DOI: 10.1111/j.1468-3083.2012.04651.x JEADV

ORIGINAL ARTICLE

# A simple scoring system for the diagnosis of palmo-plantar pigmented skin lesions by digital dermoscopy analysis

P. Rubegni,[†,*] G. Cevenini,[‡] N. Nami,[†] G. Argenziano,[§] T. Saida,[¶] M. Burroni,[†] P. Quaglino,[††] R. Bono,[‡‡] R. Hofmann-Wellenhof,[§§] M. Fimiani[†]

[†]Department of Clinical Medicine and Immunological Sciences; Dermatology Section, University of Siena, Siena, Italy
[‡]Department of Surgery and Bioengineering, University of Siena, Siena, Italy
[§]Dermatology Unit, Medical Department, Arcispedale Santa Maria Nuova, Reggio Emilia, Italy
[¶]Department of Dermatology, Shinshu University School of Medicine, Matsumoto, Japan
[††]Department of Biomedical Sciences and Human Oncology, Section of Dermatology, First Dermatologic Division, University of Turin, Italy
[‡‡]Department of Immuno-oncodermatology, Istituto Dermopatico dell'Immacolata, Rome, Italy
[§§]Department of Dermatology, Medical University Graz, Graz, Austria
*Correspondence: Pietro Rubegni. E-mail: rubegni@unisi.it

## Abstract

**Background** Many research groups have recently developed equipments and statistical methods enabling pattern classification of pigmented skin lesions. To differentiate between benign and malignant ones, the mathematical extraction of digital patterns together with the use of appropriate statistical approaches is a challenging task.

**Objective** To design a simple scoring model that provides accurate classification of benign and malignant palmo-plantar pigmented skin lesions, by evaluation of parameters obtained by digital dermoscopy analysis (DDA).

**Patients and Methods** In the present study we used a digital dermoscopy analyser to evaluate a series of 445 palmo-plantar melanocytic skin lesion images (25 melanomas 420 nevi). Area under the receiver operator curve, sensitivity and specificity were calculated to evaluate the diagnostic performance of our scoring model for the differentiation of benign and malignant palmo-plantar melanocytic lesions.

**Results** Model performance reached a very high value (0.983). The DDA parameters selected by the model that proved statistically significant were: area, peripheral dark regions, total imbalance of colours, entropy, dark area and red and blue multicomponent. When all seven model variables were used in a multivariate mode, setting sensitivity at 100% to avoid false negatives, we estimated a minimum specificity of about 80%.

**Conclusions** Simplicity of use and effectiveness of implementation are important requirements for the success of quantitative methods in routine clinical practice. Scoring systems meet these requirements. Their outcomes are accessible in real time without the use of any data processing system, thus allowing decisions to be made quickly and effectively.

Received: 22 March 2012; Accepted: 19 June 2012

## Conflict of Interest

None declared.

## Financial sources

The study was carried out with financial supports obtained through the Fellowship SIDeMaST 2009 and the 'Scholarship in memory of Giulio Ginanneschi' for study, research and therapy of melanoma.

## Introduction

In dermatology, a recent technique known as dermoscopy, which magnify skin lesions and enable them to be examined down to the dermo-epidermal junction, has led to a 5–30% increase in diagnostic accuracy of pigmented skin lesions (PSLs) with respect to simple clinical observation.[1–3] However, evaluation of the many morphological characteristics of these lesions by dermoscopy is often extremely complex and subjective.[4,5] To obviate this problem, research groups have developed equipment and methods enabling objective evaluation of the many parameters that have

emerged with decades of experience with dermoscopy.[6–10] An example is digital dermoscopy analysis (DDA) which gathers numerical data and enables pigmented lesion images to be described objectively.[11–19] After many years of using DDA to improve the diagnostic accuracy of PSLs, today we know that if we want to optimize this methodology, it is important both to screen the lesions to be submitted to automated analysis and to use reliable statistical classification methods.[20–22] In the first place, DDA is optimal when used on groups of lesions that are morphologically and biologically as homogeneous as possible [i.e. benign and malignant palmo-plantar PSLs (PPPSLs)].[20–22] Preselection by dermatologists who are expert in dermoscopy is therefore necessary to exclude clinically certain diagnoses (seborroheic keratoses, Reed nevi, basal cell carcinoma and so on) that may not be such for any instrument.[17] Another fundamental issue is the choice of statistical classification method. It is well known that the more the PSLs population is limited, the less data is available. This is the case of melanomas in acral locations, especially in Caucasian populations.[23,24] In such cases, the predictive reliability of the quantitative classification method becomes crucial. Commonly employed models, based on linear discriminant analysis, logistic regression, Bayes rule or artificial neural networks, tend to overfit training data. Their excessive complexity leads to very small error on the data used to design them but significantly greater error on new data, making these models unreliable for diagnosis.[25,26] Of the simplier classification models, those that are less accurate have a greater predictive reliability.[26,27] In the present study we propose a simple scoring model that provides accurate classification of homogeneous skin lesions, specifically benign and malignant PPPSLs, by immediate evaluation (without computer) of a few important DDA parameters. The model can also objectively identify major dermoscopic features to distinguish benign from malignant PPPSLs.

