

Clinical and Laboratory Investigations

Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions

A. BLUM, H. LUEDTKE, U. ELLWANGER, R. SCHWABE,* G. RASSNER AND C. GARBE

Departments of Dermatology and *Medical Biometry, Eberhard-Karls-University, Liebermeisterstrasse 25, 72076 Tuebingen, Germany

Accepted for publication 4 April 2004

Summary

Background Digital image analysis has been introduced into the diagnosis of skin lesions based on dermoscopic pictures.

Objectives To develop a computer algorithm for the diagnosis of melanocytic lesions and to compare its diagnostic accuracy with the results of established dermoscopic classification rules.

Methods In the Department of Dermatology, University of Tuebingen, Germany, 837 melanocytic skin lesions were prospectively imaged by a dermoscopy video system in consecutive patients. Of these lesions, 269 were excised and examined by histopathology: 84 were classified as cutaneous melanomas and 185 as benign melanocytic naevi. The remaining 568 lesions were diagnosed by dermoscopy as benign. Digital image analysis was performed in all 837 benign and malignant melanocytic lesions using 64 different analytical parameters.

Results For lesions imaged completely (diameter \leq 12 mm), three analytical parameters were found to distinguish clearly between benign and malignant lesions, while in incompletely imaged lesions six parameters enabled differentiation. Based on the respective parameters and logistic regression analysis, a diagnostic computer algorithm for melanocytic lesions was developed. Its diagnostic accuracy was 82% for completely imaged and 84% for partially imaged lesions. All 837 melanocytic lesions were classified by established dermoscopic algorithms and the diagnostic accuracy was found to be in the same range (ABCD rule 78%, Menzies' score 83%, seven-point checklist 88%, and seven features for melanoma 81%).

Conclusions A diagnostic algorithm for digital image analysis of melanocytic lesions can achieve the same range of diagnostic accuracy as the application of dermoscopic classification rules by experts. The present diagnostic algorithm, however, still requires a medical expert who is qualified to recognize cutaneous lesions as being of melanocytic origin.

Key words: computer algorithm, digital image analysis, melanocytic naevi, melanoma

Dermoscopy (dermatoscopy, epiluminescence microscopy) has been established in the last two decades as a noninvasive method for improving the early detection of malignant melanoma and for reducing unnecessary excision of benign naevi.^{1–3} Dermoscopy utilizes recognition of surface structures, pigment patterns and other features such as vessels in order to distinguish between benign and malignant skin lesions. Compared with the

clinical diagnosis of skin lesions, dermoscopy gives an improvement in diagnostic sensitivity of 10–30%.⁴ However, due to the complexity of patterns and their interpretation, the results of dermoscopic examination have limitations especially for beginners and users not trained specifically.^{5,6} Therefore, there has been much scientific endeavour aimed at obtaining an improved and consistent differentiation between benign and malignant melanocytic skin lesions by means of digital dermoscopy analysis.^{7–26} Different datasets, methods of image analysis and methods of statistical evaluation of

Correspondence: Andreas Blum.
E-mail: a.blum@derma.de

Table 1. Summary of the study groups that have developed digital dermoscopy analysis for digitized pictures⁷⁻²⁶

First author	Year	Variables (n)	Source	Statistical method	Lesions	Training set	Melanoma	Sensitivity (%)	Specificity (%)
Cascinelli ¹²	1987	NS	Slides	NS	M + NM	20	NS	NS	NS
Cascinelli ¹¹	1992	8	Patients	NS	M + NM	169	43	96	60
Schindewolf ²³	1993	23	Slides	CART	M	353	215	94	88
Schindewolf ²²	1994	NS	Patients	CART	M	309	80	89	88
Green ¹⁵	1994	22	Patients	DA	M + NM	164	18	89	89
Ercal ¹⁴	1994	14	Slides	NN	M + NM	326	136	80	86.3
Menzies ¹⁹	1997	NS	Slides	LR	M + NM	170	75	93	67
Husemann ¹⁸	1997	NS	Patients	NN	NS	215	NS	> 95	> 95
Seidenari ²⁵	1998	22	Patients	DA	M	917	65	93	95
Binder ⁹	1998	16	Slides	NN	M	120	39	90	74
Seidenari ²⁴	1999	26	Patients	DA	M	383	18	100	92
Handels ¹⁶	1999	26	Patients	NN	M	44	19	97.7	100
Andreassi ⁷	1999	13	Patients	DA	M	147	57	88	81
Blum ¹⁰	1999	3 ^a , 6 ^b	Patients	FA + LR ^{a,b}	M + NM ^{a,b}	116 ^a , 51 ^b	10 ^a , 27 ^b	90 ^a , 70.4 ^b	81.1 ^a , 70.4 ^b
Stolz ²⁶	2000	NS	Patients	LR	M	466	125	86.4	92.7
Bauer ⁸	2000	38	Patients	NN	M + NM	315	42	92.9	97.8
Elbaum ¹³	2001	13	Patients	LC + ROC	M	246	63	100	85
Rubegni ²⁰	2002	10	Patients	NN	M	147	57	93	92.8

NS, not specified; CART, classification and regression trees; DA, discrimination analysis; NN, neural network; FA, factor analysis; LR, logistic regression; LC, linear classification; ROC, receiver operating characteristics; M, melanocytic skin lesions; NM, nonmelanocytic skin lesions. ^aSmall, completely imaged lesions; ^bLarge, partially imaged lesions.

different large series of images from mostly single-centre studies have been used (Table 1).

