Featured Article

Melanoma Computer-Aided Diagnosis: Reliability and Feasibility Study

Marco Burroni,[1] Rosamaria Corona,[2] Giordana Dell'Eva,[3] Francesco Sera,[2] Riccardo Bono,[2] Pietro Puddu,[2] Roberto Perotti,[1] Franco Nobile,[3] Lucio Andreassi,[1] and Pietro Rubegni[1]

[1]Department of Dermatology, University of Siena, Italy; [2]Istituto Dermopatico dell'Immacolata, Rome, Italy; and [3]Italian Cancer League, Siena, Italy

Abstract

Background: **Differential diagnosis of melanoma from melanocytic nevi is often not straightforward. Thus, a growing interest has developed in the last decade in the automated analysis of digitized images obtained by epiluminescence microscopy techniques to assist clinicians in differentiating early melanoma from benign skin lesions.**

Purpose: **The aim of this study was to evaluate diagnostic accuracy provided by different statistical classifiers on a large set of pigmented skin lesions grabbed by four digital analyzers located in two different dermatological units.**

Experimental Design: **Images of 391melanomas and 449 melanocytic nevi were included in the study. A linear classifier was built by using the method of receiver operating characteristic curves to identify a threshold value for a fixed sensitivity of 95%. A K-nearest-neighbor classifier, a nonparametric method of pattern recognition, was constructed using all available image features and trained for a sensitivity of 98% on a large exemplar set of lesions.**

Results: **On independent test sets of lesions, the linear classifier and the K-nearest-neighbor classifier produced a mean sensitivity of 95% and 98% and a mean specificity of 78% and of 79%, respectively.**

Conclusions: **In conclusion, our study suggests that computer-aided differentiation of melanoma from benign pigmented lesions obtained with DB-Mips is feasible and, above all, reliable. In fact, the same instrumentations used in different units provided similar diagnostic accuracy. Whether this would improve early diagnosis of melanoma and/or reducing unnecessary surgery needs to be demonstrated by a randomized clinical trial.**

Received 8/14/03; revised 12/19/03; accepted 12/19/03.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Notes: Drs. Burroni and Corona contributed equally to this work.
Requests for reprints: Pietro Rubegni, Istituto di Scienze Dermatologiche, Università degli Studi di Siena, Policlinico "Le Scotte," 53100 Siena, Italy. Phone: 577-40190; Fax: 577-44238; E-mail: rubegni@unisi.it.

Introduction

The most effective management of malignant melanoma is early recognition and surgical excision of thin lesions (1), because tumor thickness is universally recognized as the primary determinant of prognosis. Despite the increasing awareness of melanoma, because of the worldwide increase of incidence reported in the last few decades (2), clinical diagnostic accuracy is still disappointing (3–9). Subsequent attempts to develop noninvasive tools to improve early diagnosis resulted in two approaches: epiluminescence microscopy (ELM) and digital image analysis. ELM, first described in 1987 (10), allows the examination of skin lesions with an incident light magnification system with oil at the skin-lens interface, increases to a great extent the lesion morphological detail, and has been reported to improve the accuracy in diagnosing cutaneous lesions, including melanoma (11). Ascierto *et al.* (12) recently compared data from patient histories and clinical evaluations with ELM-based morphological patterns to characterize skin lesions and minimize interpretation problems. From these comparisons, they proposed new guidelines for the management of pigmented skin lesions (PSL) to provide standard diagnostic and therapeutic approaches and to enhance the early identification of lesions at risk for malignant transformation. However, dermoscopic techniques require formal training and skill in image interpretation through the so-called pattern analysis (13), are highly dependent on subjective judgement, and are scarcely reproducible (14, 15). Several scoring systems and algorithms such as the ABCD rule for epiluminescence, the seven-point checklist, and the Menzies method (16, 17) have been proposed to improve the diagnostic performance of less experienced clinicians. This simplification has enabled the development of these diagnostic algorithms with good accuracy and reproducibility. However, they showed problems that have not yet been solved. The most important is that the purpose for which they were designed was not achieved, because the within- and between-observer concordance is very low, even for expert observers (18–25).

