The Dermatoscopic ABCD Rule Does Not Improve Diagnostic Accuracy of Malignant Melanoma

HENRIK LORENTZEN, KAARE WEISMANN, LENA SECHER, CARSTEN SAND PETERSEN and FREDERIK GRØNHØJ LARSEN

Department of Dermatology, Bispebjerg Hospital, Copenhagen, Denmark

The dermatoscopic ABCD rule has been suggested to improve diagnostic performance regarding cutaneous malignant melanoma. Using this rule, a total dermatoscopy score is calculated from the presence of various dermatoscopic elements. A total dermatoscopy score above 4.75 signifies possible and 5.45 probable melanoma. We compared the diagnostic accuracy of dermatoscopy with and without the use of the ABCD rule. Furthermore, receiver operating characteristic analysis was performed for the ABCD rule. The area under the receiver operating characteristic curve was 0.854 (range 0.777–0.906) demonstrating that in 85.4% of the cases, cutaneous malignant melanomas were rated higher than the non-melanoma skin lesions. Sensitivity for the melanoma diagnosis was higher for simple dermatoscopy than when the ABCD rule was used ($p < 0.05$). There was no difference in specificity when a total dermatoscopy score of 4.75 was used as cut-off point, but specificity was lower for simple dermatoscopy than when the total dermatoscopy score of 5.45 was used. Diagnostic accuracy was higher for simple dermatoscopy than for the ABCD rule ($p < 0.01$). In conclusion, the dermatoscopic ABCD rule was not superior to simple dermatoscopy, and fewer malignant melanomas were identified with this rule. *Key words: dermatoscopy; epiluminiscence microscopy; dermatoscopic ABCD; cutaneous malignant melanoma; accuracy; receiver operating characteristic.*

Dermatoscopy has been shown to increase diagnostic accuracy (1–4) and facilitate clinical decision-making in cases of cutaneous malignant melanoma (CMM).

Nachbar et al. (3) and Stolz et al. (5) proposed an ABCD rule of dermatoscopy, that has become widespread due to its obvious didactic qualities. In summary, the ABCD rule works by the calculation of a total dermatoscopy score (TDS) from the addition of scores obtained from the recognition of individual dermatoscopic findings (Table I). Other rules for identifying CMM with the dermatoscope have been suggested. Kenet et al. (6, 7) have emphasized on the importance of the pigment network of the lesions and have suggested a ranking of the findings as low-, medium- and high-risk features.

The aim of the present investigation was to evaluate the ABCD rule by receiver operating characteristic (ROC) curve analysis (8, 9).

MATERIALS AND METHODS

Nine observers assessed clinical and dermatoscopic photo-slides obtained from 232 patients that had been referred to the dermatological outpatient clinic for evaluation of a pigmented skin lesion. Observers 1–4 were experienced users of dermatoscopy and observers 5–9 were less experienced users. All observers were familiar with the literature on dermatoscopy (3, 7, 10–13). Prior to the test performance, which was divided in two sessions, the ABCD rule of Stolz et al. was recapitulated.

All pigmented skin lesions had been surgically removed and sent to pathology for histological assessment. Lesions suspicious of CMM were confirmed by immunostaining with S-100 and HMB-45.

Macroscopic clinical and dermatoscopic photo-slides of each patient case were projected to an 80×110 cm screen for approximately 3 minutes, less for simple melanocytic naevi and more for melanomas. Patient data was not presented. If requested by an observer, additional time for the assessment was provided. The observers first noted their clinical then their overall dermatoscopic diagnoses in entry forms without mutually discussing their assessments. The test setting has been reported previously by us (14).