In this study, we have developed a new computer algorithm for the diagnosis of melanocytic lesions based on the evaluation of 64 different analytical parameters for digital image analysis. The analytical parameters were developed independently of the established dermoscopic classification rules.²⁷⁻³⁰ The number of parameters taken into account in developing the diagnostic algorithm depended on the number of malignant melanomas analysed.³¹ These results were compared with established dermoscopic classification rules applied to the same image material as the diagnostic computer algorithm.

Materials and methods

Image collection, storage system and image database

In the Pigmented Lesion Clinic of the Department of Dermatology, University of Tuebingen, Germany, 837 melanocytic skin lesions were imaged prospectively with two identical colour video cameras (TeachScreen Software GmbH, Bad Birnbach, Germany) between 11 November 1998 and 2 March 2000.

The colour video camera MediCam 400 with Y/C signal exit had a 1/4-inch charged-couple device shooting element with 470 000 pixels (picture elements). The focal area for the dermoscopic pictures was

defined from 3.5 cm diameter up to infinity. The focal area for the dermoscopic pictures could be positioned continuously by zoom from 3.2 mm to ≈ 1.0 cm, corresponding to a $\times 20-70$ magnification on a 17-inch monitor. Lesions ≤ 12 mm diameter could be imaged completely. The glass plate contacting the skin was always moistened with disinfectant spray (Kodan-spray[®], B.Braun Melsungen AG, Melsungen, Germany). Constant illumination was provided by two light-emitting diodes integrated in the head of the camera. The automatic focus of the dermoscopic picture was set for white colour adjustment, magnification and focus under consideration of manual input possibilities, always on the first picture. A frame grabber was used to digitize the video signal. For archiving, images were saved in JPEG image format with a resolution of 768×576 pixels.

Images from all parts of the body were taken, with the exception of mucous membrane areas. After obtaining informed written patient consent, 269 melanocytic skin lesions were excised under local anaesthesia and the diagnosis was established by histopathology (Table 2). Of the 269 melanocytic lesions, 180 (67%) were imaged completely and 89 (33%) were imaged partially, with normal skin visible at the margin (Fig. 1, Table 3). The patients with malignant melanoma comprised 39 women and 45 men. The median Breslow thickness of all melanomas was 0.78 mm (range 0.10-3.50); for completely

Table 2. Numbers of completely and partially imaged skin lesions by histological type

Histology	Completely imaged	Partially imaged	Total
Lentigo simplex	5	–	5
Lentigo melanocytic naevus	8	–	8
Melanocytic naevus of junctional type	16	–	16
Melanocytic naevus of compound type	23	8	31
Melanocytic naevus of dermal type	5	3	8
Congenital melanocytic naevus	–	3	3
Spitz naevus	3	–	3
Reed naevus	2	1	3
Blue naevus	3	–	3
Halo naevus	1	1	2
Naevus spilus	–	1	1
Recurrent naevus	6	–	6
Dysplastic melanocytic naevus of junctional type	46	5	51
Dysplastic melanocytic naevus of compound type	37	8	45
Melanoma <i>in situ</i>	8	1	9
Superficial spreading melanoma	13	45	58
Nodular melanoma	–	6	6
Acrolentiginous melanoma	1	3	4
Lentigo maligna	3	1	4
Lentigo maligna melanoma	–	3	3
Total	180	89	269

imaged melanomas it was 0.50 mm (range 0.20–1.50) and for partially imaged melanomas it was 0.85 mm (range 0.10–3.50).

All 568 (425 completely imaged and 143 partially imaged; Fig. 1) unexcised lesions were analysed

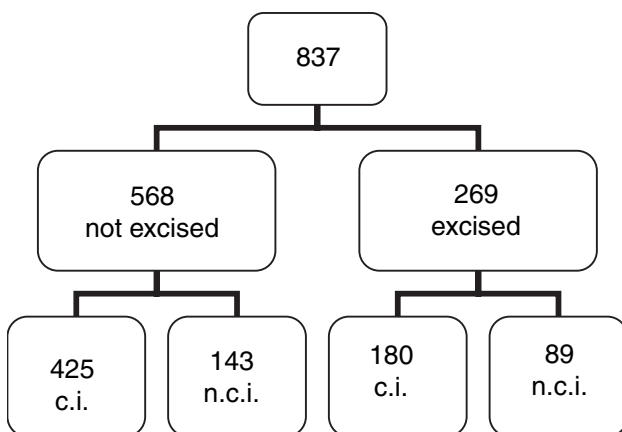


Figure 1. Of 837 recorded lesions, 568 were not excised and 269 were excised and examined histopathologically. Of the 568 unexcised lesions 425 were completely imaged (c.i.) and 143 incompletely imaged (n.c.i.). Of the 269 excised lesions 180 were completely imaged (c.i.) and 89 incompletely imaged (n.c.i.).

