

Value of the clinical history for different users of dermoscopy compared with results of digital image analysis

A Blum,†* R Hofmann-Wellenhof,‡ H Luedtke,† U Ellwanger,† A Steins,† S Roehm,† C Garbe,† HP Soyert

Departments of Dermatology, †University of Tuebingen, Liebermeisterstrasse 25, 72076 Tuebingen, Germany and ‡University of Graz, Graz, Austria.

*Corresponding author, tel. +49 7071 2984555; fax +49 7071 294561; E-mail: a.blum@derma.de

ABSTRACT

Background The clinical history of a given pigmented lesion could influence the therapeutic decision. Teledermatology and automated image analysis also hold great potential for revolutionizing dermatology services.

Aim The aim of this retrospective study was to evaluate the diagnostic accuracy of users with different experiences in dermoscopy with and without information about patients and their history compared with classification by an automated analysing system.

Setting One hundred and fifty-seven dermoscopic images of pigmented lesions, taken and proved by histopathology at the Pigmented Lesions Clinic of the Department of Dermatology of the University Tuebingen, Germany, were included.

Methods All images were viewed by three investigators with different experience: excellent (A), average (B) and beginner (C). In the first dermoscopic classification, no information was available. After 3 months the same images were once more classified by the three investigators, now with the information about the patients and their history. The melanocytic lesions were tested by the Tuebinger Mole Analyser.

Results For user A the sensitivity, specificity and diagnostic accuracy revealed no improvement on including the history (81.3% to 84.4%, 94.6% to 92.3% and 92.0% to 90.7%), whereas user B clearly improved his results (75.0% to 87.5%, 76.9% to 88.5% and 76.5% to 88.3%). No change in the sensitivity was seen by user C (84.4%), but there was a clear improvement in the specificity (69.2% to 87.7%) and diagnostic accuracy (72.2% to 87.0%). Using the computer algorithm, a sensitivity of 100%, a specificity of 76.9% and a diagnostic accuracy of 81.9% were achieved.

Conclusions The study revealed results relevant to the use of dermoscopy: (1) continuing dermoscopic education influences the diagnostic accuracy; (2) the history is helpful for averaged users and beginners in dermoscopy; (3) digital image analysis has the highest sensitivity, but a lower specificity compared to the clinicians; and (4) digital dermoscopy could be used for store-and-forward systems in teledermatology.

Key words: automated image analysis, dermatoscopy, dermoscopy, digital image, melanoma

Received: 27 March 2003; accepted 2 December 2003

Introduction

Dermoscopy of skin lesions increases the diagnostic accuracy by 10–27%.^{1,2} However, dermatologists who had not been trained in this method received no benefit by this diagnostic method.³ The clinical history of a given pigmented lesion influences further the therapeutic decision. In a multivariate model, morphological changes reported by the patients were a significant independent predictor of malignancy.⁴

Digital dermoscopy and automated image analysis could represent possible support for inexperienced clinicians.^{5–7} Digital dermoscopy also enables teledermatological consultation, which was found to provide the same diagnostic accuracy as face-to-face diagnosis.^{8,9} Both teledermatology and automated image analysis hold great potential for revolutionizing the delivery of dermatology services.^{5,10} However, further evaluation of the prerequisites and the reliability of these methods is recommended before their routine application.

In the present study we used a set of digital melanocytic images to determine whether the diagnostic accuracy of an inexperienced, an average experienced and a highly experienced investigator can be significantly improved by incorporating information about morphological changes, location of the lesion, and the age and sex of the patients. An additional objective of the study was to compare the diagnostic accuracy of the different investigators to the classification carried out by a commercially available automated analysing system.¹¹

Method and material

One hundred and sixty-two digital dermoscopic images of histologically proven pigmented skin lesions were included in the study. All lesions were excised due to suspicious clinical and/or dermoscopic features. The lesions were consecutively collected at the Pigmented Lesion Clinic of the Department of Dermatology of the University of Tuebingen, Germany, from September 1998 to March 1999.

