

INVESTIGATIVE REPORT

Viewing Exemplars of Melanomas and Benign Mimics of Melanoma Modestly Improves Diagnostic Skills in Comparison with the ABCD Method and Other Image-based Methods for Lay Identification of Melanoma

Ella CORNELL^{1#}, Karen ROBERTSON^{2#}, Robert D. McINTOSH¹ and Jonathan L. REES²

Departments of ¹Human Cognitive Neuroscience, Psychology and ²Dermatology, University of Edinburgh, UK

[#]These authors contributed equally to this paper and should be considered as first authors.

Using an experimental task in which lay persons were asked to distinguish between 30 images of melanomas and common mimics of melanoma, we compared various training strategies including the ABC(D) method, use of images of both melanomas and mimics of melanoma, and alternative methods of choosing training image exemplars. Based on a sample size of 976 persons, and an online experimental task, we show that all the positive training approaches increased diagnostic sensitivity when compared with no training, but only the simultaneous use of melanoma and benign exemplars, as chosen by experts, increased specificity and diagnostic accuracy. The ABCD method and use of melanoma exemplar images chosen by laypersons decreased specificity in comparison with the control. The method of choosing exemplar images is important. The levels of change in performance are however very modest, with an increase in accuracy between control and best-performing strategy of only 9%. Key words: skin cancer; melanoma; melanocytic nevi; seborrheic keratosis; diagnosis.

Accepted Jan 24, 2015; Epub ahead of print Jan 29, 2015

Acta Derm Venereol 2015; 95: 681–685.

Jonathan L. Rees, Prof., Department of Dermatology, Rm 4.018, Lauriston Building, Lauriston Place, Edinburgh, EH3 9HA, UK. E-mail: reestheskin@me.com

Melanoma prognosis is tightly linked to tumour (Breslow) thickness, with thinner tumours having a better prognosis than thicker tumours (1). It is widely believed that thinner tumours are at an earlier stage of development, and therefore diagnosis of these thin tumours – before they progress to thicker lesions – will result in better clinical outcomes (2, 3). Because the majority of melanomas are brought to medical attention by patients, and patient factors account for most delay in diagnosis (4–7), there has been a lot of research into how early diagnosis (or at least flagging-up of worrying lesions) by patients can be improved (7–9). The practical issue is that melanoma is relatively rare, whereas mimics of

melanoma (e.g. naevi, seborrheic keratoses) are very common. There is therefore a signal to noise issue, with both sensitivity and specificity being important given finite healthcare resource and limited patient attention (10).

Approaches to facilitating early diagnosis include general public awareness campaigns, which raise ‘concern’ with little attempt to improve specific diagnostic skills, and more targeted approaches, with the goal of improving or disseminating the skills needed to differentiate worrying lesions from benign lesions (3). Attempts to improve such diagnostic skills have usually focused on rule-based strategies such as the ABCD methodology, in which laypersons make use of a series of criteria that experts have reported to be useful in diagnosing melanoma (11, 12). These include asymmetry (A), border regularity (B) and colour variation (C) of the lesion and (D) diameter, and in some instances, information about whether the lesion is elevated or is evolving (E). A number of publications have challenged the efficacy of such ABCD(E) approaches both on theoretical and empirical grounds (13–18). Alternative approaches have made greater use of images, in which examples of melanomas (with or without benign lesions) are provided to subjects, with the hypothesis that non-experts will be able to use these exemplars to improve their ability to distinguish between melanoma and mimics of melanoma (13, 18–22). Whichever methodology is chosen, such methods might be provided as part of a prospective general educative strategy (‘health education’), or at a particular point of time, where the person seeks to check out a particular lesion they are worried about (‘just in time’).

