

Comparison of the Multi-Test II and ComforTen allergy skin test devices

Mark S. Dykewicz, M.D.,^{1,2} Kevin T. Doms, M.D.,¹ and Deborah L. Chassaing, R.N.¹

ABSTRACT

The performance of different devices for percutaneous allergy skin testing may have statistically and clinically significant differences. We compared two multihead allergy skin testing devices: Multi-Test II (Lincoln Diagnostics, Decatur, IL) and ComforTen (Hollister-Stier Laboratories, Spokane, WA). Skin tests with glycerinated histamine (6 mg/mL base) and glycerinated saline were applied to 30 adults using Multi-Test II on the volar surface of one forearm and ComforTen on the opposite forearm. For each device, data were obtained from 150 histamine sites and 90 negative control sites. Defining a positive result using cutoff wheal sizes of either 5-mm versus 3-mm inclusive, respective results for Multi-Test II sensitivity were 97% versus 100% with specificity of 100% versus 97%, whereas corresponding ComforTen sensitivity increased from 26 to 82% with specificity of 100% versus 99%. For Multi-Test II versus ComforTen, histamine mean (SD) wheal sizes were 7.47 (1.72) mm versus 3.93 (1.59) mm ($p = 0.00$), mean coefficients of variation (CV) were 23.0% versus 40.5%, and pooled estimates of variance were 1.42 versus 1.29. Comparing results at different test head positions, there was no statistically significant variation in histamine wheal sizes with either device. Multi-Test II had notably lower CV, greater wheal size, and higher sensitivity but similar specificity to ComforTen. Consequently, Multi-Test II had superior performance at both 3- and 5-mm wheal cutoffs. Because ComforTen had a low sensitivity at the 3-mm and, particularly, at the 5-mm wheal cutoff, skin testing with this device might result in underdiagnosis of allergy using either cutoff.

(Allergy Asthma Proc 32:198–202, 2011; doi: 10.2500/aap.2011.32.3441)

Although introduced as early as the 1860s, percutaneous skin testing remains a well-recommended, reliable, cost-effective technique for the diagnosis of IgE-mediated allergic diseases, despite more recent advances in the development of *in vitro* testing for specific IgE.^{1,2} Advantages of immediate hypersensitivity skin testing include rapid availability of results, high diagnostic sensitivity, technical simplicity, and low cost. Over the last several decades, successively modified and improved percutaneous skin testing devices have been introduced, with a notable trend toward greater use of multihead devices. The principal advantages of multihead devices are the ability to simultaneously

apply multiple allergen extracts to the skin, with associated time efficiency and greater acceptability by patients compared with individual needle-prick testing methods.^{3,4} Because therapeutic decisions are dependent on immediate hypersensitivity skin test results, skin test devices should reliably minimize false negative and false positive results.

In comparing percutaneous skin test devices, previous studies^{5–9} have shown statistically and clinically significant differences in performance characteristics such as the size of reactions to histamine or allergen extract and the likelihood of reactions at negative control sites. Publication of studies comparing marketed skin test devices is of particular importance to physicians considering that the U.S. Food and Drug Administration (FDA) does not require proof of efficacy or safety of skin test devices under its 510(k) submission approval process for marketing of skin test devices. The purpose of this study was to compare two multihead allergy skin testing devices that have been approved for marketing by the FDA: Multi-Test II (Lincoln Diagnostics, Decatur, IL) and ComforTen (Hollister-Stier Laboratories, Spokane, WA).

METHODS

Study Design

This study was conducted with the approval of the Saint Louis University Institutional Review Board, and written informed consent was obtained from all re-

From the ¹Section of Allergy and Clinical Immunology, Division of Immunobiology, Department of Internal Medicine, Saint Louis University School of Medicine, Saint Louis, Missouri, and ²Allergy and Immunology Unit, Section of Pulmonary Critical Care, Allergy and Immunologic Diseases, Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina. Presented in part at the annual meeting of the American Academy of Allergy, Asthma, and Immunology, March 16, 2009, Washington, D.C.

Sponsored by Lincoln Diagnostics, manufacturer of Multi-Test II, which advised the investigators in the design of the study. The conduct of the study, the data analysis, and the writing of the article were performed solely by the investigators without interference by the sponsor. The sponsor had the opportunity to review the data analysis and to provide comments about the article that were considered by the authors before they submitted it for publication. No substantive changes in the data analysis or the stated conclusions of the author-investigators occurred because of sponsor review or comments.