The dermoscopic images were documented by a computer system with a colour video camera used at a magnification of $\times 20$ (FotoFinder, TechScreen Software GmbH, Bad Birnbach, Germany). The digital dermatoscopic microscope system had a colour video camera with one charge-coupled device (CCD) chip 1/4-inch with 47 000 pixels. The illumination was realized with two light emitting diodes (LEDs) that were integrated at the camera. The maximum field was 12 mm for dermoscopic images. The computer had a Pentium III processor (500 MHz, 64 MB ram). The graphic card had a true colour mode (32 bit) with a resolution of 1.024×768 pixels. The 17-inch colour monitor (Triniton, Multiscan 200ES, Sony, Japan) had a fixed resolution of 768×576 pixels and 65 536 colours with a frequency of 75 Hz. A frame-grabber was used for digitizing the video signal. The system of Joint Photographic Expert Group (JPEG) was used for storing the images with an image size of 768×576 pixels. The storage size ranged between 125 and 230 kB per image.

Consecutive images of one lesion were not included. All patients gave their written consent for the digital documentation and the following operation under local anaesthesia. The final diagnosis was performed by histopathology.

All images were viewed by three investigators (R.H.-W., A.S. and H.P.S.) with different experiences in dermoscopy: excellent (A), average (B) and beginner (C). Pattern analysis was used by the investigators. In the first dermoscopic classification, no information about the clinical history, age, sex of the patients and location of the tumour was given (t_1). After 3 months the same digital images were once more classified by the three investigators, now with the information about the clinical history, age, sex of the patients and location of the tumour (t_2). The clinical history was scored as positive when any morphological change was recognized by the patient in the past 3 months. Morphological changes included change in size, colour or shape

or any sign of ulceration or spontaneous bleeding. Possible dermoscopic classifications were benign nevi, atypical nevi, cutaneous melanoma and other benign epithelial tumours (e.g. seborrhoeic keratosis, angioma).² Malignant epithelial tumours (basal cell carcinoma, squamous cell carcinoma) were excluded.

Digital analysis was performed on the dermoscopic images of the melanocytic lesions. With the established Tuebinger Mole Analyser six variables (symmetry, border, different types of colours and entropy) for lesions larger than 12 mm and three variables (border, colour and entropy) for lesions smaller than 12 mm were used.¹¹ If the lesions exceeded the maximum field, one sector with border and normal skin was recorded. Relative frequencies of the different features were analysed. Percentages for sensitivity, specificity and diagnostic accuracy were calculated. Differences between changed and unchanged pigmented lesions were tested by Fisher's exact test and Pearson's χ^2 -test (level 0.05, two-tailed). A *P*-value less than 0.05 was regarded as statistically significant. Statistical examinations were performed with SPSS 10.0 (SPSS Inc., Chicago, IL, USA) for Windows.

Results

One hundred and fifty-seven pigmented lesions in 86 (54.8%) females and 71 (45.2%) males were analysed. No change in the past 3 months was reported by 87 (55.4%) patients, followed by an observed change in 39 (24.8%) patients and no clear clinical history was given by 31 (19.7%) patients. The subgroup of 126 lesions with a clinical history of any or no change was composed of 16 cutaneous melanomas and 111 benign lesions. In 87.5% of these cutaneous melanomas (14/16) and in 22.5% of the benign lesions (25/111) a change was reported ($P < 0.001$).

The median age of the patients was 38.9 years (standard deviation 16.8 years; range 2–87 years). Out of 157 pigmented skin lesions, 145 (79.6%) were benign and 32 (20.4%) malignant. One hundred and two (65.0%) of the lesions were located on the trunk, followed by the extremities (38; 24.2%), face (9; 5.7%), akral (6; 3.8%) and mucosal sites (2; 1.2%). The histological diagnoses were 59 (37.6%) cases of dysplastic nevus, 53 (33.8%) nevi without any dysplasia, 32 (20.4%) cutaneous melanomas and 13 (8.3%) epithelial benign tumours. Of the 32 cutaneous melanomas, two were melanomas *in situ* and 29 were invasive melanomas with the median tumour thickness according to Breslow of 0.86 mm (standard deviation 0.54 mm; range 0.30–2.40 mm).