## Patients and methods

All clinical investigation was conducted according to the principles of the Declaration of Helsinki. Institutional approval and patient consent were obtained for all experimental procedures. The study was carried out in three Italian centres: the Department of Clinical Medicine and Immunological Sciences, Section of Dermatology, University of Siena, the *Istituto Dermopatico dell'Immacolata* in Rome and the Department of Biomedical Sciences and Human Oncology, Section of Dermatology, First Dermatology Division, University of Turin, Italy. Between January 2010 and January 2011 images of about 20 000 PSLs, 445 of which were PPPSLs, were acquired by three identical digital dermoscopy instruments, calibrated before beginning the study, in the melanoma prevention clinics of the three centres. Reproducibility was first tested on digitized images of 30 PPPSLs (5 malignant melanomas and 25 nevi) of 30 patients, acquired with the three instruments. Absolute differences between single measurements and mean values of a given lesion or parameter never exceeded 5% of the mean value.

### Patients

The 445 PPPSLs were from 416 Caucasian patients [259 females (58.2%); mean age 29.3 years; 186 males (41.8%); mean age 25.7 years], ranging in age from 18 to 82 years; 107 of the lesions were removed. The lesions were removed because of the presence of clinical and/or dermoscopic sospicious features and on the absence of any clear benignity pattern (*parallel furrow, lattice-like* or *fibrillar patterns*). Among benign removed PPPSLs 21 were junctional nevi, 52 compound nevi, 6 dermal nevi, 1 blue nevus and 2 Reed nevi. Twenty-five of the 107 PPPSLs proved to be melanomas. These 25 acral melanomas (23 plantar, 2 palmar) were from 25 Caucasian patients (16 men and 9 women; age range 35–82 years; mean age 59 years). Histological examination showed a Breslow thickness of no more than 1 mm in 18/25 lesions and greater than 1 mm in the remaining 7 lesions.

### Measurements

The lesions were imaged (magnification 16×), stored and analyzed by three computerized DB-Mips System instruments (Biomips Engineering, S.R.L., via Colleverde 15, 53100 Siena, Italy), providing a visual database and objective evaluations of PSLs. During acquisition of plantar images, patients lay prone on the examination bed. The lesions were recorded as digital signals and saved.

### Equipement

The DB-Mips System consists of a firewire/USB digital camera with 1024 × 768 or 1280 × 1024 pixels. The camera was connected to a hand-held optical system enabling a horizontal field of view of 16.5 mm. The camera was calibrated weekly using special paper for white balance. Illumination was provided by a 150 W light source at 3200°K. The components of the video signal were connected to a frame-grabber interfaced with a Pentium III 500 MHz Personal Computer having a magneto-optical drive for image storage. The system ran under Microsoft Windows, and all the software was written in language C/C++.

### Digitization and parameterization

Choice of the most useful features to extract from digital images depends on the results of epiluminescence pattern analysis. The variables we chose were dermoscopic parameters currently used in the diagnosis of PSLs. Although the system saves microscope magnifications along with texture analysis, offering an objective evaluation, the different magnifications could confuse clinicians wanting to make subjective comparisons of lesions. In this paper we only discuss images with a magnification of 16. The system used a procedure for digital image processing based on the Laplacian filter for segmentation and a zero-crossing algorithm for border automatic outline.[28] It then evaluated 49 parameters for discriminating power. As previously described, the parameters belonged to four categories: geometries, colours, textures, and islands of colour (i.e. colour clusters inside lesions).[9] All variables used are derived

from close collaboration between the clinicians and the clinical engineers. The task of the latter was to render objective/numerical, the various dermoscopic patterns typical of various skin lesions.

### Design of the score model

All cases were diagnosed histologically or clinically, and these results were used as training data for building the model classifier. The score model was designed as follows:[27]

- All quantitative predictor variables (DDA parameters) were dichotomized by ROC curve analysis, identifying cut-off values giving equal sensitivity and specificity in relation to melanoma detection.
- These obtained binary risk factors were coded 0 or 1, depending on whether the risk of malignity decreased or increased, over or under the cut-off value, respectively.
- The odds ratio of each binary factor was evaluated on the basis of the corresponding confidence interval (CI): factors with odds ratios not significantly greater than 1 were discarded.
- A forward iterative procedure, which selected stepwise the most discriminant binary factors, was applied to a data sample (training set).
- Selected factors were summed to form an integer score.
- All binary factors were reconsidered at each step, so that multiple selection of one factor gave rise to a multiple integer contribution to the score.
- At each step the risk factor providing the highest increment to model classification performance, evaluated by the area under the ROC curve (AUC), was added to the model.
- Training was stopped when the cumulative increment in AUC obtained in five consecutive steps was less than 0.5%.
- A five-fold cross-validation procedure was used to evaluate model generalization.