Digital image analysis has been found to produce objective, reliable descriptions of melanocytic lesions. Hence, a considerable interest has arisen in recent years in the development of computer-assisted, automated analysis of digitized dermoscopic images since the first study by Schindewolf *et al.* (26–34).

The aim of this study was to evaluate diagnostic accuracy provided by different statistical classifiers on a large set of pigmented skin lesions (melanomas and nevi) grabbed by four digital analyzers (DB-Mips) located in two different dermatological units.

Materials and Methods

Instrumentation and Image Acquisition

The DB-Mips System consists of a 3CCD PAL video camera with 750 lines of image resolution and 60-decibel signal noise ratio. The camera, operating in the visible spectrum, is connected to a patented hand-held optic system yielding dermo-

scopic images with a magnification power ranging from ×6 to ×40 allowing a horizontal field of view from 40 to 6 mm. The light is provided by a 3200′k source and is homogeneously distributed on the analyzing surface at all magnifications. The three separate components of the Broadcast video signal (768 × 576) are connected to an high-quality frame grabber (set to 768 × 576 resolution and 24 bit/pixel color depth) placed inside a PC. Removable magneto-optical disks (640 MB) are used for image storage.

Image Segmentation and Feature Extraction

The choice of the most useful features to extract from digital images depends on the results of epiluminescence pattern analysis. Although the system saves the microscope magnifications along with the texture analysis, offering an objective evaluation, the different magnifications could confuse clinicians wanting to make subjective comparisons of lesions. The system used a procedure for digital image processing based on the Laplacian filter for segmentation and a zero-crossing algorithm for the border automatic outline (26). It then evaluated 49 parameters for discriminant power. Reproducibility was first tested on digitized images of 100 lesions belonging to 20 subjects (one PSL for each patient recorded five times at 15-min intervals). Absolute differences between single measurements and mean values of a given lesion or parameter never exceeded 5% of the mean value. The parameters, as described previously (19), belonged to four categories: geometries; colors; textures; and islands of color (*i.e.,* color clusters inside the lesion). In brief, the geometric variables were: area; maximum and minimum diameters; radius; variance of contour symmetry; circularity; fractality of borders; and ellipsoidality. Color variables were: mean values of red, green, and blue inside the lesion; mean values of red, green, and blue of healthy skin around the lesion; deciles of red, green, and blue inside the lesion; quartiles of red, green, and blue inside the lesion, mean skin-lesion gradient, variance of border gradient, border homogeneity, and border interruptions. Texture variables were: mean contrast and entropy of lesion; and contrast and entropy fractality. The islands of color variables were: peripheral dark regions; dark area; imbalance of dark region; imbalance green area; red area; dominant green region imbalance; blue-gray area; blue-gray regions; transition area; transition region imbalance; background area; background regions imbalance; red, green, and blue multicomponent; and number of red, green, and blue percentiles inside the lesion.

The system evaluates the above variables and gives the diagnostic probability (in real time during clinical examination) at a rate of 24 checks/s.

Image Databases

The pigmented lesions were selected from the image databases of the Department of Dermatology of the University of Siena, Italy, and the Istituto Dermopatico dell'Immacolata (IDI), a research hospital for skin diseases in Rome, Italy. The selected lesions included all melanomas undergoing ELM examination before surgical excision at the Rome and Siena centers (n = 372) between 1999 and 2003 and a random sample of 449 surgically removed benign melanocytic lesions (with available histological diagnosis), including 85 histologically atypical nevi (architectural disorder and melanocytic atypia). All of the PSL were flat and impalpable. Out of a total of 372 melanomas, 70 (19%) were *in situ* and 178 (48%) were early melanomas with Breslow thickness ≤0.75 mm.

Linear Discriminant Classifier

Feature Selection. The selection of features for the linear classifier was performed on the entire set of histologically diagnosed images. As a first step in the selection process, the Pearson correlation coefficient was calculated for each possible pair of features. Among groups of highly correlated parameters ($r > 0.9$) with similar morphological meaning, we selected the one with the best discriminating power. As a second step, we only retained for the final analysis features for which there was a significant (t test) difference between melanomas and nonmelanomas and, within these diagnostic classes, no significant difference between centers. This left 10 parameters.

