First Prospective Study of the Recognition Process of Melanoma in Dermatological Practice

Julie Gachon, MD; Philippe Beaulieu, MD; Jean Francois Sei, MD; Johanny Gouvernet, MD; Jean Paul Claudel, MD; Michel Lemaitre, MD; Marie Aleth Richard, MD; Jean Jacques Grob, MD

Background: Early detection is crucial to improve melanoma prognosis. Different diagnostic guides such as the ABCD rule (asymmetry [A], irregularity of borders [B], unevenness of distribution of color [C], and diameter [D]) have been proposed to identify melanoma, but their efficacy in real life is questionable. We investigated the recognition process of melanoma by dermatologists to use as a model to improve self-detection in the general population and to train students and general practitioners.

Objectives: To understand the major principles of the recognition process of nevi and melanoma unconsciously used by dermatologists.

Design: Prospective survey recording the immediate perceptions of dermatologists of the morphologic features of the lesion and intuitive diagnostic opinion about 4036 consecutive resected nevi and melanoma.

Setting: One hundred thirty-five volunteer dermatologists in their daily practices.

Main Outcome Measures: Perceptions of the image best explaining the diagnostic opinion and best predicting the final diagnosis by univariate and multivariate analysis.

Results: The immediate diagnostic opinion of the dermatologist is mainly explained by an unconscious reference to the overall pattern compared with the common nevi, but also compared with the other nevi of the individual (the “ugly duckling sign”). The dermatologist’s ability to discriminate between nevi and melanoma relies on the assessment of the overall pattern, the ugly duckling sign, and the knowledge of a recent change. A separate or combined analysis of individual morphologic criteria such as ABCD does not seem to play a major role in this recognition process.

Conclusions: Persons most skilled at the clinical detection of melanoma seem to unconsciously rely on cognitive (overall pattern) and comparative (ugly duckling sign) processes rather than an algorithm of morphologic criteria (ABCD). These concepts could be tested in the medical training of general practitioners and education of the general population, where they might be more efficient than algorithms such as the ABCD criteria.

Arch Dermatol. 2005;141:434-438
ers and medical students to screen for MM, the challenge is to find an easier, more natural, and more efficient way to teach individuals how to recognize an MM among thousands of benign pigmented lesions. A first step is to understand how the natural recognition process of MM works, especially in people who have the best accuracy in this field, ie, the dermatologists.11-14 Very few studies have investigated the recognition process of pigmented lesions by physicians,15,16 and these were based on slide presentations rather than clinical situations.

Image recognition by humans is a complex process that ends up in the classification of an object. Commonly, recognition of an image consists of associating this image with a memory of an object called a model or a pattern, which we have built from a collection of previous images of this object stored in our visual experience.17 In the learning process, a teacher specifies a label for this object class, eg, “This is a car,” but does not need to describe it, because any individual is able to build his own recognition pattern for this object. This overall pattern recognition (OPR) is the cognitive way we learn everything around us. The identification of some objects, often those that are not a close part of our life, may involve an analytic criteria recognition (ACR) process combining a series of criteria that are listed during a supervised learning process. The teacher describes an object (eg, a rhinoceros) by a list of criteria (large animal, gray, and with horns), which unfortunately is often not specific and also applies to other objects (eg, a water buffalo). This is the principle of a checklist like ABCD. Recently, attention has been drawn to the importance of a differential recognition (DR) process in MM diagnosis, under the label of the “ugly duckling sign.”18 It is a common observation that, in a given individual, all the nevi tend to have the same pattern. A nevus that does not belong to this pattern, the ugly duckling, must be considered with suspicion, even if it does not fulfill the ABCD criteria. Conversely, a very atypical ABCD lesion may be completely normal in an individual who is covered with similar lesions.

To improve the way we teach self-detection in the population and screening among general practitioners, we hypothesized that the best solution would be to copy the efficient recognition process of experts. We thus designed a prospective survey to investigate the unconscious process of recognition used by dermatologists in their daily practice and to assess the respective importance of the OPR, ACR, and DR processes.

