Acta Dermatovenerol Croat 2014;22(4):278-283 CLINICAL ARTICLE

Pitfalls of an Automated Dermoscopic Analysis System in the Differential Diagnosis of Melanocytic Lesions

Dilek Biyik Ozkaya[1], Nahide Onsun[1], Ozlem Su[1], Hande Arda Ulusal[2], Serpil Pirmit[3]

[1]Bezmialem University, Medical Faculty, Dermatovenerology Department; [2]Hizmet Memorial Hospital, Dermatovenerology Department; [3]Arnavutkoy State Hospital, Istanbul, Turkey

Corresponding autor:
Assist. Prof. Dilek Biyik Ozakaya, MD
Bezmialem University, Medical Faculty,
Dermatovenerology Department
Adnan Menderes Bulvari Vatan Cad. 34093
Fatih/Istanbul
Turkey
dilekozkaya@gmail.com

Received: September 13, 2013
Accepted: August 1, 2014

SUMMARY Dermoscopy plays an important role in the diagnosis of pigmented lesions, particularly in the differential diagnosis of early-stage melanoma. Dermoscopy systems that aim to enable automatic "unmanned-without physician" diagnosis are becoming increasingly common. We aimed to investigate the reliability and weaknesses of diagnosis programs. Furthermore, we attempted to determine whether such programs are superior to diagnosis by a physician, compared to histopathological assessment. The images stored in the DermoGenius ultra-computerized dermoscopy system of the Dermoscopy Unit between January 2008 and December 2008 were surveyed retrospectively. Dermoscopic images made prior to excision of 77 lesions from 51 patients verified by histopathology were reviewed. Nineteen patients were men and 32 were women. Mean age was 35.5 years. Diagnosis by a clinician or automatic analysis revealed that 23 (30%) of the lesions were atypical (dysplastic) nevi, 22 (29%) were compound nevi, 10 (13%) were dermal nevi, 8 (10%) were malignant melanomas, 7 (9%) were common nevi, 6 (7%) were junctional nevi, and 1 (1%) was a blue nevus. Compared to histopathological diagnosis, considered the gold standard, the sensitivity of the automated analysis program was 96.6%, its specificity 14.9%, and its diagnostic accuracy 47%. For the clinician, the values were 100% for sensitivity, 66.7% for specificity, and 95% for diagnostic accuracy.
Based on histopathological results, the diagnostic accuracy of the physician was higher than that of the automatic analysis program. Therefore, errors are inevitable when an inexperienced physician assesses patients according to automatic program results.

KEYWORDS: computerized dermoscopy; diagnosis; melanocytic lesions; reliability

INTRODUCTION

Dermoscopy is used by dermatologists in the diagnosis and differential diagnosis of pigmented skin lesions of which diagnosis is normally difficult, as well as for clinical assessment before histopathological examination (1-5). Due to difficulties in diagnosis, many unnecessary excisions are made; nevertheless, the diagnosis of some melanomas is delayed. Dermoscopy, used with the specific criteria necessary for making a good assessment together with practical experience, is an important tool for the detection of malignancies

Acta Dermatovenerol Croat
2014;22(4):278-283

with the characteristics of a benign lesion that can easily go unnoticed in the early phase. Additionally, it can be used both to detect lesions that may be excised unnecessarily and in the follow-up of patients at risk of dysplastic nevus syndrome who have many pigmented lesions (6).

Recent developments in computer technology have increased the expectations of dermoscopy. Attempts are ongoing to create programs that will enable diagnosis of cutaneous melanoma without the need for human input by means of computer analysis.

The aim of this study was to investigate the reliability of the computer program by comparing it with assessment by a clinician according to the ABCD criteria of patients whose nevi were excised at the Outpatient Clinic for Pigmented Lesions in our institution.