Address correspondence and reprint requests to Mark S. Dykewicz, M.D., Wake Forest University Allergy and Immunology Unit, Center for Human Genomics and Personalized Medicine Research, Medical Center Boulevard, Winston-Salem, NC 27157. E-mail address: dykewicz@wfubmc.edu

Copyright © 2011, OceanSide Publications, Inc., U.S.A.

search participants. Recruitment was from the general public (predominantly from the university community of employees and students) with subjects eligible for inclusion if they were adults aged 18–65 years and not taking medications that would interfere with skin testing. There was no inclusion criterion to require that subjects be atopic, and we did not screen subjects to exclude those who might have dermatographism. Skin tests with glycerinated histamine (6-mg/mL base) and glycerinated saline were applied to 30 adults on the volar surface of one forearm using Multi-Test II and on the opposite forearm using ComforTen. Both devices were loaded by inserting device points into histamine and glycerosaline solutions contained in wells designed by the manufacturer of each product. Participants alternated between receiving Multi-Test II on the right forearm and ComforTen on the left forearm and Multi-Test II on the left and ComforTen on the right. One person highly experienced with epicutaneous and percutaneous testing administered all skin tests (D.L.C.), and a different experienced person read them (K.T.D.). Wheal sizes were calculated from the average of the longest diameter and its midpoint orthogonal diameter using the central portion of the wheal. Wheals were measured for both devices at 20 minutes, in conformance with the 15- to 20-minute interval recommended in the product insert for Multi-Test II when skin tests are performed with glycerinated extracts. ComforTen does not have a product insert that provides recommendations about when skin tests should be read. Before each session, the following medications were withheld for specified periods: oral antihistamines, 7 days; intranasal antihistamines (azelastine), 2 weeks; phenothiazines, 7 days; all antidepressants, 2 weeks (albeit skin test inhibition by antidepressants is primarily an issue with tricyclic antidepressants); and histamine 2 receptor antagonists, 24 hours.

Of the eight skin test heads of the Multi-Test II device (Fig. 1), histamine was applied with heads 1 through 5, and the negative control was applied with heads 6 through 8. This yielded five histamine skin test sites and three glycerinated saline skin test sites per participant. Of the 10 test heads of the ComforTen device, histamine was applied with heads 1 through 5, and negative control was applied with heads 6 through 8; no test solutions were applied with heads 9 and 10. This yielded the same number of histamine and glycerinated saline sites as for Multi-Test II.

Statistical Analysis

Mean wheal sizes of the two devices were compared using *t*-tests and analysis of variance. Statistical significance was determined using $\alpha = 0.05$ for all tests. For each device the variance between participants was removed to examine the pooled estimates of variance. To

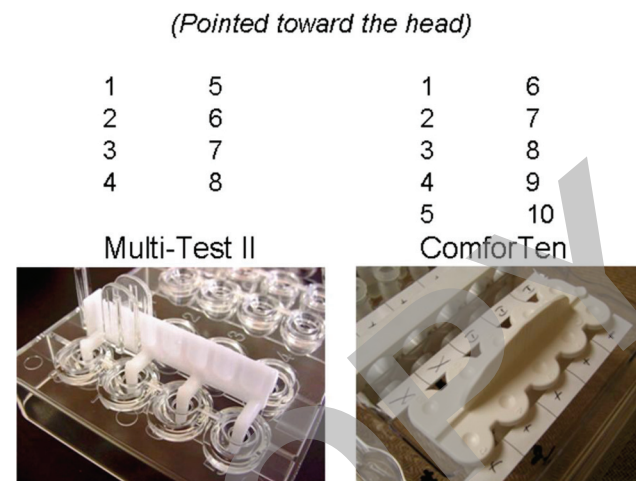


Figure 1. Configurations used for test head numbering. To permit easier comparison between the two devices, numbers were assigned to test heads in an analogous scheme for both devices, although this resulted in Multi-Test II test heads being numbered in a mirror image of that indicated in the product insert.

determine precision, intradevice variability was described by using the coefficients of variation of the mean diameters of wheal sizes of individual heads.

The Tukey honest significant difference test and the Games-Howell test were used to examine the differences in mean wheal size for each histamine position on the arm for each device separately. Additional presentation of medians with interquartile ranges for each device was produced in the form of box plots. Separate sensitivity and specificity analyses were performed using a wheal of at least 3 or 5 mm to indicate a positive result and a wheal of <3 or <5 mm to indicate a negative result. Specificity was calculated by dividing true negative results by the sum of true negative and false positive results.

