

The influence of training on the recognition of gross features of dermoscopy images

Aydin Yucel, Suhan Gunasti, Varol Lutfu Aksungur

Department of Dermatology,
Faculty of Medicine, Çukurova
University, 01330 Adana,
Turkey

Address for correspondence:
Dr. Aydin Yucel,
Department of Dermatology,
Faculty of Medicine,
Çukurova University
01330 Adana, Turkey.
E-mail: aydinyuc@gmail.com

ABSTRACT

Background: In a dermoscopic examination, besides structural components, inexperienced clinicians should also be able to recognize the gross features of the images. **Aim:** The aim of this study is, whether or not an inexperienced clinician has problems in the recognition of gross features of the images on dermoscopic examination. **Methods:** Two dermatologists, of whom one was experienced in the field of dermoscopy and the other was not, examined 161 dermoscopic images of melanocytic lesions in the gross features of their borders. Inner and outer borders were defined for each lesion. Both dermatologists separately evaluated the borders of the lesions for irregularity, asymmetry, and wideness of fading. For subjective image analysis they scored each lesion by using the four-point ordinal scale. For computerized image analysis they manually marked borders with dots, by using a computer program. We used quadratic-weighted kappa for interobserver reliability assessments for subjective scores and intraclass correlation coefficients (ICC) for automatically calculated scores. **Results:** In a subjective evaluation the inexperienced observer used a higher score than the experienced observer and the kappa values were between 0.241 – 0.286. ICC for the automatically calculated scores were between 0.357 and 0.522. According to both the outer and the inner borders, the concordance between experienced and inexperienced observers was almost perfect in measurements of diameter, perimeter, and area (ICC scores were between 0.948 and 0.990). **Conclusions:** An inexperienced person, in comparison with an experienced person, sees lesions in the same sizes, but in different shapes on dermoscopy. Therefore, it is advisable that making learners familiar with the borders of lesions should be included in the training on dermoscopy.

DOI: 10.4103/0378-6323.60550

Key words: Dermatology education, dermoscopy, melanoma-pigmented lesions, nevi

INTRODUCTION

Dermoscopy provides significant benefits to clinicians in differential diagnosis of pigmented skin lesions. Various algorithms have been recommended for dermoscopic diagnostic procedures. The major algorithms are pattern analysis, revised pattern analysis, ABCD rule of dermoscopy, seven-point check list, three-point checklist, and Menzies method.^[1,2] All of these require that the clinician be able to recognize structural components. On dermoscopic examination, an inexperienced clinician achieves less accurate diagnoses when compared to an experienced clinician.^[3-5] At first glance, one may suggest that

failure of the inexperienced clinician might be related to his inability to recognize the structural components, which are the details of a dermoscopic image. Before making a final decision on this subject, it should be investigated whether or not an inexperienced clinician also has problems in recognition of the gross features in the image. There are articles about teaching dermoscopy to residents in dermatology.^[6-8] They tried to determine the most reliable method (pattern analysis, ABCD rule, seven-point checklist) for nonexperts to diagnose melanocytic lesions. In their manuscript, they did not consider that an inexperienced clinician may also have problems in recognition of the gross features of the images.

How to cite this article: Yucel A, Gunasti S, Aksungur VL. The influence of training on the recognition of gross features of dermoscopy images. *Indian J Dermatol Venereol Leprol* 2010;76:132-7.

Received: March, 2009. **Accepted:** April, 2009. **Source of Support:** Nil. **Conflict of Interest:** None declared.

In this study, two dermatologists, of whom one was experienced in the field of dermoscopy and the other was not, examined 161 dermoscopic images of melanocytic lesions with respect to the gross features of their borders. Concordance between these observers was evaluated.

METHODS

Digital Dermoscopic Images

Dermoscopic images of 161 melanocytic lesions from 91 patients (47 females, 44 males; mean age \pm SD, 32.8 ± 15.5 years; range, 8 – 75 years) were included in this study. Dermoscopic examinations of these lesions had been done with Mole Max II from 2000 to 2006. In all of these lesions, diagnoses were confirmed by histopathological examination. These diagnoses were 57 common acquired nevocmelanocytic nevi, 81 atypical melanocytic nevi, 13 blue nevi, six Spitz nevi, and four malignant melanomas (all the four were invasive superficial spreading melanomas). Of the common acquired nevocmelanocytic nevi, 10 were junctional; and 47, compound. Of the atypical melanocytic nevi, 13 were junctional; and 68, compound. The dermoscopic images represented a 30-fold magnification of the lesions. These images were digital (24-bit color; 640 x 480 pixels). In these images, 52 pixels corresponded to 1 mm. These images were evaluated on the same computer by the two observers.