MATERIALS AND METHODS

In this study, the records of patients referred to the Outpatient Clinic for Pigmented Lesions of the Clinic of Dermatological and Venereal Diseases of the Vakıf Gureba Research and Training Hospital (later renamed the Medical Faculty of Bezmi Alem University) in 2008 whose nevi were excised were retrospectively studied within the period September 1 to September 30, 2009 by computer analysis. Dermoscopic data were surveyed using the computerized DermoGenius ultra (Rodenstock Prazisionsoptik, Linos Photonics GmbH & Co., Munich, Germany).

Dermoscopic images included in the study were taken as part of routine procedure during the clinical examination, and were stored after verbal approval had been obtained. Written consent was obtained from the patients by informing them before a skin biopsy was taken to confirm a suspicion or diagnosis of malignant lesions. No biopsy was taken from patients who did not provide consent. Seventy-seven lesions from 51 patients from which melanocytic lesions were excised were included in the study. Those patients with malignant melanoma but without dermoscopic images were excluded. Patient data, including age, gender, skin type, work environment, exposure to sun, use of a sun block, age of the lesion, dimension of the lesion, clinical type of the lesion, and localization of the lesion were obtained from the records of the Dermoscopy Unit.

Nevi were clinically examined using the ABCD criteria. The lesion is assessed according to its asymmetry (A), border (B), color (C) and components (D). Each criterion is multiplied by a predetermined coefficient. The result obtained is the total dermoscopy score (TDS). A TDS ≤4.75 is considered benign, 4.75 to 5.45 suspicious, and ≥5.45 malignant.

In the DermoGenius ultra system, a value termed the digital standardized dermatoscopic point (DSDP) accuracy is computed by means of variables calculated using the classical ABCD rule. DSDP accuracy lies between 3 and 6. When this accuracy is compared with a histological database, it is possible to compute the probability of a lesion being malignant or benign. This accuracy is presented as: 1) DSDP numerical value; 2) representation of the DSDP value in bar chart form; 3) comparison of the DSDP value with the database. By virtue of the normalization of the picture, the skin color is eliminated, and hence, the contrast becomes more apparent, allowing better observation of the structural elements. For assessment of ABCD, eight criteria are used: 1) color asymmetry; 2) shape asymmetry; 3) border; 4) color variability; 5) color homogeneity; 6) structural alterations; 7) structural asymmetry; 8) structural homogeneity (7). SPSS version 15.0 (IBM, Armonk, New York, USA) was used for data analysis and the chi-square test was applied for comparisons. A *P* value of <0.05 was considered significant. Sensitivity was defined as the ability of the test to accurately identify malignant lesions. Specificity was defined as the ability of the test to accurately identify benign lesions. Diagnostic accuracy (percentage of correct diagnoses established) was defined as the probability of accurate discrimination of malignant and benign lesions.

RESULTS

Of the 51 patients, 19 (37.3%) were men and 32 (62.7%) were women. The ages of the patients varied between 12 and 85 years, with a mean age of 35.5 years. The most common lesion location in men and women was the trunk (72%). In male patients, 82% of the lesions were on the trunk, 9% on the upper extremity, 6% on the head, and 3% on the lower extremity. In female patients, 66% of the lesions were on the trunk, 25% on the upper extremity, 7% on the head, and 2% on the lower extremity. When the skin types of the male and female patients included in our study were classified according to the Fitzpatrick Scale, the most common was type 3 in 26 patients (51%), followed by type 2 in 24 patients (47%), and type 1 in

Table 1. Comparison of computer diagnoses and diagnoses by clinicians

Diagnosis by		N	%
Computer	Malignant	48	62
	Suspicious	4	6
	Benign	25	32
Clinician	Malignant	12	13
	Suspicious	8	10
	Benign	57	77