In addition, the observers recorded the presence or absence of the singular elements of Stolz' ABCD rule by ticking off in the entry form. The observers did not perform calculation of the TDS, and did not use this score in deciding whether a case was suspicious of CMM or not. All assessments were later entered into a database and the TDS was calculated. The validity of the ABCD rule and the TDS to distinguish CMM from other pigmented skin lesions was tested by ROC curve analysis. ROC curves are constructed by plotting the sensitivity of a test, e.g. the dermatoscopic ABCD rule, against 1 specificity for varying decision thresholds, e.g. the TDS, is used to discriminate between a diseased and disease-free population. If the threshold is set very low most of the diseased population will have a positive test result, i.e. the sensitivity is high, but many disease-free cases will have a false positive test-result, hence the specificity will be low. When the decision threshold is gradually increased, the sensitivity gradually diminishes as the specificity increases. ROC curve analysis allows the determination of a reasonable or acceptable trade-off point between sensitivity and specificity so that the number of false positive cases (e.g. resulting in disfiguring surgery) and false negative cases (e.g. resulting in tumour progression) are minimized. The shape of the ROC curve provides information about the test's ability to identify diseased and disease-free cases, respectively. The $X = Y$ line is called "the line of no information" because a test following this line is unable to distinguish a diseased population from a healthy one. A test with a ROC curve passing through the upper left corner, i.e. sensitivity $= 1$ and specificity $= 1$ is an ideal test, which perfectly separates the diseased from the healthy population. The area under a ROC curve represents the probability of correctly ranking a patient as diseased or normal. The area under the ROC curve where $X = Y$ is 0.5, i.e. the same as "tossing a coin", whereas the area for a ROC curve through the upper left corner equals 1. Most tests fall between these two extremes. From the ROC analysis sensitivities and specificities for TDS $= 4.75$ and TDS $= 5.45$ (possible and probable CMM (4)) were calculated. The sensitivity and specificity obtained by simple dermatoscopy were plotted in the matching ROC-space. The Wilcoxon test was used for comparing the diagnostic accuracy (true positives/[true positives + false positives + false negatives]), the sensitivity and specificity obtained by the TDS of the ABCD rule and dermatoscopy performance.

Interobserver variation for the 2 groups of observers (experts and trainee level) was calculated using the method described by Schouten

(15) for a fixed group of observers and where missing values may occur. Diagnoses were dichotomized into probable melanoma (TDS > 5.45) and other pigmented skin lesions (TDS < 5.45), i.e. including possible melanoma (4.75 < TDS < 5.45). After a 1-year interim, 3 of the expert observers went through the patient cases again in order to establish intra-observer variation using Cohen's kappa-coefficient (16). The interim was intended in order to minimize bias from simple recollection of patient cases. For the intra-observer study non-melanocytic lesions were not included. The following categorization of the kappa-coefficient into an agreement scale was made: Kappa < 0: poor agreement; between 0 and 0.20: slight; 0.21–0.40: fair; 0.41–0.60: moderate; 0.61–0.80: substantial and 0.81–1: almost perfect agreement.

RESULTS

Fig. 1 shows a representative ROC-curve obtained by one of the observers using the ABCD rule. The median AUC for the ROC-curves of all observers was 0.854 (range 0.777–0.906), demonstrating that the ABCD rule yielded an AUC significantly different from 0.5. By using a TDS of 4.75, suggested by Stolz et al. (3, 13) as a cut-off score for lesions suspicious of CMM, the median sensitivity was 0.59 (range 0.49–0.81) and the median specificity was 0.92 (range 0.79–0.97). When a cut-off point of TDS 5.45 was used, which indicates a high degree of CMM suspicion (3, 13), the median sensitivity was 0.41 (0.29–0.75) and the median specificity was 0.98 (0.91–1). For simple dermatoscopy without ABCD scoring, the sensitivity was 0.72 (0.58–0.92) which was significantly higher than obtained with TDS = 4.75 ($p < 0.05$) and TDS = 5.45 ($p < 0.01$). Specificity for simple dermatoscopy was 0.94 (0.78–0.97), which was not different from the specificity with a TDS of 4.75 but lower than the specificity of a TDS of 5.45 ($p < 0.01$).

