For consideration for publication in
*as an Original article*

# Achieving clinical pictures of melanocytic skin neoplasms before the histopathologic examination does not introduce a bias in the diagnostic process

*Gerardo Ferrara,[a] MD, Giorgio Annessi,[b] MD, Zsolt Argenyi,[c] MD, Giuseppe Argenziano,[d] MD, Helmuth Beltraminelli,[e] MD Rino Cerio,[f] MD, Lorenzo Cerroni,[g] MD, Carlo Cota,[h] MD, Stefano Simonetti,[i] MD, Catherine M. Stefanato,[j] MD FRCPath, Iris Zalaudek,[g] MD, Harald Kittler,[k] MD, H. Peter Soyer,[l] MD*

Conflicts of interst: none to be declared
Role of funding source: none
Word count: 1807
Figures: 2
Tables: 1
References: 28

*Running title:* Clinicopathologic correlation
*Keywords:* Melanocytic skin neoplasms - Histopathologic diagnosis - Clinicopathologic correlation - Clinical information - Dermoscopy

[a]Anatomic Pathology Unit, Gaetano Rummo General Hospital, Benevento, Italy
[b]Dermatopathology Unit, Istituto Dermopatico dell'Immacolata, Rome, Italy
[c]Department of Dermatology, University of Washington, Seattle, Seattle, U.S.A. (Prof Z Argenyi, MD),
[d]Department of Dermatology, Arcispedale 'S. Maria Nuova' Reggio Emilia, I
[e]Department of Dermatology, University of Bern, CH
[f]Department of Dermatology, University of London, London, U.K. (Prof R Cerio, MD),
[g]Department of Dermatology, Medical University of Graz, Graz, Austria
[h]Dermatopathology Unit, San Galligano Institute, Rome, Italy
[i]Department of Dermatology, University of Perugia, Perugia, Italy
[j]Department of Dermatopathology, St John's Institute of Dermatology, St Thomas' Hospital, London, U.K.
[k]Department of Dermatology, Division of General Dermatology, Medical University of Vienna, Austria
[l]Department of Dermatology Group, School of Medicine, The University of Queensland, Brisbane, Australia
**Correspondence to:** Gerardo Ferrara – Department of Oncology - Anatomic Pathology Unit – 'Gaetano Rummo' General Hospital – Via dell'Angelo 1 – I82100 Benevento – ITALY
Phone +39082457315
Fax +39082457334
e-mail: gerardo.ferrara@libero.it

## Abstract

**Background –** The *a priori* knowledge of the clinical picture of melanocytic skin neoplasms (MSN) introduces a potential bias in the histopathologic examination.

**Methods –** Histologic slides from 99 MSN were circulated among ten histopathologists with clinical expertise: five histopathologists had clinical images available after a 'blind' examination (Group 1); the other five had clinical images available before microscopic examination (Group 2). Data from the two groups were compared regarding consensus diagnosis (CD: a diagnosis in agreement by ≥4 histopathologists/group), chance-corrected interobserver agreement (Fleiss' *k*), and level of diagnostic confidence (LDC: a 1 to 5 arbitrary scale indicating 'increasing reliability' of any given diagnosis).

**Results** – Compared with Group 1 histopathologists, Group 2 achieved a lower number of CD (84 vs. 90) but a higher *k* value (0.74 vs. 0.69) and a greater mean LDC value ( 4.57 vs 4.32). The same CD was rendered by the two groups in 81/99 cases. Spitzoid MSN were most frequently controversial for both groups.

**Conclusions –** The histopathologic interpretation of MSN seems to be not influenced by the knowledge of the clinical picture before histopathologic examination..