The first objective was to assess how much the diagnostic opinion about a nevus is based on an OPR, an ACR, or a DR process and how much this opinion can be modulated by subjective data provided by the patient. The second objective was to identify which among these different processes is the most relevant to predicting the final histological diagnosis.

DATA COLLECTION

The following data were collected on a standardized questionnaire form that each volunteer dermatologist agreed to complete immediately whenever he or she decided to remove a nevus or MM for any reason (eg, suspicion of MM, aesthetics, comfort, prevention). First, the dermatologists had to record on the form their diagnostic opinion, ie, their intuitive immediate estimation of the probability that the lesion was an MM at the very time of the removal decision, with no possibility of changing. This opinion was scored on a visual analog scale from 0% (MM totally excluded) to 100% (complete certainty of MM). This score is designated as the immediate estimation of probability of MM (IEMP) in the rest of the study.

To assess the role of the OPR process, the dermatologist was first invited to classify the nevus by reference to those nevi stored in his or her experience with the following: Referring to your experience of nevi, how would you qualify this lesion (0-100 on a visual analog scale)? Responses included completely regular to totally irregular for a melanocytic lesion; simple aspect to highly complex aspect for a melanocytic lesion; and symmetric to completely asymmetric for a melanocytic lesion.

To assess the role of the ACR process, the dermatologist was asked to quote the A, B, and C criteria on 3 visual analog scales from 0 to 100. Responses ranged from perfectly symmetric to completely asymmetric shape (A), perfectly regular to completely irregular borders (B), and unique and perfectly homogeneous to multiple and totally uneven color (C).

To assess the role of the intranindividu comparative process (ugly duckling sign), the dermatologist was asked to assess how different the lesion was from the other nevi in the same individual. Responses included completely similar to the other nevi of the patient, some similarities with several nevi of the patient, somewhat different from most nevi of the patient, and completely different from the other nevi of the patient.

To assess the role of subjective information provided by the patient, the dermatologists had to note whether or not the nevus had changed, according to the patient. Finally, they had to fulfill a detailed description of the nevus (eg, size, shape). Later, after receiving the pathological report, a second part of the form was completed.

Several precautions were taken to limit biases in the responses. To ensure the reliability and to avoid temptation by the dermatologist to correct the first immediate impression after receiving the histopathological report, dermatologists had to be convinced that an individual assessment would not be possible; therefore, they were informed that the name of the physician no longer appeared, once it had been sent to the data center. As spontaneity in the response was crucial, the dermatologists were not aware of the exact target of the study. They were informed that the purpose was to assess diagnostic accuracy, but they did not know that the diagnostic process itself was evaluated. The thematic field (OPR, ACR, and DR) explored by each question was not mentioned. To limit interference between the different responses, the questions always had to be answered in the same order, ie, first, immediate diagnostic presumption and reasons for removal; second, overall assessment (OPR); third, record of criteria (ACR); fourth, comparison with other nevi (DR); fifth, recent change; and sixth, detailed description.

METHODS

DESIGN AND OBJECTIVE

This prospective study was conducted in volunteer dermatologists during their daily work. We investigated how dermatologists establish their diagnostic opinion about a pigmented nevus or MM. This study was thus conceived to collect their daily practice and to assess the respective importance to predicting the final histological diagnosis.

The following data were collected on a standardized questionnaire form that each volunteer dermatologist agreed to complete immediately whenever he or she decided to remove a nevus or MM for any reason (eg, suspicion of MM, aesthetics, comfort, prevention). First, the dermatologists had to record on the form their diagnostic opinion, ie, their intuitive immediate estimation of the probability that the lesion was an MM at the very time of the removal decision, with no possibility of changing. This opinion was scored on a visual analog scale from 0% (MM totally excluded) to 100% (complete certainty of MM). This score is designated as the immediate estimation of probability of MM (IEMP) in the rest of the study.

To assess the role of the OPR process, the dermatologist was first invited to classify the nevus by reference to those nevi stored in his or her experience with the following: Referring to your experience of nevi, how would you qualify this lesion (0-100 on a visual analog scale)? Responses included completely regular to totally irregular for a melanocytic lesion; simple aspect to highly complex aspect for a melanocytic lesion; and symmetric to completely asymmetric for a melanocytic lesion.