The pigmented lesions were classified into the four dermoscopic categories (benign, atypical nevi, cutaneous melanoma and benign epithelial tumours) specified by the three investigators in Tables 1–4. For user A, a distinct improvement concerning the history in the diagnosis of the cutaneous melanoma was achieved, but in benign nevi without any dysplasia, atypical nevi and epithelial benign tumours a distinct decrease in the correct diagnosis was observed. In the group of benign nevi without any dysplasia the correct classification was improved

Table 1 Dermoscopic classification of the benign nevi ($n = 53$) by the three investigators (A = excellent; B = average; C = beginner) at the two examinations (t_1 and t_2)

Classification	A (t_1)	A (t_2)	B (t_1)	B (t_2)	C (t_1)	C (t_2)
Correct	50	49	41	46	35	47
Not correct	3	4	12	7	18	6

Table 2 Dermoscopic classification of the atypical nevi ($n = 59$) by the three investigators (A = excellent; B = average; C = beginner) at the two examinations (t_1 and t_2)

Classification	A (t_1)	A (t_2)	B (t_1)	B (t_2)	C (t_1)	C (t_2)
Correct	56	54	44	53	42	50
Not correct	3	5	15	6	17	9

Table 3 Dermoscopic classification of the cutaneous melanomas ($n = 32$) by the three investigators (A = excellent; B = average; C = beginner) at the two examinations (t_1 and t_2)

Classification	A (t_1)	A (t_2)	B (t_1)	B (t_2)	C (t_1)	C (t_2)
Correct	26	27	24	28	27	27
Not correct	6	5	8	4	5	5

Table 4 Dermoscopic classification of the benign epithelial tumours ($n = 13$) by the three investigators (A = excellent; B = average; C = beginner) at the two examinations (t_1 and t_2)

Classification	A (t_1)	A (t_2)	B (t_1)	B (t_2)	C (t_1)	C (t_2)
Correct	13	12	11	12	8	12
Not correct	0	1	2	1	5	1

for user B from 77.4% to 86.8% and for user C from 66.0% to 88.7% by the clinical history (Table 1). For the atypical nevi the improvement for user B was from 74.6% to 89.8% and for user C from 71.2% to 84.7% (Table 2). For the cutaneous melanoma the improvement for user B was from 75.0% to 87.5% and for user C there was no change (Table 3). Finally, for the epithelial benign tumours the improvement for user B was from 84.6% to 92.3% and for user C from 61.5% to 92.3% (Table 4).

In the subgroup of cutaneous melanomas with a reported clinical history in the past 3 months, a distinct decrease for user A but a clear increase for user B and a distinct increase for

Table 5 Results of the three investigators (A = excellent; B = average; C = beginner) for the subgroup of cutaneous melanomas with a reported clinical history in the past 3 months at the two examinations (t_1 and t_2)

Classification	A (t_1)	A (t_2)	B (t_1)	B (t_2)	C (t_1)	C (t_2)
Correct	13	12	9	13	12	13
Not correct	1	2	5	1	2	1

Table 6 Results of the three investigators (A = excellent; B = average; C = beginner) for the subgroup of benign melanocytic and epithelial skin tumours with a reported clinical history in the past 3 months at the two examinations (t_1 and t_2)

Classification	A (t_1)	A (t_2)	B (t_1)	B (t_2)	C (t_1)	C (t_2)
Correct	22	23	18	21	17	21
Not correct	3	2	7	4	8	4

user C were observed (Table 5). An improved diagnosis was also realized for the three users in the group of benign melanocytic and epithelial skin tumours (Table 6). In total, a change in the diagnosis was observed in 17 cases (six of the malignant and 11 of the benign group of skin tumours).

For user A a distinct improvement in the sensitivity was achieved by the history of the patients (from 81.3% to 84.4%), but also a distinct reduction in the specificity (from 94.6% to 92.3%); the diagnostic accuracy changed from 92.0% to 90.7%. The sensitivity of user B was clearly improved by the history (from 75.0% to 87.5%) and also the specificity (from 76.9% to 88.5%); the diagnostic accuracy improved from 76.5% to 88.3%. No change in the sensitivity was seen by user C (84.4%), but a clear improvement in the specificity by the history was found (from 69.2% to 87.7%); the diagnostic accuracy improved from 72.2% to 87.0%. No significant difference was seen between the two groups of melanocytic and non-melanocytic lesions in the analysis of sensitivity and specificity.