## Introduction

Whether any diagnostic test should be read together with clinical information has been debated since 1963, when Schreiber[1] suggested the clinical information as a way to improve the accuracy of chest X-ray evaluation. Nonetheless, the impact of the clinical information in diagnostic cytohistopathology has been addressed by very few studies.[2-5] A commonly used argument against the 'interdisciplinary' approach is that clinical information may bias the reading:[6-7] in order to minimize such a bias, the microscopic examination should be carried out first blind to the clinical information and then in light of them:[4,5,8] thus, *perception* (identification of abnormal areas and their features) would result unbiased by the clinical information, whereas the latter should help the final *interpretation* (attribution of the abnormalities to an entity). By following such a methodology, two recent studies performed in the field of dermatopathology have shown that the dermatopathologic diagnosis starts as a perception of microscopic criteria which can work as such, but is finally a clinically-aided interpretation.[4-5] A major problem, however, is the potential bias born by an *a priori* knowledge of the clinical picture, as happens when clinical dermatologists evaluate the histologic specimens from their own patients.[7,9-11]

In a previous study,[4] the histologic slides from 99 clinically atypical melanocytic skin neoplasm (MSN) were submitted to ten histopathologists, five of these with clinical expertise ('clinical histopathologists'); in order to evaluate the diagnostic impact of the single clinical data, clinical information were given with a five-step procedure (no information; age/sex/location; clinical diagnosis; clinical image; dermoscopic image). Steps 1-3 of such a procedure excluded the bias of the knowledge of the clinical picture before histopathology. We now introduced such a bias by submitting the same dataset from the previous study[4] to five clinical histopathologists who were requested to look at the histologic slides only after having all pertinent clinical information. The new data were compared with those given by the former five clinical histopathologists.

## Methods

After an *ad hoc* authorization obtained by the patients or their guardians, 99 consecutive cases of clinically/dermoscopically atypical MSN were submitted to two groups of clinical histopathologists:

***Group 1*** (ZA, RC, LC, HK, HPS): histopathologists having a stepwise access to clinical information in the course of their microscopic evaluation;

***Group 2*** (GAn, HB, CC, SS, CMS): histopathologists having full knowledge of the clinical picture before the microscopic examination.

Each panelist is almost equally involved in routine clinical and histopathologic work. Only two of the above panelists (LC and HPS) had worked for several years at the same Institution.

For each case, a single hematoxylin-eosin stained slide, accurately checked for its technical and diagnostic adequacy, was provided to each panelist.

All the clinical information concerning the selected cases was included into a FileMaker Pro 7$^{TM}$ (FileMaker Inc.)-generated database. For each case, Group 1 histopathologists were requested to evaluate the microscopic slide according to a five-step procedure: i) no information; ii) knowledge of age and gender of the patient and location of the lesion; iii) clinical diagnosis; iv) clinical image; v) dermoscopic image. Group 2 histopathologists had all information available before microscopic examination.

The influence derived on the diagnosis of MSN from the knowledge of the full clinical information before the histopathologic examination was checked by comparing the data provided by Group 2 histopathologists with the data provided by Group 1 histopathologists. The parameters evaluated were: the consensus diagnosis (CD), the chance-corrected interobserver agreement, and the mean level of diagnostic confidence (LDC). For statistical analysis, all the diagnoses were grouped into two ratings: ‘melanoma’ and ‘nevus’.

CD was defined as a diagnosis made in agreement by at least four out of five panelists per group. This have been recently proposed as a 'surrogate' gold standard when follow-up data can give little information (as in this case: see the Results section).[12,13] Unanimous diagnoses were a subgroup of CDs; the comparison between the number of both unanimous diagnoses and CDs given by Group 1 and Group 2 histopathologists was performed with McNemar's test.[14] The given *p*-value is one-tailed and a *p*-value of <0.05 indicates statistical significance.

Case by case, the histopathologists were asked to log into the database also a *level of diagnostic confidence* (LDC), namely, the probability, as scored according to an arbitrary scale, that they 'subjectively' attributed to the given diagnosis.[15] The LDC scale was structured into five levels:

***LDC 1*** – No diagnostic certainty: no diagnosis can be made.

***LDC 2*** – Low diagnostic certainty: a diagnosis is felt as slightly more likely.

***LDC 3*** – Moderate diagnostic certainty: a diagnosis is favoured, but with some elements of doubt.

***LDC 4*** – High diagnostic certainty: a diagnosis is strongly favoured.

***LDC 5*** – Absolute diagnostic certainty: no other diagnosis is possible.

The interobserver agreement among the observers was calculated using the k statistics for multiple ratings introduced by Fleiss.[16-18] Given $n$=5 as the number of panelists, $k$ values range between +1 (perfect agreement) and <0 (perfect disagreement); values greater than 0.75 represent an excellent agreement; values lower than 0.40 a poor For the comparison of LDC values a Normal z-test for dependent samples was used.[19] The given *p*-value is 2-tailed and a *p*-value <0.05 indicates statistical significance.