The world wide web (WWW) is now a major source of health advice, and disease-related material (23, 24). The ease with which the Internet can be used to present images to the public allows strategies based on images to be both developed and empirically tested. In a previous study using a web browser type interface, we compared the ability of volunteers to distinguish between test images of melanomas and mimics of melanoma using 2 strategies: the rule-based ABCD approach, or by providing subjects with a set of melanoma images to act as exemplars (18). We failed to find any

difference between these approaches, and the modest improvements in accuracy seen with either method (as compared with a no intervention control arm) would not justify widespread use. The sample size was small ($n = 72$) however, and our study left unanswered a number of key questions, including whether combining the ABCD methods and image training, or the simultaneous use of exemplars of both melanomas and mimics might improve performance. Given the range of appearance of melanomas (and of melanoma mimics), it is also an open question as to how you select images to use in any training set. Should you select training images randomly, or should you choose particular sets of images with the aim of covering the range of morphology seen in the clinic? If the latter, how do you decide which to choose?

In the present study we have used the Internet to undertake a much larger study using social media and other tools to recruit subjects. We have also tested different methods of choosing image exemplars, and the role of both negative and positive exemplars, as well as combining image exemplars with ABCD type rule-based strategies.

METHODS

We created an online melanoma identification task in which we were able to systematically manipulate the type of training given to a large number of participants. We devised 6 study conditions to compare: no training condition (control); rule-based training using the written ABC criteria (ABC only); image training using melanoma examples chosen by experts (MEL(EXP)); image training using melanoma examples statistically selected from judgements made by laypeople (MEL(LAY)); image training with examples of both melanoma and benign lesions (MEL+BEN); and training using a combination of both rule-based (ABC) and melanoma images (ABC+MEL). Note, that we did not aim to compare all possible combinations of the various training conditions.

Written ABC information

The written ABC information was compiled from the most commonly used descriptions of the ABC(D) criteria available on websites such as the British Association of Dermatologists (BAD), The American Academy of Dermatology (AAD), and Cancer Research UK (CRUK). As justified in our previous paper (18), we excluded 'D' for diameter because the images used in the study were not presented as life size on the computer monitor. No images were used alongside the descriptions to avoid the potential effect of incidental image learning; some prior work suggests that using images as visual anchors for the ABCD method does not, however, improve performance (25).

Lesion images

Photographs of 80 melanoma, 300 seborrhoeic keratoses and 300 benign naevi were obtained from the image database of the Department of Dermatology, University of Edinburgh, that comprises over 5,000 images, collected prospectively using the same photographic set-up: Canon EOS 350D 8. IMP cameras, Sigma 70 mm f2.8 macro lens and Sigma EM-140 DG Ring Flash at a fixed distance of 50 cm (16). The database is a research resource and, as far as possible, image collection was

based on sequential patients rather than being based on selection of 'interesting' cases. Many lesions were not the index lesion a patient was referred to hospital with, and we believe the database is likely to be representative of the various lesion classes. Each lesion was cropped from the original digital image to an image of 300×300 pixels with the lesion positioned centrally.

Expert image sets

We wished to compare different strategies of choosing batches of exemplar images, on the basis that different sets of exemplars may perform differently, and that any ideal set has to encompass the range of morphology seen in any diagnostic class. We therefore chose and compared a set of melanoma exemplars based on images chosen by expert dermatologists, and a set related to layperson perceptions (explained below).

Expert melanoma training set. Two consultant dermatologists selected 8 melanoma images (out of 80) that they deemed to be typical and illustrative of important clinical features. Four of the 8 images were common between both experts, and 2 of the other remaining choices made by each were used (after discussion)

Expert benign training set. The same 2 consultants chose representative examples of 16 seborrhoeic keratoses and 16 benign naevi (out of 300 per class), of which a set of 8 exemplars for each diagnostic group were chosen after discussion. The benign example set was randomly selected from these chosen lesions each time it was used, but always contained 4 images of seborrhoeic keratoses and 4 of benign naevi.

Layperson-selected melanoma set. To create an alternate set of melanoma training examples, we statistically extracted 8 melanoma images (out of 80) based on similarities observed by a sample of 34 laypersons. These 34 participants were presented with a stack of 80 photographic examples of melanoma, and were given 15 min to sort the cards into 4–7 groups based on visual similarity. Each possible pairing of lesions received a score of one (if the subject placed them in a group together), or zero (if they were placed in different groups). Across subjects, these scores were averaged to produce a relatedness score between 0–1 for each image pair. This matrix was treated as a correlation matrix, and a principal component analysis carried out with Oblimin Rotation to estimate the underlying factors. The scree plot indicated that decreases in the eigenvalues from one to the next levelled off at 8, so we extracted 8 factors. These factors constituted an empirically-derived sorting of the library, which reflected the average perception of similarity between lesions, with the 'typicality' of each lesion within each factor given by its loading for that factors. We assembled our final group of 8 melanoma training images by selecting the melanoma lesion that loaded highest on each of the 8 factors.