Observers

Two dermatologists from our clinic participated in this study. The first (AY) was a specialist in dermatology for eight years and had a practical experience in dermoscopy for seven years, after a formal training. The second (SG) was a specialist in dermatology for one-and-a-half-years and had neither practical experience nor formal training in dermoscopy. Both of them evaluated the borders of the lesions in the images separately and also manually marked borders with dots, using a computer program developed by one of us (VLA). None of the observers were aware of the diagnosis of the lesions.

Subjective Image Analysis

Two borders were defined for each lesion: (1) the inner border outside of which the melanocytic lesion starts to fade, and (2) the outer border outside of which the lesion fades completely [Figure 1].

Both observers scored each image with respect to

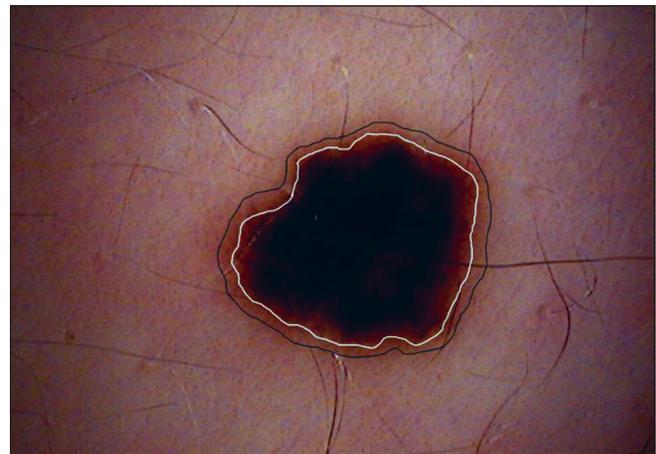


Figure 1: The outer and the inner borders of a melanocytic lesion in a dermoscopic image

three features of the outer border. They used a four-point ordinal scale: 0 to 3 (absent, mild, moderate, and severe). The three features were irregularity, asymmetry, and wideness of fading. For irregularity, the degree of scalloping, notching, ragging, and / or blotching of the outer border was evaluated. Evaluation of asymmetry was done in a similar manner to the ABCD rule.^[9,10] However, only shape was taken into consideration for asymmetry, whereas, pigment distribution and structural components were disregarded. For evaluation of wideness of fading, the outer border was taken into consideration together with the inner border. If these borders almost completely overlapped, the wideness of fading was considered to be absent. The greater the distance between the borders, the higher the score (0 to 3). The scores were called subjective irregularity (sub-I) score, subjective asymmetry (sub-A) score, and subjective fading (sub-F) score.

Computerized Image Analysis

In order to determine both the inner and outer borders of each lesion, the observers marked at least 20 dots (pixels) by clicking on them. Average marked dots on the lesions were 55 for the experienced and 70 for the inexperienced dermatologist. The coordinates of these pixels on the x-axis and y-axis were automatically stored in a database. Subsequently, the maximum diameter, the perimeter, and the area of the lesion were automatically calculated in pixels, separately, for the inner and outer borders.

Evaluation of irregularity according to the pixels of a border was based on the postulate [Figure 2a]: “If

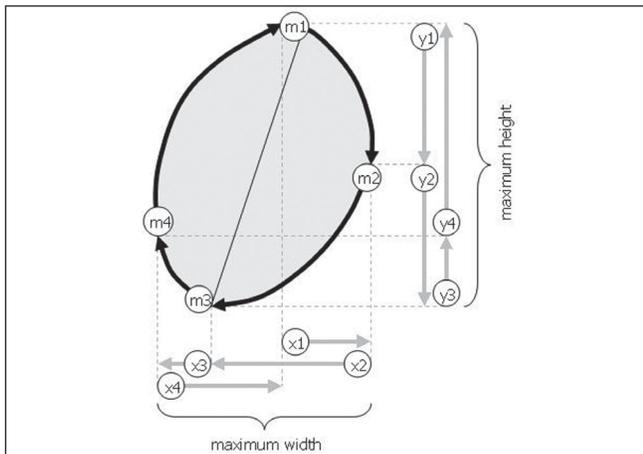


Figure 2a: During movements labeled with m1, m2, m3, and m4, over the border of the shape, the ways taken on the x-axis have been labeled with x1, x2, x3, and x4; and the ways on the y-axis, with y1, y2, y3, and y4. As shown, the total way on the x-axis, in other words sum of x1, x2, x3, and x4, is equal to two-fold the maximum width; and the total way on the y-axis, two-fold the maximum height