Table 3. Overview of all melanocytic skin lesions (n[overall]) as well as their division into two groups: n[partial], partially imaged lesions, and n[complete], completely imaged lesions. These groups were further divided into two equal random subgroups n_1 and n_2 for validation of the determined parameters from the factor analysis for the groups of small and large melanocytic skin lesions. The numbers of malignant lesions in each group are shown

		n	Benign	Malignant
All lesions	n[overall]	837	753	84
	n[partial]	232	173	59
	n[complete]	605	580	25
Subgroup n_1	n_1 [overall]	419	375	44
	n_1 [partial]	109	77	32
	n_1 [complete]	310	298	12
Subgroup n_2	n_2 [overall]	418	378	40
	n_2 [partial]	123	96	27
	n_2 [complete]	295	282	13

independently by two of the investigators (A.B., at least three times in 6 months; C.G., at least twice with a minimum follow-up period of 6 months) on the basis of dermoscopic criteria (pattern analysis, ABCD rule, Menzies' score, seven-point checklist and seven features for melanoma).^{27–30,32,33} These lesions were classified as benign without any suspicion of malignancy by dermoscopic criteria, and follow-up records for at least 6 months showed no evidence of malignancy.³⁴

Digital image analysis

Our philosophy for the development of the diagnostic algorithm was to establish the computer as a second independent diagnostic system not based on already established dermoscopic classification rules, i.e. we did not follow a preformed strategy (e.g. ABCD rule). Instead we applied a large number of algorithms of vision algebra to the pictures, as reported in the standard textbook on vision algebra by Ritter and Wilson.³⁵

Artefacts (hairs and air bubbles) which influenced the determination of margins were eliminated, and geometric parameters, colours and parameters of texture of the lesion were determined (Fig. 2a,b).³⁵ Afterwards, 64 analytical parameters were calculated for the entire set of melanocytic lesions. These included a large number of morphological parameters such as margin (Fig. 2c,d), geometric parameters (surface area, extent, largest diameter and largest orthogonal diameter), invariant moments, symmetry, colours (red, green, blue and grey value), texture (energy, entropy, correlation, inverse difference moment and inertia),

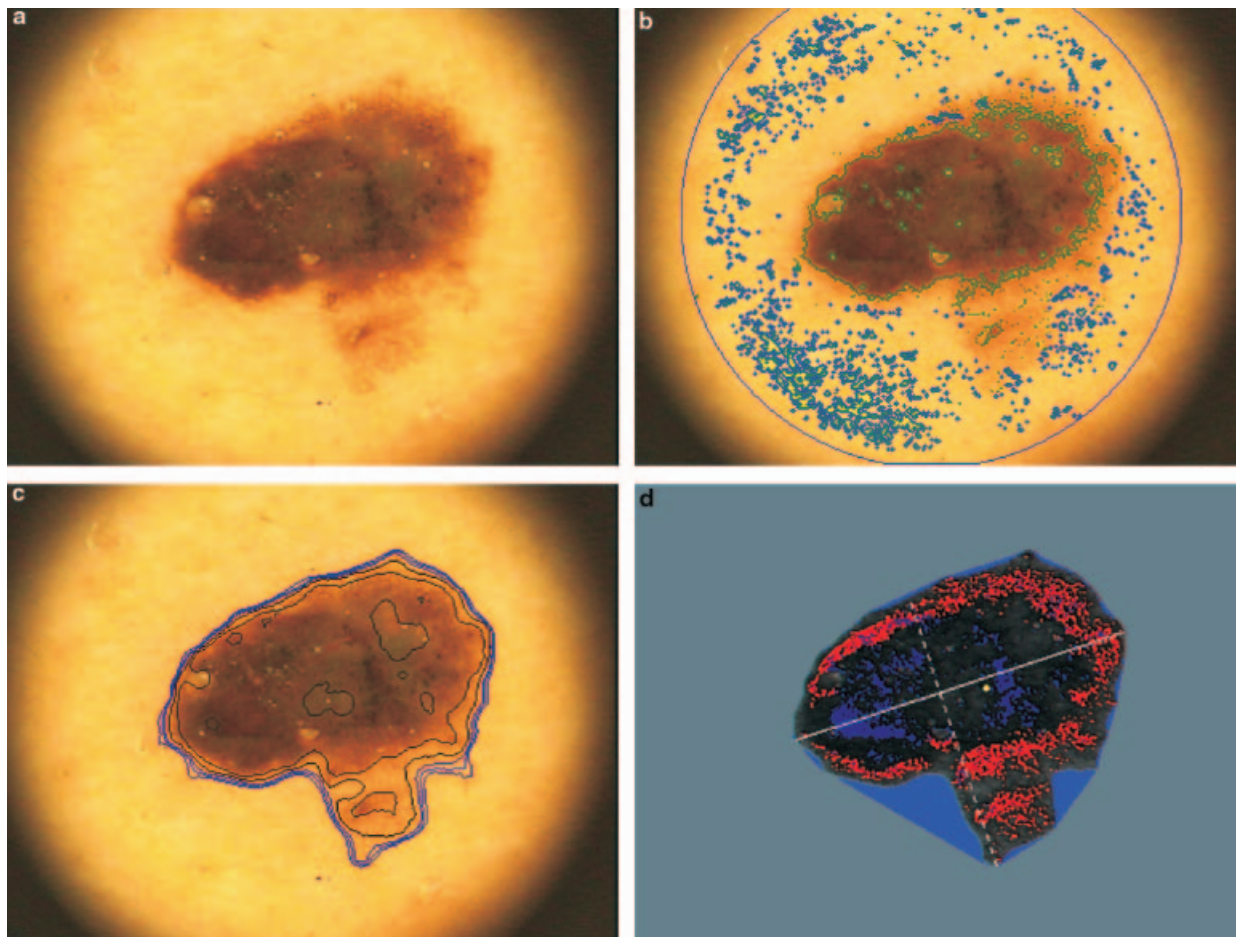


Figure 2. (a) Small, completely imaged melanocytic lesion with partially visible air bubbles (superficial spreading melanoma, tumour thickness 0.4 mm, level of invasion II). (b) Analysis of melanocytic lesion for air bubbles and hair, and their elimination. (c) Determination of margin, textures and number of regions of the melanocytic lesion. (d) Determination of circumference, area, greatest diameter and greatest orthogonal diameter.