A high generalization power is of fundamental importance for predictive models. The optimal model is the simplest possible model able to supply the highest possible performance on training data and to maintain the same predictive error on any other set of testing data, equally representative of the same population. Excessively complex models tend to overfit, i.e. give significantly lower errors on the training data than on the testing data. Overfitting produces data storage rather than learning of prediction rules. Models must be designed to avoid overfitting, so as to improve generalization, through efficient cross-validation procedures to control of the training process. The most common type of cross-validation procedure is k-fold, where the original sample is randomly partitioned into k subsamples, one of which is used as testing set and the other k–1 as training set. The process is then repeated k times, changing the testing set each time so that all available data both for training and testing the model. A usual choice is to assume k = 5 in which the procedure takes four-fifths of the data as training set and the remaining fifth as testing set, then proceeds with five training sessions by rotating subsamples.

The bias-corrected and accelerated bootstrap method was applied to estimate 95% CIs of model sensitivity, specificity and accuracy, i.e. AUC, using 2000 bootstrapped samples.[27]

The total score, S, associated with a generic lesion, was expressed as:

$$S = \sum_{i=1}^{d} p_i s_i$$

where d is the number of DDA predictive parameters included in the model, $p_i$ the binary value of the $i^{th}$ predictor, and $s_i$ its model-identified associated score. Computer analysis was performed using the MATLAB package.[29]

## Results

The DDA-Mips system enabled acquisition of high quality images in real time, making it possible to examine all the features revealed by epiluminescence. The digital images were subjectively almost as good as conventional photographs. Image resolution was 45 pixels/mm at a magnification of 16. The system was easy to use. Objective evaluations were also performed automatically in real time; no modification by the operator was necessary. Graphic windows showed the objective results, which were readily understood. The operators were able to use the program without any special training.

Table 1 shows descriptive statistics and the results of the univariate *F* test for group comparison of DDA data. The variables are listed in descending order of separation power, that is, in order of decreasing values of *F*. Seventeen variables reached a statistical significance of at least 95% ($F > 3.84$, $P < 0.05$).

### Score model

Figure 1 illustrates the design of the diagnostic score model. Model performance is shown by the step-by-step increase in AUC on training data. The AUC reached a very high value (0.983) at the final (eleventh) step, when the model was complete. The end-step value of testing data AUC was very close to that of the training data, validating model prediction ability and ensuring high diagnostic reliability of the model. The DDA parameters selected by the model are also indicated; those that proved statistically significant were: area, peripheral dark regions, total imbalance of colours, entropy, dark area, and red and blue multicomponent. Our multivariate score model was optimized for discriminant purposes by selecting an optimal minimum set of DDA parameters capable of optimally distinguishing benign from malignant lesions. Thus the model did not include all parameters that singly (univariately) have some discriminant power, but only a number of them that minimize classification error multivariately.

Table 2 describes the scoring system. For each PPPSL, the total score, S, was obtained by considering the seven DDA parameters selected, comparing their values with the corresponding cut-offs and adding the contributions of the associated risk scores to S when the cut-off conditions were met.

**Table 1** Descriptive statistics and univariate *F* test of differences between palmoplantar melanomas and benign palmoplantar nevi. Only statistically significant DDA parameters are reported ($P < 0.05$), sorted by descendent *F* values