RESULTS

The trial accumulated data from 30 subjects (aged 20–64 years, 12 men and 18 women) for collectively 150 histamine sites and 90 negative control sites for each device. For Multi-Test II versus ComforTen, histamine mean (SD) wheal sizes were 7.47 (1.72) mm versus 3.93 (1.59) mm ($p = 0.00$), mean coefficients of variation were 23.0% versus 40.5%, and pooled estimates of variance were 1.42 versus 1.29. Compared with ComforTen, Multi-Test II histamine mean wheal sizes were significantly larger ($p = 0.000$; Table 1; Fig. 2). For each device, histamine wheal sizes did not statistically vary with test head position (Tukey honest significant differences test, Table 2). Glycerosaline mean wheal sizes were 0.01 mm (range, 0–4 mm) for Multi-Test II and 0 mm (range, 0–1 mm) for ComforTen.

Table 1 Histamine wheal sizes for both allergy skin test devices

Device	No. Test Sites	Wheal diameter, Mean (95% CI) (mm)	SD	CV (mean, %)	Pooled Estimate of Variance
Multi-Test II	150	7.47	1.72	23.0	1.421
ComforTen	150	3.93	1.59	40.5	1.289

CI = confidence interval; CV = coefficient of variation.

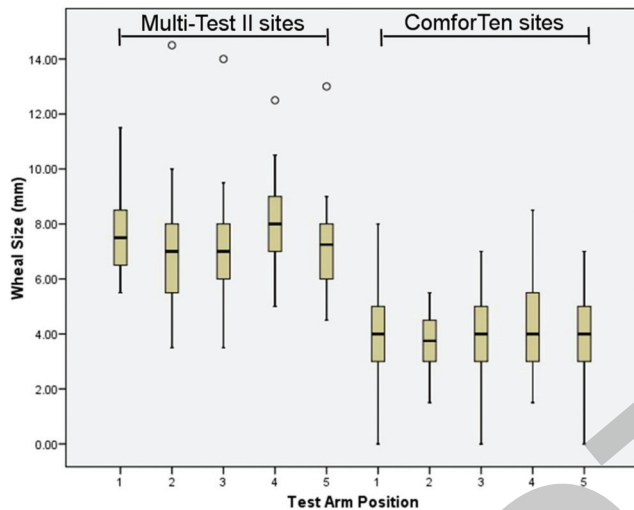


Figure 2. Box plot of histamine mean wheal diameters (in millimeters) by device and test head position (MT2 on left and CT on right). Heavy line indicates median; box indicates interquartile range (25th–75th percentile) of the data; and the whiskers from the 25th and 75th percentile to the furthest observation or 1.5 times the length of the box, whichever is shorter. Outliers (circles) lay within 1.5–3 box lengths from the edge of the box.

Table 2 Histamine wheal diameter by device and test head position*

	Test Head Position				
	1	2	3	4	5
Multi-Test II mean (SEM)	7.62 (0.25)	7.15 (0.37)	7.22 (0.33)	8.12 (0.29)	7.25 (0.29)
ComforTen mean (SEM)	3.92 (0.30)	3.75 (0.18)	3.98 (0.33)	4.33 (0.30)	3.68 (0.33)

*Data are given as standard error of the mean (SEM) in millimeters.

Defining a positive result using cutoff wheal sizes of either 5-mm versus 3-mm inclusive, respective results for Multi-Test II sensitivity were 97% (95% confidence interval [CI], 93–99%) versus 100% (95% CI, 97–100%) with of specificity 100% (95% CI, 95–100%) versus 97% (95% CI, 90–100%), whereas ComforTen sensitivity increased from 26% (95% CI, 19–34%) to 82% (95% CI,

75–88%) with specificity of 100% (95% CI, 95–100%) versus 99% (95% CI, 93–100%).

DISCUSSION

This prospective, head-to-head trial compared two multihead allergy skin testing devices approved by the FDA for marketing: Multi-Test II and ComforTen. Our results in this study of adults found statistically significant differences among various performance measures between the two devices when applied to the forearms. Compared with ComforTen, Multi-Test II performed better, with a lower mean coefficient of variance, a larger mean histamine wheal size, and higher sensitivity. Previous studies have reported that skin testing devices producing smaller wheals with histamine are more likely to yield false negative reactions.^{5,6} This is consistent with our findings that Multi-Test II had higher sensitivity than ComforTen, whether defining a positive result using cutoff wheal sizes of 5 or 3 mm inclusive. Several previously published studies have compared Multi-Test and Quintip (Hollister-Stier Laboratories),^{7,8} the latter device having essentially similar test heads to ComforTen although in a linear arrangement of five test heads. These studies found that Quintip produced smaller histamine wheals than Multi-Test.

