

Automated measurement of skin prick tests: an advance towards exact calculation of wheal size

Wöhrl S, Vigl K, Binder M, Stingl G, Prinz M. Automated measurement of skin prick tests: an advance towards exact calculation of wheal size.

Exp Dermatol 2006; 15: 119–124. © 2006 The Authors. Journal compilation © 2006 Blackwell Munksgaard

Background: Using the Khoros[®] image processing software environment, we developed a software-based system capable of extracting wheal size from skin prick tests (SPTs) in mm² for research and routine purposes.

Methods: At 20 min, the outlines of up to 20 wheals were marked with a pen and transferred with translucent adhesive tapes to a white paper form carrying predefined markings. The form was scanned at 200 dpi. The software automatically analysed the scanned image and calculated the sizes of the wheals. In a pilot study, serial SPTs with histamine in increasing dilutions were performed in 12 healthy volunteers in duplicate on both volar forearms. We matched the application results with a reference created from the scanned pictures.

Results: Bland–Altman analysis showed reference and software calculation reaching very high agreement. The comparison of reference/software resulted in a low centred coefficient of variation (COV) of 11.9%. This was superior to the conventional measurement of horizontal (COV 37.9%) or maximal/minimal diameter (COV 25.9%).

Conclusion: We present an accurate tool for exactly calculating SPT wheal size in mm².

**Stefan Wöhrl¹, Kornelia Vigl¹,
Michael Binder², Georg Stingl¹
and Michael Prinz³**

¹Department of Dermatology, Division of Immunology, Allergy and Infectious Diseases (DIAID), Medical University of Vienna, Vienna, Austria

²Department of Dermatology, Division of General Dermatology, Medical University of Vienna, Vienna, Austria;

³Core unit for Medical Statistics and Informatics, Section on Medical Computer Vision, Medical University of Vienna, Vienna, Austria

Key words: allergy – diagnosis – methodology – morphometry – wheal size

Stefan Wöhrl, MD, MSc
Medical University of Vienna
Department of Dermatology
Division of Immunology
Allergy and Infectious Diseases (DIAID)
Währinger Gürtel 18-20
A-1090 Vienna, Austria
Tel.: +43 1 40400 7704
Fax: +43 1 4031900
e-mail: stefan.woehrl@meduniwien.ac.at

Accepted for publication 26 October 2005

Introduction

Skin prick tests (SPTs) are a standard method for diagnosing type 1 allergy. Positive SPTs result in wheals. For routine and most study settings, the result is given as an estimated average diameter in mm. A nearly circular shape is a prerequisite for a diameter to represent an area. However, wheals often do not present as circles. Sometimes, they have pseudopodia, and measurement of the diameter must lead to error-prone results (Fig. 1a, first line). To avoid this, some authors measured the average of the maximum and the minimum diameter, which gives good results with nearly circular shapes (1,2) (Fig. 1a, second line). As diameters can be measured directly whereas the direct

calculation of wheal area needs a more sophisticated approach, the error-prone measurement of diameters has been accepted for most study settings (Fig. 1a, third line).

Wheals wane quickly and documentation must take place at the time of the reading. One way of doing so is taking digital photographs at the time of reading. The human visual interpretation of these was shown to cause significant amount of variation (3).

One study reported semiautomated readings of SPTs (4). In this study, human beings redrew the outer border of a digital picture of the wheal with a CAD software. Although their setting resulted in highly reliable and reproducible results, it involved the time-consuming manual picture analysis by a human being. This fact had limited this method for scientific purposes.

We wanted to develop an easy and fast tool for the calculation of wheal size in mm² irrespective

Abbreviations: CI, confidence interval; dpi, dots per inch; PC, personal computer; SD, standard deviation; SPT, skin prick test.