Acta Dermatovenerol Croat
2014;22(4):278-283

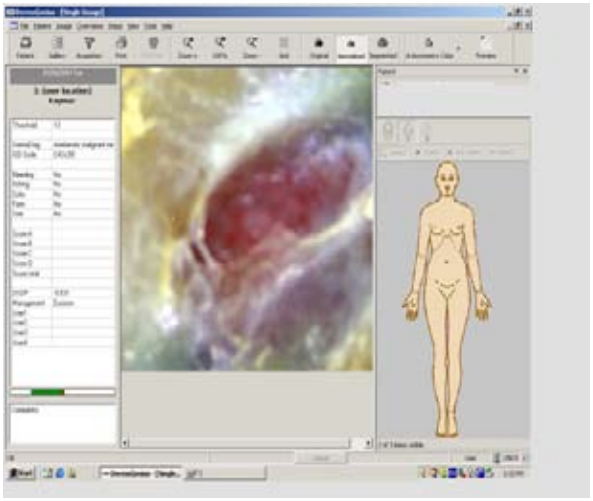

Figure 1. Melanoma diagnosed by the computer as a benign lesion.

one patient (2%). In terms of lesion histopathology, 23 (30%) were found to be atypical (dysplastic) nevi, 22 (29%) were compound nevi, 10 (13%) were dermal nevi, 8 (10%) were malignant melanomas, 7 (9%) were common nevi, 6 (7%) were junctional nevi, and one (1%) was a blue nevus. When considering the atypical nevus group suspicious, 46 (60%) of the excised nevi were found to be benign, 23 (30%) were suspicious, and 8 (10%) were malignant. The most common location of a dysplastic nevus was the trunk (68%), followed in descending frequency by the lower extremity (18%), the upper extremity (9%), and the scalp (4%). The youngest patient with a dysplastic nevus was 12 years old, while the oldest was 74 years old, the mean age being 34 years. While the computer assessed 48 (62%) lesions as malignant, 4 (6%) as suspicious, and 25 (32%) as benign, the clinician assessed 12 (13%) lesions as malignant, 8 (10%) as suspicious, and 57 (77%) as benign (Table 1).

Table 2. Comparison of computer diagnoses with histopathology as the gold standard.

P=0.140		Computer		Total
		Benign	Malignant	
Pathology	Benign	29	40	69
	Malignant	42.0%	58.0%	100%
		96.7%	85.1%	89.6%
		1	7	8
		12.5%	87.5%	100%
		3.3%	14.9%	10.4%
Total		30	47	77
		39.0%	61.0%	100%
		100%	100%	100%

When computer analysis and the analysis of the clinician were compared with histopathological results as the gold standard, the assessments made by the clinician were found to be more accurate than the computer in terms of sensitivity, specificity, and diagnostic accuracy (Tables 2, 3).

Eight melanocytic lesions were diagnosed pathologically as malignant melanomas. Three of these were from the same patient, with one diagnosed as a nodular malignant melanoma and the other two as malignant melanoma metastases. The computer designated one of the melanoma lesions as green (DSDP: –0.533), which implied that this was a non-malignant melanoma (Table 4). The histopathological diagnosis of the melanoma the computer missed indicated that it was an amelanotic malignant melanoma (Fig. 1). When the eight melanocytic lesions were assessed clinically using ABCD, no melanoma was missed.

One patient diagnosed with melanoma at another center was referred to our clinic for a follow-up assessment of their atypical nevus; both the computer