a geometrical shape has neither indentations nor protrusions, during a movement over its border starting from a point and ending at the same point, the ways taken on the x-axis and y-axis were two-fold the maximum width and two-fold the maximum height of the shape, respectively. If the shape has indentations or protrusions, these ways exceed two folds proportionately with the amount of indentations and protrusions.” Therefore, the horizontal and vertical ways during a movement over a border of the lesion were automatically calculated in pixels. These ways were divided to the maximum width and height of the lesion, respectively. Next, the mean of the quotients was taken, in order to obtain the automatically calculated irregularity (auto-I) score. This score was calculated separately for the inner and outer borders.

The image was rotated until the maximum diameter of the lesion was parallel to the x-axis [Figure 2b]. So, the diameter divided the lesion into upper and lower parts. For each point on the diameter, the pixels of the upper part and the pixels of the lower part were counted vertically. The absolute differences between the results of these countings were calculated. Then, the sum of the differences was divided by the area of the lesion, in order to obtain the automatically calculated asymmetry (auto-A) score. This score was calculated separately for the inner and outer borders.

For calculation of the automatically calculated fading (auto-F) score, first, the difference between the area

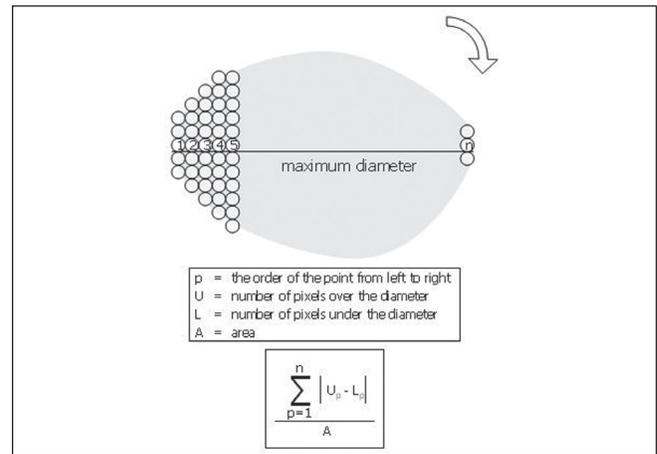


Figure 2b: In order to calculate an asymmetry score (auto - A score), first, the absolute differences between number of pixels over the diameter and number of pixels under the diameter are determined for each point on the diameter. Then, their sum is divided by the area of the shape

enclosed by the outer border and the area enclosed by the inner border was determined. Then, this difference was divided by the area encircled by the outer border. By definition, this score is unique.

Statistical Analysis

Interobserver reliability assessments for subjective scores were performed by computing quadratic-weighted kappa values using MedCalc - version 9.3.0.0, and those for automatically calculated scores, by computing intraclass correlation coefficients (ICC) (two-way mixed-effect model consistency definition) using SPSS (version 14.0). Kappa value and ICC greater than 0.80 indicated “almost perfect”, 0.61 to 0.80 “substantial”, 0.41 to 0.60 “moderate”, 0.21 to 0.40 “fair”, and 0 to 0.20 “poor” agreement.

RESULTS

Interobserver Reliability for Subjective Scores

The kappa value for sub-I score was 0.241. The experienced observer used scores from “0” to “3” at rates of 14.3, 59.0, 23.6, and 3.1%, respectively. For the inexperienced observer, these rates were 1.2, 37.3, 49.1, and 12.4%, respectively. In other words, the experienced observer used the most common score “1”, while the inexperienced observer used score “2”.

The kappa value for the sub-A score was 0.253. The experienced observer used scores from “0” to “3” at rates of 41.0, 53.4, 5.0, and 0.6%, respectively. For the inexperienced observer, these rates were 1.9, 49.7,

36.0, and 12.4%, respectively. In other words, the experienced observer used the least common score “3”, while the inexperienced observer used score “0”. More interestingly, there were lesions, in which the inexperienced observer used a higher score than the experienced observer, while there were no lesions for which the experienced observer used a higher score than the inexperienced observer.

