Comparison of test devices for skin prick testing

Warner W. Carr, MD,[a] Bryan Martin, DO,[a] Robin S. Howard, MA,[b] Linda Cox, MD,[c] Larry Borish, MD,[d] and the Immunotherapy Committee of the American Academy of Allergy, Asthma and Immunology *Silver Spring, Md, Fort Lauderdale, Fla, and Charlottesville, Va*

Background: Allergy skin testing guides developing avoidance plans and writing an immunotherapy prescription. The goal for the allergist is to apply allergen skin testing to the appropriate patient population by using a device that minimizes both false-negative and false-positive findings while minimizing patient discomfort. New skin testing devices continue to be developed with a trend toward production of multiheaded devices. Data on the performance of these devices in a head-to-head prospective fashion are limited.
Objective: Our goal was to study 8 commonly used devices to compare their performance in a head-to-head fashion.
Methods: In a prospective, double-blind fashion, the performance of 8 skin test devices was evaluated. Devices were tested with histamine and saline on both the arms and back of each subject. Devices were rotated over 4 testing sessions, at least a week apart, so each device was tested in each anatomic testing location. Performance elements examined included wheal, flare, pain, sensitivity, specificity, and intradevice variability.
Results: We found significant differences in all areas of device performance among all devices examined. Multiheaded devices also demonstrated significant intradevice variability and were more painful than single devices. Furthermore, multiheaded devices had larger reactions on the back, whereas single devices had larger reactions on the arms.
Conclusion: Statistically significant differences exist among all devices tested. Providers should consider this data when choosing a device that suits their practice setting and ensure that technicians are sufficiently trained on the correct use of that device. (J Allergy Clin Immunol 2005;116:341-6.)

Key words: *Skin testing, device, performance, variability, pain*

The US Joint Council of Allergy, Asthma and Immunology[1] and the European Academy of Allergology and Clinical Immunology[2] recommend percutaneous testing as the primary test for diagnosis of IgE mediated allergic disease. Skin testing is also the preferred method for selecting allergens to be included in immunotherapy.[3] Given this, the findings on the initial skin test panel are very important clinical data. If a particular device is too sensitive (resulting in false-positive findings), the patient may receive an antigen that is not required to achieve clinical benefit. On the other hand, a high false-negative rate for a particular device will result in a patient not receiving a needed antigen while undergoing immunotherapy. The goal for the allergist is to perform allergen skin testing in an appropriate patient population by using a device that minimizes both false-negative and false-positive findings. In addition, it is desirable to use a device that results in minimal patient discomfort. Previous studies comparing devices for skin prick (ie, prick and puncture) testing have revealed significant differences in the size of wheal and flare reactions. These differences have been seen at both positive (allergen extract or histamine) and negative (saline) sites.[4-7] In these studies, the difference appeared to result from the degree of trauma imparted to the skin by the device, an interpretation that was reinforced by the fact that those producing larger wheals also caused more patient discomfort.[6]

New skin devices continue to be developed, with a current trend toward devices that allow for application of several antigens simultaneously, referred to as *multiheaded*. This may limit technician time and increase efficiency. In addition, multiheaded devices have increasing popularity in children, in whom the acceptance of a few multiple test devices tends to be better than many individually applied devices. In a recent letter to the editor, Nelson et al[7] compared 3 new multidevices with previously reviewed devices. In this communication, significant differences were noted with the smallpox needle on the back and the Greer Track (Greer Labs, Lenoir, NC) on the arm. Given this, we reviewed 4 devices that allow

From [a]the Department of Allergy and Immunology and [b]the Department of Clinical Investigations, Walter Reed Army Medical Center, Silver Spring; [c]private practice, Fort Lauderdale; and [d]the Asthma and Allergic Disease Center, University of Virginia Health System.

Supported by the Walter Reed Army Medical Center, Department of Clinical Investigations.

Disclosure of potential conflict of interest: L. Borish has consultant arrangements with PDL, Syngenta, and Sepracor. No other relevant conflicts of interest to disclose.

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

Received for publication September 3, 2004; revised March 28, 2005; accepted for publication March 28, 2005.

Available online May 16, 2005.