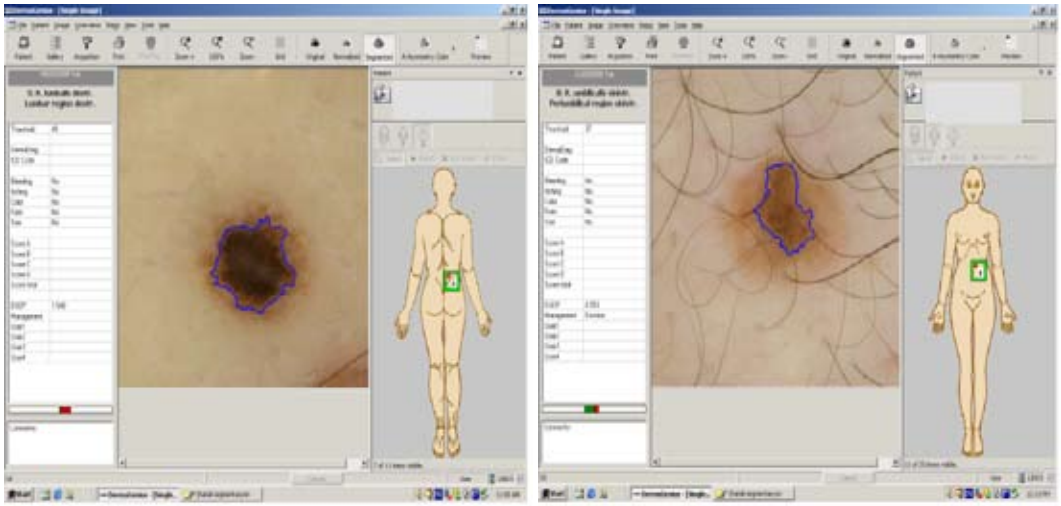

Figure 2. Incorrect segmentation of dysplastic nevi.

Table 3. Comparison of diagnoses by clinicians with histopathology as the gold standard

P<0.001		ABCD		Total
		Benign	Malignant	
Pathology	Benign	65	4	69
	Malignant	94.2%	5.8%	100%
		100%	33.3%	89.6%
		0	8	8
		0%	100%	100%
		0%	66.7%	10.4%
Total		65	12	77
		84.4%	15.6%	100%
		100%	100%	100%

and the clinician assessed it as benign. The lesions of three patients with a history of malignant melanoma whose lesions were found histopathologically to be atypical nevi were assessed as benign by both the computer and the clinician. Among the 48 lesions identified as malignant by the computer, 7 were classified as malignant melanomas, 18 as dysplastic nevi, 10 as compound nevi, 7 as dermal nevi, and 1 as a junctional nevus. Of the 12 lesions identified as malignant by the clinician using the ABCD method, 8 were classified as malignant melanomas, 3 as dysplastic nevi, and 1 as a compound nevus. The errors in diagnosis by the automatic analysis system originated from erroneous segmentation (inability to determine lesion boundaries) and insufficiencies in the assessment of hairy regions. This clearly indicated the need to shave hairy regions before assessment (Fig. 2 a, b).

In a dermal nevus found on the scalp, the computer segmented only the hair located in the middle of the nevus. The lesion was assessed as high-risk while the hair was present but as risk-free after shaving (Fig. 3 a, b).

The results of the automatic analysis system were negatively affected upon addition of fluid. Additionally, no assessment of large nevi could be made. Two nevi in close proximity were segmented as a single nevus (Figure 4).

DISCUSSION

The ABCD dermoscopy rules were defined by Stolz *et al.* in 1994. This system is based on categorization of melanocytic lesions into three groups; benign, malignant, and suspicious (8). Lesions are assessed according to the following criteria: asymmetry (A); border (B); colour (C); and components (D) (9,10). Each criterion is then multiplied by a predefined coefficient. The result obtained is termed the total dermoscopy score (TDS). A TDS ≤4.75 is considered benign, a TDS between 4.75 and 5.45 suspicious, and a TDS ≥5.45 malignant. This method is learned more easily than pattern analysis by clinicians who are not experienced in dermoscopy (10).