| DDA parameters | | Benign lesions | | Malignant lesions | | *F* |
|---|---|---|---|---|---|---|
| **Name** | **Category** | **Median** | **Range** | **Median** | **Range** | |
| Area | Geometries | 0.2 | 0.05–1 | 0.61 | 0.34–1 | 296.3 |
| Maximum diameter | Geometries | 0.17 | 0.03–0.85 | 0.62 | 0.28–1 | 277.6 |
| Minimum diameter | Geometries | 0.24 | 0.05–1 | 0.68 | 0.35–1 | 252.5 |
| Perimeter | Geometries | 14 | 1.5–127 | 71 | 26–196 | 250.7 |
| Blue multicomponent | Islands of c. | 5.6 | 1.5–19 | 12 | 0.01–19 | 130.2 |
| Red multicomponent | Islands of c. | 3.2 | 1.0–112 | 6.4 | 0.01–14 | 126.3 |
| Green multicomponent | Islands of c. | 27 | 7.7–225 | 60 | 31–208 | 94.7 |
| Entropy | Texture | 3.0 | 2.2–4.1 | 3.6 | 2.6–4.4 | 71.8 |
| Dark region imbalance | Islands of c. | 0.03 | 0–0.2 | 0.07 | 0.01–0.17 | 67.2 |
| Contrast | Texture | 1.1 | 0.76–2.9 | 1.4 | 0.75–2.7 | 41.2 |
| Variance of border gradient | Islands of c. | 82 | 9.9–99 | 64 | 33–96 | 39.7 |
| Border homogeneity | Islands of c. | 2 | 0–14 | 5 | 0–13 | 30.9 |
| Peripheral dark regions | Islands of c. | 0.12 | 0.01–3.3 | 0.35 | 0.02–0.65 | 28.0 |
| Transition region imbalance | Islands of c. | 3 | 0–29 | 13 | 0–33 | 22.2 |
| Background region imbalance | Islands of c. | 0.13 | 0.02–0.45 | 0.22 | 0.07–0.37 | 16.3 |
| Mean values of red inside the lesion | Colours | 0 | 0–0.32 | 0 | 0–0.46 | 13.4 |
| Mean values of blue inside the lesion | Colours | 0 | 0–0.17 | 0 | 0–0.81 | 6.96 |
| Total imbalance of colours | Islands of c. | 2 | 1–14 | 5 | 1–12 | 4.03 |

DDA, digital dermoscopy analysis; Islands of c., Islands of colours.

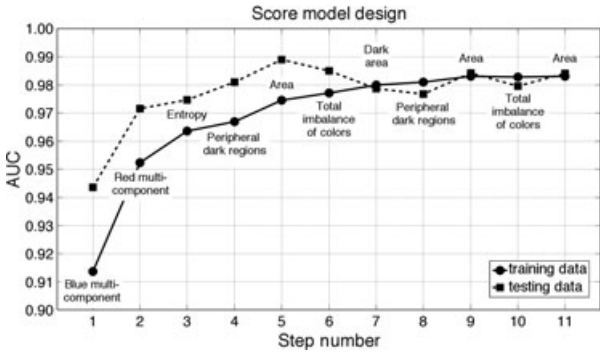


**Figure 1** Stepwise procedure to train and test the integer score model. At each step the digital dermoscopy analysis (DDA) parameter entered contributes one unit to the total score.

Table 3 shows the frequency distribution of classification results provided by the scoring system. Model scores ranged from 0, when none of seven DDA parameters was at risk, to 11, when all parameters indicated risk. All melanomas obtained a total score of more than 4. Setting a score threshold T = 4 and removing lesions having a score S > T avoided false negatives, giving a sensitivity of 100% and a specificity of 83.1% (71 false positives). To eliminate the risk of false negatives in our data sample, we proposed the borderline score S = 4 as an indicator of the need for follow-up. In our sample, 30 benign lesions (about 7%) were identified by the score model for follow-up.

**Table 2** Cut-off conditions applied to digital dermoscopy analysis (DDA) parameters included in the model and corresponding integer risk score values to be added to the total score

| DDA parameter | Cut-off | Risk score |
|---|---|---|
| Area ($mm^2$) | >30 $mm^2$ | 3 |
| Peripheral dark regions (%) | >0.3 | 2 |
| Total imbalance of colours (%) | >4 | 2 |
| Entropy (0–4) | >3 | 1 |
| Dark area (%) | >0.1 | 1 |
| Red homogeneity (%) | >0.3 | 1 |
| Blue homogeneity (%) | >0.4 | 1 |

DDA, digital dermoscopy analysis.

At score threshold T = 4, the 95% confidence interval of specificity was estimated at 79.3–86.5% by the bootstrap technique. This means that in the worst condition at the lower bound we would obtain about 21% false positives. In our sample, this would be at most 87 benign lesions out of 420, which the model would indicate for removal. If we also count the 25 melanomas, the highest expected percentage of acral skin lesions that the model would indicate for removal would be about 26% of the total.

Figure 2 shows the score model ROC curve and the 95% CI estimated from our data sample using the bootstrap method. The CI cancelled out at 100% sensitivity that matched about 80% specificity. This means that to avoid false negatives, the score model wrongly indicates at least 20% benign acral lesions for removal.