Reprint requests: Warner W. Carr, MD, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, 503 Robert Grant Ave, Silver Spring, MD 20910. E-mail: wmcarr@comcast.net.

0091-6749/$30.00

doi:10.1016/j.jaci.2005.03.035

Abbreviation used
CV: Coefficient of variation

application of multiple antigens at once (multiheaded) and 4 devices that allow application of only 1 antigen at a time (single devices).

Our goal was to study a cohort of devices to compare their performance in a head-to-head fashion. We specifically intended to determine sensitivity, specificity, variability, and pain. With these results, providers will be able to determine which device is best suited for their practices.

METHODS

Study design

The study was a prospective, double-blind clinical trial and was reviewed and approved by the Walter Reed Army Medical Center Clinical Investigation Committee and the Human Use Committee. All subjects enrolled into the study voluntarily agreed to participate and gave written informed consent. Each subject underwent testing in 4 sessions, with each at least 1 week apart. Each device was tested both on the arm and the back, with histamine (10 mg/mL; Hollister-Stier, Spokane, Wash) and glycerol-saline (Hollister-Stier) during each session. During the course of 4 sessions, the locations on the arm and back were rotated to ensure each device was tested on the upper and lower arm and upper and lower back. Therefore, each session yielded 4 test sites per device: 2 histamine tests (1 back and 1 arm) and 2 glycerol-saline tests (1 back and 1 arm). Fig 1 illustrates the back and left arm test regions for one test session. Over the course of the study, sites were rotated to give an equal number of tests in all areas to offset any differences in reactivity.[6] At the end of the study, each device had been tested on the upper and lower back and the upper and lower arms, with an even distribution between left and right. All heads of a multiheaded device were tested with histamine at the histamine site, and all heads were tested with saline at the saline site. At the conclusion of the fourth session, a mean result was determined for each head of the multiheaded devices, and from this, intradevice variability was determined. Single device test sites were spaced at least 30 mm apart, and multiheaded spacing was fixed at 20 mm to 30 mm on the basis of the design of the device. With the devices examined, this resulted in 132 individual pricks per subject per session. The total number of individual pricks over the course of the study for each subject was 528. With 13 subjects completing the study, this yielded 6864 individual prick sites for examination. Before each session, antihistamines were withheld for at least 1 week, and H2 antagonists and leukotriene antagonists were withheld for 72 hours.

To maintain objectivity, the technician who performed all of the tests was blind to the contents of the test solution, either histamine or saline. A second technician who was not in the room during application of each device recorded the results. This technician was blind to the device used as well as to the solution used. Before the study was initiated, a representative of the manufacturer trained the technician who performed the skin tests on each device. This step was taken to achieve the best possible results by using the manufacturer's recommended skin testing technique.

Pain assessment was performed by using the Wong-Baker FACES pain rating scale[8] immediately after application of each skin test device (measured on a scale from 0-10). On the basis of this scale, a level of 1 to 2 is considered minimal pain. The greatest reported pain was recorded for that particular test site and session. Pain was recorded within seconds of application of the skin test device to minimize the influence of histamine on pain perception. Results were recorded for pain sensation on the arm and on the back.

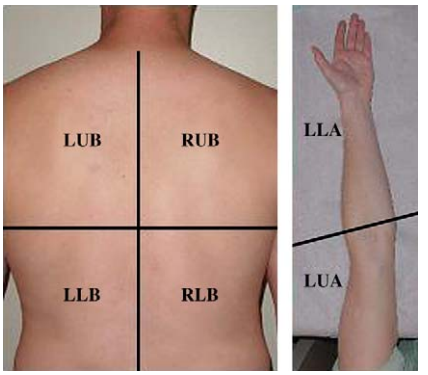

FIG 1. The *left* image illustrates the 4 test zones of the back, and the image on the *right* represents the 2 left arm test zones. Right arm test zones are not shown but mirror those of the left arm. *LLA,* Left lower arm; *LLB,* left lower back; *LUA,* left upper arm; *LUB,* left upper back; *RLB,* right lower back; *RUB,* right upper back.

Subjects

Male or female subjects age 18 to 70 years, with or without allergies, were eligible for the study. Subjects were excluded if they had dermatographism, severe atopic dermatitis, or asthma, or were taking antidepressants. Antihistamines were withheld for 1 week before testing. Leukotriene antagonists and H2 antagonists were withheld for 72 hours before testing.