number of regions, focus and difference of the lesion and its convex cover.³⁵

Dermoscopic diagnosis according to established classification rules

According to the established dermoscopic classification rules (ABCD rule, Menzies' score, seven-point checklist and seven features for melanoma) the lesions were prospectively classified as benign or malignant melanocytic lesions by the principal investigator (A.B.).^{27–30}

Statistical analysis

The median, 2.5 and 97.5 quartile and *P*-value for the Mann–Whitney *U*-test were investigated for all analyt-

ical parameters of the digital dermoscopy analysis. For this purpose, benign melanocytic naevi and malignant melanomas were separated as endpoints of the analysis. Afterwards, the analytical parameters of the digital dermoscopy analysis were reduced by means of factor analysis. In a second step, the impact of the different parameters was examined by logistic regression analysis. The number of parameters included in the multivariate analysis was limited in relation to the number of malignant melanomas: in the sample of large, partially imaged lesions it was restricted to six parameters and for small, completely imaged lesions it was limited to three parameters (Table 4).³¹ The emphasis (*B*), SEM, *P*-value and constants are indicated for the respective parameters. Finally, for both groups, sensitivity, specificity and diagnostic accuracy were examined:

Table 4. Results of the logistic regression for the parameters investigated, showing the emphasis (*B*), SEM, *P*-value and constant for all melanocytic lesions imaged and for the groups of partially and completely imaged lesions³⁵

Lesions	Analytical parameter	<i>B</i>	SEM	<i>P</i> -value
All	Entropy	- 1.285	0.406	< 0.005
	Greatest orthogonal diameter	0.031	0.003	< 0.0001
	Standard deviation black-white	0.183	0.025	< 0.0001
	Correlation	- 3.019	1.553	0.0519
	Mean axis symmetry	- 13.85	2.435	< 0.0001
	Red distance	3.419	1.273	0.0076
	Constant	11.60	3.466	< 0.001
Partial	Inverse difference moment	22.00	12.03	0.0675
	Correlation	3.073	1.434	0.0032
	Difference black-white	0.075	0.013	< 0.0001
	Margin	0.011	0.002	< 0.0001
	Mean axis symmetry	- 11.61	3.836	< 0.005
	Red distance	1.242	1.633	0.4470
	Constant	- 10.83	3.704	0.0086
Complete	Standard deviation black-white	0.156	0.030	< 0.0001
	Entropy	- 2.062	0.617	< 0.001
	Margin	0.006	0.001	< 0.0001
	Constant	5.139	4.432	0.2463

sensitivity = no. of correctly classified malignant melanocytic skin lesions
 \div no. of all malignant melanocytic skin lesions

specificity = no. of correctly classified benign melanocytic skin lesions
 \div no. of all benign melanocytic skin lesions

diagnostic accuracy = no. of correctly classified malignant and benign melanocytic skin lesions
 \div no. of all malignant and benign melanocytic skin lesions

The respective receiver operating characteristics curves and the area under the curves were likewise calculated.

For validation of the described procedure the complete collection was divided into two equal random subgroups n_1 and n_2 (Table 5). While the parameters determined by the factor analysis were kept fixed the threshold values for the small and large melanocytic skin lesions were determined by fitting a new logistic regression to the data of subgroup n_1 (training set). The quality of the algorithm derived was assessed by calculating the corresponding sensitivity, specificity and diagnostic accuracy in the subgroup n_2 (test set).

For the results of dermoscopy, the frequency of criteria was calculated and examined for statistically significant differences using Fisher's exact test for 2×2 tables and Pearson's χ^2 test with α level 0.05 (two-tailed). $P < 0.05$ was considered as statistically significant. Also, if possible, odds ratios and their 95% confidence intervals were calculated as criteria for malignancy. For the classification of benign and malignant lesions, sensitivity, specificity and diagnostic accuracy were calculated. The data were evaluated with SPSS 9.0 for Windows (SPSS, Chicago, IL, U.S.A.).

Results

Development of a diagnostic computer algorithm for digital dermoscopy analysis

For development of the diagnostic computer algorithm, 837 melanocytic lesions were included (Table 3), of which 605 (72%) were imaged completely and 232 (28%) partially (Fig. 1). By use of a factor analysis from the 64 analytical parameters generated by vision algorithms, six parameters were identified that were important for all lesions (753 benign and 84 malignant melanocytic lesions).

Next, the collection was divided into two groups: (i) n [partial], incompletely imaged skin lesions > 12 mm in diameter, and (ii) n [complete], completely imaged skin lesions (Table 3). It was not possible to use the same number of analytical parameters for both groups, as they contained different numbers of malignant lesions. Thus, from the 64 analytical parameters evaluated in two factor analyses six parameters were

Group		<i>n</i>	Melanoma	Sensitivity (%)	Specificity (%)	Diagnostic accuracy (%)
All lesions	Small	605	25	80.0	82.4	82.3
	Large	232	59	88.1	82.7	84.1
Subgroup <i>n</i> ₁	Small	310	12	83.3	86.9	86.8
	Large	109	32	90.6	85.7	87.2
Subgroup <i>n</i> ₂	Small	295	13	53.9	87.2	85.8
	Large	123	27	92.6	76.0	79.7

Table 5. Overview of the groups of all melanocytic skin lesions (*n* = 837) as well as the subgroups *n*₁ and *n*₂ showing sensitivity, specificity and diagnostic accuracy

taken into account for the group *n*[partial] (with 59 malignant lesions) and three parameters were taken into account for the group *n*[complete] (with 25 malignant lesions).