Robert *et al.* did a study to evaluate the performance of dermoscopists in diagnosing small pigmented skin lesions (diameter <6mm) compared with an automatic multispectral computer-vision system. In their results, dermoscopists were able to correctly identify small melanomas with an average diagnostic sensitivity of 39% and specificity of 82%, and recommended small melanomas for biopsy with sensitivity of 71% and specificity of 49%, with only fair inter-observer agreement (k=0.31 for diagnosis and 0.34 for biopsy). In comparison, in recommending biopsy to rule out melanoma, the computer-vision system achieved 98% sensitivity and 44% specificity. In their conclusion, they said computer-vision systems can facilitate early detection of small melanomas and may limit the number of biopsies to rule out melanoma and performed on benign lesions (11). Ryan Wells *et al.* performed a study about comparison of diagnostic and management sensitivity to melanoma between dermatologists and MelaFind. Estimated biopsy sensitivity was 22 out of 23 (0.96;95% LCB, 0.83) for MelaFind, and ranged from 0.48 to 1.00 for the

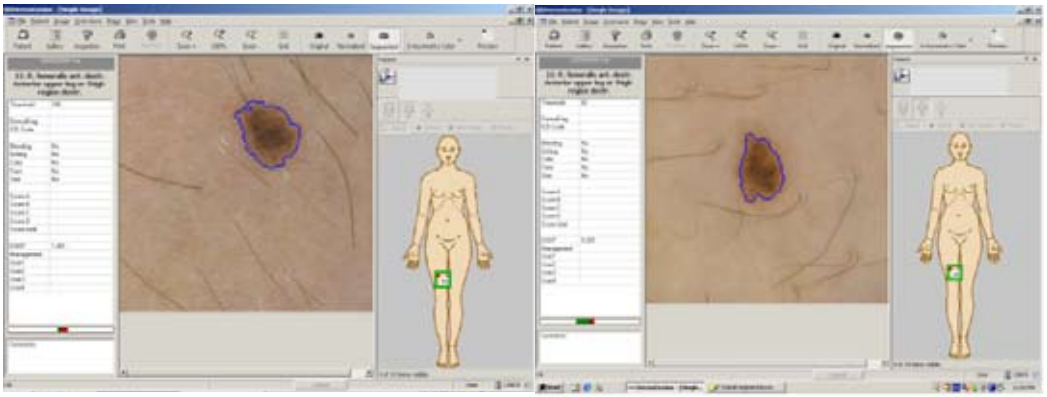

Figure 3. Lesions were examined with or without hair present.

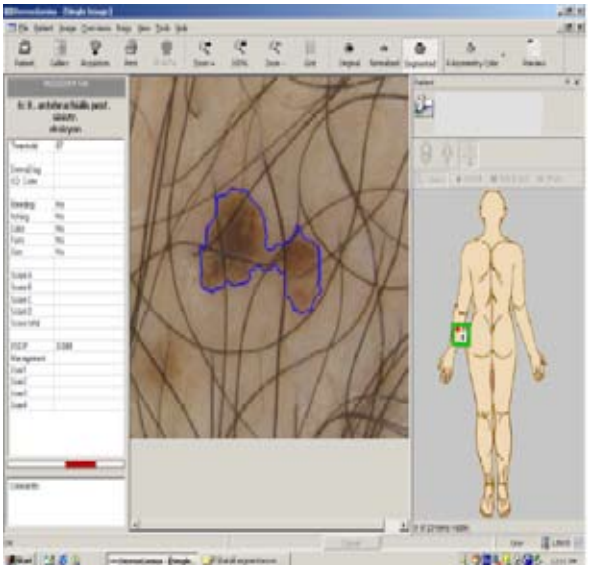

Figure 4. Two nevi in close proximity were segmented as a single nevus.

dermatologist. Average biopsy sensitivity of the dermatologist among the 23 melanomas was 0.80 (95% CI, 0.72-0.87).Estimated biopsy specificity was 2 of 24 (0.08;95% CI, 0.01-0.25) for MelaFind and ranged from 0.04 to 0.71 for the dermatologist. Average biopsy specificity of study dermatologists was 0.43. In their study, MelaFind performed with a high sensitivity but a low specificity in recommending biopsy for melanomas (12).

