Limitations of Dermoscopy in the Recognition of Melanoma

Hans Skvara, MD; Ligia Teban, MD; Manfred Fiebiger, MD; Michael Binder; Harald Kittler, MD

Objective: To compare dermoscopic features of melanocytic nevi with those of early melanomas that were not excised initially because of their uncharacteristic clinical and dermoscopic appearance.

Design: Retrospective study of the baseline images of 325 melanocytic skin lesions that were observed by digital dermoscopy and finally excised because of changes over time.

Setting: A dermatologic clinic and a dermatologic department at a university hospital.

Main Outcome Measures: Comparison of baseline images of melanomas and melanocytic nevi by pattern analysis, the ABCD rule of dermoscopy, and the 7-point checklist.

Results: Baseline dermoscopic images of 262 melanocytic nevi and 63 melanomas from 315 patients were included in the analysis. The patterns of dermoscopic features observed in the baseline images of melanocytic lesions finally diagnosed as melanomas during follow-up did not differ substantially from the patterns observed in the baseline images of melanocytic nevi. Pattern analysis, the ABCD rule of dermoscopy, and the 7-point checklist failed to achieve adequate diagnostic accuracy for melanoma. In retrospect, no dermoscopic feature or pattern of features could be identified that reliably differentiated between melanomas and melanocytic nevi at the time of the first presentation.

Conclusion: Dermoscopy depends on the appearance of classic dermoscopic features and is therefore limited in the diagnosis of very early and mainly featureless melanomas.

Arch Dermatol. 2005;141:155-160

EARLY RECOGNITION OF cutaneous melanoma is an ongoing challenge in dermatology.1-4 Dermoscopy (dermatoscopy and epiluminescence microscopy) has evolved over 20 years and is now widely used for the examination of pigmented skin lesions. According to previous studies, including 2 meta-analyses, this technique increases the sensitivity and specificity for the diagnosis of early melanoma in comparison with examination with the unaided eye.5,6 The sensitivity of dermoscopy has been reported to range from 60% to 100%, depending on, among other factors, the level of experience of the examiners and the diagnostic difficulty of the evaluated lesions.7-10

Although dermoscopy improves the diagnostic accuracy for melanoma, it cannot replace histopathologic examination.11-16 Some lesions, especially early melanomas, may lack specific dermoscopic features and are difficult to diagnose even with dermoscopy.17 It has been demonstrated that early recognition of these “featureless” melanomas can be enhanced by follow-up with digital dermoscopy. This technique adds diagnostic information by allowing observation of dermoscopic changes over time, which may be the only clue to establish the correct diagnosis.12-18,22

However, the question remains whether melanomas ultimately excised owing to changes over time initially displayed any morphologic features that might help differentiate them in the early stages from melanocytic nevi. We therefore retrospectively reviewed cases of melanoma initially missed by dermoscopy and examined whether dermoscopic criteria can be established to differentiate them from melanocytic nevi.
The sample included in this retrospective study consisted of 345 consecutive lesions that showed changes over time during follow-up digital dermoscopy and were excised from July 1996 to April 2002 in the Vienna clinic and September 1996 to July 2002 in the Bad Dürrnberg facility. After exclusion of palmar, plantar, and facial lesions and lesions that exceeded the maximum field of view of the electronic camera, the final sample included 325 melanocytic skin lesions. Standard histopathologic examination of all excised lesions identified 262 melanocytic nevi and 63 melanomas.

PRESENTATION OF IMAGES

Dermoscopic images were stored in Windows bitmap (BMP) format, a common graphics file format that stores digital images without compression. Each of the digital dermoscopic images was presented on a computer screen at 640 × 480 pixel resolution and 24-bit color depth.

The baseline images of the 325 lesions were presented to 2 experienced dermatologists who were unaware of the final histologic diagnosis. The follow-up images, which led to the decision to excise the lesions, were not shown. The dermatologists examined the baseline images for dermoscopic criteria (pattern analysis) in accordance with the consensus Internet meeting on dermoscopy held in 2001. Additionally, the baseline images of the 63 melanomas and a random sample of 63 melanocytic nevi were evaluated by 2 blinded investigators using the ABCD rule of dermoscopy and the 7-point checklist for dermoscopy.

STATISTICAL ANALYSIS

Sensitivity and specificity were calculated according to standard formulas. The χ² test and Fisher exact test were used for the comparison of proportions. Receiver operating characteristic (ROC) analysis was performed using the ROCKIT software provided by Charles Metz, MD, University of Chicago, Chicago, Ill (available at http://ray.bsd.uchicago.edu/cgi-bin/roc_software.cgi). All reported P values are 2-tailed, and P < .05 indicated statistical significance.