Web interface

A basic web interface was created in-house to fit the parameters of the study and housed on University of Edinburgh servers. The study could be accessed online (<http://tinyurl.com/melanomastudy>). Participants were recruited over a one-month period via email and social media websites, and the study URL was posted on both the University of Edinburgh Dermatology and Psychology web sites.

An introductory page provided information regarding the aims and content of the study, and subjects were required to confirm that they were at least 18 years of age before being allowed to proceed. Self-reported age and sex were collected, and whether the individual had completed the study before. The instruction pages for each of the 6 conditions contained the same information about melanoma and general instructions on how to complete the task, but differed in the explanation of how to

use the training in each specific condition. The test interface consisted of 2 side panels (left and right) that were varied based on the 6 conditions as follows (see Fig. S1¹): (i) Control: The participant received no training and only basic instruction in performing the experimental task. There was no information in either side panel. (ii) ABC only: The left panel contained a description of the ABC criteria, and there was nothing in the right panel. (iii) MEL(EXP): The left panel contained 8 images of melanomas selected by dermatologists under the heading "Examples of Melanoma." There was nothing in the right panel. (iv) MEL(LAY): The left panel contained 8 images of melanomas selected by laypeople under the heading "Examples of Melanoma." There was nothing in the right panel. (v) ABC+MEL: The left panel contained written ABC information and the right panel contained the 8 dermatologist-selected melanoma images under the heading "Examples of Melanomas." and (vi) MEL+BEN: The left panel contained 8 dermatologist-selected melanomas under the heading "Examples of Melanomas," and the right panel contained 8 dermatologist-selected benign lesions under the heading "Examples of Harmless Skin Lesions."

The test image was always presented in the centre of the page, with the instruction "State whether or not you think this image:" with radio buttons below which read, "IS a melanoma" or "is NOT a melanoma." For each image, participants selected one or other of the 2 buttons. Subjects evaluated 30 test images (10 each of melanoma, seborrhoeic keratoses and benign naevi), which were randomly selected from the total pool, providing a ratio of 1:2 melanoma:benign lesions. The order of test lesions was randomly assigned, and for each image condition the melanoma and/or benign lesions used in the training panel(s) were excluded from the pool of images from which the test lesions were randomly drawn. Once the participant had completed the task, they were directed to a final page which thanked them for their participation and provided a link to the CRUK website for further information on melanoma (<http://www.cancerresearchuk.org>).

We did not perform any formal power calculations but were aiming at close to a 1,000 respondents, accepting that some more trials would be needed as some are likely to be incomplete. The decision to close the study preceded any statistical analysis of the accrued data.

RESULTS

In total, 1,151 persons visited the website in a 3-week period, of whom 976 completed the study. Incomplete datasets were discounted, and for subjects who attempted the study more than once, only the first attempt was accepted. Of those who contributed valid datasets, 640 were females and 336 males, age range 18–79 years (mean 39.16, SD 14.51). A summary of the age and sex demographic across the 6 conditions is shown in Table S1¹. The age distribution was skewed towards youth (A histogram of age is available as Fig. S2¹). There was no significant difference in age ($p=0.49$), or sex ($p=0.78$) in allocation to study groups.

Each response was classed as positive if the participant identified the lesion as a melanoma, and negative if they did not. Depending on whether the test lesion was in fact a melanoma or benign lesion, each response was therefore either a true positive (TP), false positive

(FP), true negative (TN) or false negative (FN). For each participant, outcome measures of sensitivity (TP/TP+FN), specificity (TN/TN+FP), and accuracy (TP/TP+FP+FN) were calculated across the 30 test lesions (10 melanomas and 20 benign lesions). A summary of the mean percentage value for each outcome variable can be seen in Table I.