Devices

Four single-headed devices and 4 multiheaded devices were tested. Single headed devices included the Greer Pick (Greer Labs), Accuset (ALK-Abelló, Inc, Round Rock, Tex), Sharptest (Panatrex, Inc, Placentia, Calif), and Quintip (Hollister-Stier). Multiheaded devices tested were the Quintest (Hollister-Stier), Quantitest (Panatrex, Inc), Greer Track, and Multi-Test II (Lincoln Diagnostics, Inc, Decatur, Ill; Fig 2).

Skin testing

All testing was performed first on the arms, and once results were obtained and recorded, testing proceeded on the back. The wheal and flare results were recorded at 15 minutes by obtaining the longest orthogonal diameters. Mean diameters were used for statistical analyses. Pain was recorded immediately after application of each skin test device. Positive test solution consisted of 10 mg/mL histamine (Hollister-Stier), with standard glycerol saline (Hollister-Stier) used as a negative solution.

Statistical analysis

Results were analyzed by using repeated-measures ANOVA, with the within-subject factors body site (upper arm, lower arm, upper back, lower back) and device. Thirteen subjects were needed to power the study adequately to detect a minimum difference of 2 mm between each device. When calculating sensitivity and specificity, a true-positive result was considered a histamine wheal of 3 mm or greater, and a true-negative result was a glycerol-saline wheal less than 3 mm. A result was considered false-negative if a histamine

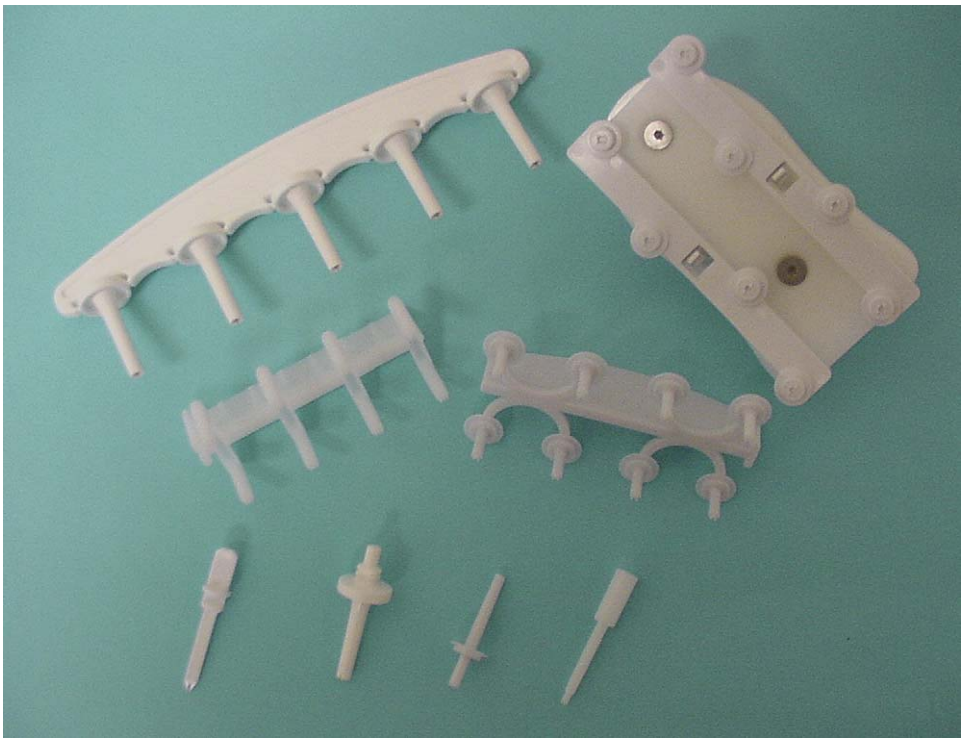

FIG 2. Skin test devices investigated. Multiheaded devices from *top left to right* followed by *midleft to right*: Quintest, Greer Track, Multi-Test II, and Quantitest. Single devices from *bottom left to right*: Accuset, Quintip, Sharptest, Greer Pick.

wheal was less than 3 mm, and a result was considered false-positive if the glycerol-saline site was 3 mm or greater. Results are presented as the means ± SDs, and for multiheaded devices, the average of all heads was used in the calculation of sensitivity and specificity. Sensitivity and specificity of each device are presented as proportions with 95% CIs, and devices were compared by using the Fisher exact test (2-tailed). Sensitivity was calculated by dividing true-positive results by the sum of true-positive and false-negative results. Specificity was calculated by dividing true-negative results by the sum of true-negative and false-positive results.