In considering potential reasons for the performance differences found in this study between Multi-Test II and ComforTen, one can speculate that differences in test head point structure, length, and test solution loading characteristics are relevant. Each Multi-Test II test head contains nine tines that are arrayed in an area measuring 2 × 2 mm, with tines being 2 mm in length. As shown in Fig. 3, test solutions adhere to the injection molded acrylic test heads by apparent strong capillary action that helps distribute solution to the tine tips and promote solution delivery at the point of skin contact. In contrast, each ComforTen test head contains a single stainless steel point attached to a plastic leg, with each point being only half as long (1 mm) as the Multi-Test II tines. With this ComforTen test point structure, capillary action pulls test solution to accumulate at the plastic base rather than at the steel point tip (Fig. 4).

ComforTen does not contain a specific recommendation in its product insert to define a positive reaction



Figure 3. Multi-Test II testing head charged with glycerinated allergy testing solution. Capillary action helps distribute solution to the nine tine tips and promote solution delivery at the point of skin contact.

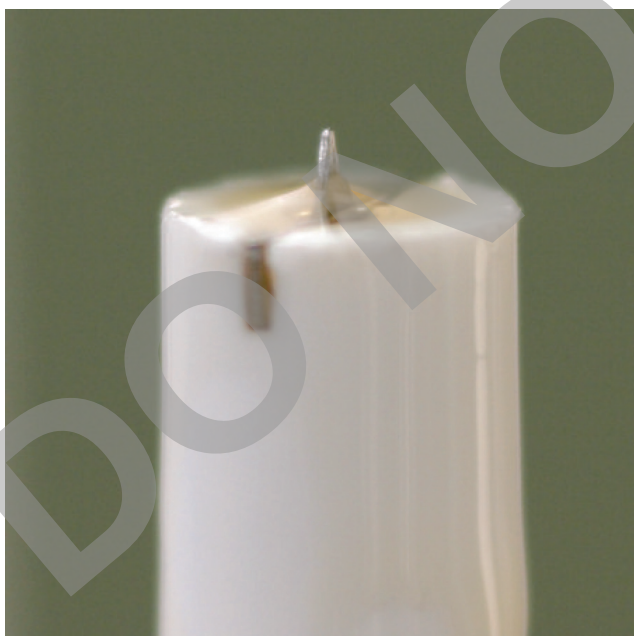


Figure 4. ComforTen testing head charged with glycerinated allergy testing solution. Capillary action pulls test solution to accumulate at the plastic base away from the steel tip at the point of skin contact.

based on wheal diameter. Our results indicate that using a 3-mm rather than 5-mm inclusive wheal size would lead to better sensitivity with no appreciable loss of specificity when this device is used. Although Nelson *et al.*⁷ have stated that devices that produce

larger reactions at histamine sites can also produce larger reactions at negative controls that could be read as false positive, for Multi-Test II, our results found high sensitivity and specificity using either 3- or 5-mm inclusive wheal sizes, but our data do support the recommendation in the Multi-Test II product insert to use a 5-mm wheal cutoff to define a positive reaction.

In this study, we read skin tests at 20 minutes, in conformance with the 15- to 20-minute interval recommended in the product insert for Multi-Test II when skin tests are performed with glycerinated extracts. Wheals from histamine may peak earlier than 20 minutes,² particularly with aqueous histamine extracts, which may peak as early as 10 minutes after application.¹⁰ However, in current clinical practice the wheals from glycerinated histamine and allergen extracts are often read simultaneously for comparison, and a 20-minute reading interval reflects a time when histamine wheals are commonly measured.

This was an unblinded prospective observational study because it is impossible to perform a blinded study to completely minimize bias when comparing two multi-head skin test devices that have different numbers of skin test heads and create distinctively different patterns of test head impressions on the skin. To reduce operator bias, one experienced person administered all skin tests, but a different experienced person read them. Before the study, the person administering skin tests was highly experienced using Multi-Test II but had limited previous experience using ComforTen. Although she administered ComforTen properly and in keeping with the manufacturer's instructions for use and had undergone performance testing with ComforTen just before beginning the study, we can not completely exclude the possibility that her greater previous experience using Multi-Test II may have contributed to the superior results observed with that device.