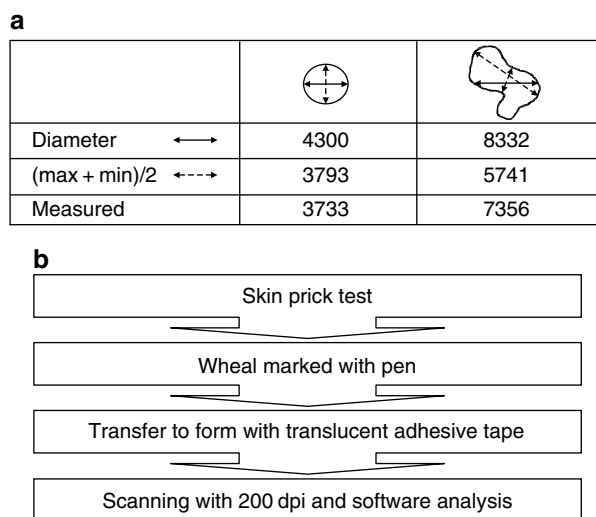


Figure 1. (a) Demonstration of errors due to inadequate estimation of wheal size in two hypothetical wheals, an ellipsoid on the left, a polygonal shape on the right. First line: Measurement of diameter. Calculation of area assuming a circle: $\text{area} = (\text{diameter}/2)^2 \times \pi$. Second line: Measurement of maximal and minimal diameter. Calculation of area assuming a circle: $\text{area} (\text{mm}^2) = (\text{max} + \text{min diameter})/(2)^2 \times \pi$. Third line: The wheals were transformed to pictures at a resolution of 200 dpi, and the area was directly measured with the IMAGEJ software (see *Methods*). All measurements are given as [pixels]. (b) Work flowchart. Standard SPTs with histamine were read at 20 min. The wheals' outlines were marked with a skin-friendly pen; a transparent adhesive tape was applied once to this marking. The tape with the marked wheal was then transferred to a standard form (Fig. 2). Next, the form was scanned at 200 dpi with a standard scanning device, and the wheals were analysed with the software application.

of the geometric shape. The system should analyse all wheals of an SPTs series in a single patient at once. Routine purposes demand reliable documentation and easy use. Therefore, we chose the common way of documentation by drawing the wheals' outlines with a colour pen. The outlines were then transferred with a translucent adhesive tape to a machine-interpretable standard form (Fig. 1b). We developed a software application able to calculate the wheal sizes of multiple SPTs after a single scanning of the form on a standard scanning device. We tested the application's clinical potential in a pilot study.

Methods

Integrating an algorithm into a clinical setting

Our approach employed the following standard conditions (visualized in Fig. 1b). SPTs were performed with standardized SPT lancets (Allergopharma, Reinbek, Germany) on the volar forearms. At the time of reading at 20 min, the wheal's outline was marked with a blue pen (Viomedex VX 100, Buxted, UK). Standard, transparent adhesive tapes were applied to the skin once and then transferred to a white DIN-A4 paper form with the predefined markings (Fig. 2). The markings had been applied to the sheet for allowing the software to adjust the position of the

scanned document automatically and thus obtaining a perfectly positioned form. The form was then scanned with a resolution of 200 dpi on a standard scanning device (HP ScanJet 7400 C, Hewlett-Packard Company, Palo Alto, CA, USA). The Khoros[®] Pro 2001 image processing software environment with the MMach-package for Khoros (Khoros Inc., Albuquerque, NM, USA) was used to develop an algorithm for the automated measurement of the inner area of the marked wheals (meanwhile, Khoros has been taken over by Accusoft Corporation and has been renamed to VisiQuest, URL: <http://www.accusoft.com>).

The algorithm

The algorithm automatically corrected the position of the form (Fig. 2) when the form had been scanned distortedly or displaced, extracted the regions of the applied wheal markings, determined the outer border of the wheals and measured their inner area in pixels (Fig. 3). The extraction of the wheal markings was performed by applying a sequence of mathematical morphological operators and dynamical thresholding. Our algorithm detected the border of the wheal, then smoothed the shape and finally counted the pixels within the determined border.

The accuracy of the algorithm was nearly independent of the size of the wheal's area and the shape of the wheal contour, as the difference between calculated and measured area remained at the same over the whole range of areas and shapes. Similar to Pijnenborg et al. (5), we originally observed slightly higher deviations from the actual size at small wheals $<5 \text{ mm}^2$.

Allergieambulanz Abt. f. Immun Dermatologie u. Infektiöse Hautkrankheiten Leiter: Univ.-Prof. Dr. G. Stingl Univ.-Klinik f. Dermatologie, AKH-Wien Währinger Gürtel 18-20, A-1090 Wien		Date:	
Control		5 left 1 : 100	5 right 1 : 100
1 left undiluted		1 right undiluted	
2 left 1 : 3		2 right 1 : 3	
3 left 1 : 10		3 right 1 : 10	
4 left 1 : 30		4 right 1 : 30	

Figure 2. Machine-readable standard form used for scanning. The markings on the top left and the right bottom had been applied for that the software could automatically correct the position of the form when the form had been scanned distortedly or displaced.