Using the computer algorithm,¹¹ 144 of 157 were included as melanocytic lesions in the digital analysis. For all melanocytic lesions, a sensitivity of 100%, a specificity of 76.8% and a diagnostic accuracy of 81.9% were achieved (Table 7). If the lesions was smaller than 12 mm ($n = 102$), all cutaneous melanomas and 72.8% of the benign melanocytic lesions were correctly classified. For the lesions larger than 12 mm ($n = 47$), all melanomas and 95% of the benign lesions were correctly classified (Table 7).

Table 7 Results of the digital analysis of all the malignant and benign: melanocytic lesions ($n = 144$), those smaller than 12 mm ($n = 97$) and those larger than 12 mm ($n = 47$) (NS = non-significant)

Classification	All lesions		Smaller than 12 mm		Larger than 12 mm	
	Benign	Malignant	Benign	Malignant	Benign	Malignant
Correct	86	32	67	5	19	27
Not correct	26	0	25	0	1	0
	$P < 0.005$		NS		NS	

Discussion

The present study revealed the following results of relevance to the use of dermoscopy in daily practice. (1) The level of continuing dermoscopic education influences the diagnostic accuracy. (2) The clinical history of the patients is helpful for average users and beginners in dermoscopy for the benign and malignant diagnoses. (3) Digital image analysis has the highest sensitivity but, with a known history, a lower specificity compared to the clinicians. (4) Digital dermoscopy is an excellent basis for store-and-forward teledermoscopy.

The level of continuing dermoscopic education influences the diagnostic accuracy. Between the three users clear differences were revealed depending on their experience in dermoscopy. User A (excellent) has the highest diagnostic accuracy with a high specificity of more than 90%. The differences between users A, B (average) and C (beginner) is clearer in the results of the first analysis of images (92.0%, 76.5% and 72.7%, respectively). These results were published recently by Kittler *et al.* based on a meta-analysis of 27 studies.² The diagnostic accuracy depends significantly on the degree of experience of the user of dermoscopy. Therefore, continuing dermoscopic education for average users and beginners is necessary.³

The clinical history of the patients is helpful for average users and beginners in dermoscopy for the benign and malignant diagnoses. For user A the clinical history slightly improved the sensitivity and slightly decreased the specificity. A clear improvement was seen for user B in both sensitivity and specificity. For user C, no improvement was seen in sensitivity, but there was improvement in specificity. Less experienced users were able to improve their diagnostic accuracy with the help of the clinical history. Possible associated problems are the time to take the history and its credibility. Kittler *et al.* chose any change in the past 12 months for the history.⁴ In our study we asked for the history of the past 3 months in which the memory of any change could be higher. If the history of the patient is accurate, it could be helpful in the diagnosis of benign and malignant skin tumours.^{1,2}

Digital image analysis has the highest sensitivity, but, by knowing the history, a lower specificity compared to the clinicians. The high sensitivity clearly led to a lower specificity. More benign melanocytic lesions were omitted, but all included melanomas *in situ* and melanomas were detected as malignant. Compared to some other groups who are developing systems for image analysis, in the present study only melanocytic lesions could be included for the automated image analysis.^{5,12} Therefore, the user of this system must be familiar with the differentiation between these two groups. One advantage of the automated image analysis is independence of time, concentration and number of images. However, in the cases when lesions are not easy to classify, an experienced clinician must check these dermoscopic images again. This applies in particular to the small melanocytic lesions. Particular hints could be also given to the

dermatopathologist.¹³ The integration of clinical data, especially of the history, could increase the diagnostic accuracy.¹⁴ In addition, the follow-up of pigmented and in particular of melanocytic lesions is easier and has benefits for both the patient and the clinician.^{15–17} In the future, digital image analysis systems should have the possibility of including the history of the patients, for example any change in the past 3 months (yes, no, and no available information), which could be a part of the diagnostic algorithm.