The results of the logistic regression for the six and three analytical parameters, respectively, are given in Table 4. For each group, the emphasis (*B*) of the respective parameter and the constant for the determination of the analytical algorithm were determined.

The diagnostic computer algorithm thus comprises two different scores for completely and incompletely imaged lesions. The calculation of the scores was performed with Equations 1 and 2 based on the results of the multifactorial logistic regression. The calculation of the respective parameters is described in detail in Ritter and Wilson.³⁵

Score for the partially imaged lesions

$$\begin{aligned}
 &= 1.0 / (1.0 + \exp(-(-10.83 + 22.0 \\
 &\quad \times \text{inverse difference moment} + 3.073 \\
 &\quad \times \text{correlation} + 0.075 \times \text{difference}_{[\text{black}-\text{white}]} \\
 &\quad + 0.011 \times \text{margin} - 11.61 \times \text{mean}_{[\text{axis symmetry}]} \\
 &\quad + 1.242 \times \text{difference}_{[\text{red}]})) \quad (\text{Eqn1})
 \end{aligned}$$

Score for the completely imaged lesions

$$\begin{aligned}
 &= 1.0 / (1.0 + \exp(-(5.139 - 2.062 \\
 &\quad \times \text{entropy} + 0.156 \\
 &\quad \times \text{standard deviation}_{[\text{black}-\text{white}]} \\
 &\quad + 0.006 \times \text{margin}))) \quad (\text{Eqn2})
 \end{aligned}$$

Evaluation of the diagnostic sensitivity, specificity and accuracy

Testing this analytical algorithm for all melanocytic skin lesions, 648 of the 753 benign lesions were classified correctly, as were 74 of the 84 malignant melanomas. Sensitivity was found to be 88.1%, specificity 86.1% and diagnostic accuracy 86.3% (Table 6,

Fig. 3a). The diagnostic computer algorithm for the large, incompletely imaged melanocytic skin lesions classified 143 of 173 benign lesions and 52 of 59 malignant lesions correctly. Sensitivity was 88.1%, specificity 82.7% and diagnostic accuracy 84.1% (Table 6, Fig. 3b). The diagnostic computer algorithm for the small, completely imaged melanocytic skin lesions correctly classified 478 of 580 benign lesions and 20 of 25 malignant lesions. Sensitivity was 80%, specificity 82.4% and diagnostic accuracy 82.3% (Table 6, Fig. 3c).

*Validation of the diagnostic computer algorithm using subgroups *n*₁ and *n*₂ of the data*

For validation of the formula generated by the logistic regression for groups of small and large melanocytic skin lesions, the results for subgroup *n*₁ (training set) were compared with the results for subgroup *n*₂ (test set) and with the results for the total collection. For the large, incompletely imaged skin lesions, the logistic regression was performed again with subgroup *n*₁ and the six analytical parameters investigated by factor analysis (Eqn 1). Similarly, this analysis was performed with the small, completely imaged melanocytic skin lesions and the three parameters determined (Eqn 2).

Next, the algorithm determined using subgroup *n*₁ was applied to subgroup *n*₂. Sensitivity, specificity and diagnostic accuracy were calculated for the group of all melanocytic skin lesions as well as for the subgroups *n*₁ and *n*₂ (Table 5). In subgroup *n*₁, of the large skin lesions, 66 of 77 benign and 29 of 32 malignant lesions were classified correctly, and of the small skin lesions, 259 of 298 benign and 10 of 12 malignant lesions were classified correctly. The diagnostic algorithm generated the following results in subgroup *n*₂: of the large skin lesions, 73 of 96 benign and 25 of 27 malignant lesions were classified correctly, and of the small skin lesions, 246 of 282 benign and seven of 13 malignant lesions were classified correctly.