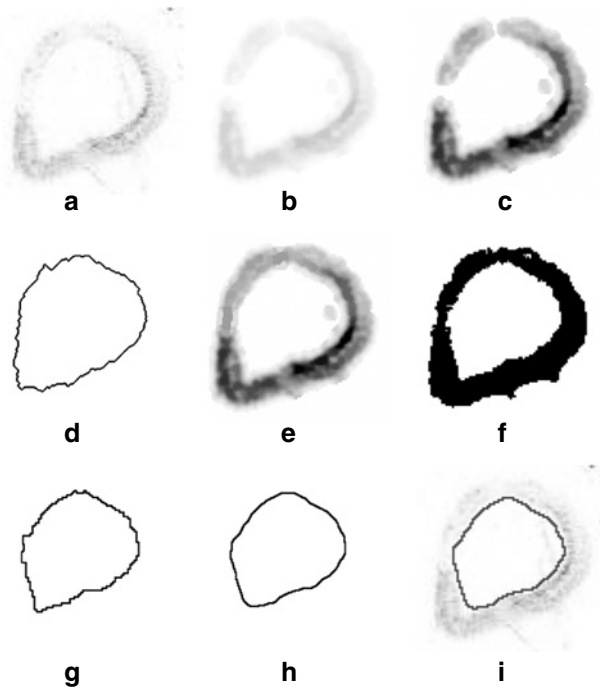


Figure 3. Sample wheal image illustrating the detection and reconstruction of the wheal's inner border. (a) Original image of the wheal contour. (b) Extraction of the image's blue colour channel and conversion to a greyscale image. (c) Stretching of the greyscale image's histogram to increase the contrast. (d) Watershed contour which resembles the wheal contour's middle contour. (e) Combination of image 'c' and dilated 'd' to close gaps in wheal contour. (f) Binary image obtained by thresholding image 'e'. (g) Extraction of the inner contour. (h) Inner contour's shape is smoothed. (i) Superimposition of reconstructed contour onto the original wheal contour.

However, we observed that the algorithm behaves precisely at larger wheals. Therefore, we enhanced the performance of our original algorithm by enlarging images containing small wheals by a factor of 1.5–3.0. After reconstructing the wheal contours accurately, the images were reduced to their original size and the size of the wheal's actual area was measured automatically.

The algorithm behaved very stable on analysing varying contrasts and intensities of the blue-coloured contour on the adhesive tape. By using mathematical morphological operations, even larger gaps in the contour were very well reconstructed. The exact description of the algorithm is published elsewhere (6). All calculations were performed in pixels and converted to mm^2 in a final step [for 200 dpi: (pixels) $\times$ 0.01613 = (mm^2)].

The algorithm generated a matrix, where each calculation of wheal size was assigned to the labelling according to its position on a standard form (Fig. 2). As a prerequisite for this automatic assignment, the investigator doing the SPTs (KV) always applied the wheals in the same order in all subjects. Empty fields were automatically interpreted as 'zero' (no reaction). A manual diagonal pen-stroke from the lower left to the upper right – drawn by the investigator who had read the SPTs at the time of reading – meant that the test in this dilution and location had not been performed, which led to the automatic interpretation as 'missing value'.

Pilot study

For the evaluation of this method's usefulness, we performed a pilot study with 12 healthy, Caucasian volunteers (six

women, six men, mean age 28.1 ± 3.6 years), who gave informed consent for taking part in this study. We produced 'standardized' wheals with SPTs with histamine. The SPTs were performed under standard conditions with standardized lancets and read after 20 min. We tested histamine for SPT (ALK-Scherax, Hamburg, Germany) in duplicate on the volar side of the left and right forearms in the following dilutions: undiluted (10 mg/ml), diluted in 0.9% saline 1:3, 1:10, 1:30, 1:100. One male individual had suffered from a non-allergic local reaction of Ø 5 cm after a yellow-jacket stinging on one of the upper limbs, and he was the only individual in whom SPTs were only performed on one arm. Summing up, there were 115 SPTs.