The kappa value for sub-F score was 0.286. The experienced observer used scores from “0” to “3” at rates of 19.9, 68.3, 11.2, and 0.6%, respectively. For the inexperienced observer, these rates were 52.8, 42.9, 4.3, and 0.0%, respectively. In other words, the experienced observer used the most common score “1”, while the inexperienced observer used score “0”.

All the kappa values are given in Table 1.

Interobserver Reliability for Sizes of Lesions

According to the marking of the outer border by the experienced versus the inexperienced observers, the mean maximum diameter, the mean perimeter, and the mean area were 7.9 versus 7.6 mm, 22.3 versus 21.6 mm, and 37.1 versus 34.5 mm square, respectively. Although measurements of the inexperienced observer were somewhat smaller than those of the experienced

observer, the interobserver concordance was almost perfect. According to the outer border, ICC for the maximum diameter was 0.985; for the perimeter, 0.986, and for the area, 0.985 [Table 1].

According to marking of the inner border by the experienced versus the inexperienced observers, the mean maximum diameter, the mean perimeter, and the mean area were 7.5 versus 7.2 mm, 22.0 versus 22.0 mm, and 32.9 versus 29.7 mm square, respectively. The difference in the mean diameter was greater than 0.1 mm in 118 lesions between the observers. Although measurements of the inexperienced observer, other than the perimeter, were somewhat smaller than those of the experienced observer, the interobserver concordance was almost perfect. According to the inner border, ICC for the maximum diameter was 0.983, for the perimeter, 0.962, and for the area, 0.978 [Table 1].

Interobserver Reliability for Automatically Calculated Scores

According to the marking of the outer border by the experienced versus the inexperienced observers, the mean auto-I score and the mean auto-A score were 2.051 versus 2.059, and 0.155 versus 0.152, respectively [Table 1]. According to the marking

Table 1: Kappa and ICC values; and 95% confidence intervals between observers for the subjective scores, sizes of lesions, and the automatically calculated scores

	Mean \pm SD in measurements of		Kappa value or ICC	95% Confidence interval	P value
	Experienced observer	Inexperienced observer			
Subjective scores					
Sub-I ^a	-	-	0.241	0.126 – 0.356	< 0.001
Sub-A ^b	-	-	0.253	0.176 – 0.329	< 0.001
Sub-F ^c	-	-	0.286	0.162 – 0.409	< 0.001
Sizes					
Diameter (outer) (mm)	7.9 \pm 2.5	7.6 \pm 2.5	0.985	0.979 – 0.989	< 0.001
Diameter (inner) (mm)	7.5 \pm 2.5	7.2 \pm 2.4	0.983	0.977 – 0.988	< 0.001
Perimeter (outer) (mm)	22.3 \pm 6.8	21.6 \pm 6.8	0.986	0.981 – 0.990	< 0.001
Perimeter (inner) (mm)	22.0 \pm 7.1	22.0 \pm 7.6	0.962	0.948 – 0.972	< 0.001
Area (outer) (mm square)	37.1 \pm 20.0	34.5 \pm 19.4	0.985	0.980 – 0.989	< 0.001
Area (inner) (mm square)	32.9 \pm 18.7	29.7 \pm 17.9	0.978	0.970 – 0.984	< 0.001
Automatically calculated scores					
Auto-I ^d (outer)	2.051 \pm 0.061	2.059 \pm 0.059	0.522	0.400 – 0.626	< 0.001
Auto-I (inner)	2.128 \pm 0.128	2.218 \pm 0.163	0.490	0.363 – 0.599	< 0.001
Auto-A ^e (outer)	0.155 \pm 0.087	0.152 \pm 0.078	0.419	0.283 – 0.538	< 0.001
Auto-A (inner)	0.159 \pm 0.087	0.183 \pm 0.105	0.514	0.391 – 0.619	< 0.001
Auto-F ^f	0.125 \pm 0.078	0.160 \pm 0.109	0.357	0.214 – 0.484	< 0.001

^asubjective irregularity; ^bsubjective asymmetry; ^csubjective fading; ^dautomatically calculated irregularity; ^eautomatically calculated asymmetry; ^fautomatically calculated fading

of the inner border by the experienced versus the inexperienced observers, the mean auto-I score and the mean auto-A score were 2.128 versus 2.218, and 0.159 versus 0.183, respectively. Auto-F score was 0.125 for the experienced observer and 0.160 for the inexperienced observer. In measurements of the inexperienced observer, means of all scores other than auto-A score of the outer border were higher. Moreover, interobserver concordance was fair or moderate. ICC for the auto-I score of the outer border was 0.522, for the auto-I score of the inner border, 0.490, for the auto-A score of the outer border, 0.419, for the auto-A score of the inner border, 0.514, and for the auto-F score, 0.357 [Table 1].