Lesion Classification. For lesion classification, a discriminant analysis approach was used (35). Starting with the two classes of lesions, *i.e.,* melanomas and melanocytic nevi, we calculated a score

$$z = a_0 + \sum_{i=1}^{10} a_i X_i$$

for each lesion, where the weights a were obtained maximizing the distance between the means of melanomas and of nevi in the training set in unidimensional space with standardized variability.

Three classifiers were designed using the training samples of the Siena and Rome centers: (a) training set of IDI, Rome; (b) training set of Siena University; and (c) pool of Rome and Siena training sets. Their accuracies were measured in terms of their performance on the test samples. The discriminant linear function was calculated for each lesion in the test set, and the lesion was assigned to the melanoma group if the value was above the threshold value. This classification was then compared with the histopathological classification (gold standard), and classifier performance was measured as sensitivity and specificity. The method of receiver operating characteristic curves (36) was used to identify the threshold value for a fixed sensitivity of 95%.

K-Nearest-Neighbor Classifier

The K-nearest-neighbor classifier (37) is a nonparametric method of pattern recognition used to determine the class of an object by its features vector. For a lesion belonging to the test set (query vector), it finds the K vectors closest to the query vector in the training set. The unclassified sample is then assigned to the class represented by the majority of the K closest neighbors. This method uses the nonparametric Bayes decision rule, which does not require prior knowledge of the distribution but instead relies on a training set of objects with known class membership to make decisions on the membership of an unknown object. In other words, it assigns an unknown object to the class with the highest *a posteriori* probability, using an Euclidean metric. *A posteriori* probabilities are computed after estimating class-conditional densities. Accurate estimates of class-conditional densities require a large volume of training data. If there are

Table 1 Description of the image databases of histologically diagnosed pigmented lesions of the skin

		Melanocytic nevi			Melanomas				
		Total	Common	Dysplastic	Total	*In situ*	Thickness ≤0.75	Thickness >0.75	n.a.[a]
IDI, Rome	Training set	99	83	16	77	10	37	23	7
	Test set	92	75	17	78	9	38	25	6
	Total	191	158	33	155	19	75	48	13
Siena University	Training set	121	92	29	107	24	54	17	12
	Test set	137	114	23	110	27	49	24	10
	Total	258	206	52	217	51	103	41	22

[a] n.a., not available.

enough members in the training set, the probability of error for the K-nearest-neighbor classifier is sufficiently close to the Bayes (optimal) probability of error.

Training Set. The training set for the K-nearest-neighbor classifier consisted of 1081 histologically diagnosed skin lesions, including 428 (40%) melanomas and 653 benign pigmented lesions, 58 of which were atypical nevi. These lesions were selected from the image databases of several institutions using the same ELM instrumentation, *i.e.*, IDI-Rome; Siena University Dermatology Clinic; IDI-Capranica; and the Italian Cancer League Clinics of Grosseto, Livorno, Arezzo, Trento, and Siena.

Lesion Classification. K-nearest-neighbor classifier accuracy was estimated on the histologically diagnosed lesions of the Rome and Siena centers by the jackknife procedure. The prevalence of melanomas among the first 100 closest neighbors was determined, and the lesion was assigned to the melanoma group if the prevalence was higher than a threshold value T_{100}. The method of receiver operating characteristic curves was used to identify the T_{100} value necessary for a sensitivity of 98%.

Results

Description of the Image Sets

Table 1 shows the distribution of the lesions included in the training and test sets by histological diagnosis and center. The Siena University image database included a higher proportion of *in situ* melanomas. For each center, the random allocation of lesions to the training and test sets yielded two groups of lesions including melanomas of similar thickness and the same proportion of *in situ* melanomas and atypical nevi.