For descriptive purposes, the mean rate of positive responding (TP + FP) has also been included in Table I. In the control condition (no training) positive responding was close to 50%, which may suggest that the binary choice between radio buttons encouraged an implicit assumption that half of the target lesions were melanomas. Notably, the only condition in which positive responding was not increased above this control level was the MEL+BEN training condition in which (benign) counter-examples were provided in addition to positive diagnostic information.

The effect of training condition was statistically analysed in terms of the formal outcome variables of sensitivity, specificity and accuracy, illustrated in Fig. 1. A MANOVA, showed an overall effect of training condition [$F(10,1938) = 17.94$, Wilk's Lambda = 0.84, $p < 0.0005$, partial $\eta^2 = 0.09$]. The univariate tests confirmed that training condition influenced sensitivity [$F(5,970) = 19.03$, $p < 0.0005$, partial $\eta^2 = 0.09$], specificity [$F(5,970) = 17.75$, $p < 0.0005$, partial $\eta^2 = 0.08$] and accuracy [$F(5,970) = 16.41$, $p < 0.0005$, partial $\eta^2 = 0.08$].

These main effects were investigated further using the Tukey procedure. For sensitivity, the control condition was found to produce significantly lower sensitivity than all other conditions ($p < 0.0005$ in all cases), amongst which there were no significant differences. For specificity, the MEL+BEN condition produced significantly higher specificity than every other condition ($p < 0.01$ in all cases), whilst the ABC+MEL and MEL(LAY) conditions both produced significantly poorer specificity than control ($p < 0.01$ in both cases). Finally, the expert selected melanoma lesions (MEL(EXP)) led to significantly greater specificity than those selected by laypersons (MEL(LAY)) ($p < 0.005$). In terms of overall accuracy (which is a weighted combination of sensitivity and specificity), the MEL+BEN condition

Table I. Summary of sensitivity, specificity and accuracy by intervention

Condition	Positive responding ^a	Sensitivity	Specificity	Accuracy
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Control	48% (15.0)	58% (21.0)	57% (17.1)	57% (11.5)
ABC only	56% (15.1)	73% (18.9)	52% (17.1)	59% (10.6)
ABC+MEL	58% (13.9)	76% (16.1)	50% (16.8)	59% (10.7)
MEL(EXP)	54% (14.8)	72% (17.5)	54% (17.3)	60% (10.7)
MEL(LAY)	59% (13.7)	72% (17.0)	48% (16.90)	56% (11.4)
MEL+BEN	48% (9.7)	71% (16.1)	63% (12.7)	66% (10.2)

^aRefers to the percentage of test images that respondents scored as melanomas (true positives and false positives).

¹<http://www.medicaljournals.se/acta/content/?doi=10.2340/00015555-2058>

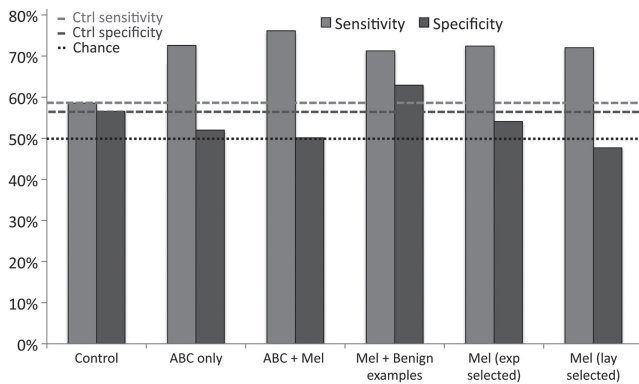


Fig. 1. Comparison of mean sensitivity and specificity between training conditions. 'Mel' refers to melanoma examples, and 'exp selected' refers to images selected by experts, and 'lay selected' to images chosen by laypersons. Dotted lines indicate sensitivity and specificity scores, and the line for 'Chance' refers to the expected score given the binary choice of melanoma or benign lesion and a random response.

outperformed every other condition ($p < 0.0005$ in all cases), and was the only condition producing greater mean accuracy than the no-training control condition. The expert-selected melanomas (MEL(EXP)) produced significantly greater overall accuracy than did those selected by laypersons (MEL(LAY)) ($p < 0.005$).