When multiheaded devices were analyzed, intradevice variability was described by using the coefficients of variation (CVs; presented as medians with the interquartile range) for each device. For each multiheaded device, the wheal produced by each head was compared by using repeated-measures ANOVA.

Pain scores were compared among devices by using the Wilcoxon signed-rank test: median pain scores were presented as well as the proportion of pain scores above a value of 2 (representing mild pain on the Wong-Baker FACES pain rating scale). For interdevice comparisons of pain, wheal, and flare size within the single-headed or multiheaded groups, there are 6 possible pairwise analyses. Using a Bonferroni correction of the overall experimental P value of .05, a P value of .008 (.05/6) or less is considered significant.

RESULTS

Twenty subjects were recruited for the study, and 7 subjects did not complete because of pregnancy (1) and military operational requirements (6). Eight men and 5 women completed the study. The mean age was 32.2 years (range, 22-57), and 7 subjects had a history of atopy.

Interdevice comparisons

Histamine and saline reactions are presented in Table I. Controlling for site (arm vs back), there was a significant difference in histamine wheal size among all devices in each device group ($P < .008$ for all comparisons), except for no significant difference between the Accuset and the Quintip ($P = .28$) and the Multi-Test II and Quantitest ($P = .27$). The largest reactions to histamine base were found with 2 single devices, Sharptest and Greer Pick. There were no significant differences in saline wheal reactions. In addition, all mean histamine flares were greater than 10 mm, and mean saline flares were below 5 mm. Table II gives the number of results that exceeded the limits for positive and negative reactions set for histamine and saline. For histamine wheal reactions, the Greer Pick gave the lowest number of false-negative results (2/208 or 0.96%); the range for single devices was 0.96% to 3.8% (Accuset). The range for multiheaded devices was 57/1664 (3.4%) with the Multi-Test II and 366/1664 (22%) with Greer Track.

Single devices demonstrated a high degree of reproducibility, with CVs ranging from 0.22 to 0.37 (Table I). The CV reported in Table I for the multiheaded devices represents a CV of the mean of all heads. For intradevice variability, or differences between each head of a multiheaded device, see Table III.

TABLE I. Outcome measures for 8 devices*

	Histamine wheal, mean ± SD	Histamine flare, mean ± SD	CV	Saline wheal, mean ± SD	Saline flare, mean ± SD	Sensitivity % (95% CI)	Specificity % (95% CI)
Single devices							
Sharptest	7.1 ± 1.7	31.6 ± 8.4	0.22	0.03 ± 0.3	3.2 ± 2.8	97 (91-99+)	99 (94-99+)
Greer Pick	6.6 ± 1.8	33.3 ± 9.5	0.37	0.0 ± 0.0	2.7 ± 2.4	98 (93-99+)	100 (97-100)
Accuset	5.1 ± 1.9	24.3 ± 10.7	0.34	0.1 ± 0.5	1.5 ± 2.4	92 (85-97)	98 (93-99+)
Quintip	4.8 ± 1.7	22.6 ± 9.3	0.36	0.0 ± 0.0	1.1 ± 2.6	95 (89-99)	100 (97-100)
Multidevices							
Multi-Test II	5.9 ± 1.3	26.0 ± 5.7	0.23	0.02 ± 0.2	3.3 ± 1.5	93 (91-95)	99 (98-99+)
Quantitest	5.7 ± 1.6	25.6 ± 7.3	0.34	0.01 ± 1.6	3.2 ± 1.8	89 (86-91)	99 (98-99+)
Quintest	4.3 ± 1.4	19.9 ± 7.8	0.25	0.0 ± 0.0	0.8 ± 1.5	86 (82-89)	100 (99-100)
Greer Track	3.2 ± 1.3	16.5 ± 6.4	0.42	0.012 ± 0.1	3.4 ± 1.4	56 (52-60)	99 (98-99+)

*Values for wheal and flare expressed in millimeters.