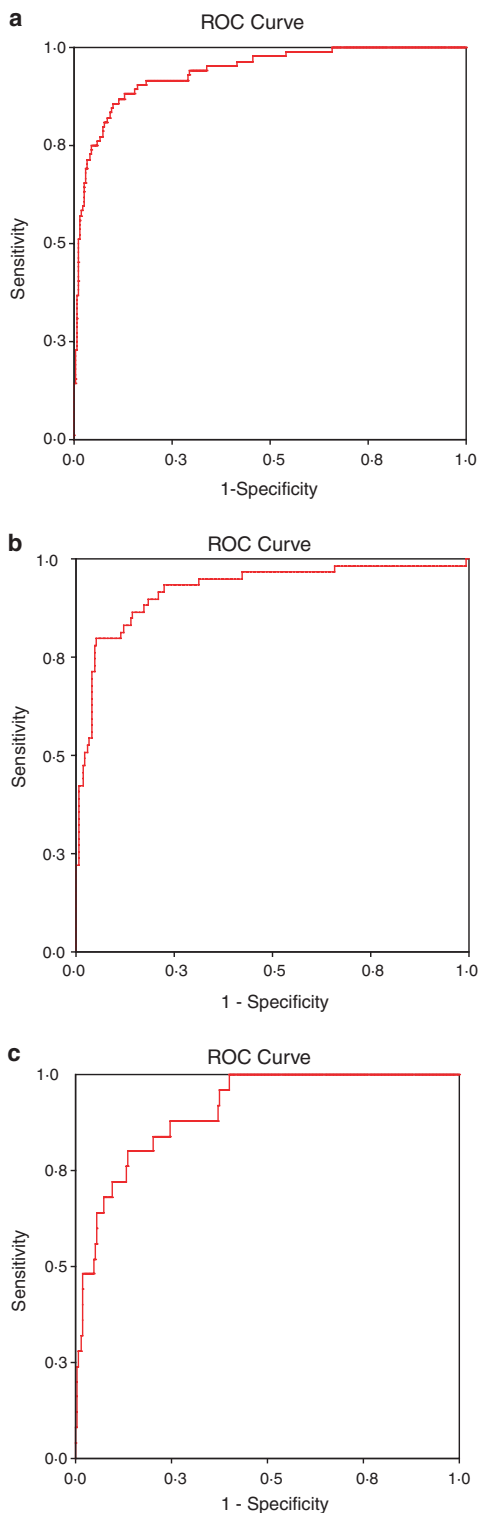


Figure 3. (a) Receiver operating characteristics (ROC) curve for all melanocytic skin lesions (area under the curve, $AUC = 0.938$). (b) ROC curve for the large, partially imaged melanocytic skin lesions ($AUC = 0.923$). (c) ROC curve for the small, completely imaged melanocytic skin lesions ($AUC = 0.906$).

Classification of benign and malignant melanocytic lesions by established dermoscopic scores

The results of the ABCD rule, Menzies' score, seven-point checklist, seven features for melanoma and the digital dermoscopy analysis are given in Table 6.^{27–30} Of the benign lesions, 72.4% were classified as not suspicious, 21.6% as suspect and 5.9% as malignant according to the ABCD rule.³⁰ Corresponding figures for the malignant lesions were 9.5%, 9.5% and 81.0%, respectively. By Menzies' score 77.8% of the benign lesions were classified as benign and 22.2% as malignant.²⁹ Of the malignant skin lesions, 4.8% were classified as benign and 95.2% as malignant. By the seven-point checklist 87.0% of the benign lesions were classified as benign and 13.0% as malignant,²⁷ while 9.5% of the malignant lesions were classified as benign and 90.5% as malignant. Using the seven features for melanoma algorithm 74.6% of the benign lesions were classified as benign and 25.4% as malignant,²⁸ while 6.0% of the malignant lesions were classified as benign and 94.0% as malignant.

Discussion

We report the development of a new diagnostic algorithm for digital dermoscopy analysis based on factor analysis and logistic regression analysis. The sensitivity for small completely imaged lesions was 10–14% lower than that of established dermoscopic classification rules, but the specificity was 7–10% higher than in three of four algorithms compared. Therefore, no significant difference in diagnostic accuracy was found between the established classification rules of dermoscopy and the computer algorithm. A remarkable feature was that the sensitivity in small lesions was slightly lower than that for large lesions. In small malignant lesions, typical features (size, margin, colour and differential structures) are more likely to be missing or indistinct than in larger melanomas, which may explain these results. In dermoscopic examination of small lesions the differentiation between benign and malignant melanocytic lesions can be quite difficult.^{4,36}

In previous reports, the sensitivity of computer algorithms for digital dermoscopy analysis has ranged between 80% and 100% (Table 1).^{7–26} The images analysed by the different groups included melanocytic and nonmelanocytic lesions. Therefore the dataset of images differed from that used in the present study, making comparisons difficult. In particular, seborrhoeic

Table 6. Results given by the digital dermoscopy analysis (completely and partially imaged skin lesions) and the algorithms and scores examined^{27–30}

Method	Sensitivity (%)	Specificity (%)	Diagnostic accuracy (%)
Digital dermoscopy analysis (all lesions)	88.1	86.1	86.3
Digital dermoscopy analysis (completely imaged lesions)	80.0	82.4	82.3
Digital dermoscopy analysis (partially imaged lesions)	88.1	82.7	84.1
ABCD rule	90.5	72.4	78.1
Menzies' score	95.2	77.8	83.3
Seven-point checklist	90.5	87.0	88.1
Seven features for melanoma	94.0	74.6	80.7

keratoses could affect the results of digital dermoscopy analysis. As nonmelanocytic pigmented lesions are mostly simple to distinguish from melanoma by use of dermoscopy we concentrated on the differential diagnosis of exclusively melanocytic lesions. This reflects the main practical problem with which the physician is confronted in melanoma diagnosis. As only melanocytic lesions were included for the development of the diagnostic algorithm in the present study, it is a precondition for the use of this algorithm that melanocytic and nonmelanocytic lesions must first be differentiated.

In the development of the present diagnostic computer algorithm we did not follow any already established dermoscopic rules for melanoma diagnosis and we did not try to translate such rules into vision algorithms of image algebra. Our approach was to use already existing vision algorithms of image algebra as described by Ritter and Wilson.³⁵ Thus, 64 analytical parameters were selected for this study. We considered that the multivariate analysis might identify important analytical parameters that had not yet been included in the established dermoscopic classification rules for melanoma diagnosis, and that the diagnostic approach by use of digital image analysis might be different from the human approach using dermoscopic rules. Interestingly, the results of the factor analysis followed by the logistic regression analysis identified geometric structures, lesion margin, differences in colour and differential structures as the main diagnostic parameters. These criteria are similar to those used in the dermoscopic ABCD rule.³⁰ Less than 10% of the parameters examined were included in the final diagnostic algorithm. Depending on the number of melanomas in the respective samples, three analytical parameters were used for small, completely imaged

melanomas and six parameters for large, partially imaged melanomas.³¹ In large melanomas the features of the ABCD rule were nearly all represented, but not in the small melanomas (Table 4). Small melanomas could still be symmetrical and, probably for this reason, the geometric area was not included. Also, the typical dermoscopic features of melanoma are not so clearly visible in small melanomas, which may explain the difference in sensitivity for small vs. large melanomas (80.0% vs. 88.1%). On the other hand, no difference was seen in the diagnostic specificity between these groups (82.4% vs. 82.7%).