DISCUSSION

The limitation of this study is that we have only one pair of observers (experienced and inexperienced), if we had more pairs of observers we might see more details of inexperienced deficiency with regard to dermoscopy. Another limitation is the small number of malignant melanoma in our study, in spite of the fact that in practice, the main use of dermoscopy is to diagnose or to exclude malignant melanoma.

According to both the outer and inner borders, the concordance between our experienced and inexperienced observers was almost perfect in measurements of diameter, perimeter, and area. This finding suggests that a person, who has not had enough experience in dermoscopy, but has enough experience in clinical dermatology, is probably able to see the pigmented lesions in the dermoscopic images in almost the same sizes as an experienced person sees them.

In subjective evaluation, the concordance between observers was fair on all scores, namely irregularity, asymmetry, and wideness of fading. These results are not surprising, as in order to evaluate the dermoscopic images properly, knowledge and experience is necessary.^[11,12] In computer-assisted evaluation, this concordance increased only slightly. This finding suggests that presumably, an inexperienced person, in comparison with an experienced person, sees lesions in same sizes, but in different shapes on dermoscopy. In this study almost perfect agreement was detected between the observers regarding the diameter of the lesion (ICC score for diameter-outer was 0.985). Not being able to recognize the small pigmented structures

such as dots may lead to failure in recognizing the shape of the lesion. Dots are defined as pigmented structures, 0.1 mm in diameter, which can be peripherally localized. We found that the difference in mean diameter was greater than 0.1 mm in 118 lesions between the observers, and the agreement for asymmetry and irregularity was moderate. We may assume that the inexperienced observer, not observing the correct shape of the lesion, was not able to classify those structures as being part of it or as not being part of it. Therefore, it is advisable that making learners familiar with borders of lesions should be included in the training on dermoscopy. There are several studies about automated dermoscopy shape feature extraction that may be helpful for such training.^[13-16]

Inexperienced observers' inability to recognize the shape of the lesion may be due to not having a formal training on dermoscopy. There are several studies on training in dermoscopy, as also which algorithmic method in dermoscopy is more useful for the inexperienced person.^[3,6,7,17,18] It is certain that training and algorithms are beneficial for all, but especially for the inexperienced physicians. As a result of our study we may suggest that in dermoscopic examination, failure of the inexperienced person might be with regard to the recognition of the shapes of the lesion, besides the recognition of the structural components. In order to evaluate the lesion properly, besides recognizing the specific characteristics of the lesion, one should also be able to recognize the gross features. This is not only necessary for pigmented lesions; by this way, one should also be able to recognize the nonpigmented lesions such as amelanotic melanoma, nonpigmented Bowen disease, and nonpigmented basal cell carcinoma, which may appear clinically as red, scaly, and ill-defined, defying accurate diagnosis.^[19]

Not only in clinical evaluation, but also in dermoscopic evaluation, border irregularity and asymmetry are usually considered to be bad signs for pigmented lesions, suggestive of either dysplastic nevi or melanomas. Abrupt cut-off is also a bad sign. Wideness of fading is an opposite variable to abrupt cut-off. In other words, if wideness of fading decreases, it should be taken as a bad sign. In our study, the inexperienced observer used higher scores for subjective evaluation of irregularity and asymmetry, although lower scores for wideness of fading. This was not a discrepancy. Instead, both attitudes were consistent, as both were

toward the aforementioned bad signs. This finding suggests that inexperienced persons might have a tendency to interpret even indistinct changes as bad.