To assess the role of the ACR process, the dermatologist was asked to quote the A, B, and C criteria on 3 visual analog scales from 0 to 100. Responses ranged from perfectly symmetric to completely asymmetric shape (A), perfectly regular to completely irregular borders (B), and unique and perfectly homogeneous to multiple and totally uneven color (C).

To assess the role of the intranindividu comparative process (ugly duckling sign), the dermatologist was asked to assess how different the lesion was from the other nevi in the same individual. Responses included completely similar to the other nevi of the patient, some similarities with several nevi of the patient, somewhat different from most nevi of the patient, and completely different from the other nevi of the patient.

To assess the role of subjective information provided by the patient, the dermatologists had to note whether or not the nevus had changed, according to the patient. Finally, they had to fulfill a detailed description of the nevus (eg, size, shape). Later, after receiving the pathological report, a second part of the form was completed.

Several precautions were taken to limit biases in the responses. To ensure the reliability and to avoid temptation by the dermatologist to correct the first immediate impression after receiving the histopathological report, dermatologists had to be convinced that an individual assessment would not be possible; therefore, they were informed that the name of the physician no longer appeared, once it had been sent to the data center. As spontaneity in the response was crucial, the dermatologists were not aware of the exact target of the study. They were informed that the purpose was to assess diagnostic accuracy, but they did not know that the diagnostic process itself was evaluated. The thematic field (OPR, ACR, and DR) explored by each question was not mentioned. To limit interference between the different responses, the questions always had to be answered in the same order, ie, first, immediate diagnostic presumption and reasons for removal; second, overall assessment (OPR); third, record of criteria (ACR); fourth, comparison with other nevi (DR); fifth, recent change; and sixth, detailed description.
the use of the Mann-Whitney or Kruskal-Wallis test.

than 80%. We assessed comparisons of scores of 1 to 100 with

among community dermatologists working in a private setting,

According to the patient. Multivariate analysis included an ascending stepwise logistic regression to predict an IEPM of greater

best account for the subjective diagnostic opinion of the derma-
tologist, univariate and multivariate analyses were conducted for

Table 1. Ascending Stepwise Multivariate
Linear Model to Predict the IEPM*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Standardized Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without weight</td>
<td></td>
</tr>
<tr>
<td>Overall complexity</td>
<td>0.29</td>
</tr>
<tr>
<td>Overall irregularity</td>
<td>0.24</td>
</tr>
<tr>
<td>“Ugly duckling sign”</td>
<td>0.12</td>
</tr>
<tr>
<td>C criterion</td>
<td>0.11</td>
</tr>
<tr>
<td>B criterion</td>
<td>0.11</td>
</tr>
<tr>
<td>Weight†</td>
<td></td>
</tr>
<tr>
<td>Overall irregularity</td>
<td>0.36</td>
</tr>
<tr>
<td>Ugly duckling sign</td>
<td>0.18</td>
</tr>
<tr>
<td>C criterion</td>
<td>0.16</td>
</tr>
<tr>
<td>Overall asymmetry</td>
<td>0.09</td>
</tr>
<tr>
<td>Overall complexity</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Abbreviations: ABCD criteria, asymmetry (A), irregularity of borders (B), unevenness of color (C), and diameter (<6 or ≥6 mm) (D); IEPM, immediate estimation of probability of melanoma.

*Candidate variables were overall irregularity, overall asymmetry, overall complexity, the ugly duckling sign, the ABCD criteria, and evolution according to the patient.
†Calculated by the following formula: \[(\text{IEPM}/10 + 1) \times (\text{IEPM} - 50)^2\].

DATA MANAGEMENT AND STATISTICS

Clinical Perceptions Accounting for the Diagnostic Opinion of Dermatologists

To determine which of the data available at clinical examination best account for the subjective diagnostic opinion of the dermatologist, univariate and multivariate analyses were conducted for prediction of the IEPM with the following variables: overall irregularity; overall asymmetry; overall complexity; the ugly duckling sign; intensity of the A, B, and C criteria and diameter (<6 and ≥6 mm) (D criterion); and evolution according to the patient. Multivariate analysis included an ascending stepwise linear regression, with a weight of 1 and a weight equal to \[(\text{IEPM}/10 + 1) \times (\text{IEPM} - 50)^2\] to force a good prediction of marked suspicions (high and low value of IEPM) that were few, and an ascending stepwise logistic regression to predict an IEPM of greater than 80%. We assessed comparisons of scores of 1 to 100 with the use of the Mann-Whitney or Kruskal-Wallis test.