To evaluate the accuracy of our application, we created a manual reference analysis. One of the authors (SW) redrew the borders of the scanned images in a blinded fashion three times with a graphic tablet (Wacom ET-0405-R, Wacom Co., Kitasaitama-Gun, Japan) using the JavaTM-based (Sun Microsystems Inc., Santa Clara, CA, USA) public domain CAD software IMAGEJ at a magnification of 400% (download from URL: <http://rsb.info.nih.gov/ij/>). Redrawing with a CAD software had been shown to be a reproducible approach for calculating wheal size of SPTs (4,7).

The same conditions were applied for measuring the horizontal and the maximum/minimum diameter.

Statistics

We used the software package SPSS 12.0.1 for Windows (SPSS Inc., Chicago, IL, USA) for statistical analysis. Bland–Altman plots (8–10) were used to visualize agreement between the various measurements and were produced with the MEDCALC 8.0 for Windows software (Medcalc Software, Mariakerke, Belgium).

Results

Only 109 of 115 SPTs resulted in wheals and were therefore available for analysis (Table 1). Not all individuals developed wheals to all concentrations of histamine.

Evaluation of the application

Accuracy. We created manual reference measurements of the digital pictures as a reference for the software calculation, an approach already used by others for the calculation of the area of skin lesions (4,7,11). The reference measurement and the software reached highly corresponding results. The ratio of the reference (mean from three independent measurements) to software-calculated area was very close to 1, exactly 1.023 [99% (confidence interval) CI: 0.993–1.053] (Table 1). The mean centred coefficient of variation (COV) between the computed and the reference calculations was at a low 11.9%.

Confirming our null hypothesis, the software-computed area and the mean area of the three manual measurements were highly correlated with each other (Pearson's two-sided $R^2 = 0.987$, $P < 0.001$). Correlation as a statistical method was, however, shown to overestimate agreement

Table 1. Detailed results of calculated wheal size and the mean of a triple manual analysis of the five dilutions of histamine

		Wheal size (mm ²)			Calculated/measured		
	Positive skin prick tests	Calculated	3 × Measured			99% confidence interval	
			Mean	Standard deviation		Upper	Lower
Undiluted	23/23	25.48	24.71	8.10	1.033	0.983	1.083
1 : 3	23/23	22.63	22.10	12.02	1.030	0.990	1.070
1 : 10	22/23	16.81	16.45	8.41	1.039	0.973	1.104
1 : 30	20/23	11.64	11.77	7.33	0.981	0.929	1.034
1 : 100	21/23	9.49	9.71	7.23	1.029	0.896	1.161
Overall	109/115	17.51	17.23	10.84	1.023	0.993	1.053

The ratio of calculated/measured wheal size is depicted on the right-hand column.

between two methods measuring the same item (9). We therefore performed Bland–Altman plots, which is an adequate way of visualizing agreement between two measurements (8–10). The Bland–Altman plot of the mean of the three manual measurements and the software calculation of wheal area showed very good agreement [arithmetic mean: 0.28 mm²; standard deviation (SD): ± 1.72] (Fig. 4a).

Time consumption. At the moment, the processing speed of the wheal detection procedure still has to be optimized. The overall processing of an entire form with 20 SPTs took about 15 min on a standard 800 MHz personal computer (PC) due to the fact that the chosen method of mathematical operators for reconstructing and detecting the wheal borders is very calculation intensive. However, this processing is fully automated and does not need any user input and can therefore be performed over night or when the processor is idle.

Although the analysis itself is semiautomated, there remained two time-consuming manual preparative steps: first, marking of the wheal's border; second, transferring it to a sheet with a translucent tape (Fig. 1b). However, wheals have to be documented in study as well as routine settings anyway. Therefore, these steps do not require extra work force. Also, the software analyses all wheals at once, which is time saving compared to measuring the diameter of every single wheal. Visual analysis alone was shown to be insufficient for the estimation of wheal size (12).

Automated calculation is superior to conventional ways of wheal size measurement

The software calculation was superior to the conventional ways of measurement:

1. reference measurement/horizontal diameter 1.051 (99% CI: 0.951–1.151), COV 37.9%.

Figure 4(b) visualizes the much less agreement in another Bland–Altman plot (arithmetic mean: 0.14 mm²; SD: ± 5.99);

2. reference measurement/minimal and maximal wheal diameter, was 1.076 (99% CI: 1.006–1.146), COV 25.9%, Bland–Altman plot in Fig. 4(c) (arithmetic mean: 0.62; SD: ± 3.38 mm²).