**Table 3** Number of palmoplantar pigmented skin lesions associated with each score, estimated percentage risk and proposed decision (biopsy)

| S | Benign lesion | Melanoma | Risk% | Biopsy |
|---|---|---|---|---|
| 0 | 67 | 0 | 0% | No |
| 1 | 98 | 0 | 0% | No |
| 2 | 110 | 0 | 0% | No |
| 3 | 44 | 0 | 0% | No |
| 4 | 30 | 0 | 0% | No (follow-up) |
| 5 | 28 | 3 | 20% | Yes |
| 6 | 24 | 3 | 40% | Yes |
| 7 | 11 | 2 | 60% | Yes |
| 8 | 3 | 6 | 80% | Yes |
| 9 | 3 | 4 | 90% | Yes |
| 10 | 2 | 2 | 99% | Yes |
| 11 | 0 | 5 | 100% | Yes |
| Total | 420 | 25 | | |

S = risk score

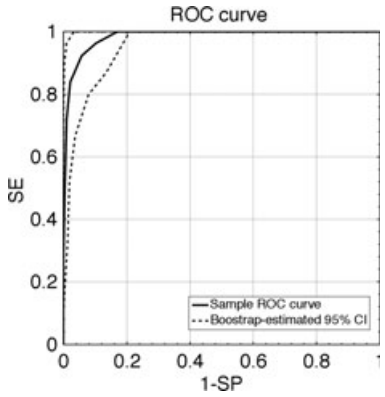


**Figure 2** ROC curve of the score model estimated from sample data and the corresponding 95% confidence interval (CI). SE, sensitivity; SP, specificity.

Thus, the threshold score S = 4 does not completely exclude risk of melanoma and in this case it would seem appropriate to recommend periodic follow-up of the nevus.

Table 4 shows an example of application of our score model to a benign lesion (nevus), a clinically atypical lesion that was removed (nevus) and a melanoma (Fig. 3).

## Discussion

Acral melanoma is a rare but distinctive subtype of melanoma. It is reported to constitute 60–75% of all cutaneous melanomas in blacks, 43–49% in Asians and 5–7% in whites.[30,31] However, absolute incidences of acral melanoma seem to be almost the same in all races.[30,31]

Although the proportion of acral melanoma is low in Caucasians, pigmented skin lesions of the hands and feet are very common in this population.[32] Thus, acral nevi are a source of confusion that challenges physicians and the lay public in the differentiation of typical vs. atypical acral nevi and acral melanoma.[23,33,34] Univariate analysis enabled us to identify certain numerical variables which were significant for differentiating palmo-plantar melanomas from benign nevi (Table 1). We observed that a significant parameter for distinguishing benign PPPSLs and melanomas is 'size' (area, maximum and minimum diameter and perimeter). Indeed, the importance of these three variables was recently also pointed out by Koga and Saida,[35] who proposed the '*revised three-step algorithm*' based on the simple distinction of lesions greater or less than 7 mm in diameter, in the absence of dermoscopic features typical of palmo-plantar melanomas, such as the *parallel ridge pattern*. This value seems also confirmed by our data sample, where the smallest maximum diameter among melanomas was 6.96 mm. We also found that seven variables related to multicomponent pattern (green, blue, red multicomponent, entropy, dark region imbalance, contrast, background region imbalance and total imbalance of colours) were also significant. In this regard, the combination of many elements with variable shades from tan to black and their disorganized disposition within a PSL is termed

**Table 4** Values and relative risk scores of the three palmp-plantar pigmented skin lesions in Fig. 3

| DDA parameter | Cut-off | Risk Score | Lesion 1 (nevus) | Score | Lesion 2 (nevus) | Score | Lesion 3 (melanoma) | Score |
|---|---|---|---|---|---|---|---|---|
| Area ($mm^2$) | >30 | 3 | 23.36 | 0 | 19.24 | 0 | 36.47 | 3 |
| Peripheral dark regions (%) | >0.3 | 2 | 0.15 | 0 | 0.10 | 0 | 0.09 | 0 |
| Total imbalance of colours (%) | >4 | 2 | 6 | 2 | 6 | 2 | 5 | 2 |
| Entropy (0–4) | >3 | 1 | 2.89 | 0 | 2.82 | 0 | 3.62 | 1 |
| Dark area (%) | >0.1 | 1 | 0.12 | 1 | 0.259 | 1 | 0.19 | 1 |
| Red homogeneity (%) | >0.3 | 1 | 0.24 | 0 | 0.32 | 1 | 0.68 | 1 |
| Blue homogeneity (%) | >0.4 | 1 | 0.28 | 0 | 0.16 | 0 | 0.34 | 0 |
| Total risk score<br>No follow-up (<4)<br>Follow-up (=4)<br>Biopsy (>4) | | | 3 | | 4 | | 8 | |
| Decision | | | No follow-up | | Follow-up | | Biopsy | |

DDA, digital dermoscopy analysis.

