Concordance analysis of dermoscopic features between five observers in a sample of 200 dermoscopic images

Dear Editor,

Dermoscopy is a non-invasive technique that has been proven in recent metanalysis to increase the accuracy in the diagnosis of cutaneous melanoma compared to the naked eye examination. However, its application can be considered subjective and too observer-dependent, with a heavy influence by previous experience. To this date, few reproducibility studies focusing on interobserver agreement of dermoscopic images have been published. Most of them are based on the evaluation of <50 cases by each observer. Furthermore, since the Internet Consensus Meeting of 40 dermoscopy experts in 2003, a few novel dermoscopic features have been described that were not evaluated.

The aim of this project was to analyze and determine the reproducibility of the analysis of both classic and novel dermoscopic features for the diagnosis of melanoma in 200 dermoscopic images by five blinded observers (E.H.T., L.N.B., JAAI, BLM, ERL). Previous experience on dermoscopy was >10 years (1 observer), >5 years (2 observers), and <5 years (2 observers).

A retrospective evaluation of dermoscopic images collected from the database of the Melanoma Unit in our department was performed. Images were obtained using a digital polarized dermoscopy system (DermLite Photo II Pro HR® [3Gen®, San Juan Capistrano, CA, USA] and an Olympus E-420® camera [Olympus, Tokyo, Japan]). Lesion diameter had to fit in the picture to be included. Cases without histopathological confirmation, melanoma metastases, or melanomas of special sites (facial, acral, nail, genital, or mucosal melanoma) were excluded. The distribution of diagnosis in our sample is shown in Table 1.

The following dermoscopic features were analyzed: colors (light brown, dark brown, black, blue/gray, red/pink, white), asymmetry of color/structures, atypical, pigmented network, irregular globules, streaks, irregular blotches, shiny white streaks, negative pigment o brown areas, blue-black pigmentation, milky-red areas, rainbow pattern, pseudolacunae, ulceration, and irregular vessels. Data were analyzed using SPSS version 22.0 (Chicago, IL, USA). Interobserver agreement between the five dermoscopists was evaluated using the Fleiss’ Kappa statistic test.

The results of the concordance analysis are shown in Table 2. Most of the dermoscopic features ranged Kappa values between 0.3–0.5, which can be considered fair to moderate. Asymmetry of color and structures (yes/no) showed a moderate agreement (0.46–0.49), slightly higher than the same evaluation considering one or two axes. The presence of more than three colors presented the moderate agreement as well (0.46). The melanoma-specific structures that showed more consistency were the shiny white streaks (up to 0.55), while structureless brown areas had the worst results (0.05). A few features that have been recently described, such as prominent skin markings and blue-black pigmentation, showed a very poor correlation (0.23 and 0.18, respectively) between all observers.

Previous studies have determined that features regarding the overall organization, colors, and symmetries have a higher agreement and discriminatory power than many well-known diagnostic structures such as atypical pigment networks or irregular blotches. Recent dermoscopic algorithms, such as CASH and TADA algorithms, are based on this idea. This finding has been confirmed in our concordance analysis and supports this approach. On the other hand, the low values observed in novel features such as prominent

### Table 1 Distribution of diagnosis in the 200 cases for evaluation.

| Diagnosis                              | n  | %    |
|----------------------------------------|----|------|
| Melanoma                               | 99 | 49.5 |
| Melanocytic nevus                      | 45 | 22.5 |
| Melanocytic nevus with histopathological atypia | 22 | 11.0 |
| Blue nevus                             | 5  | 2.5  |
| Spitz nevus                            | 4  | 2.0  |
| Basal cell carcinoma                   | 14 | 7.0  |
| Seborrheic keratosis                   | 5  | 2.5  |
| Dermatofibroma                         | 2  | 1.0  |
| Squamous cell carcinoma                | 2  | 1.0  |
| Other neoplasms                        | 2  | 1.0  |
| **Total**                              | 200| 100  |

© Corresponding author.

*Corresponding author.*

**E-mail:** mariana@hafner.med.br (M.F. Hafner).

Received 15 August 2020; accepted 5 December 2020

Available online 17 March 2022

https://doi.org/10.1016/j.abd.2020.12.012

0365-0596/ © 2022 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Dermatologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

---

10. Baeck M, Chemelle JA, Terreur R, Drieghe J, Goossens A. Delayed hypersensitivity to corticosteroids in a series of 315 patients: clinical data and patch test results. Contact Dermatitis. 2009;61:163–75.