Our study showed that in dermoscopic examination, an inexperienced observer did not have enough ability to recognize the gross features of the lesions. Irregularity, asymmetry, and wideness of fading are all important findings, and point to the malignant or benign characteristic of lesions. In our opinion, in dermoscopy training, besides teaching the structural components, to begin with, one should inform the learner about what lesion borders are and where they start and end. We also want to state that as we only evaluated melanocytic lesions, conclusions of this study could only be considered valid for the dermoscopic evaluation of melanocytic lesions and not for its use in general practice.

REFERENCES

- Braun RP, Rabinovitz HS, Oliviero M, Kopf AW, Saurat JH. Dermoscopy of pigmented skin lesions. *J Am Acad Dermatol* 2005;52:109-21.
- Zalaudek I, Argenziano G, Soyer HP, Corona R, Sera F, Blum A, *et al.* Three-point checklist of dermoscopy: an open internet study. *Br J Dermatol* 2006;154:431-7.
- Piccolo D, Ferrari A, Peris K, Diadone R, Ruggeri B, Chimenti S. 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.
- Lorentzen H, Weismann K, Petersen CS, Larsen FG, Secher L, Skødt V. Clinical and dermoscopic diagnosis of malignant melanoma. Assessed by expert and non-expert groups. *Acta Derm Venereol* 1999;79:301-4.
- Barzegari M, Ghaninezhad H, Mansoori P, Taheri A, Naraghi ZS, Asgari M. Computer-aided dermoscopy for diagnosis of melanoma. *BMC Dermatol* 2005;6:5-8.
- Dolianitis C, Kelly J, Wolfe R, Simpson P. Comparative performance of 4 dermoscopic algorithms by nonexperts for the diagnosis of melanocytic lesions. *Arch Dermatol* 2005;141:1008-14.
- Pagnanelli G, Soyer HP, Argenziano G, Talamini R, Barbati R, Bianchi L, *et al.* Diagnosis of pigmented skin lesions by dermoscopy: web-based training improves diagnostic performance of non-experts. *Br J Dermatol* 2003;148:698-702.
- Carli P, Quercioli E, Sestini S, Stante M, Ricci L, Brunasso G, *et al.* Pattern analysis, not simplified algorithms, is the most reliable method for teaching dermoscopy for melanoma diagnosis to residents in dermatology. *Br J Dermatol* 2003;148:981-4.
- Braun RP, Rabinovitz H, Oliviero M, Kopf AW, Saurat JH, Thomas L. Dermoscopy of pigmented lesions. *Ann Dermatol Venereol* 2002;129:187-202.
- Nachbar F, Stolz W, Merkle T, Cognetta AB, Vogt T, Landthaler M, *et al.* The ABCD rule of dermoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *J Am Acad Dermatol* 1994;30:551-9.
- Troyanova P. A beneficial effect of a short-term formal training course in epiluminescence microscopy on the diagnostic performance of dermatologists about cutaneous malignant melanoma. *Skin Res Technol* 2003;9:269-73.
- 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;131:286-91.
- Day GR. How Blurry is that Border? An Investigation into Algorithmic Reproduction of Skin Lesion Border Cut-off. *Comput Med Imaging Graph* 2000;24:69-72.
- Lee TK, McLean DI, Atkins MS. Irregularity Index: A New Border Irregularity Measure for Cutaneous Melanocytic Lesions. *Med Image Anal* 2003;7:47-64.
- Schmid-Saugeon P. Symmetry Axis Computation for Almost-Symmetrical and Asymmetrical Objects: Application to Pigmented Skin Lesions. *Med Image Anal* 2000;4:269-82.
- Celebi ME, Kingravi HA, Uddin B, Iyatomi H, Aslandogan YA, Stoecker WV, *et al.* A Methodological Approach to the Classification of Dermoscopy Images. *Comput Med Imaging Graph* 2007;31:362-73.
- Pellacani G, Grana C, Seidenari S. Algorithmic reproduction of asymmetry and border cut-off parameters according to the ABCD rule for dermoscopy. *J Eur Acad Dermatol Venereol* 2006;20:1214-9.
- Argenziano G, Soyer HP, Chimenti S, Talamini R, Corona R, Sera F, *et al.* Dermoscopy of Pigmented Skin Lesions: Results of a Consensus Meeting via the Internet. *J Am Acad Dermatol* 2003;48:679-93.
- Seidenari S, Pellacani G, Grana C. Asymmetry in Dermoscopic Melanocytic Lesion Images: a Computer Description Based on Colour Distribution. *Acta Derm Venereol* 2006;86:123-8.