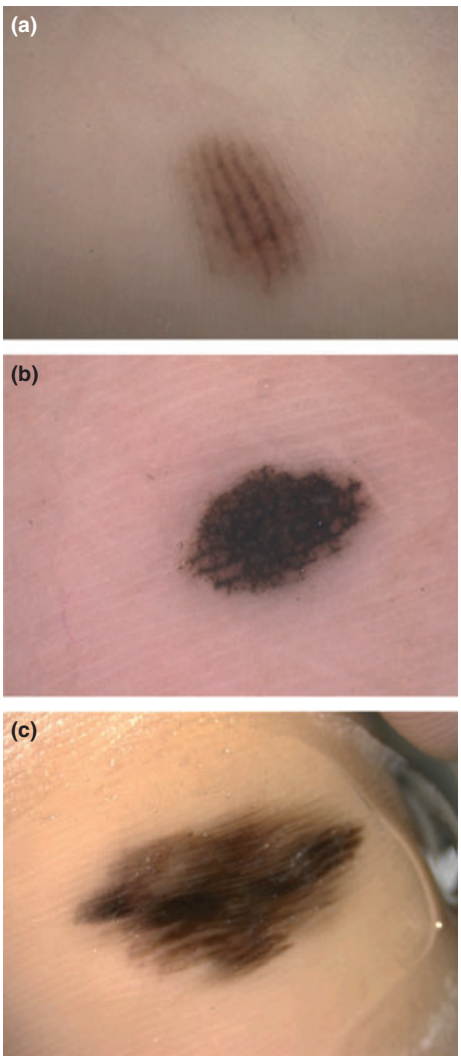


**Figure 3** Three palmo-plantar skin lesions. (a) The plantar nevus at the top dermoscopically exhibits a typical *parallel furrow pattern* and has a total risk score of 3 (Total imbalance of colours = 2; Dark area = 1) and therefore according to our model does not warrant follow-up. (b) The plantar nevus at the centre of the figure shows a *parallel furrow pattern* at its periphery and *reticular-globular pattern* at the centre; its total risk score is 4 (total imbalance of colours = 2; dark area = 1; red multicomponent = 1), which according to our model warrants follow-up. (c) The plantar lesion at the bottom is a melanoma *in situ* with typical dermoscopic *parallel ridge pattern*, and has a total risk score of 8 (area = 3; total imbalance of colours = 2; entropy = 1; dark area = 1; red homogeneity = 1).

*irregular diffuse pigmentation* by Saida and coworkers.[31,36] This dermoscopic feature is found in 85% of palmoplantar melanomas, being the second most typical malignant dermoscopic pattern of melanomas in these sites after the *parallel ridge pattern,* and seems to suggest a greater degree of invasion.[37,38] Other variables regarding border features (variance of border gradient, border homogeneity, peripheral dark regions and transition region imbalance) showed statistically significant differences between PSLs and melanomas. This finding is sustained by many dermoscopy studies that have demonstrated the significance/importance of border morphology for differential diagnosis between benign and malignant PSLs.[39] Indeed, it is well known that benign PSLs have borders that fade evenly, whereas malignant PSL often have quite different features.[39] Finally, Table 1 shows that the amounts of red and blue in the lesion are statistically different in benign and malignant lesions: a frequent lack of pigmentation (pink colour) is common and well documented, and seems to be a factor in delayed diagnosis of acral melanoma, like advanced age of patient, hidden site and unusual presentation.[40] On the other hand, bluish colour is usually associated with delayed diagnosis, which in some studies ranged from one to 3.7 years, and a consequent significant thickness of the melanoma.[40] This is also documented by our series, in which mean depth of invasion was 1.8 mm.

The stepwise procedure for score model design only selected seven parameters (Fig. 1). It is interesting to note that most discriminant capacity was obtained with the two variables of colour homogeneity (blue multicomponent, red multicomponent and entropy) (AUC about 97%). Area was the only shape parameter included in the model. Although it gave the highest score, it only entered the model at the fifth step, proving to be an important parameter but not decisive on its own for diagnosis of palmoplantar melanoma. Indeed, lesions having the sole feature of large size (in our model, area >30 $mm^2$, see Table 2) score only 3, which in our model indicates benignity.