TABLE II. Number of tests that exceed 3 mm for saline wheal and 10 mm for saline flare, and number of tests that are below 3 mm for histamine wheal and 10 mm for histamine flare

		Histamine wheal, mm		Histamine flare, mm		Saline wheal, mm		Saline flare, mm	
	Total test	<3	Range	<10	Range	>3	Range	>10	Range
Single devices									
Sharptest	208	3	0-10	1	9-60	0	0-3	1	0-15
Greer Pick	208	2	0-12	2	7-75	0	0	0	0-10
Accuset	208	8	0-8	11	0-50	1	0-4	1	0-12
Quintip	208	5	0-10	9	0-40	0	0	2	0-15
Multidevices									
Multi-Test II	1664	57	0-12	42	0-50	4	0-4	4	0-20
Quantitest	1664	94	0-11	85	0-62	1	0-5	3	0-32
Quintest	1040	73	0-11	93	0-50	0	0	1	0-15
Greer Track	1664	366	0-11	361	0-47	2	0-5	1	0-22

TABLE III. Intradevice variability for multiheaded devices expressed as CV

Multidevices	Histamine wheal,* mean ± SD	CV† (interquartile range)
Multi-Test II	5.9 ± 1.3	0.20 (0.14-0.44)
Quantitest	5.7 ± 1.6	0.23 (0.14-0.50)
Quintest	4.3 ± 1.4	0.25 (0.18-0.59)
GreerTrack	3.2 ± 1.3	0.93 (0.70-1.39)

*Values expressed in millimeters.
†CV median (interquartile range), intradevice variability.

Sensitivity and specificity

The results of device sensitivity and specificity are listed in Table I. All single devices and the Multi-Test II had sensitivities >90%, and there was no significant difference in sensitivity among the single devices and the Multi-Test II. The Multi-Test II was more sensitive compared with all other multiheaded devices ($P < .002$). The Quintest was less sensitive than the Greer Pick, Sharptest, and Multi-Test II. The Greer Track was less sensitive than all other devices ($P < .0005$).

Arm versus back comparisons

There was a significant difference in histamine wheal sizes between the arms and backs for all devices ($P < .0005$; Fig 3). Histamine wheals for all single devices were significantly larger on the arms ($P < .05$ for all comparisons), and wheals for all multiheaded devices (except the Quintest) were larger on the back ($P < .0013$). The Quintest device was larger on the back, but this difference did not reach statistical significance ($P = .17$). There was no significant difference between upper and lower arm. There also was no significant difference between upper and lower back.

Multidevices: intradevice variability

Intradevice variability reactions are presented in Table III. Analyzing the multiheaded devices for intradevice variability, there were significant differences in the wheal sizes between the various heads for each device ($P < .009$ for all devices). Fig 4 illustrates the intradevice variability for the Greer Track. With the 8-headed devices (Greer Track, Multi-Test II, and Quantitest), the greatest degree of variability was found comparing the interior heads (S2, S3, S6, S7) with the corner heads (S1, S4, S5, S8) for all of them.

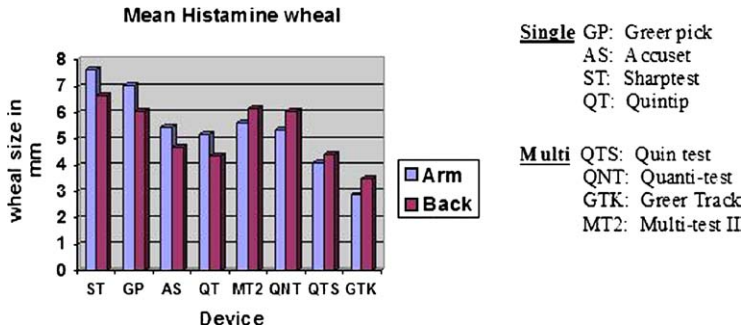

FIG 3. Mean histamine wheal size in millimeters for all devices.