Digital dermoscopy is an excellent basis for store-and-forward teledermoscopy. Independent of time and location, and when or where the images are recorded, they could be stored and immediately analysed or sent by e-mail or the internet to a centre in which an automated analysing system is available.^{8,10} If the images are analysed in a centre with a time delay, patients could have missed the immediate decision of the treating doctor.¹⁸ Digital images could also be used for continuing dermoscopic education in telemedicine.¹⁹ With teledermoscopy and good evaluated analysing systems, a cost-minimization could be realized in the care of patients, especially in areas with low frequency of specialists.^{20,21} Guidelines and standards for recording the digital images and telemedicine are therefore necessary.^{5,22}

References

- 1 Mayer J. Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma. *Med J Aust* 1997; **167**: 206–210.
- 2 Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. *Lancet Oncol* 2002; **3**: 159–165.
- 3 Binder M, Poespoeck Schwarz M *et al.* Epiluminescence microscopy of small pigmented skin lesions: short-term formal training improves the diagnostic performance of dermatologists. *J Am Acad Dermatol* 1997; **36**: 197–202.
- 4 Kittler H, Seltenheim M, Dawid M *et al.* Morphologic changes of pigmented skin lesions. A useful extension of the ABCD rule for dermatoscopy. *J Am Acad Dermatol* 1999; **40**: 558–562.
- 5 Menzies SW. Automated epiluminescence microscopy. Human vs machine in the diagnosis of melanoma. *Arch Dermatol* 1999; **135**: 1538–1540.
- 6 Ganster H, Pinz A, Rohrer R *et al.* Automated melanoma recognition. *IEEE Trans Med Imaging* 2001; **20**: 233–239.
- 7 Dreiseitl S, Ohno-Machado L, Kittler H *et al.* A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *Biomed Inform* 2001; **34**: 28–36.
- 8 Piccolo D, Smolle J, Wolf IH *et al.* 'Face-to-face' versus remote diagnosis of pigmented skin tumors: a teledermoscopic study. *Arch Dermatol* 1999; **135**: 1467–1471.
- 9 Piccolo D, Smolle J, Argenziano G *et al.* Teledermoscopy – results of a multicentre study on 43 pigmented skin lesions. *J Telemed Telecare* 2000; **6**: 132–137.
- 10 Eedy DJ, Wootton R. Teledermatology: a review. *Br J Dermatol* 2001; **144**: 696–707.

- 11 Blum A, Ellwanger U, Lüdtke H, Garbe C. Digital image analysis of pigmented lesions: the Tübinger Mole Analyser. *Skin Res Technol* 1999; **5**: 127.
- 12 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**: 207–218.
- 13 Bauer J, Metzler G, Rassner G *et al*. Dermatoscopy turns histopathologist's attention to the suspicious area in melanocytic lesions. *Arch Dermatol* 2001; **137**: 1338–1340.
- 14 Binder M, Kittler H, Dreiseitl S *et al*. Computer-aided epiluminescence microscopy of pigmented skin lesions: the value of clinical data for the classification process. *Melanoma Res* 2000; **10**: 556–561.
- 15 Menzies SW, Gutenev A, Avramidis M *et al*. Short-term digital surface microscopic monitoring of atypical or changing melanocytic lesions. *Arch Dermatol* 2001; **137**: 1583–1589.
- 16 Kittler H, Binder M. Risks and benefits of sequential imaging of melanocytic skin lesions in patients with multiple atypical nevi. *Arch Dermatol* 2001; **137**: 1590–1595.
- 17 Braun RP, Calza AM, Krischer J, Saurat JH. The use of digital dermoscopy for the follow-up of congenital nevi: a pilot study. *Pediatr Dermatol* 2001; **18**: 277–281.
- 18 Williams T, May C, Esmail A *et al*. Patient satisfaction with store-and-forward teledermatology. *J Telemed Telecare* 2001; **7**: 45–46.
- 19 Aas I, Monrad H. Learning in organizations working with telemedicine. *J Telemed Telecare* 2002; **8**: 107–111.
- 20 Jacklin P, Roberts J. Social cost–benefit analysis of teledermatology. Costs were understated. *Br Med J* 2000; **321**: 896–897.
- 21 Loane MA, Oakley A, Rademaker M *et al*. A cost-minimization analysis of the societal costs of realtime teledermatology compared with conventional care: results from a randomized controlled trial in New Zealand. *J Telemed Telecare* 2001; **7**: 233–238.
- 22 Loane M, Wootton RA. Review of guidelines and standards for telemedicine. *J Telemed Telecare* 2002; **8**: 63–71.