The dermatoscopy sensitivities and specificities were plotted in the corresponding ROC space for each observer (Fig. 1).
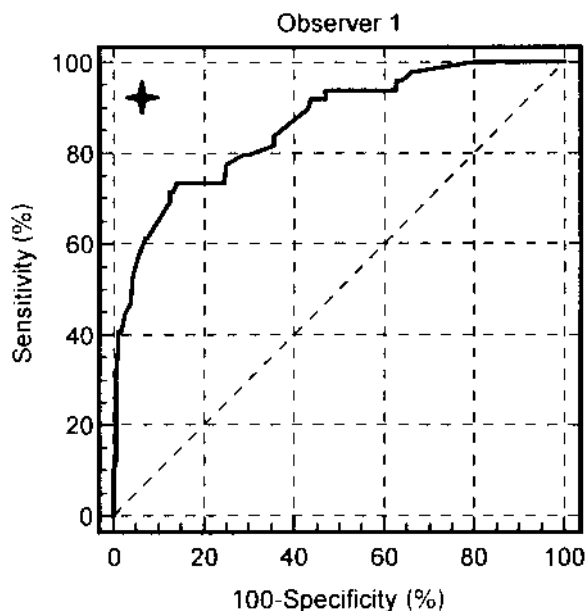
For most observers this point is located to the upper left of the ROC curve. Diagnostic accuracy was higher for simple dermatoscopy than for the ABCD rule, both with TDS of 4.75 ($p < 0.01$) and TDS of 5.45 ($p < 0.05$).

The TDS (median and inter quartile range) for CMM *in situ*, superficially spreading CMM, nodular CMM and lentigo maligna melanoma are plotted in Fig. 2, together with the TDS of various skin lesions. Basal cell carcinomas had relatively high TDS (median = 4.3, IQR 3.3, 4.6).

Observed agreement in the expert group was 0.90 and interobserver variation was kappa = 0.83 indicating almost perfect agreement. In the "non-expert" group the observed proportion of agreement was 0.88 and the group kappa-coefficient was 0.80, indicating substantial/almost perfect agreement. Intra-observer variation demonstrated an observed proportion of agreement of 0.85; 0.90 and 0.92 and a kappa coefficient of 0.54; 0.58 and 0.70 (moderate to substantial agreement) for the 3 expert observers, respectively.

DISCUSSION

Early recognition of CMM is a challenge to the clinician, as the incidence is still increasing. For this reason, the ability to discriminate between CMM and benign pigmented skin lesions is critical. In the hands of an experienced investigator, dermatoscopy increases the diagnostic accuracy of various pigmented skin lesions, especially CMM. The ABCD rule was proposed by Stolz et al. (5) to improve the dermatoscopic diagnosis of CMM. The system has not been studied by ROC analysis before. This is an efficient way of analysing the diagnostic performance of a quantitative or semi-quantitative test result that falls on a scale dichotomized into disease positives and negatives. For most laboratory, paraclinical and clinical tests there is an overlap of the diseased and disease-free populations. This is also the case for the ABCD rule demonstrating ROC curves with areas less than 1. We found a median AUC of 0.854, which can be read as an overall probability of correctly ranking the patient cases, i.e. 85.4% of CMM would have a higher TDS than other pigmented skin lesions.

In the material used in this investigation, most non-CMM had TDS below the cut-off value of 4.75, which is reflected in the high specificities obtained. In general, high TDS were
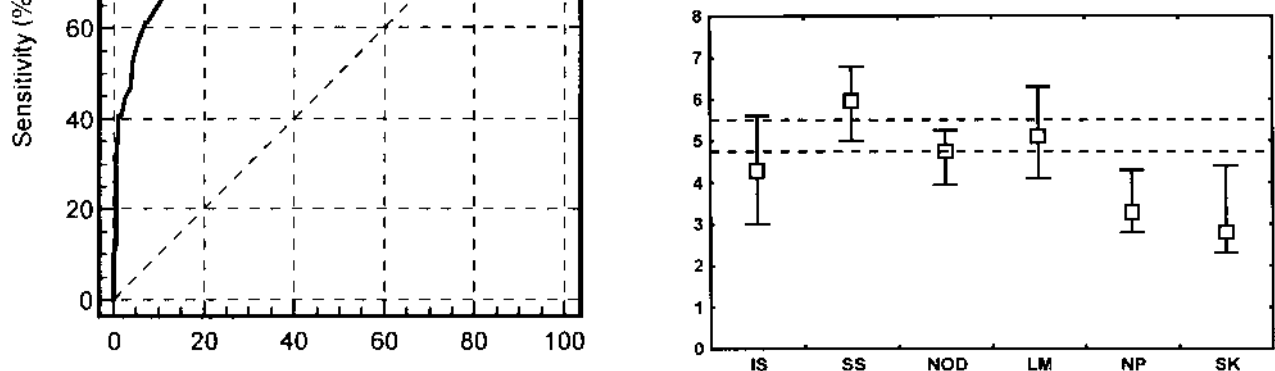