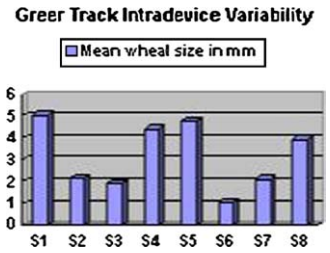

FIG 4. Mean intradevice variability of the Greer Track. Sites are labeled S1 through S8, with sites *S1, S4, S5, and S8* representing the corners and *S2, S3, S6, and S7* representing the interior heads.

The greatest degree of intradevice variability was found within the Greer Track.

TABLE IV. Pain outcome measures for 8 devices

	Mean pain	Median pain	Pain %* (95% CI)
Single devices			
Sharptest	1.17	1	13% (7-22)
Greer Pick	0.88	1	5% (1-11)
Accuset	0.94	1	9% (4-16)
Quintip	1.0	1	7% (2-14)
Multidevices			
Multi-Test II	1.62	1	26% (17-36)
Quantitest	1.74	1	26% (17-36)
Quintest	1.45	1	17% (10-26)
Greer Track	2.04	2	34% (23-43)

*Percentage of values above 2 on the Wong-Baker FACES pain rating scale (percentage of values interpreted as greater than mild pain).

Pain

Median pain scores for all of the devices was 1.0, except for the Greer Track, with a median of 2.0 (Table IV). Reports of pain were considered minor, with only 1 pain rating reported above 6 on a scale from 0 to 10 on the Wong-Baker FACES pain rating scale.[8] The highest pain rating was for the Greer Track (34% of pain scores above 2), and the minimum pain reported was for the Greer Pick (5% of pain scores above 2). All single devices were significantly less painful than the multiheaded devices ($P < .0005$). Comparing the single devices, Sharptest pain scores were significantly higher than Greer Pick ($P < .0005$) and Accuset ($P = .001$). For the multidevices, Greer Track scores were significantly higher than all other devices ($P < .0005$ for all comparisons), and the Quantitest was more painful than the Quintest ($P = .001$). In addition, pain was negatively associated with sensitivity ($r = -0.77$; $P = .027$), because the devices with greater sensitivity had lower pain scores. For the Greer Track multidevice with 56% sensitivity, 34% of pain scores were above 2. For the Greer Pick single device with 98% sensitivity, only 5% of pain scores were above 2.

DISCUSSION

We have concluded a head-to-head prospective comparative study of 8 skin test devices and found that there are statistically significant differences among virtually all devices tested. One device that stands out with the lowest performance in all areas is the Greer Track. This device was the most painful, had the smallest mean histamine wheals and flares, was the least sensitive, and had the greatest degree of intradevice variability. These statistical findings of performance may very well equate to clinically significant differences in performance. Excluding the Greer Track, it is unknown whether the statistical differences among the remaining 7 devices will equate to clinically significant differences in performance. Of the remaining 7 devices, all had mean histamine wheals greater than 3 mm and mean histamine flares greater than 10 mm, with sensitivities from 86% to 97%. In addition, all of the remaining 7 devices had specificities of 98% or greater. Therefore, each individual provider should determine which device is best for that provider's practice. Keep in mind that these studies were performed under the best of circumstances, with all tests conducted by 1 technician who was certified by a representative of the manufacturer on the proper use of each device. We would recommend that technicians within a given practice undergo this same type of training before using a given device. In addition, these findings may not be directly applicable to allergen skin testing, because we looked only at histamine and glycerol-saline responses. A separate study may be required to compare devices for this purpose.

When choosing a skin test device, a few points are worth consideration. First, in our study, single devices had larger reactions on the arm, and multidevices had larger reactions on the back. Historically, it has been thought that the back has been more reactive than the arms, and although our multiheaded device data concur with this, our single-headed device data do not. There are several possibilities for this observation, including intraoperator variability, inadvertent operator bias, or a true difference. With regard to the difference between single and multiheaded devices, we think that this difference is related to the back being a flatter surface; therefore, better contact is made with all of the test sites on a multiheaded device. The arms are technically more challenging when placing a multiheaded device, given natural curvatures. To compensate for the curvatures and to ensure contact with all of the device heads, manufacturers have recommended a rocking motion. With this motion, contact between all heads on a multiheaded device and the skin is achieved. However, we think that this rocking motion is responsible for the differences noted between individual test sites within a given multiheaded device, or intradevice variability (Fig 4). With this rocking motion, more pressure is exerted on the skin from the corner test sites. It is important to note that single devices also have differences in the recommended technique of application. The Quintip and Sharptest use a simple downward perpendicular pressure, and both of these devices have a depth control feature. Manufacturer-recommended techniques for the Greer Pick and Accuset are slightly more complicated. The skin surface is penetrated at an angle, and then a flick, prick-not-puncture technique is used. Neither of these last 2 devices has a depth control feature. Given the technique and lack of depth control, the Greer Pick and Accuset may result in greater intertechnician variability if care is not taken to control for these features. However, with correct technique, and by using 1 technician, all single devices had sensitivities greater than 90% while maintaining specificities of 98% or greater.