Description of the Features Selected for the Linear Discriminant Classifier

The means and SDs of the 10 selected parameters by diagnosis are reported in Table 2. For all parameters, a clear-cut difference between melanomas and nevi was observed. As an additional check of the feature selection procedure, a logistic regression model including all of the selected features was run to assess the association of each feature with melanoma, after adjusting for all of the others. All but two geometric parameters,

Table 2 Mean and SD of 10 digital image parameters selected for the linear classifiers, by diagnosis

Selected digital image parameters (dermoscopic correlates)	Melanomas (n = 391) Mean	Melanomas (n = 391) SD	Nevi (histological diagnosis) (n = 449) Mean	Nevi (histological diagnosis) (n = 449) SD	P value (t test)
Geometry					
Area inside the outline (size)	4.097	0.815	3.001	0.699	0.000
Variance of contour symmetry with respect to 180° axes (symmetry of lesion layout)	3.952	1.858	4.207	1.779	0.004
Fractality of borders (border indenting)	0.786	0.104	0.767	0.086	0.043
Colors and Texture					
Skin lesion gradient (mean sharpness of lesion border)	23.785	14.015	10.138	5.997	0.000
Variance of skin lesion gradient histogram (variance of sharpness along lesion border)	60.152	15.575	77.872	7.690	0.000
Texture entropy (network analysis)	3.367	0.342	3.191	0.289	0.000
Islands (Clusters of colors)					
Imbalance of transition regions between lesion and healthy skin (imbalance with respect to "center of gravity" of skin-lesion transition regions)	0.324	0.276	0.213	0.166	0.000
Imbalance of blue-gray areas (imbalance with respect to "center of gravity" of areas of lesion tending to gray-blue color)	0.197	0.215	0.055	0.093	0.000
Gradient of the dark areas from lesion center to periphery (peripheral dark areas)	0.330	0.207	0.148	0.115	0.000
Number of border abruptions in red band (border abruptions)	6.115	2.815	4.866	2.440	0.000

Table 3 Performance of four classifiers on different sets of lesions, expressed as sensitivity (percentage of melanomas correctly classified) and specificity (percentage of melanocytic nevi correctly classified)

		Melanomas					Melanocytic nevi				
Center	Sets of lesions	*n*	Linear classifier 1[a]	Linear classifier 2[a]	Linear classifier 3[a]	K-NN[b] classifier	*n*	Linear classifier 1[a]	Linear classifier 2[a]	Linear classifier 3[a]	K-NN[b] classifier
IDI Rome	Training	77	95%				99	83%			
	Test	78	95%		95%		92	83%		84%	
	Total	155		93%		98%	191		81%		82%
Siena University	Training	107		95%			121		78%		
	Test	110		96%	95%		137		71%	72%	
	Total	217	94%			98%	258	73%			76%
IDI Rome + Siena University	Training	184			95%		220			78%	79%

[a] Linear classifier 1, constructed on the Rome IDI training set; linear classifier 2, constructed on the Siena University training set; linear classifier 3, constructed on the pooled Rome and Siena University training sets.
[b] K-NN, K-nearest-neighbor classifier.

i.e., fractality of the border and variance of the contour symmetry, were independently associated with melanoma (data not shown).

Classifier Performance

Linear Discriminant Classifiers. Table 3 shows the performance of three linear discriminant classifiers, two of which were trained on separate sets derived from the Rome and Siena centers, whereas the third was built on the pooled Siena and Rome training sets. The three classifiers were then independently tested on the Rome and Siena sets of lesions.

The first linear classifier, constructed on the Rome training set, with a fixed sensitivity of 95% reached a specificity of 83% on the Rome test set. When tested on the whole set of lesions belonging to the Siena Dermatology Department, a substantially stable performance was observed in terms of sensitivity (94%), whereas the specificity was 73% (Table 3). Similar results were obtained with the second classifier, constructed on the Siena University training set, which yielded a sensitivity of 93% and a specificity of 81% on the set of lesions from the Rome center.

K-Nearest-Neighbor Classifier. Table 3 shows also the performance of the K-nearest-neighbor classifier on the same sets of lesions. With a fixed sensitivity of 98%, a mean specificity of 79% was obtained on all sets of histologically diagnosed benign lesions, comparable with that obtained by the linear classifiers.

Discussion

Since the development of digital ELM, which allowed the acquisition and processing of high-quality images of pigmented skin lesions, there has been a growing interest in developing computerized image analysis ("machine vision") and proper algorithms to distinguish with high accuracy subtle differences, unperceived by the human eye, between cutaneous melanoma and benign melanocytic lesions (26–34, 38–45). Thus, several research efforts have focused in these last few years on the possibility of introducing into daily clinical practice computer-aided classification or automatic machine vision to increase the accuracy of melanoma diagnosis. In fact, although dermoscopy seems to have a discriminant power significantly higher than clinical examination in classifying pigmented lesions, as documented in a recent meta-analysis (17), the accuracy of dermoscopy is highly variable across different studies and is still far from the desirable levels of 100% sensitivity and high specificity. Sources of variation are likely to arise from differences in sample sizes, proportion of melanomas in the sample, type of instrument used, dermoscopic criteria used, and, last but not least, human variability in feature recognition and coding.