Wheal size decreased on histamine dilution

All individuals had wheals to undiluted and 1 : 3 diluted histamine (Table 1). The amount of positive reactions decreased on dilution: 22, 20 and 21 of 23 SPTs resulted in wheals at the dilutions of 1 : 10, 1 : 30 and 1 : 100, respectively. The average calculated wheals covered 25.5 ± 8.6 , 22.6 ± 12.3 , 16.8 ± 8.3 , 11.6 ± 7.6 and 9.4 ± 6.8 mm² for undiluted, 1 : 3, 1 : 10, 1 : 30 and 1 : 100 diluted histamine (Table 1). Wheal area and dilution were negatively correlated (Pearson's two-sided $R^2 = -0.567$, $p < 0.001$).

A circle with a diameter of the standard cut-off of 3 mm (4) covers an area of 7 mm². Even the highest dilution of histamine had resulted in SPTs above cut-off (Table 1).

Notably, the standard deviations did not differ between the different dilutions that varied depending on the histamine concentration (one-sided ANOVA, $p = 0.647$). Hence, the performance of the algorithm did not depend on the wheal size.

Discussion

Our application calculated histamine-induced wheals with high accuracy. Bland–Altman plot analysis demonstrated a very good correlation between the triple manual and the application-calculated wheal areas (Fig. 4a). As for the requirements of this initiating study, the forms were scanned separately on a standard scanning

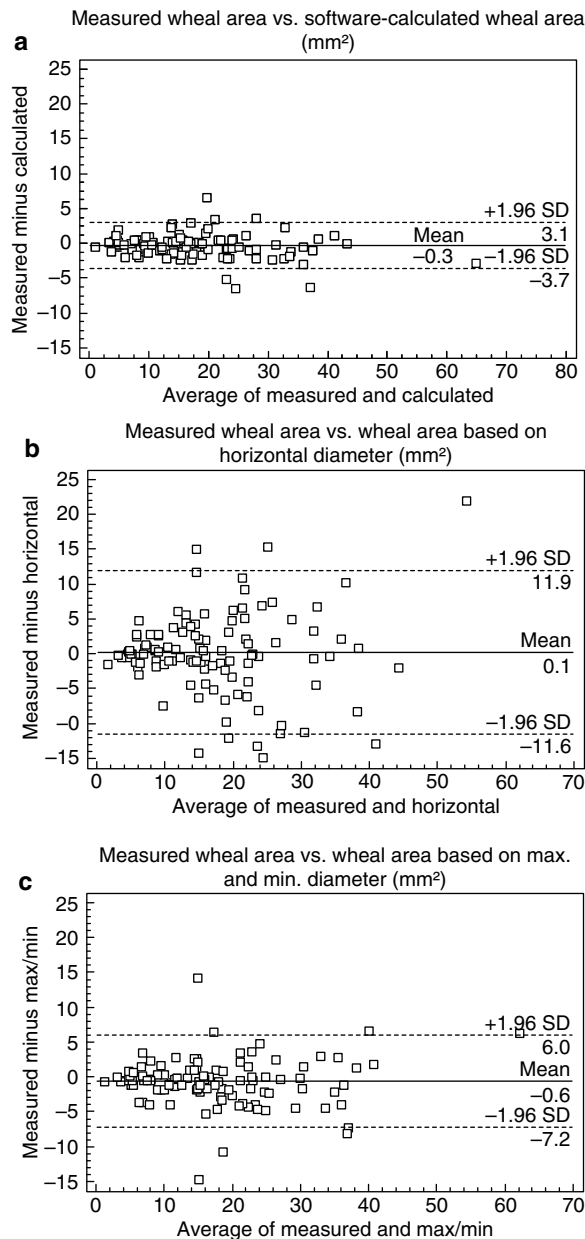


Figure 4. Bland-Altman plots with the triple manual measurement of the wheal size as reference vs. other methods of calculating wheal area. (a) Strong agreement with the software-calculated wheal area. (b) Mediocre agreement with the wheal area after measuring the horizontal diameter (according to Fig. 1a, first line). (c) Acceptable agreement with measuring maximum and minimum diameter (according to Fig. 1a, second line).

device. As can be seen in Fig. 2, a space is left blank for a patient's identification label. Adding a software routine capable of interpreting the scanned patient's identification label will allow the implementation into routine settings. For example, the forms could be collected over the day and analysed over night after a single mass scanning with a scanner with an automatic paper feed. The data could be allocated automatically to

the patients' electronic file according to the recognized patients' identification. This should spur high-quality documentation into routine settings.