Clinical Perceptions Best Predicting the Final Diagnosis

To determine which of the data available at clinical examination were the most informative to make a correct diagnosis of MM, we conducted an ascending stepwise logistic regression to predict a histologically confirmed MM with the following candidate variables: overall irregularity; overall asymmetry; overall complexity; ugly duckling sign; intensity of the A, B, and C criteria and diameter (<6 or ≥6 mm) (D criterion); and evolution according to the patient.

RESULTS

DESCRIPTIVE DATA

Among the 200 volunteers, 135 dermatologists actually participated in the study (mean age, 47 years). Most were community dermatologists working in a private setting, and only 2 were academic dermatologists. They removed 4036 successive melanocytic nevi and MM, which represent a mean of 3.75 nevi per week (range, 1-16) and 4.31 MM per year (range, 1-30). The reasons for removal of nevi were as follows: 1634 (40.7%) because of aesthetic or functional reasons, 535 (13.3%) “only to reassure the patient,” 1199 (29.7%) because they were considered suspicious by the dermatologist, and 869 (21.5%) because they were considered as precursors by the dermatologist. These different reasons accounted for 2 (1.3%), 9 (6.0%), 141 (94.6%), and 14 (9.4%), respectively, of the 149 MM during the study. The final diagnoses of the 4036 lesions at histopathological examination included 149 MMs (3.7%), including 36 (24.1%) of 149 MM in situ and other invasive lesions with a median Breslow thickness of 0.60 mm; 3629 nevi (89.9%); 4 uncertain MMs/nevi (0.1%); and 254 nonmelanocytic lesions clinically considered to be nevi or MMs (6.3%).

UNIVARIATE ANALYSIS

We sought to answer what best accounts for the clinical opinion of dermatologists in a univariate study with the data available at clinical examination. The IEPM was shown to increase with increasing degree of each of the morphologic analytic criteria, including asymmetry (A) (P<.001), irregularity of borders (B) (P<.001), unevenness of distribution of color (C) (P<.001), and the dimension (<6 or ≥6 mm in diameter) (P<.001). The IEPM also tended to increase with an increasing impression of overall irregularity (P<.001), overall complexity (P<.001) (data not shown), and overall asymmetry (P<.001) (data not shown). The IEPM increased with an increasing impression of difference between the lesion removed and most of the other nevi in the same subject (ugly duckling sign) (P<.001) and with the knowledge of a recent change of the lesion according to the patient (P<.001).

MULTIVARIATE ANALYSIS

The results of the multivariate linear model to predict the IEPM show that the estimation of overall complexity and overall irregularity were retained first; when we used a weight to have a better prediction of low and high values of the IEPM, overall irregularity was retained far before the ugly duckling sign, unevenness of color, and impression of overall complexity (Table 1).

We used the multivariate logistic model to predict the best accounting for a strong presumption or conviction of MM, defined by an IEPM of greater than 80%. In this analysis, the impression of overall irregularity was also retained as the most informative variable, before the ugly duckling sign, a diameter of 6 mm or greater, and the A and D criteria (Table 2). When the individual A, B, C, and D criteria were forced into the model separately or in combination, the same variables were retained first (data not shown).

We used a multivariate logistic model to determine the best predictors of the final pathological diagnosis. Overall irregularity, the ugly duckling sign, and recent change according to the patient were retained first (Table 3). Indicative predictive values of each of the variables retained by the model are given in Table 4. When the lo-
This study provides the first insight in the recognition tool operating in the brain of an expert facing a pigmented lesion in daily practice. The intuitive diagnostic opinion of a dermatologist who is examining a nevus with the underlying question—nevus or MM?—is explained much more by references to an OPR than an ACR process (Table 1). Among the clinical perceptions that induce a high suspicion of MM in experts, the impression of an overall irregularity is by far the most important, before the ugly duckling sign and the size (Table 2), again underlining the role of the OPR and DR processes. Although analytic criteria such A, B, C, and D had a close link with the dermatologist’s opinion, they were not retained separately or in combination in the first ranks in the multivariate predictive models.