Fig. 1. Representative receiver operating characteristic curve. ✚: actual combination of sensitivity and specificity obtained by the observer using dermatoscopy without application of the ABCD rule and total dermatoscope score.

Fig. 2. Median total dermatoscopy score (TDS) distributed after type of pigmented skin lesion. IS: *in situ* CMM; SS: superficially spreading CMM; NOD: nodular CMM; LM: lentigo maligna; NP: naevus pigmentosus; SK: seborrhoeic keratoses. (Bars: 25 and 75 percentiles).

Table I. *The dermatoscopic ABCD-rule (5)*

	Weighting factor	Possible scores	Min. – Max.
Asymmetry	1.3	0 – 2	0 – 2.6
Border	0.1	0 – 8	0 – 0.8
Colour	0.5	1 – 6	0.5 – 3
Light brown			
Dark brown			
Red			
White			
Bluish			
Black			
Differential structures	0.5	0 – 5	0 – 2.5
Dots			
Globules			
Network			
Streaming			
Homogenous areas			
Total dermatoscopy score			0.5 – 8.9

associated with CMM, but not all had a TDS above 4.75, particularly not the nodular melanomas (Fig. 2), consequently yielding a low sensitivity. The observers generally obtained a higher sensitivity and accuracy without ABCD scoring, so the observers would have failed to diagnose some melanomas if they had adhered strictly to the system. Other investigators have reported higher sensitivity and accuracy than we have found for the ABCD rule (3, 5). The discrepancy between the dermatoscopic performance of the observers and that obtained with the ABCD rule may have plural causes: the selection of the patient cases, the experience of the observers in recognizing the ABCD elements and deficiencies of the ABCD and/or TDS when used in a clinical setting as intended in our study.

For creation of the ABCD rule the investigators used colour prints of the lesions and only melanocytic naevi and predominantly early CMM less than 1 mm Breslow thickness were used (5). Some of the characteristic findings of thin CMM are not encountered in thick CMM: pigment network, radial streaming and white areas (17). The expertise of the observers ranged from 5 years of daily work with dermatoscopy ("expert level") to 2 years of interest in and training in dermatoscopy ("trainee level"). A different diagnostic accuracy of the TDS might have been achieved if the individual patient cases had been evaluated as colour prints by consensus making of presence or not of ABCD findings. This was not our intention, since we wanted to study a practical clinical and more everyday-like use of dermatoscopy.

The correction factors used for the calculation of the TDS were determined by multivariate statistical methods with a high weighting on asymmetry and different colours, whereas the impact of the more specific disturbances in the pigment network is relatively small (5). The presence of radial streaming or pseudopods, which histologically correlate with a malignant growth pattern (10), is a strong predictor of malignancy (12), but in the ABCD rule it only scores 0.5. Coarseness and irregularity of the pigment network are only assessed indirectly through scoring of asymmetry and as part of the "differential structures". Kenet & Fitzpatrick (6) base their dermatoscopic strategy on disturbances in the pigment network. In their system a singular high-risk factor is regarded as indicative of CMM. A comparison of Stolz' and Kenet's different diagnostic strategies will be performed.

In the ABCD rule different colours signifies a risk of CMM. In the majority of cases, colour variation at the same time is registered as an asymmetry factor. Each of the colours is equally weighted. This could explain why the heavily black pigmented nodular melanomas had sub-threshold TDS in our study. Nilles et al. (11) demonstrated that black-brown pigmentation was a predictor of CMM, whereas dark brown pigmentation was not. We found almost perfect inter-observer agreement in the expert group and substantial agreement in the "non-expert" group. In the "non-expert" group missing values amounted 12.5%. At least 3 possible causes to the missing values must be considered: (i) random missing values, which do not affect the kappa-coefficient as the distribution of the missing values would equal the assessed cases; (ii) cases not fitting the system; and (iii) cases the observers found difficult to diagnose. In this case the missing values would impose a bias on the kappa coefficient towards higher values so the kappa coefficients should be considered a maximum estimate in the "non-expert" group.