## Results

The study included 99 cases from 96 patients (M:F=0.6:1; age range: 10-78 years; mean age: 43.3 years; median age: 42 years). The most common location was the back (40 cases), followed by the lower limbs (16 cases) and by the upper limbs (12 cases). The referring histopathologist had diagnosed 54 cases as ‘nevus’ and 45 cases as ‘melanoma’. Follow-up (range: 24-96 months; mean: 54.5 months) was available in 65 cases, 30 of which originally diagnosed as melanoma. Ten of these cases underwent sentinel node biopsy, which was negative in all instances; an adverse clinical outcome was recorded in five cases (four sentinel node negative cases in which distant metastases developed; one 0.80 mm-thick melanoma in which a nodal metastasis detected 40 months after excision). All the metastasizing cases received a CD of melanoma from both groups of observers.

Within Group 1 histopathologists, 65 cases had a unanimous diagnosis and 25 cases had only one discrepant diagnosis, with a CD obtained in 90/99 cases. The *k* value was 0.69 and the LDC 4.32±0.59.

Group 2 histopathologists diagnosed 75/99 MSN in unanimity and 9 MSN with only one discrepant diagnosis: therefore, a CD was achieved in 84/99 cases. The *k* value was 0.74; the LDC was 4.57±0.39. Thus, compared with data from Group 1, there was a lower number of CD, but a greater chance-corrected interobserver agreement and a greater mean LDC value. Remarkably, the differences among the number of unanimous diagnoses as well as the number of CDs generated by the two groups was not statistically significant (McNemar’s test: $p$=0.078 for the differences in the number of unanimous diagnoses; $p$=0.18 for the differences in the number of CDs), whereas the difference between the mean LDC values was highly significant (Normal z-test $p$<0.001).

By comparing case by case the CD generated by the two groups of observers, 81 cases resulted with the same CD; five cases were controversial for both groups; six cases

were controversial only for Group 1; and seven cases were controversial only for Group 2. The five most controversial cases showed Spitzoid (3/5; Fig. 1) or lentiginous (2/5; Fig. 2) features. Overall, the most common morphologic patterns of controversial cases (Table 1) were Spitzoid (8/18 cases), lentiginous (3/18), and regression(-like). Among these controversial cases, there was no association between a given morphologic pattern and the protocol of microscopic observation followed.

## Discussion

This study demonstrates that the histopathologist who knows the clinical picture of MSN before the histopathologic examination is not biased in the final interpretation and is actually strengthened in the diagnostic process.

It has been demonstrated that full clinical information provided after a first microscopic examination accomplished blind to the clinical data can aid the final interpretation for both MSN[4] and inflammatory dermatoses.[5] It is still disputable, however, whether a bias is introduced in the diagnostic process when the full knowledge of the clinical picture is achieved before histopathologic examination. This issue also involves the neverending debate regarding the histopathologic practice by dermatologists and the level of training required them to sign out dermatopathology cases.[7,20]

We have compared some parameters (CDs, chance-corrected interobserver agreement, mean LDC) provided in a series of 99 MSN by two groups of histopathologists with clinical expertise: one of these (Group 1) had clinico-dermoscopic images available before the microscopic examination, whereas the other (Group 2) had all clinical information available *a priori*. Within the Group 2 there was a lower number of CDs: this finding, although not statistically significant, could suggest that a *perception* influenced by the knowledge of the clinical picture does necessarily translate into a more homogeneous *interpretation*. Therefore, the extreme scenario of a clinical prejudice forcing the histopathologic diagnosis[21] is unlikely. Instead, the *k* value generated by Group 2 histopathologists was greater than Group 1: this means that the influence of the clinical pictures is similar, regardless the moment in which they are introduced into the diagnostic process.

