirksen A, Mosbech H, Søborg M, Biering I. Comparison of a new lancet and a hypodermic needle for skin prick test. Allergy 1983;38:359-62.
20. Holgersson M, Stalenheim G, Dreborg S. The precision of skin prick test with PHAZET™, the Østerballe needle and the bifurcated needle. Allergy 1985;40(suppl 4):64-5.
21. Basomba A, Sastre A, Pelaez A, et al. Standardization of the prick test: a comparative study of three methods. Allergy 1985;40:395-9.
22. Perrin LF, Déchamp CH, Deviller P, Joly P. Reproducibility of skin tests: a comparative study of the Pepys prick test and the Morrow Brown needle and their correlation with the serum IgE level. Clin Allergy 1984;14:581-8.
23. Vohlonen I, Terho EO, Koivikko A, et al. Reproducibility of the skin prick test. Allergy 1989;44:525-31.
24. Pepys J. Types of allergic reactions. Clin Allergy 1973;3: 491-9.
25. Nordic Council on Medicines. Registration of allergen preparations. 2nd ed. Uppsala, Sweden: Nordiska Läkemedelsnämnden, NLN Publication No. 23, 1989.
26. Malling HJ. Skin prick testing and the use of histamine references. Allergy 1984;39,596-601.
27. Dreborg S. Biological standardization: a comparison between three methods. Allergy 1985;39:27-30.
28. Belin L. Biological standardization—the HEP unit. In: Sisk C, Shaeffer M, Going H, et al. Proceeding of the Fourth International Paul Ehrlich Seminar. Stuttgard: Gustav Fisher Verlag, 1986:145-53.
29. Lee RE, Smolenski MH, Leach CS, McGoven JP. Circadian rhythms in the cutaneous reactivity to histamine and selected antigens, including phase relationships to urinary cortisol excretion. Ann Allergy 1977;38:231-6.
30. Bousquet J. In vivo methods to study allergy: skin tests. In: Middleton E Jr, Ellis EF, Reed CE, Adkinson NF Jr, Yunginger JW, eds. Allergy: principles and practice. 3rd ed. St. Louis: CV Mosby, 1988:414-36.
31. Terho EO, Husman K, Kivekäs J, Riihimäki H. Histamine control affects the weal produced by the adjacent diluent control in skin prick tests. Allergy 1989;44:30-2.
32. Watson HK, Harris RI. Histamine control affects the wheal produced by the adjacent diluent control in skin prick tests [Letter]. Allergy 1989;44:442.
33. Garibaldi E, Slavin RG. Positive Multi-Test reactions do not cause false positive reactions at adjacent sites. Ann Allergy 1990;65:481-4.
34. Swain HH, Becker EL. Quantitative studies in skin testing. V. The whealing reactions of histamine and ragweed pollen extract. J ALLERGY 1952;23:441-6.
35. Clarke CW, Mitchell J, Nunn AJ, Pepys J. Reproductivility of skin prick tests to five common allergens. Clin Allergy 1982;12:1-8.