Our study gives an important contribution to this research area for several reasons. First, it is, to our knowledge, the second study on computer classification of pigmented lesions that compares the performance of different automatic classifiers on independent test sets of lesions (46). Second, this is the only study that assessed the performance of the classifiers on distinct test sets of lesions taken by different instruments in different times and locations belonging to patients from two different population groups. Third, our study highlights the importance of factors such as classifier design and feature selection in computer-aided diagnosis that are generally overlooked in the previously published studies. We adopted a very conservative procedure of feature selection for the linear classifier to obtain a relatively small set of robust parameters to discriminate melanoma from benign melanocytic lesions. This strategy and the use of a hold-out (separate training and test sets) design allowed performance estimates that were likely conservatively biased (47–49). Although conservatively biased, the performances of the linear classifiers were remarkably accurate, with a mean sensitivity of 95% and a mean specificity of 77% and highly stable on sets of lesions derived from different dermatology centers, where the referral criteria for patients with pigmented lesions and the operating conditions of the instruments could have been different.

The most critical requirement of the K-nearest-neighbor classifier is to have a training set including enough examples of each class of pigmented lesions to adequately represent the full range of measurements that can be expected from each class. The use of a training set of 1081 lesions in our study allowed accurate computations of *a posteriori* probabilities, after estimating class-conditional densities, and an estimate of the Bayes error rate. With a misclassification rate for the K-nearest-neigh-

bor classifier of 12.5%, the Bayes error rate was greater than 6.25% and below 12.5%.

Comparing the performances of the two classifiers, the Bayes nonparametric approach, yielding a sensitivity of 98% and a specificity of 79%, seemed to give results similar to the geometrical linear discriminant approach (sensitivity of 95% and specificity of 77%). Optimizing the procedures of feature selection and weight definition could additionally improve the performance of the K-nearest-neighbor classifier.

In conclusion, our study suggests that computer-aided differentiation of melanoma from benign pigmented lesions obtained with DB-Mips is feasible and, above all, reliable. In fact, the same instrumentation used in different units on different data sets provided similar diagnostic accuracy. Although the bottom line in the diagnosis of melanoma is likely to continue to depend on the clinical insight of the physician and on the expertise of the pathologist, computer-aided diagnosis could provide clinicians an objective second opinion, at expert level, based on consistently extracting and analyzing image features. To what extent the combination of human and machine-based diagnoses would affect the decision-making process in the management of patients with pigmented lesions by improving the detection of early melanoma and/or decreasing unnecessary surgery remains to be evaluated by well-designed, randomized clinical trials in the field.

Acknowledgments

We thank Drs. M. Biagioli and P. Taddeucci from the Department of Dermatology, University of Siena, Siena, Italy; Drs. S. Pallotta, and G. Ferranti from the Istituto Dermopatico dell'Immacolata, Rome, Italy; and Drs. M. Nudo, G. Rossi, F. Carlesimo, and F. Lechiancole from the Istituto Dermopatico dell'Immacolata, Capranica, Italy.

References

1. NIH Consensus Conference. Diagnosis and treatment of early melanoma. J Am Med Assoc 1992;268:1314–9.

2. Rigel DS, Carucci JA. Malignant melanoma: prevention, early detection, and treatment in the 21st century. CA-Cancer J Clin 2000;50: 215–36.

3. DeCoste SD, Stern RS. Diagnosis and treatment of nevomelanocytic lesions of the skin: a community-based study. Arch Dermatol 1993;129: 57–62.

4. Grin CM, Kopf AW, Welkovich B, Bart RS, Levenstein MJ. Accuracy in the clinical diagnosis of malignant melanoma. Arch Dermatol 1990;126:763–6.