RESULTS

GENERAL DATA

A total of 325 melanocytic skin lesions of 297 patients were included in the analysis (mean patient age, 39 years; 56% female patients). All lesions were excised because of changes over time, which were identified during follow-up with digital dermoscopy. Histologically, 262 melanocytic lesions turned out to be melanocytic nevi, and 63 lesions were diagnosed as melanomas. In 9 (14.3%) melanomas, an association with a preexisting melanocytic nevus was found. Of the 63 melanomas, 31 (49.2%) were in situ melanomas. The median invasion thickness of the 32 invasive melanomas was 0.4 mm (range, 0.17-2.8 mm). Only 1 melanoma had an invasion thickness greater than 1 mm.

DERMOSCOPIC CHARACTERISTICS OF BASELINE IMAGES

We compared the frequencies of global and local dermoscopic patterns between baseline images of melanocytic nevi and melanoma. The distribution of global patterns is summarized in Table 1. We observed a trend toward a higher frequency of the multicomponent pattern in melanomas (24/63, 38.1%) compared with melanocytic nevi (21.4%), but the overall distribution of global patterns did not differ significantly between melanocytic nevi and melanomas (P = .10). When the le-
sions were analyzed according to the dermoscopic type, we found no statistically significant difference between melanocytic nevi and melanomas (Table 1). There was a tendency toward a higher frequency of the peripheral hyperpigmented type in melanomas (9/63, 14.3%) than in melanocytic nevi (6.9%), but this difference was not statistically significant.

The distribution of local features is summarized in Table 2. We observed no differences between melanocytic nevi and melanomas with respect to the presence or absence of streaks, regression, blue veil, blotches, hypopigmented areas, and reticular depigmentation or with respect to the presence or distribution of brown globules. There was also no statistically significant difference between melanocytic nevi and melanomas regarding the presence and the quality of vascular patterns. Melanocytic nevi and melanomas differed with respect to the pigment network, the presence and distribution of black dots, and the presence of pseudopods. The presence of an irregular pigment network had the highest sensitivity for melanoma (23.8%). The presence of black dots at the periphery had a sensitivity of 11.1% and the presence of pseudopods had a sensitivity of only 7.9%.

**ABCD RULE AND 7-POINT CHECKLIST**

The baseline images of all 63 melanomas and of 63 randomly selected melanocytic nevi were analyzed using the ABCD rule for dermoscopy and the 7-point checklist. As shown in Figure 1, neither algorithm achieved adequate diagnostic accuracy for melanoma. The area under the curve was 0.67 (95% confidence interval, 0.57-0.76) for the ABCD rule and 0.64 (95% confidence interval, 0.52-0.75) for the 7-point checklist. At a specificity of 90%, the ABCD rule had a sensitivity of 23%, and the 7-point checklist had a sensitivity of 27%.

| Table 2. Distribution of Local Dermoscopic Features* |
|-----------------|-----------------|----|
| Characteristic  | Melanocytic nevi | Melanoma | P Value† |
| Pigment network |                 |     |     |
| Absent          | 152 (58.0)       | 30 (47.6) | .02 |
| Regular         | 82 (31.3)        | 18 (28.6) |     |
| Irregular       | 28 (10.7)        | 15 (23.8) |     |
| Streaks         | 26 (9.9)         | 11 (17.5) | .09 |
| Pseudopods      | 5 (1.9)          | 5 (7.9)  | .03 |
| Globules        |                 |     | .57 |
| Absent          | 178 (67.9)       | 37 (58.7) |     |
| Diffuse         | 34 (13.0)        | 11 (17.5) |     |
| Peripheral      | 44 (16.8)        | 13 (20.6) |     |
| Central         | 6 (2.3)          | 2 (3.2)  |     |
| Blue veil       | 2 (0.8)          | 3 (4.8)  | .05 |
| Regression      | 4 (1.5)          | 3 (3.8)  | .14 |
| Hypopigmentation| 74 (28.2)        | 22 (34.9) | .29 |
| Blotches        | 96 (36.3)        | 21 (33.3) | .62 |
| Vascular pattern|                 |     | .23 |
| Absent          | 223 (85.1)       | 51 (81.0) |     |
| Regular         | 22 (8.4)         | 4 (6.3)  |     |
| Irregular       | 17 (6.5)         | 8 (12.7) |     |
| Reticular depigmentation | 42 (16.0) | 11 (17.5) | .78 |
| Black dots      |                 |     | .03 |
| Absent          | 220 (84.0)       | 43 (68.3) |     |
| Diffuse         | 28 (10.7)        | 6 (12.7) |     |
| Peripheral      | 8 (3.1)          | 7 (11.1) |     |
| Central         | 6 (2.3)          | 5 (7.9)  |     |

*Unless otherwise indicated, data are number (percentage) of lesions. Using the χ² test (with appropriate degrees of freedom) or Fisher exact test (for proportions with 2 categories).

Figure 1. Receiver operating characteristic (ROC) curves for the diagnostic accuracy for melanoma of the ABCD rule for dermoscopy (A) and for the 7-point checklist (B).
Using the conventional thresholds of the ABCD rule (ABCD score $\geq 4.75$) and the 7-point checklist (7-point score $\geq 2$), the sensitivity of the ABCD rule was 31.7% with a corresponding specificity of 87.3%, and the sensitivity of the 7-point checklist was 11.1% with a corresponding specificity of 95.2%.