Any comparison of the results reported here with the studies published in the literature is difficult. Between 1987 and 2002 various groups developed diagnostic systems for recorded images (using slides or digital cameras) (Table 1). The datasets in the literature included both melanocytic and nonmelanocytic lesions. The melanomas examined had varying median Breslow tumour thicknesses. Furthermore, varying recording systems were used. No study has yet been performed using a standardized calibration.^{37,38} The exact algorithms utilized for the diagnostic algorithm have not yet been disclosed in the literature. Also, the respective sets of images have never been analysed by different approaches. Therefore, this detailed report of our study with disclosure of our algorithm may represent a basis for future comparison of different diagnostic algorithms validated on the basis of a common set of images of melanocytic lesions.

Nevertheless, based on the results of this study and the literature so far, digital dermoscopy analysis of recorded images is at least equal to the clinical and dermoscopic diagnostic sensitivity and specificity obtained by a trained expert in dermoscopy.⁶ Advantages of a digital dermoscopy diagnostic system are that this instrument works independently of time and is not affected by different levels of attention as are humans. It might also be a useful tool particularly in centres without expertise in dermoscopy. On the other hand, it is not likely that the digital system will completely replace the expert in dermoscopy because trained users will recognize certain significant details in melanocytic lesions which lead to the diagnosis of a disease with malignant potential.^{32,33,39,40} Details just visible in one small area of the entire lesion may not have the impact to change the classification of the lesion from benign to malignant by the computer algorithm.

Computer diagnostic algorithms could also be used in the follow-up of patients with atypical moles.^{34,36} The comparison of images recorded at different times is

helpful in these patients. In addition, the results using computer diagnostic algorithms could be useful for decisions in clinical management of patients with atypical mole syndrome.⁶

In conclusion, a diagnostic algorithm for digital image analysis of melanocytic lesions was developed which achieves the same range of diagnostic accuracy as the application of dermoscopy with established dermoscopic classification rules interpreted by trained experts. Taking the great morphological variability of melanocytic naevi and malignant melanoma into account, this represents a success for digital image analysis. Thus, the computer could be of substantial help for the physician in establishing the final diagnosis. Similar diagnostic algorithms could probably be developed for the interpretation of endoscopic pictures or X-rays and may likewise be helpful for the diagnosis of other types of cancer. However, digital image analysis will probably not replace the physician. The diagnostic algorithm developed by us still requires the physician to distinguish between melanocytic and nonmelanocytic lesions. The role of digital image analysis will be supplementary in order to gain a higher level of certainty for the final establishment of the diagnosis of malignant melanoma.