Some previous studies^{8,9} have found that with some multihead skin test devices and application techniques, wheal sizes produced by "inner" test heads may be statistically smaller than wheals produced by "corner/outer" test heads. Several factors have been proposed to contribute to this issue, including application of too much pressure on corner test heads after a rocking motion that is too vigorous,⁸ and other factors that may be device specific.⁹ In this study, we found that histamine wheal sizes were not significantly different in relation to head test position for either ComforTen or Multi-Test II. So that we could analyze an equal number of test heads when comparing the 8-head Multi-Test II and 10-head ComforTen, we analyzed histamine responses at five positions and glycerol-saline at three positions for each device. For ComforTen, histamine response was measured at two of four outer positions (positions 1 and 5), an identical approach used by us in a previous study comparing the 10-head SkinTestor

Omni to Multi-Test II.⁹ In the previous study, our analytic approach clearly showed that there was significant intradevice variability in histamine wheal sizes at different test head positions and smaller wheals at inner heads with the SkinTestor Omni device, but no statistically significant difference in wheal size in relation to test head position for Multi-Test II.⁹

In conclusion, we completed a prospective, head-to-head trial comparing two FDA-approved multihead allergy skin testing devices: Multi-Test II and ComforTen. Of the two devices, Multi-Test II performed better, with a larger mean histamine wheal size and higher sensitivity. Specificity and intradevice variability were similar. In terms of how our results would have clinical relevance to skin testing with allergen extracts, one example is that when a device has lower sensitivity as assessed by responses to histamine controls, an unacceptable number of patients may have negative histamine controls, precluding reliable interpretation of allergen skin test results. Although one might speculate that application of skin tests with ComforTen might lead to underdiagnosis of allergy when tested extracts are applied with test heads prone to produce smaller wheal sizes with other performance characteristics such as lower sensitivity, our results may not be directly applicable to testing for allergen skin test responses because this study only looked at responses to histamine and glycerinated saline. However, our methodology is a standard, well-established approach that has been used to compare the performance of different skin test devices.⁵⁻⁹ We did not exclude patients with known dermatographism from the study, but because none of our study population showed significant dermatographism that confounded skin test interpretation, our results may not apply to patients with significant dermatographism. Because our study examined skin

test responses on the forearm only in adults, our results might differ from skin testing performed on other areas of skin such as the back or performed in children.

ACKNOWLEDGMENTS

The authors thank Patrick Ercole of the Saint Louis University Center for Outcomes Research (SLUCOR) for performing statistical analysis.

REFERENCES

1. Taylor G, and Walker J. Charles Harrison Blackley, 1820–1900. *Clin Allergy* 3:103, 1973.
2. Bernstein IL, Li JT, Bernstein DI, et al. American Academy of Allergy, Asthma and Immunology; American College of Allergy, Asthma and Immunology. Diagnostic testing: An updated practice parameter. *Ann Allergy Asthma Immunol* 100: S1–S148, 2008.
3. Rizzo MC, Naspitz CK, and Sole D. Comparative performance for immediate hypersensitivity skin testing using two skin prick test devices. *J Allergy Clin Immunol* 5:354–356, 1995.
4. Huber B, and Berdel D. A comparison of the Multi-Test with the prick test and Multi-Test 2. *Clin Allergy* 13:467–472, 1983.
5. Adinoff AD, Rosloniec DM, McCall LI, and Nelson HS. A comparison of six epicutaneous devices in the performance of immediate hypersensitivity skin testing. *J Allergy Clin Immunol* 84:168–174, 1989.
6. Nelson HS, Rosloniec DM, McCall LI, and Ikle D. Comparative performance of five commercial prick skin test devices. *J Allergy Clin Immunol* 92:750–756, 1993.
7. Nelson HS, Lahr J, Buchmeier A, and McCormick D. Evaluation of devices for skin prick testing. *J Allergy Clin Immunol* 101: 153–156, 1998.
8. Carr WW, Martin B, Howard RS, et al. Immunotherapy Committee of the AAAAI. Comparison of test devices for skin prick testing. *J Allergy Clin Immunol* 116:341–346, 2005.
9. Dykewicz MS, Lemmon JK, and Keaney DL. Comparison of the Multi-Test II and Skintestor Omni allergy skin test devices. *Ann Allergy Asthma Immunol* 98:559–562, 2007.
10. Aas K. Some variables in skin prick testing. *Allergy* 67:64–67, 1980. □

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.