This study demonstrates the limitations of dermoscopy in the detection of early melanomas that lack specific dermoscopic criteria. Dermoscopic algorithms such as pattern analysis,\textsuperscript{27,28} the ABCD rule,\textsuperscript{24,25} and the 7-point checklist\textsuperscript{26} could not reliably differentiate these types of melanomas from melanocytic nevi. The results of our study are clinically important because the melanocytic skin lesions that are typically examined by dermoscopy are usually small and doubtful lesions, difficult to diagnose with the unaided eye. Many critics of dermoscopy argue that dermoscopy is only accurate in cases where the diagnosis of melanoma is easily established even with the unaided eye. At first glance, our study supports this view.
While most dermoscopy studies have demonstrated an improvement in the diagnosis of pigmented skin lesions with dermoscopy compared with examination with the unaided eye, this was not the case in the present study. In fact, we found that dermoscopy achieved a lower diagnostic accuracy for melanoma than in any other published study, a level of accuracy only slightly better than chance. This suggests that the diagnostic accuracy of dermoscopy has been overestimated.

The major difference between the present study and former studies is the way the lesions were selected. Previous studies included lesions only when they had been selected on the basis of histologic evidence. The histologic report is regarded as the gold diagnostic standard in our field, and usually scientific journals will not accept a manuscript on dermoscopy that includes lesions for which the diagnosis has not been proven histologically. The problem with this approach is that the decision to excise a lesion is based on the presence of certain dermoscopic features. Lesions that do not display these features are not excised. This leads to a bias toward lesions that show suggestive dermoscopic features and patterns and to an overestimation of the sensitivity of dermoscopy.

Our study includes images of melanomas, which are usually not included in other studies. Owing to their uncharacteristic dermoscopic appearance at first presentation, these melanomas were not excised initially. Before the arrival of digital documentation, most of these cases would have been missed initially and would have been excised later in the course of the disease when the criteria for melanoma became more obvious.

There are 2 groups of melanomas not included in this study: (1) those that can be easily diagnosed with the unaided eye, probably the largest group; and (2) those that are missed by examination with the unaided eye but can be diagnosed by dermoscopy—this group is responsible for the gain in diagnostic accuracy achieved with dermoscopy. The first group is used by proponents of dermoscopy to argue that dermoscopy is unnecessary. The second group is used by proponents of dermoscopy to argue that dermoscopy is necessary. Logic tells us that both views are incomplete and selective.

Our data demonstrate that there is a third group of melanomas, those targeted by the present study, which can be diagnosed neither with the unaided eye nor dermoscopically. This group of melanomas is especially important because it includes a high proportion of early lesions. Our data suggest that these early and featureless melanomas can be recognized only by observing their dynamic evolution over time (Figure 2). In our cases, the decision for excision was made after follow-up with digital dermoscopy, in line with other reports by Kittler et al,12 Menzies et al,18 and Robinson and Nickoloff.29 It is not difficult to foresee that this group of melanomas will be used by proponents of digital dermoscopy to argue that follow-up with digital dermoscopy is necessary.30

There are 3 possible limitations of our study. First, although we argued as if we knew from first presentation that all melanomas in our sample were melanomas, some of them might have been melanocytic nevi that developed into melanomas. In fact, only 9 of our 63 melanomas histologically showed an association with a preexisting melanocytic nevus, whereas 54 did not. It is therefore likely that most of the melanomas in our sample developed de novo, and it should be assumed that these excised melanomas were malignant when first imaged.

Second, it may be argued that image quality of the digital system may have influenced the results of this analysis. Dermoscopic features may exist to differentiate melanomas from melanocytic nevi at first presentation, but these might not have been identified because of the limitations in the optical resolution of the system. However, this is very unlikely because many lesions, the images of which were stored with our digital dermatoscope, were also examined with a classic, handheld dermatoscope by the dermatologist who examined the patient at the first visit.

Third, it may be argued that diagnostic algorithms for dermoscopy, which were not tested in this study (eg, the Menzies rule22), might provide a higher sensitivity than pattern analysis, the 7-point checklist, or the ABCD rule. We find this very unlikely because all known algorithms are more or less combinations of certain dermoscopic criteria, none of which was found to be useful for the identification of the early melanomas included in our study.

We conclude that even though dermoscopy is a very useful adjunct in the diagnosis of melanoma, its application is dependent on the appearance of classic dermoscopic features and is therefore limited in the diagnosis of very early and mainly featureless melanomas. Short-term follow-up of these lesions may be a promising approach to recognize these types of melanoma as early as possible.

Accepted for Publication: November 2, 2004.

Correspondence: Harald Kittler, MD, Department of Dermatology, University of Vienna Medical School, Wahringerguertel 18-20, A-1090 Vienna, Austria (harald.kittler@meduniwien.ac.at).

Funding/Support: This study was supported by grant 9952 from the Austrian Federal Bank, Vienna, Austria.

REFERENCES