Among the different perceptions of a lesion by a dermatologist, those most relevant to making an accurate diagnosis of MM seem to be the feeling that this lesion is overall irregular by reference to the usual nevi, the perception of an ugly duckling by reference to the other nevi in the subject, and, to a lesser degree, the knowledge that the lesion has recently changed (Table 3). Although other perceptions of the tumor’s morphologic features, such as the analytic A, C, or D criteria, had rather good predictive values, the multivariate model shows that their inference to the correct diagnosis, regardless of whether they are used separately or in a combination, was lower than the perception of overall irregularity and the ugly duckling sign. The keys for an efficacious recognition of MM are thus probably the OPR and DR processes and the ability of the physician to listen to the patient.

**Table 2. Ascending Stepwise Multivariate Logistic Model to Predict an IEPM Greater Than 80%**

<table>
<thead>
<tr>
<th>Variable</th>
<th>P Value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall irregularity</td>
<td>&lt;.001</td>
<td>10.42 (4.18-25.60)†</td>
</tr>
<tr>
<td>“Ugly duckling sign”</td>
<td>&lt;.001</td>
<td>3.30 (1.74-6.26)</td>
</tr>
<tr>
<td>D criterion (≥6 mm)</td>
<td>&lt;.001</td>
<td>3.42 (1.64-7.12)</td>
</tr>
<tr>
<td>A criterion</td>
<td>.001</td>
<td>3.27 (1.49-7.46)‡</td>
</tr>
<tr>
<td>Evolution according to the patient</td>
<td>.02</td>
<td>1.77 (1.07-2.19)</td>
</tr>
</tbody>
</table>

**Table 3. Ascending Stepwise Multivariate Logistic Model to Predict Histologically Confirmed Melanoma**

<table>
<thead>
<tr>
<th>Variable</th>
<th>P Value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall irregularity†</td>
<td>&lt;.001</td>
<td>3.27 (1.64-6.45)†</td>
</tr>
<tr>
<td>“Ugly duckling sign”†</td>
<td>&lt;.001</td>
<td>3.50 (2.18-5.62)</td>
</tr>
<tr>
<td>Evolution according to the patient†</td>
<td>&lt;.001</td>
<td>2.69 (1.81-4.01)</td>
</tr>
<tr>
<td>C criterion</td>
<td>.002</td>
<td>2.00 (1.16-3.44)‡</td>
</tr>
<tr>
<td>D criterion (≥6 mm)†</td>
<td>.01</td>
<td>1.74 (1.11-2.72)</td>
</tr>
<tr>
<td>A criterion</td>
<td>.03</td>
<td>1.91 (1.05-3.27)‡</td>
</tr>
</tbody>
</table>

**Table 4. Indicative Predictive Values for the Criteria Retained in the Multivariate Logistic Model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall irregularity</td>
<td>0.107*</td>
<td>0.987*</td>
</tr>
<tr>
<td>“Ugly duckling sign”</td>
<td>0.090</td>
<td>0.972</td>
</tr>
<tr>
<td>Evolution according to the patient</td>
<td>0.074</td>
<td>0.982</td>
</tr>
<tr>
<td>C criterion</td>
<td>0.092*</td>
<td>0.997*</td>
</tr>
<tr>
<td>D criterion (≥6 mm)</td>
<td>0.080</td>
<td>0.985</td>
</tr>
<tr>
<td>A criterion</td>
<td>0.110*</td>
<td>0.986*</td>
</tr>
</tbody>
</table>

These results are not really surprising for 2 reasons. First, the general principles of image recognition by humans are verified in this specific example. Recognition of an object involves associating this image with a memory of a pattern, which is a personal template that any individual builds from his or her visual experience. Second, the discriminative properties of separate criteria have been rather optimistically assessed, by testing ABCD rules in artificial subgroups in which the incidence of MM is rather optimistically assessed, by testing ABCD rules in artificial subgroups in which the incidence of MM is much higher than in the general population. Most experts know that the ABCD rule is not specific or sensitive for MM. Many atypical nevi and even seborrheic keratoses respond to ABC criteria, whereas nodular MM, for instance, rarely fulfill these criteria. Third, computer-based diagnostics that use a combination of criteria are unable to be more accurate than physicians are, and new research on artificial intelligence is trying to approach the cognitive recognition process.