References

- Argenziano G, Soyer HP. Dermoscopy of pigmented skin lesions—a valuable tool for early diagnosis of melanoma. *Lancet Oncol* 2001; **2**: 443–9.
- Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. *Lancet Oncol* 2002; **3**: 159–65.
- Argenziano G, Soyer HP, Chimenti S *et al*. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. *J Am Acad Dermatol* 2003; **48**: 679–93.
- Mayer J. Systematic review of the diagnostic accuracy of dermoscopy in detecting malignant melanoma. *Med J Aust* 1997; **167**: 206–10.
- Binder M, Kittler H, Seeber A *et al*. Epiluminescence microscopy-based classification of pigmented skin lesions using computerized image analysis and an artificial neural network. *Melanoma Res* 1998; **8**: 261–6.
- Piccolo D, Ferrari A, Peris K *et al*. Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study. *Br J Dermatol* 2002; **147**: 481–6.
- Andreassi L, Perotti R, Rubegni P *et al*. Digital dermoscopy analysis for the differentiation of atypical nevi and early melanoma. *Arch Dermatol* 1999; **135**: 1459–65.
- Bauer P, Cristofolini P, Boi S *et al*. Digital epiluminescence microscopy: usefulness in the differential diagnosis of cutaneous pigmentary lesions. A statistical comparison between visual and computer inspection. *Melanoma Res* 2000; **10**: 345–9.
- Binder M, Schwarz M, Winkler A *et al*. Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Arch Dermatol* 1995; **131**: 286–91.
- Blum A, Ellwanger U, Lüdtke H, Garbe C. Digital image analysis of pigmented lesions: the Tübingen Mole Analyser. *Skin Res Technol* 1999; **5**: 127.
- Cascinelli N, Ferrario M, Bufalino R *et al*. Results obtained by using a computerized image analysis system designed as an aid to diagnosis of cutaneous melanoma. *Melanoma Res* 1992; **2**: 163–70.
- Cascinelli N, Ferrario M, Tonelli T, Leo E. A possible new tool for clinical diagnosis of melanoma: the computer. *J Am Acad Dermatol* 1987; **16**: 361–7.
- Elbaum M, Kopf AW, Rabinovitz HS *et al*. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study. *J Am Acad Dermatol* 2001; **44**: 207–18.
- Ercal F, Chawla A, Stoecker WV *et al*. Neural network diagnosis of malignant melanoma from color images. *IEEE Trans Biomed Imag* 1994; **41**: 837–45.
- Green A, Martin N, Pflitzner J *et al*. Computer image analysis in the diagnosis of melanoma. *J Am Acad Dermatol* 1994; **31**: 958–64.
- Handels H, Ross T, Kreusch J *et al*. Computer-supported diagnosis of melanoma in profilometry. *Methods Inf Med* 1999; **38**: 43–9.
- Horsch A, Stolz W, Neiss A *et al*. Improving early recognition of malignant melanomas by digital image analysis in dermoscopy. *Stud Health Technol Inform* 1997; **43**: 531–5.
- Husemann R, Tölg S, Seelen W *et al*. Computerised diagnosis of skin cancer using neural networks. In: *Skin Cancer and UV-Radiation* (Altmeyer P, Hoffmann K, Stücker M, eds). Berlin: Springer-Verlag, 1997; 1052–63.
- Menzies SW, Bischof LM, Peden G *et al*. Automated instrumentation for the diagnosis of invasive melanoma: image analysis of oil epiluminescence microscopy. In: *Skin Cancer and UV-Radiation* (Altmeyer P, Hoffmann K, Stücker M, eds). Berlin: Springer-Verlag, 1997; 1064–70.
- Rubegni P, Burrioni M, Cevenini G *et al*. Digital dermoscopy analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: a retrospective study. *J Invest Dermatol* 2002; **119**: 471–4.
- Rubegni P, Ferrari A, Cevenini G *et al*. Differentiation between pigmented Spitz naevus and melanoma by digital dermoscopy and stepwise logistic discriminant analysis. *Melanoma Res* 2001; **11**: 37–44.
- Schindewolf T, Schiffner R, Stolz W *et al*. Evaluation of different image acquisition techniques for a computer vision system in the diagnosis of malignant melanoma. *J Am Acad Dermatol* 1994; **31**: 33–41.
- Schindewolf T, Stolz W, Albert R *et al*. Classification of melanocytic lesions with color and texture analysis using digital image processing. *Anal Quant Cytol Histol* 1993; **15**: 1–11.
- Seidenari S, Pellacani G, Giannetti A. Digital videomicroscopy and image analysis with automatic classification for detection of thin melanomas. *Melanoma Res* 1993; **9**: 163–71.
- Seidenari S, Pellacani G, Pepe P. Digital videomicroscopy improves diagnostic accuracy for melanoma. *J Am Acad Dermatol* 1998; **39**: 175–81.
- Stolz W, Abmayr W, Pompl R *et al*. Computergestützte dermatoskopische Diagnose. In: *Dermatologie an der Schwelle zum neuen Jahrtausend. Aktuelle Stand von Klinik und Forschung* (Plettenberg A, Meigel WN, eds). Berlin: Springer-Verlag, 2000; 34–7.
- Argenziano G, Fabbrocini G, Carli P *et al*. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin

- lesions—comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Arch Dermatol* 1998; **134**: 1563–70.
- 28 Dal Pozzo V, Benelli C, Roscetti E. The seven features for melanoma: a new dermoscopic algorithm for the diagnosis of malignant melanoma. *Eur J Dermatol* 1999; **9**: 303–8.
 - 29 Menzies SW, Ingvar C, Crotty KA, McCarthy WH. Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. *Arch Dermatol* 1996; **132**: 1178–82.
 - 30 Stolz W, Riemann A, Cagnetta AB *et al.* ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. *Eur J Dermatol* 1994; **4**: 521–7.
 - 31 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; **15**: 361–87.
 - 32 Pehamberger H, Steiner A, Wolff K. *In vivo* epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions. *J Am Acad Dermatol* 1987; **17**: 571–83.
 - 33 Steiner A, Pehamberger H, Wolff K. *In vivo* epiluminescence microscopy of pigmented skin lesions. II. Diagnosis of small pigmented skin lesions and early detection of malignant melanoma. *J Am Acad Dermatol* 1987; **17**: 584–91.
 - 34 Menzies SW, Gutenev A, Avramidis M *et al.* Short-term digital surface microscopic monitoring of atypical or changing melanocytic lesions. *Arch Dermatol* 2001; **137**: 1583–9.
 - 35 Ritter GX, Wilson JN. *Handbook of Computer Vision Algorithms in Image Algebra*. Boca Raton: CRC Press, 1996.
 - 36 Kittler H, Binder M. Follow-up of melanocytic skin lesions with digital dermatoscopy: risks and benefits. *Arch Dermatol* 2002; **138**: 1379.
 - 37 Menzies SW. Automated epiluminescence microscopy: human vs machine in the diagnosis of melanoma. *Arch Dermatol* 1999; **135**: 1538–40.
 - 38 Rosado B, Menzies S, Harbauer A *et al.* Accuracy of computer diagnosis of melanoma: a quantitative meta-analysis. *Arch Dermatol* 2003; **139**: 361–7.
 - 39 Hofmann-Wellenhof R, Blum A, Wolff IH *et al.* Dermoscopic classification of atypical melanocytic nevi. *Arch Dermatol* 2001; **137**: 1575–80.
 - 40 Blum A, Soyer HP, Garbe C *et al.* The dermoscopic classification of atypical melanocytic naevi (Clark naevi) is useful to discriminate benign from malignant melanocytic lesions. *Br J Dermatol* 2003; **149**: 1159–64.