When all seven model variables were used in a multivariate mode to evaluate the percentage of correct classification between melanomas and PSLs, the results were very interesting. In particular, setting sensitivity at 100% to avoid false negatives (false negative ratio = 0), we estimated a minimum specificity of about 80% (see the lower bound of 95% CI in Fig. 2), meaning that only one-fifth (false positive ratio) of benign lesions were indicated for removal. However we have to remind that for a dermatologist it is much more dramatic to loose a melanoma than to excise benign lesions. It is not simple to compare these results with those of previous studies, because the latter calculated sensitivity and specificity with respect to the capacity of a non-automatic device to detect single dermoscopic patterns, such as *parallel ridge pattern* and *irregular diffuse pigmentation*, without making an overall assessment of the possibility of statistically predicting the malignity or benignity of the lesions examined.[31,36–38,41] However, it can be said that the best dermoscopy experts in this field obtained a sensitivity of about 95% and a specificity only slightly better than those observed in our series.[36,38,41]

Scoring systems are a very attractive family of clinical predictive models, because a patient's score can be calculated without using any data processing system. Many quantitative methods for assess-

ing the risk of melanoma and/or other diseases have been developed recently, with the aim of providing objective and accurate diagnostic information.[21,42,43] Experience has shown that simplicity of use and effectiveness of implementation are important requirements for the success of quantitative methods in routine clinical practice, allowing decisions to be made quickly and effectively.[43,44] Other important benefits of score models are their easy updating and customisation to local practices.

In conclusion, our study suggests that DDA together with a simple scoring system could be useful for the differentiation of palmo-plantar melanoma from benign lesions. On these basis, in our opinion, computer-aided diagnosis could be introduced in dermatological practice to provide clinicians a second opinion at expert level, or, in non-expert settings, it could be a useful screening tool.

## Author contributions

PR, GC, NN, GA and TS equally contributed to the paper and drafted the manuscript; PR, GC and NN performed study concept and design; PR, RB and PQ helped in acquisition of data; GC, MB, GA and TS analyzed and interpreted data; MF critically revised the manuscript for important intellectual content; GC and MB performed statistical analysis; PR, NN and MF Obtained funding; PR and NN provided administrative, technical and material support; PR and MF supervised the study.

## References


1 Soyer HP, Smolle J, Kerl H, Steiner H. Early diagnosis of malignant melanoma by surface microscopy. *Lancet* 1987; **2**: 803.

2 Ascierto PA, Palmieri G, Celentano E *et al.* Sensitivity and specificity of epiluminescence microscopy: evaluation on a sample of 2731 excised cutaneous pigmented lesions. The Melanoma Cooperative Study. *Br J Dermatol* 2000; **142**: 893–898.

3 Vestergaard ME, Macaskill P, Holt PE, Menzies SW. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *Br J Dermatol* 2008; **159**: 669–676.

4 Rosendahl C, Tschandl P, Cameron A, Kittler H. Diagnostic accuracy of dermatoscopy for melanocytic and nonmelanocytic pigmented lesions. *J Am Acad Dermatol* 2011; **64**: 1068–1073.

5 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**: 286–291.

6 Seidenari S, Pellacani G, Pepe P. Digital videomicroscopy improves diagnostic accuracy for melanoma. *J Am Acad Dermatol* 1998; **39**: 175–181.

7 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**: 1459–1465.

8 Dreiseitl S, Ohno-Machado L, Kittler H *et al.* A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *J Biomed Inform* 2001; **34**: 28–36.

9 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. *J Invest Dermatol* 2002; **119**: 471–474.

10 Isasi AG, Zapirain BG, Zorrilla AM. Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. *Comput Biol Med* 2011; **41**: 742–755.

11 Ercal F, Chawla A, Stoecker WV *et al.* Neural network diagnosis of malignant melanoma from color images. *IEEE Trans Biomed Eng* 1994; **41**: 837–845.

12 Ganster H, Pinz A, Roehrer R *et al.* Automated melanoma recognition. *IEEE Trans Med Imaging* 2001; **20**: 233–239.

13 Elbaum M, Kopf AW, Rabinovitz HS *et al.* Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study. *J Am Acad Dermatol* 2001; **44**: 207–218.

14 Rubegni P, Cevenini G, Burroni M *et al.* Automated diagnosis of pigmented skin lesions. *Int J Cancer* 2002; **101**: 576–580.

15 Hoffmann K, Gambichler T, Rick A *et al.* Diagnostic and neural analysis of skin cancer (DANAOS). A multicentre study for collection and computer-aided analysis of data from pigmented skin lesions using digital dermoscopy. *Br J Dermatol* 2003; **149**: 801–809.

16 Blum A, Luedtke H, Ellwanger U *et al.* Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. *Br J Dermatol* 2004; **151**: 1029–1038.

17 Rubegni P, Burroni M, Andreassi A, Fimiani M. The role of dermoscopy and digital dermoscopy analysis in the diagnosis of pigmented skin lesions. *Arch Dermatol* 2005; **141**: 1444–1446.

18 Celebi ME, Kingravi HA, Uddin B *et al.* A methodological approach to the classification of dermoscopy images. *Comput Med Imaging Graph* 2007; **31**: 362–373.

19 Abbas Q, Celebi ME, Fondón García I, Rashid M. Lesion border detection in dermoscopy images using dynamic programming. *Skin Res Technol* 2011; **17**: 91–100.