Previously, other authors had undertaken the laborious way of exactly measuring wheal areas, e.g. Eigenmann and Sampson (13) directly scanned wheals with a hand-held scanning device. They did, however, not pursue this technique, as they did not find a superior predictive value of the scanned wheals for detecting food allergy in comparison with the ≥ 3 -mm wheal diameter criterion. Poulsen et al. (14,15) also used a hand-held scanning device for directly measuring wheals that had been hand-copied to plastic foils. High precision and reproducibility of their approach was confirmed in another study by Pijnenborg et al. (5). Both methods, however, were very laborious and did not meet the needs of routine documentation.

In contrast, our approach aimed at scientific as well as routine settings. The required equipment is cheap and probably already available at most units: a standard Windows- or Linux-PC available at a maximum of 800 Euro and a scanning device at not more than 250 Euro. However, the Accusoft® Runtime-Environment needs to be licensed. The whole procedure can be integrated into places already using electronic patient data management.

We have shown the reproducibility and reliability of our approach; the relevance of the results with histamine, however, needs validation with SPTs to allergen in well-characterized type 1 allergic patients.

Conclusion

We present a reliable tool for exactly determining multiple SPT wheal sizes in mm² in one session. With a few additions, it can be integrated into routine settings when already using electronic patient data management.

Acknowledgements

Kornelia Vigl was supported by a grant from MSD Austria, Vienna.

References

1. Perrin L F, Dechamp C, 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: 581-588.
2. Maccario J, Oryszczyn M P, Charpin D, Kauffmann F. Methodologic aspects of the quantification of skin prick test responses: the EGEA study. *J Allergy Clin Immunol* 2003; 111: 750-756.

3. McCann W A, Ownby D R. The reproducibility of the allergy skin test scoring and interpretation by board-certified/board-eligible allergists. *Ann Allergy Asthma Immunol* 2002; 89: 368–371.
4. Nogueira J M, de Almeida M M, Santa Marta C et al. Quantitative skin prick tests and specific IgE (CAP System) for *D. pteronissynus* – correlation of results in a paediatric population. *Allerg Immunol (Paris)* 1994; 26: 102–106.
5. Pijnenborg H, Nilsson L, Dreborg S. Estimation of skin prick test reactions with a scanning program. *Allergy* 1996; 51: 782–788.
6. Prinz M, Vigl K, Wöhrl S. Automatic measurement of skin wheals provoked by skin prick tests. *Stud Health Technol Inform* 2005; 116: 441–446.
7. Ownby D R. Computerized measurement of allergen-induced skin reactions. *J Allergy Clin Immunol* 1982; 69: 536–538.
8. Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
9. Bland J M, Altman D G. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 1995; 346: 1085–1087.
10. Bland J M, Altman D G. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8: 135–160.
11. Baran R, Sparavigna A, Setaro M, Mailland F. Computerized image analysis of nails affected by fungal infection: evaluation using digital photographs and manually defined areas. *J Drugs Dermatol* 2004; 3: 489–494.
12. de Weck A L, Derer T. Critical evaluation of the use of skin tests and cellular tests in standardization of allergens. *Arb Paul Ehrlich Inst Bundesamt Sera Impfstoffe Frankf A M* 1994; 87: 89–114; discussion 114–117.
13. Eigenmann P A, Sampson H A. Interpreting skin prick tests in the evaluation of food allergy in children. *Pediatr Allergy Immunol* 1998; 9: 186–191.
14. Poulsen L K, Liisberg C, Bindslev-Jensen C, Malling H J. Precise area determination of skin-prick tests: validation of a scanning device and software for a personal computer. *Clin Exp Allergy* 1993; 23: 61–68.
15. Poulsen L K, Bindslev-Jensen C, Rihoux J P. Quantitative determination of skin reactivity by two semiautomatic devices for skin prick test area measurements. *Agents Actions* 1994; 41: C134–C135.