Morales *et al.* reported results of a study involving 200 excised dysplastic nevi from 166 patients with a mean age of 33.7 years. In our study, the mean age of patients with a dysplastic nevus was 34 years (13). Seidenari *et al.* compared diagnosis by a dermatologist trained in dermoscopy, one not trained in dermoscopy, and a computer (14). The computer was found to be significantly superior in terms of sensitivity. In terms of specificity, the computer was equal to the experienced dermatologist, while it was significantly superior to the inexperienced dermatologist. While the inexperienced dermatologist assessed six melanomas as benign, the computer diagnosed only two melanomas incorrectly. Sboner *et al.* investigated whether use of a computerized system facilitated diagnosis by dermatologists (15). Eight dermatologists assessed the images on screen without computer support, and later reassessed these images with computer support. An 11% increase in sensitivity and a 6% decrease in specificity were found when assisted by computers. The main reason for the increased sensitivity was the computer, with the dermatologists establishing correct diagnosis of different melanomas. In the study of Bauer *et al.,* a clinician made no errors in differentiation of malignant and benign lesions when using a computer, although errors had previously been made by the clinician unassisted by the computer system (16). These nevi were re-investigated using pathology as the gold standard. More significant results were achieved when assessments were made with computer assistance; thus the authors concluded that assessment with computer assistance was superior (16). Callaghan *et al.* investigated the clinical dermoscopic and histopathological correlations of 200 atypical nevi from 166 patients (13). They found that atypical characteristics could not be determined by the naked eye and that dermoscopy facilitated diagnosis. The authors claimed that symmetry did not represent a significant difference between atypical and non-atypical nevi, and that border irregularities and color multiplicity were more compatible with atypical histopathology (13). Cristofolini *et al.* reported that assessment by a clinician was superior to the computer in terms of sensitivity and specificity (17). In the studies by Bono *et al.* (18) and Horsch et al. (19), the sensitivity and specificity of assessment by the clinician were statistically significantly superior to that by the computer system.

In our study, the results generated by the computer were compared to those by the clinician using ABCD, with histopathology as the gold standard. For the computer the values obtained were: sensitivity 96.6%; specificity 14.9%; and diagnostic accuracy (percentage of establishing correct diagnosis) 47%. The clinician assessment provided a sensitivity of 100%; specificity of 66.66%; and diagnostic accuracy of 95%. Therefore, assessment by the clinician was significantly superior in all three categories. The errors made by the computer system are discussed below.

The computer tended to generate erroneous results for nevi on which hair was present or around which hair was growing. Seidenari *et al.* reported that hairy areas should be examined by a computer system after shaving (14). In larger lesions, erroneous results can arise due to the computer system being able to generate segmentation of only a limited area. If two lesions are in close proximity, the computer may assess these as a single nevus. If no fluid is applied before dermoscopy or if an air bubble is formed on the lesion during the procedure, the lesion may be assessed as risky; however, when the same lesion is assessed using bubble-free fluid it will likely be classified as benign. Gewirtzman *et al.* investigated the fluids and application techniques used in dermoscopy, and concluded that the "roll-on" technique resulted in the minimum false assessments (artifacts), independently of the type of fluid applied (20). This study included fluids that inhibit formation of air bubbles during application.

CONCLUSIONS

The inability of the computer system to identify non-melanocytic pigmented lesions constitutes a limitation of this study. Pigmented seborrheic keratosis can be missed even by expert dermatologists. The absence of these forms of lesion led to a suboptimal assessment of the computer program.

Investigation of lesions by computerized dermoscopy has many advantages. The stable and homogeneous light source and three CCD video cameras resulted in higher quality color and the ability to visualize small lesions in greater detail. During follow-up of patients with many nevi, it was possible to compare the colors, structural elements, and dimensions with the previous status using archived images, thus reducing the number of unnecessary excisions. Archived images also facilitate training of clinicians who wish to gain experience with this technique. Computerized dermoscopy will allow consultation between centers with respect to suspicious lesions. Although we believe that computerized dermoscopy offers many benefits, it is, however, not itself sufficient for establishment of a diagnosis.