Another observation from our study is that skin testing is not a painful procedure on average. The mean pain scores for all devices ranged from 0.88 to 2.04. Using the Wong-Baker FACES pain rating scale,[8] this is considered mild pain. The degree of pain was significantly associated with the type of device used, with the multiheaded devices more painful than the single devices. However, when making this comparison, it is noteworthy that with a minimal increase in pain, as many as 8 times more tests are applied. Therefore, with a pain score of 0.88, the Greer Pick applied 1 skin test, and with a pain score of 1.62, the Multi-Test II applied 8 skin tests. It is unclear whether these observed statistical differences in pain would equate to significant clinical differences, because all pain was considered mild.

In contrast with previous studies, we did not find a clear relationship between pain and wheal size.[6] In fact, the device that resulted in the greatest degree of pain had the smallest mean histamine wheal size. Again, we think it is up to the individual practitioner to consider these differences when using a given device in practice.

Finally, when comparing sensitivity and specificity, there are few differences among devices, with 2 exceptions. The Greer Track was less sensitive than all other devices, and the Quintest was less sensitive than the Greer Pick, Sharptest, and Multi-Test II. With the 6 remaining devices, there was no significant difference among sensitivities. In addition, we found no significant differences in specificity among devices. Overall, we found a very low false-positive rate in all devices when using the manufacturer's recommended skin testing technique.

CONCLUSION

We have completed a prospective, head-to-head comparison of the performance of 8 skin test devices. This study was performed under the best of clinical circumstances, with 1 technician, trained by a representative of the manufacturer, who performed all skin testing, and another technician who read all of the results. We have found significant differences among all devices tested. Whether this equates to clinical differences is yet to be determined. Overall, skin testing is associated with minimal pain, and individual providers should choose a skin test device on the basis of their own practice setting. As new devices are being produced, this study suggests the need for continued evaluation of these devices in a prospective, nonbiased fashion.

REFERENCES

1. Li JT, Lockey RF, Bernstein IL, Portnoy JM, Nicklas RA. Allergen immunotherapy: a practice parameter. Ann Allergy Asthma Immunol 2003;90(suppl):1-40.
2. Position paper: immunotherapy. (EAACI) The European Academy of Allergology and Clinical Immunology. Allergy 1993;48(suppl 14):7-35.
3. Spector SL, Nicklas RA. Practice parameters for the diagnosis and treatment of asthma. J Allergy Clin Immunol 1995;96:707-870.
4. Adinoff AD, Rosloniec DM, McCall LI, Nelson HS. A comparison of six epicutaneous devices in the performance of immediate hypersensitivity skin testing. J Allergy Clin Immunol 1989;84:168-74.
5. Nelson HS, Rosloniec DM, McCall LI, Ikle D. Comparative performance of five commercial prick skin test devices. J Allergy Clin Immunol 1993; 92:750-6.
6. Nelson HS, Lahr J, Buchmeier A, McCormick D. Evaluation of devices for skin prick testing. J Allergy Clin Immunol 1998;101:153-6.
7. Nelson HS, Kolehmainen C, Lahr J, Murphy J, Buchmeier A. A comparison of multiheaded devices for allergy skin testing. J Allergy Clin Immunol 2004;113:1218-9.
8. Wong-Baker FACES Pain Rating Scale. In: Wong DL, Hockenberry-Eaton M, Wilson D, Winkelstein ML, Schwartz P. Wong's essentials of pediatric nursing. 6th ed. St Louis: Elsevier; 2001. p. 1301-2.