5. Koh HK, Caruso A, Gage I, et al. Evaluation of melanoma/skin cancer screening in Massachusetts: preliminary results. Cancer 1990; 65:375–9.

6. Curley RK, Cook MG, Fallowfield ME, Marsden RA. Accuracy in clinically evaluating pigmented lesions. Br Med J 1989;299:16–8.

7. Friedman RJ, Rigel DS, Kopf AW. Early detection of malignant melanoma: the role of physician examination and self-examination of the skin. CA-Cancer J Clin 1985;35:130–51.

8. MacKie RM. Clinical recognition of early invasive malignant melanoma. Br Med J 1990;301:1005–6.

9. Healsmith MF, Bourke JF, Osborne JE, Graham-Brown RA. An evaluation of the revised seven-point checklist for the early diagnosis of cutaneous malignant melanoma. Br J Dermatol 1994;130:48–50.

10. McGovern TW, Litaker MS. Clinical predictors of malignant pigmented lesions: a comparison of the Glasgow seven-point checklist and the American Cancer Society's ABCDs of pigmented lesions. J Dermatol Surg Oncol 1992;18:22–6.

11. du Vivier AW, Williams HC, Brett JV, Higgins EM. How do malignant melanomas present and does this correlate with the seven-point check-list? Clin Exp Dermatol 1991;16:344–7.

12. Ascierto PA, Palmieri G, Botti G, et al. Early diagnosis of malignant melanoma: proposal of a working formulation for the management of cutaneous pigmented lesions from the Melanoma Cooperative Group. Int J Oncol 2003;22:1209–15.

13. Whited JD, Grichnik JM. Does this patient have a mole or a melanoma? J Am Med Assoc 1998;279:696–701.

14. Steiner A, Pehamberger H, Wolff K. Improvement of the diagnostic accuracy in pigmented skin lesions by epiluminescent light microscopy. Anticancer Res 1987;7:433–4.

15. Steiner A, Pehamberger H, Wolff K. *In vivo* epiluminescence microscopy of pigmented skin lesions. II. Diagnosis of small pigmented skin lesions and early detection of malignant melanoma. J Am Acad Dermatol 1987;17:584–91.

16. Pehamberger H, Binder M, Steiner A, Wolff K. *In vivo* epiluminescence microscopy: improvement of early diagnosis of melanoma. J Invest Dermatol 1993;100:356S–62S.

17. Bafounta ML, Beauchet A, Aegerter P, Saiag P. Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests. Arch Dermatol 2001;137:1343–50.

18. Binder M, Schwarz M, Winkler A, et al. Epiluminescence microscopy: a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch Dermatol 1995;131:286–91.

19. Soyer HP, Argenziano G, Chimenti S, Ruocco V. Dermoscopy of pigmented skin lesions. In: An atlas based on the Consensus Net Meeting on Dermoscopy 2000. Milan, Italy: EDRA Medical Publishing; 2001.

20. Carli P, De Giorgi V, Naldi L, Dosi G. Reliability and interobserver agreement of dermoscopic diagnosis of melanoma and melanocytic naevi: dermoscopy panel. Eur J Cancer Prev 1998;7:397–402.

21. Stanganelli I, Burroni M, Rafanelli S, Bucchi L. Intraobserver agreement in interpretation of digital epiluminescence microscopy. J Am Acad Dermatol 1995;33:584–9.

22. Nachbar F, Stolz W, Merkle T, et al. The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. J Am Acad Dermatol 1994;30:551–9.

23. Argenziano G, Fabbrocini G, Carli P, De Giorgi V, Sammarco E, Delfino M. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. Arch Dermatol 1998;134:1563–70.

24. Menzies SW, Ingvar C, Crotty KA, McCarthy WH. Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. Arch Dermatol 1996;132:1178–82.

25. Pizzichetta MA, Talamini R, Piccolo D, et al. The ABCD rule of dermatoscopy does not apply to small melanocytic skin lesions. Arch Dermatol 2001;137:1376–8.

26. Schindewolf T, Stolz W, Albert R, Abmayr W, Harms H. Classification of melanocytic lesions with color and texture analysis using digital image processing. Am J Epidemiol 1993;15:1–11.