References

1. Argenziano G, Soyer HP, Chimenti S, Argenziano G, Ruocco V. Impact of dermoscopy on the clinical management of pigmented skin lesions. Clin Dermatol 2002;20:200-2.
2. Grin CM, Friedman KP, Grant-Kels JM. Dermoscopy: a review. Dermatol Clin 2002;20:641-6.
3. Soyer HP, Argenziano G, Chimenti S, Ruocco V. Dermoscopy of pigmented skin lesions. Eur J Dermatol 2001;11:270-6.
4. Onsun N, Atılganoglu U. Dermoskopi Atlası, birinci baskı, 5G Matbaacılık Ltd.Sti., İstanbul 1997.
5. Braun RP, Rabinovitz HS, Oliviero M, Kopf AW, Saurat JH. Dermoscopy of pigmented skin lesions. J Am Acad Dermatol 2005;52:109-21.
6. Soyer HP, Argenziano G, Talamini R, Chimenti S. Is dermoscopy useful for the diagnosis of melanoma? Arch Dermatol 2001;137:1343-50.
7. Dermogenius Linos systems Kullanım Kılavuzu.
8. Braun PR, Rabinovitz SH, Oliviero M, Kopf WA, Saurat HJ, Thomas L. Dermoscopic Examination. In: Soyer HP, Argenziano G, Hofmann-Wellenhof R, Johr R, eds. Color Atlas of Melanocytic Lesions of the Skin. Berlin:Springer-Verlag, 2007. p. 3:19.
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.
10. Nachbar F, Stolz W, Merkle T, Cogneta AB, Vogt T, Lanthaler M, *et al*. The ABCD rule of dermatoscopy. J Am Acad Dermatol 1994, 551-9.
11. Friedman RJ, Gutkowicz-Krusin D, Farber MJ, Warycha M, Schneider-Kels L, Papastathis N, *et al*. The diagnostic performance of expert dermoscopist vs a computer –vision system on small-diamater melanomas. Arch Dermatol 2008;144:476-82.
12. Wells R, Gutkowicz-Krusin D, Veledar E, Toledano A, Chen SC. Comparison of diagnostic and management sensitivity to melanoma between dermatologist and MelaFind: A pilot study. Arch Dermatol 2012 ;148:1083-4
13. Callaghan AM, Sanz J, Garcia G, Martinez PI. Correlation between clinical, dermatoskopic, and histopathologic variables in atypical melanocytic nevi. Actas Dermosifiliogr 2008;99:380-9.
14. Seidenari S, Pellacani G, Pepe P. Digital videomicroscopy improves diagnostic accuracy for melanoma. J Am Acad Dermatol 1998;39:175-81.
15. Sboner A, Bauer P, Zumiani G, Eccher C, Blanzieri E, Forti S, *et al*. Clinical validation of an automated system for supporting the early diagnosis of melanoma. Skin Res Technol 2004;10:184-92.
16. Bauer P, Cristofolini P, Boi S, Burroni M, Dell'Eva G, Micciolo R, *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.
17. Cristofolini M, Bauer P, Boi S, Micciolo R, Sicher M. Diagnosis of cutaneous melanoma: accuracy of a computerized images analysis system (Skin View). Skin Res Technol 1997;3:23-7.
18. Bono A, Tomatis S, Bartoli C, Cascinelli N, Clemente C, Cuperta C, *et al*. The invisible colours of melanoma: a telespectrophotometric diagnosis approach on pigmented skin lesions. Eur J Cancer 1996;32:727-9.
19. Horsch A, Stolz W, Neiss A, Abamayr W, Pompl R, Bernklav V, *et al*. Improving early recognition of malignantantant melanomas by digital image analysis in dermatoscopy. Stud Health Technol Inform 1997;43:531-5.
20. Gewirtzman AJ, Saurat JH, Braun RP. An evaluation of dermoscopy fluids and application techniques. Br J Dermatol 2003;149:59-63.