A further interesting finding is the greater LDC found within Group 2 histopathologists, a finding which was statistically significant. This could imply that the clinical picture, albeit being unable to bias the final *interpretation*, can aid the *perception*

and give greater strength to the diagnostic process. Parenthetically, among controversial cases, there was no association between a given morphologic pattern and the protocol of microscopic observation followed. As expected, controversies in Spitzoid MSN[4,22] (an example is given in Fig. 1) were found to be sizable, independent from the moment in which the histopathologist is aware of the clinical picture. Along with the results of our previous study,[4] we also found some diagnostic controversy regarding lentiginous MSN (an example is given in Fig. 2): all the disputable lesions of this category were removed from the back of middle-aged patients (data not shown). It has been previously demonstrated that the differential diagnosis between lentiginous 'dysplastic' nevus and lentiginous melanoma can be significantly aided by the clinico-dermoscopic digital monitoring of the lesions, since lentiginous melanoma continuously, but slowly grows and remains 'in situ' over years or even decades.[23-25] In the present study, however, data regarding the 'E' (=evolution) criterion of the ABCDE clinical alphabet of melanoma[26] were not provided, and this must be underlined as a limitation in the clinicopathologic evaluation made by the panelists. Lesions with regression(-like) features[24,27,28] have been found somewhat controversial, but to a lesser extent than expected. This can be due to the fact that all the MSN included in the present study had undergone macroscopic sampling according to their dermoscopic features, a procedure which can help the diagnostic evaluation of these lesions by highlighting their most atypical features.[24]

In conclusion, there is evidence that clinical pictures of MSN can aid their *interpretation* and that the latter is not biased if the clinical pictures are available before the microscopic examination. Good lines of communication between dermatologists and histopathologists are always desirable; when the dermatologist and the histopathologist are the same person, the communication is best and, if intellectually honest, is probably not misleading.

## References

1. Schreiber MH. The clinical history as a factor in roentgenogram interpretation. JAMA 1983;185:399.

2. Abbey LM, Kaugars GE, Gunsolley JC, et al. The effect of the clinical information on the histopathologic diagnosis of oral epithelial dysplasia. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1998;85:74.

3. Raab SS, Oweity T, Hughes JH, et al. Effect of clinical history on diagnostic accuracy in the cytologic interpretation of bronchial brush specimens. Am J Clin Pathol 2000;114:78-83.

4. Ferrara G, Argenyi Z, Argenziano G, et al. The influence of the clinical information in the histopathologic diagnosis of melanocytic skin neoplasms. PLoSONE 2009; 4: e5375. doi:10.1371/journal.pone.0005375. Epub 2009 Apr 30.

5. Cerroni L, Argenyi Z, Cerio R, et al. Influence of evaluation of clinical pictures on the histopathologic diagnosis of inflammatory skin disorders. J Am Acad Dermatol 2010; 63:647.

6. Shitabata PK. Do clinical histories submitted on biopsy requisition sheets bias the Pathologists? http://www.dermpathmd.com

7. Ackerman AB (2005) Dermatologist not equal dermatopathologist: no place in a profession for pretenders. J Am Acad Dermatol 2005; 53:698.

8. Griscom NT (2002) A suggestion: look at the images first, before you read the history. Radiology 223:9-10.

9. Moy R, The reason that dermatologists should not send all their slides to dermatopathologists is a scope of practice argument, not an ethical argument. J Am Acad Dermatol 2005, 53:700.

10. Glogau RG, Collegiality and dermatology. J Am Acad Dermatol 2005, 53:701.

11. Ferrara G, Crisman G. F1000 Prime Recommendation of [Cerroni L, Argenyi Z, Cerio R, et al. Influence of evaluation of clinical pictures on the histopathologic diagnosis of inflammatory skin disorders. J Am Acad Dermatol 2010; 63:647]. http://www.f1000.com/prime/5557956#eval5525054

12. Coggon D, Martyn C, Palmer KT, Evanoff B. Assessing case definitions in the absence of a diagnostic gold standard. Int J Epidemiol 2005;34:949.