Mariana de Figueiredo Silva Hafner a,*, Amanda Ivanchuk Lopes a, Nathalie Mie Suzuki a, Rosana Lazzarini b a Clinica de Dermatologia da Santa Casa de São Paulo, São Paulo, SP, Brazil b Faculdade de Ciências Médicas da Santa Casa de São Paulo, São Paulo, SP, Brazil

---

© Study was conducted using data from patients obtained from the Melanoma Unit of the Hospital General Universitario Gregorio Marañón, Madrid, Spain.
Table 2 Interobserver agreement of colors and dermoscopic structures (Fleiss’ Kappa statistical test).

| Structure                     | $\kappa$ |
|-------------------------------|----------|
| Three or more colors          | 0.46     |
| Color asymmetry (yes/no)      | 0.46     |
| Structure asymmetry (yes/no)  | 0.49     |
| Color asymmetry (0-2 axes)    | 0.35     |
| Structure asymmetry (0-2 axes)| 0.36     |
| Shiny white streaks           | 0.55     |
| Polymorphous vascular pattern | 0.48     |
| Regression                    | 0.47     |
| Streaks/pseudopods            | 0.45     |
| Atypical pigment network      | 0.42     |
| Pseudolacunae                 | 0.41     |
| Negative pigment network      | 0.40     |
| Irregular globules            | 0.38     |
| Irregular vessels             | 0.37     |
| Irregular blotches            | 0.37     |
| Blue-white veil               | 0.35     |
| Rainbow pattern               | 0.34     |
| Milky red areas               | 0.30     |
| Prominent skin markings       | 0.23     |
| Blue-black pigmentation       | 0.18     |
| Hypopigmented areas           | 0.16     |
| Structureless brown areas     | 0.05     |

skin markings and blue-black pigmentation are noteworthy. Although these features have been well-defined, and their identification may be useful to increase the suspicion of certain lesions, their interobserver agreement might seem too low to be reliable in clinical practice. This variability is expected to be more striking, especially between non-expert dermatoscopy, limiting their efficacy on a diagnosis.

Dermoscopy might be a subjective diagnostic tool, and its efficacy in detecting malignant lesions dependable on previous experience. We consider that basic algorithms for non-experts should include only dermoscopic structures that are easily identified by most dermatologists. Limitations of our study are its retrospective and single-institution design.

Financial support

None declared.

Authors’ contributions

Enrique Rodríguez-Lomba: Approval of the final version of the manuscript; critical literature review; data collection, analysis, and interpretation; effective participation in research orientation; intellectual participation in propaedeutic and/or therapeutic; management of studied cases; critical manuscript review; preparation and writing of the manuscript.

Lula María Nieto-Benito: Approval of the final version of the manuscript; data collection, analysis, and interpretation; intellectual participation in propaedeutic and/or therapeutic; management of studied cases; critical manuscript review.

Elisa Hernández de la Torre: Approval of the final version of the manuscript; data collection, analysis, and interpretation; intellectual participation in propaedeutic and/or therapeutic; management of studied cases; critical manuscript review.

Ricardo Suárez-Fernández: Approval of the final version of the manuscript; data collection, analysis, and interpretation; intellectual participation in propaedeutic and/or therapeutic; management of studied cases; critical manuscript review.

José Antonio Avilés-Izquierdo: Approval of the final version of the manuscript; critical literature review; data collection, analysis, and interpretation; effective participation in research orientation; intellectual participation in propaedeutic and/or therapeutic; management of studied cases; critical manuscript review; preparation and writing of the manuscript; statistical analysis; study conception and planning.

Conflicts of interest

None declared.

References

1. 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.
2. Dolianitis C, Kelly J, Wolfe R, Simpson P. Comparative performance of 4 dermoscopic algorithms by non-experts for the diagnosis of melanocytic lesions. Arch Dermatol. 2005;141:1008–14.
3. Carrera C, Marchetti MA, Dusza SW, Argenziano G, Braun RP, Halpern AC, et al. Validity and reliability of dermoscopic criteria used to differentiate nevi from melanoma. A Web-Based International Dermoscopy Society Study. JAMA Dermatol. 2016;152:798–806.
4. Rogers T, Marino ML, Dusza SW, Bajaj S, Usatine RP, Marchetti MA, et al. A Clinical Aid for Detecting Skin Cancer: The Triage Amalgamated Dermoscopic Algorithm (TADA). J Am Board Fam Med. 2016;29:694–701.
5. Henning JS, Dusza SW, Wang SQ, Marghoob AA, Rabinovitz HS, Polsky D, et al. The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. J Am Acad Dermatol. 2007;56:45–52.