Whether we can really identify several different intellectual processes of image recognition such as OPR, ACR, and DR through a simple questionnaire is worth discussing. We recorded immediate individual perceptions of the image. For example, responses to questions scoring the degree of overall irregularity and complexity of a lesion may not mean the same thing for each dermatologist, but they are likely to always refer to the same general concept of OPR, a cognitive approach of the whole lesion. It is also...
likely that 2 dermatologists will not perceive, understand, and rate colors, borders, or asymmetry in the same way. However, whatever their conception, their response will refer to the same concept of ACR, an analysis of separate morphologic criteria. A concept such as overall irregularity probably combines color, shape, and borders, but also will include many unexpressed features, and one may suspect that OPR is more or less the recognition of a combination of criteria. Our data discard this hypothesis. Indeed, a multivariable model predicting IEMP shows that the variable overall irregularity was still retained, even when A, B, C, and D criteria were forced into the model. Finally, the different parts of our questionnaire reliably reflect the following 3 very different intellectual diagnostic processes: cognitive assessment of the whole lesion (OPR), algorithms combining analytic criteria (ACR), and a direct comparative process (DR). For simplification of study design, our study focused on the discrimination between nevi and MM and did not take into account the differential diagnosis with seborrhoeic keratoses. It is likely that the general rules of image recognition would also apply to this similar but generally easier problem.

The overall recognition process, mimicking experts in their daily practice, could be useful in the field of education at MM detection. Indeed, a learning process with photographs, based on a global cognitive approach, is worth being assessed. In this regard, experiences have shown that performance in the discrimination of pigmented skin lesions on photographic slides was better with pictorial examples than with written descriptions and did not decline when viewing time was reduced, suggesting that judgments could be made on the general appearance of lesions, a concept similar to OPR. This global recognition used by dermatologists is not a specific technique dedicated to experts but a quite natural learning procedure that each of us has been using from birth. Indeed, we have learned how to recognize most of the basic objects around us by constructing our patterns from images. The quality of overall pattern assessment is certainly dependent on the number of images stored in our memory, and, by definition, nonexperts will not rapidly become experts. Whether this diagnostic process used by experts can be applied to nonexperts is debatable. The possible use of OPR and the ugly duckling sign in the education of the general population and in the training of medical students, general practitioners, or even nonphysician health care providers must be assessed by comparing it with the ABCD rule or the 7-point checklist.

Our study also confirms how much a recent change is important for a reliable diagnosis of MM. In this regard, the fact that an MM was found in 9 (1.68%) of the 535 nevi that were removed only to reassure the patient should lead physicians to listen carefully to patients who want a mole resected, because patients may perceive subtle changes that they may be unable to communicate.

Finally, these principles of image recognition are likely to be applicable to similar fields. In dermoscopy, OPR has already been shown to be more efficacious than the dermoscopy-specific ABCD criteria. In the computer-assisted diagnosis of pigmented lesions, recent technology has already integrated the fact that the machine must build its own recognition pattern (OPR) and no longer use an algorithm of criteria derived from the ABCD rule. However, each pigmented lesion is analyzed separately and thus overlooks the ugly duckling sign. These machines will probably not be able to compete with the accuracy of experts until they computerize not only the suspicious lesion but also the image of all the other nevi in a given individual.

Accepted for Publication: June 11, 2004.

Correspondence: Jean Jacques Grob, MD, Service de Dermatologie Hôpital Ste Marguerite, 270 Boulevard Ste Marguerite, 13009 Marseille, France (jgrob@mail.ap-hm.fr).

Acknowledgment: We thank all the volunteer dermatologists from the Fédération Française de Formation Continue en Dermatologie-Vénéréologie, Pontoise, France, who actively participated in this study.

REFERENCES