27. Green A, Martin N, Pfitzner J, O'Rourke M, Knight N. Computer image analysis in the diagnosis of melanoma. J Am Acad Dermatol 1994;31:958–64.

28. Binder M, Steiner A, Schwarz M, Knollmayer S, Wolff K, Pehamberger H. Application of an artificial neural network in epiluminescence microscopy pattern analysis of pigmented skin lesions: a pilot study. Br J Dermatol 1994;130:460–5.

29. Binder M, Kittler H, Seeber A, Steiner A, Pehamberger H, Wolff K. Epiluminescence microscopy-based classification of pigmented skin lesions using computerized image analysis and an artificial neural network. Melanoma Res 1998;8:261–6.

30. Andreassi L, Perotti R, Rubegni P, et al. Digital dermoscopy analysis for the differentiation of atypical nevi and early melanoma: a new quantitative semiology. Arch Dermatol 1999;135:1459–65.

31. Rubegni P, Cevenini G, Burroni M, et al. Automated diagnosis of pigmented skin lesions. Int J Cancer 2002;101:576–80.

32. Seidenari S, Pellacani G, Giannetti A. Digital videomicroscopy and image analysis with automatic classification for detection of thin melanomas. Melanoma Res 1999;9:163–71.

33. Binder M, Kittler H, Dreiseitl S, Ganster H, Wolff K, Pehamberger H. Computer-aided epiluminescence microscopy of pigmented skin lesions: the value of clinical data for the classification process. Melanoma Res 2000;10:556–61.

34. Bauer P, Cristofolini P, Boi S, 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.

35. Mardia KV, Kent JT, Bibby M. Multivariate analysis. New York: Academic Press; 1979.

36. Swets JA, Pickett RM. Evaluation of diagnostic systems: methods from signal detection theory. New York: Academic Press; 1992.

37. Cover T, Hart P. Nearest neighbor pattern classification. IEEE Trans Information Theory 1967;13:21–7.

38. Schindewolf T, Stolz W, Albert R, Abmayr W, Harms H. Classification of melanocytic lesions with color and texture analysis using digital image processing. Am J Epidemiol 1993;15:1–11.

39. Green A, Martin N, Pfitzner J, O'Rourke M, Knight N. Computer image analysis in the diagnosis of melanoma. J Am Acad Dermatol 1994;31:958–64.

40. Ercal F, Chawla A, Stoecker WV, Lee HC, Moss RH. Neural network diagnosis of malignant melanoma from color images. IEEE Trans Biomed Eng 1994;41:837–45.

41. Gutkowicz-Krusin D, Elbaum M, Szwajkowski P, Kopf AW. Can early malignant melanoma be differentiated from atypical melanocytic nevi by *in vivo* techniques? Part II: Automatic machine vision classification. Skin Res Technol 1997;3:15–22.

42. Seidenari S, Pellacani G, Giannetti A. Digital videomicroscopy and image analysis with automatic classification for detection of thin melanomas. Melanoma Res 1999;9:163–71.

43. Hall PN, Claridge E, Smith JD. Computer screening for early detection of melanoma: is there a future? Br J Dermatol 1995;132:325–38.

44. Rubegni P, Burroni M, Cevenini G, et al. Digital dermoscopy analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: a retrospective study. Invest Dermatol 2002;119:471–4.

45. Sober AJ, Burstein JM. Computerized digital image analysis: an aid for melanoma diagnosis—preliminary investigations and brief review. J Dermatol 1994;21:885–90.

46. Dreiseitl S, Ohno-Machado L, Kittler H, Vinterbo S, Billhardt H, Binder MA. Comparison of machine learning methods for the diagnosis of pigmented skin lesions. J Biomed Inform 2001;34:28–36.

47. Sahiner B, Chan HP, Petrick N, Wagner RF, Hadjiiski L. Feature selection and classifier performance in computer-aided diagnosis: the effect of finite sample size. Med Phys 2000;27:1509–22.

48. Chan HP, Sahiner B, Wagner RF, Petrick N. Classifier design for computer-aided diagnosis: effects of finite sample size on the mean performance of classical and neural network classifiers. Med Phys 1999;26:2654–68.

49. Bellman R. Adaptive control processes: a guided tour. Princeton, NJ: Princeton University Press; 1961.