Enrique Rodríguez-Lomba*, Belén Lozano-Masedemont*, Lula María Nieto-Benito*, Elisa Hernández de la Torre*, Ricardo Suárez-Fernández*, José Antonio Avilés-Izquierdo*

*Department of Dermatology, Hospital General Universitario Gregorio Marañón, Madrid, Spain
Curettage and electrocoagulation versus surgical excision in the treatment of low-risk basal cell carcinoma—Postoperative follow-up and satisfaction at three months: randomized clinical trial

Dear Editor,

Several therapeutic approaches are currently described in the treatment of basal cell carcinoma (BCC), such as curettage and electrocoagulation (C&E), conventional surgery, Mohs micrographic surgery, topical or intralesional agents, radiotherapy, and photodynamic therapy. The choice of method depends on the BCC subtype, considering the size, location, clinicopathological pattern, and the patient clinical condition.1,2

C&E has the advantages of shorter surgical time and easier approach to the lesion, requiring less facilities and material costs for it to be performed. However, the procedure has a longer healing or recovery time and may generate more unaesthetic scars when compared to conventional surgical procedures.3,4

Although C&E has been classically described in the treatment of BCC, there are few studies prospectively comparing the technique with surgical excision. However, in observational studies, local recurrence rates for curettage and electrocoagulation have been shown to be comparable to those of conventional surgery for low-risk lesions.2,3

Therefore, the present study performed a prospective, randomized, open, and controlled comparison between C&E interventions and conventional surgery for low-risk BCC in relation to surgical complications, postoperative recovery, aesthetic appearance, and patient satisfaction three months after the procedure.

Between February 2018 and March 2019, immunocompetent patients without coagulation disorders attended at the Outpatient Clinic of the Dermatology Service of Faculdade de Medicina de Botucatu—UNESP with a diagnosis of low-risk, clinically evident BCC, measuring up to 10 mm in diameter were included in the study. The study was approved by the Research Ethics Committee of the institution (Counsel number: 4101709).

Low-risk lesions were those with well-defined edges, not a recurrence, outside of previously irradiated areas, located on the cervical region, trunk, extremities (except hands and feet), or low-risk regions on the face (forehead and cheek), with a pattern of nodular or superficial expansive growth.2

The lesions were outlined with a surgical pen, with a safety margin of 3 mm with the aid of a dermatoscope.

In the C&E group, the central portion of the lesion was excised with a tangential surgical blade incision (shaving) and sent for anatomopathological examination. The remainder of the tumor tissue was curetted until complete elimination on inspection, in addition to the surrounding skin up to a previous demarcation of three millimeters, or to the limit of friable tissue. Subsequently, the floor of the curetted area was electrocoagulated. The C&E process was repeated for a total of two cycles.4,5

In the surgical excision group, the lesion was removed using a spindle-shaped incision, following the skin tension lines, guided by the demarcation of the three-mm safety margin, closed with direct sutures, using mononylon 4.0 (trunk and limbs) and 5.0 (face) using simple interrupted stitches.6

The patients were instructed to maintain local wound care by cleaning and dressing the wound with antibiotic ointment and to have the stitches removed, in the case of conventional surgery, at the basic health unit.

The studied characteristics were: sex; age; years of schooling; smoking status; phototype; clinical type of BCC; lesion diameter; signs of tumor recurrence; postoperative infection requiring local or systemic care; suture dehiscence; significant bleeding requiring health care; clinically evident scar retraction; pain intensity according to the analog scale from 1 to 10; pruritus according to the analogous scale from 1 to 10; the impact of postoperative care on the daily routine using a Likert-type question with the options almost never, a few times, many times, almost always; satisfaction with the surgical procedure using a Likert-type question with the options: none, a little, moderate, a lot and total; and aesthetic scar result at three months using the Patient and Observer Scar Assessment Scale and an analog scale, ranging from 0 to 10.7

The sample size calculation resulted in 98 lesions and was based on a difference in maximal satisfaction rate with the expected procedure of 60% versus 85% for a power test of 80% and p ≤ 0.05, two-tailed.

The characteristics with p < 0.2 in the bivariate analyses were analyzed by multivariate mixed models adjusted for age, sex, phototype, years of schooling, smoking status, size and location of lesions, with the intervention used as the dependent variable. Two-tailed values of p ≤ 0.05 were considered significant.

A total of 116 lessons from 82 patients were included, 49% of which were female, with a mean of 1.4 lesions per patient (Table 1).