13. Ferrara G, Tomasini C, Argenziano G, Zalaudek I, Stefanato CM. Small-diameter melanoma: toward a conceptual and practical reappraisal. J Cutan Pathol 2012; 39:721.

14. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages". *Psychometrika* 1947;12: 153.

15. Meghini C, Sebastiani F, Straccia U, Thanos C. A model of information retrieval based on a terminological logic. In: Proceedings of SIGIR-93, 16th ACM International Conference on Research and Development in Information Retrieval. Pittsburgh, PA, 1993;298.

16. Fleiss J L, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. Psychol Bull 1969; 72:323.

17. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull 1971; 76:378.

18. Fleiss JL. Statistical Methods for Rates and Proportions . 2nd ed. New York, Wiley, 1981:2.

19. D'Ambra L. Verifica delle ipotesi. In: Lezioni di Inferenza Statistica. Rocco Curto Editore, Naples, 2007;232.

20. Glusac EJ. Under the microscope: doctors, lawyers, and melanocytic neoplasms. J Cutan Pathol 2003;30:287.

21. Grant-Kels JM. The whys and the wherefores of who reads

dermatopathologic slides. J Am Acad Dermatol 2005; 53:703.

22. Barnhill R The Spitzoid lesion: the importance of atypical variants and risk assessment. Am J Dermatopathol 2006; 28:75.

23. Ferrara G, Zalaudek I, Argenziano G. Lentiginous melanoma: a distinctive clinicopathological entity. Histopathology 2008; 52: 523.

24. Ferrara G, Argenziano G, Giorgio CM, Zalaudek I, Kittler H. Dermoscopic-pathologic correlation: apropos of six equivocal cases. Semin Cutan Med Surg 2009; 28: 157.

25. Terushkin V, Dusza SW, Scope A, et al. Changes observed in slow growing melanomas during long-term dermoscopic monitoring. Br J Dermatol 2012; 166:1213.

26. Abbasi NR, Shaw HM, Rigel DS. Early diagnosis of cutaneous melanoma. Revisiting the ABCD criteria. JAMA 2004; 292: 2771.

27. Ferrara G, Argenziano G, Soyer HP, et al. Dermoscopic and histopathologic diagnosis of equivocal melanocytic skin lesions. An interdisciplinary study on 107 cases. Cancer 2002;95:1094.

28. Zalaudek I, Argenziano G, Ferrara G, et al. Clinically equivocal melanocytic skin lesions withy features of regression: a dermoscopic-pathological study. Br J Dermatol 2004; 150:64.

290

291 Table 1 – Morphologic features of the most controversial cases of the present study
292 distributed according to the group

293

| **Panelists** | **Main morphologic pattern** | | | | | | |
|---|---|---|---|---|---|---|---|
| | *Spitzoid* | *Lentiginous* | *Regression(-like)* | *Nested* | *Halo* | *Congenital nevus-like* | Total |
| All | 3 | 2 | 0 | 0 | 0 | 0 | 5 |
| Group 1 | 2 | 0 | 2 | 1 | 1 | 0 | 6 |
| Group 2 | 3 | 1 | 1 | 1 | 0 | 1 | 7 |
| Total | 8 | 3 | 3 | 2 | 1 | 1 | 18 |

294

## Figure legends

**Figure 1:** A melanocytic lesion of the thigh in a 58-year-old woman. Dermoscopy is characterized by a striking asymmetry, with a reticular depigmentation associated with a dotted vascular pattern in its upper portion (A). Histopathologically, the neoplasm is compound, with a moderate epidermal hyperplasia (B), lack of circumscription (C), and with a Spitzoid cytomorphology (D). The differential diagnosis is between a Spitz nevus and an early invasive Spitzoid melanoma.

**Figure 2:** A flat melanocytic lesion of the back in a 58-year-old man. Dermoscopically there is an atypical pigment network with features of regression (A); histopathologically, le lesion is large, with partially preserved retiform epidermal hyperplasia (B), a prevailingly single cell proliferation at the junction, involvement of the adnexal epithelium (C) and cytologic features of ‘severe dysplasia’ (D). The differential diagnosis is between a ‘severely dysplastic’ (Clark) nevus and a lentiginous melanoma in situ.[23]