20 Burroni M, Corona R, Dell'Eva G *et al.* Melanoma computer-aided diagnosis: reliability and feasibility study. *Clin Cancer Res* 2004; **10**: 1881–1886.

21 Rubegni P, Cevenini G, Burroni M *et al.* Objective follow-up of atypical melanocytic skin lesions: a retrospective study. *Arch Dermatol Res* 2010; **302**: 551–560.

22 Rubegni P, Burroni M, Nami N *et al.* Objective melanoma progression. *Skin Res Technol* 2011; **17**: 69–74.

23 Kuchelmeister C, Schaumburg-Lever G, Garbe C. Acral cutaneous melanoma in caucasians: clinical features, histopathology and prognosis in 112 patients. *Br J Dermatol* 2000; **143**: 275–280.

24 Altamura D, Altobelli E, Micantonio T *et al.* Dermoscopic patterns of acral melanocytic nevi and melanomas in a white population in central Italy. *Arch Dermatol* 2006; **142**: 1123–1128.

25 Hodgman SB. Predictive modeling and outcomes. *Prof Case Manag* 2008; **13**: 19–23.

26 Barbini E, Cevenini G, Scolletta S *et al.* A comparative analysis of predictive models of morbidity in intensive care unit after cardiac surgery – Part I: model planning. *BMC Med Inform Decis Mak* 2007; **7**: 35.

27 Cevenini G, Barbini P. A bootstrap approach for assessing the uncertainty of outcome probabilities when using a scoring system. *BMC Med Inform Decis Mak* 2010; **10**: 45.

28 Fang S, Gupta P. Zero-crossing iso-surfaces in volume datasets. *Conf Proc IEEE Eng Med Biol Soc* 2005; **5**: 5108–5111.

29 Cai JJ, Smith DK, Xia X, Yuen KY. MBEToolbox: a MATLAB toolbox for sequence data analysis in molecular biology and evolution. *BMC Bioinformatics* 2005; **6**: 64.

30 Hudson DA, Krige JE. Melanoma in black South Africans. *J Am Coll Surg* 1995; **180**: 65–71.

31 Saida T, Miyazaki A, Oguchi S *et al.* Significance of dermoscopic patterns in detecting malignant melanoma on acral volar skin: results of a multicenter study in Japan. *Arch Dermatol* 2004; **140**: 1233–1238.

32 Palicka GA, Rhodes AR. Acral melanocytic nevi: prevalence and distribution of gross morphologic features in white and black adults. *Arch Dermatol* 2010; **146**: 1085–1094.
33 Feibleman CE, Stoll H, Maize JC. Melanomas of the palm, sole, and nailbed: a clinicopathologic study. *Cancer* 1980; **46**: 2492–2504.
34 Jimbow K, Takahashi H, Miura S *et al.* Biological behavior and natural course of acral malignant melanoma: clinical and histologic features and prognosis of palmoplantar, subungual, and other acral malignant melanomas. *Am J Dermatopathol* 1984; **6**: 43–53.
35 Koga H, Saida T. Revised 3-step dermoscopic algorithm for the management of acral melanocytic lesions. *Arch Dermatol* 2011; **147**: 741–743.
36 Saida T, Koga H. Dermoscopic patterns of acral melanocytic nevi: their variations, changes, and significance. *Arch Dermatol* 2007; **143**: 1423–1426.
37 Ishihara Y, Saida T, Miyazaki A *et al.* Early acral melanoma in situ: correlation between the parallel ridge pattern on dermoscopy and microscopic features. *Am J Dermatopathol* 2006; **28**: 21–27.
38 Saida T, Koga H, Uhara H. Key points in dermoscopic differentiation between early acral melanoma and acral nevus. *J Dermatol* 2011; **38**: 25–34.
39 Blum A, Rassner G, Garbe C. Modified ABC-point list of dermoscopy: a simplified and highly accurate dermoscopic algorithm for the diagnosis of cutaneous melanocytic lesions. *J Am Acad Dermatol* 2003; **48**: 672–678.
40 Rex J, Paradelo C, Mangas C *et al.* Management of primary cutaneous melanoma of the hands and feet: a clinicoprognostic study. *Dermatol Surg* 2009; **35**: 1505–1513.
41 Iyatomi H, Oka H, Celebi ME *et al.* Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. *J Invest Dermatol* 2008; **128**: 2049–2054.
42 den Boer S, de Keizer NF, de Jonge E. Performance of prognostic models in critically ill cancer patients – a review. *Crit Care* 2005; **9**: 458–463.
43 Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. *Crit Care* 2010; **14**: 207.
44 Cei M, Bartolomei C, Mumoli N. In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: a prospective study. *Int J Clin Pract* 2009; **63**: 591–595.