The lower intra- than inter-observer agreement may be explained by selection of cases but also by a shift in the observers' attention towards pigment network disturbances in the 1-year period between the assessments. Finally, Cohen's kappa-coefficient is highly dependent on the prevalence of the disease in question and a higher proportion of melanomas would result in higher kappa coefficients.

REFERENCES

1. Binder M, Schwarz M, Winkler A, Steiner A, Kaider A, Wolff K, et al. Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch Dermatol 1995; 131: 286 – 291.
2. Cristofolini M, Zumiani G, Bauer P, Cristofolini P, Boi S, Micciolo R. Dermatoscopy: usefulness in the differential diagnosis of cutaneous pigmentary lesions. Melanoma Res 1994; 4: 391 – 394.
3. Nachbar F, Stolz W, Merkle T, Cognetta AB, Vogt T, Landthaler M, et al. The ABCD rule of dermatoscopy. J Am Acad Dermatol 1994; 30: 551 – 559.
4. Binder M, Puespoeck-Schwartz M, Steiner A, Kitler H, Muellner M, Wolff K, et al. Epiluminiscence microscopy of small pigmented skin lesions: short-term formal training improves the diagnostic performance of dermatologists. J Am Acad Dermatol 1997; 36: 197 – 202.
5. Stolz W, Rieman A, Cognetta AB, Pillet L, Abmayr W, Hölzel D, et al. ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur J Dermatol 1994; 4: 521 – 527.
6. Kenet RO, Fitzpatrick TB. Reducing mortality and morbidity of cutaneous melanoma: a six year plan B) identifying high and low risk pigmented lesions using epiluminiscence microscopy. J Dermatol 1994; 21: 881 – 884.
7. Kenet RO, Kang S, Kenet BJ, Fitzpatrick TB, Sober AJ, Barnhill

RL. Clinical diagnosis of pigmented lesions using digital epiluminiscence microscopy. Arch Dermatol 1993; 129: 157–174.

8. Langley FA, Buckley CH, Tasker M. The use of ROC curves in histopathologic decision making. Anal Quant Cytol Histol 1985; 7: 167–173.

9. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993; 39: 561–577.

10. Menzies SW, Crotty KA, McCarthy WH. The morphologic criteria of the pseudopodes in surface microscopy. Arch Dermatol 1995; 131: 436–440.

11. Nilles M, Boedeker R, Schill W. Surface microscopy of naevi and melanomas- clues to melanoma. Br J Dermatol 1994; 130: 349–355.

12. Steiner A, Binder M, Schemper M, Wolff K, Pehamberger H. Statistical evaluation of surface microscopy criteria for melano-

cytic pigmented skin lesions. J Am Acad Dermatol 1993; 29: 581–588.

13. Stolz W, Braun-Falco O, Bilek P, Landthaler M, Cognetta AB. Color atlas of dermatoscopy. Oxford: Blackwell Science, 1994.

14. Lorentzen H, Weismann W, Sand Petersen C, Grønhøj Larsen F, Secher L, Skødt V. Clinical and dermatoscopic diagnosis of malignant melanoma-assessed by an expert and non-expert group. Acta Derm Venereol 1999; 79: 301–304.

15. Schouten HJA. Statistical measurement of interobserver agreement–analysis of agreements and disagreements between observers. Thesis, Rotterdam 1985.

16. Cohen J. A coefficient of agreement for nominal scales. Educational and psychological measurement 1960; 20: 37–46.

17. Argenziano G, Fabbrocini G, Carli P, De Giorgi V, Delfino M. Epiluminiscence microscopy: criteria of cutaneous melanoma progression. J Am Acad Dermatol 1997; 37: 68–74.