DISCUSSION

Within the constraints of the experimental approach we have chosen (the limitations of which we discuss below) our results appear clear, and are likely more statistically robust than our earlier smaller study (18). The provision of any sort of positive training information (ABC rules, or positive images of melanomas), whether in combination or alone, increased the rate of positive responding, thereby making people more likely to say that any lesion was a melanoma. This increase in sensitivity was, however, accompanied by a reduction in specificity, except where image training involving both images of expert selected melanoma and benign lesions was provided. Only by combining these expert-selected melanoma images with (expert-selected) examples of benign lesions were we able to promote parallel increases in sensitivity and specificity: this was the only experimental intervention that increased diagnostic accuracy. Contrary to some studies we found no additive value to providing images and written (ABC) information (20). However, we did not examine the value of the ABC method in addition to providing images of both melanoma and benign exemplars, a training condition that would have required a different interface design.

Our two key findings, that any sort of training increases sensitivity, and that provision of images of both melanomas and counter images of benign lesions improves specificity are perhaps not too surprising. Vigilance may be increased by any sort of intervention

and explanations and text about melanoma may increase subject concern non-specifically, leading to more false positives. However, increasing sensitivity alone is not of necessity useful if it is accompanied by no change in specificity (10). That the provision of examples and counter examples improves performance (with or without learning) is in keeping with findings in some other cognitive domains (26). The difference between expert and layperson chosen exemplars is worthy of follow up, but at present we interpret it as support for the idea that exactly which images are chosen may be critical for test performance – use of any images that are 'to hand' in public health campaigns may be sub-optimal. Similarly, the number of exemplar images used, may influence the effectiveness of any intervention.

The absolute increase in accuracy is very modest, and needs to be judged in the light of several limitations of our experimental approach. First, the age distribution of the test subjects was not representative of the general population, nor of those with the highest incidence of melanoma (3). This we assume relates to the methods used to recruit subjects. This may have underestimated intervention effects, as we have previously shown that older persons perform better on similar tasks (18). On the other hand we know that younger people are disproportionately represented in melanoma diagnostic clinics, so they remain a key target group (3). Targeting older people may require a different approach. Second, almost inevitably, and in keeping with virtually all work in this domain, we are testing individuals in a way that does not closely match the real world. For instance, we have shown subjects a large number of test images, whereas in a clinical setting a subject is concerned with only a single lesion. In addition, there is evidence that stress may alter (and worsen) performance in such tasks, a factor we are unable to easily model (13).

Third, caution is needed in interpreting the summary measures used in such studies. Sensitivity and specificity are key measures of test performance in many clinical situations, but in the sort of experiment described, subject assumptions about the exact prevalence of positive diagnoses may alter subject performance in ways that are not relevant in other domains. If a subject 'assumed' that half the lesions were melanomas (whereas the real rate was 33.3%) decision making may have been different if a different prevalence of test images had been used. The figures for accuracy are influenced by the experimentally determined prevalence of positive diagnoses in the test set. In the real world, the base-rate of melanoma is at least several orders of magnitude lower than the one we used experimentally, and therefore extrapolating summary measures to the 'real world' is problematic. Of course, screening tests need a high sensitivity, but specificity is also critical where health provision resource is finite and where patient attention to health-related tasks is limited.

If we are to put our work in a broader context, we would make several points. Diagnosis of suspicious skin lesions is known to be very difficult, requiring many years of clinical training. It is not therefore too surprising that attempts to improve the accuracy of laypersons have had limited success. Of course a larger proportion of melanoma patients present with thinner lesions than was the case historically in most developed nations. This may reflect many factors, including an increase in health care provision, and many non-specific attempts at increasing awareness of skin cancer. The exact mechanisms by which awareness has been increased may be hard to codify, or even improve upon, although we think our work suggests ways in which current patient education campaigns might be improved. Against this, and subject to the limitations we have highlighted, some interventions such as showing particular images on campaign information – however intuitively sensible they may seem – may have negative as well as positive effects.

ACKNOWLEDGEMENTS

We thank Wendy Johnson for assistance with the Principal Components Analysis, Lisa Naysmith for help choosing exemplars, and Cedric MacMartin for assistance with the web interface. Collection of the images used in this report was supported by the Wellcome Trust, grant number 083928/Z/07/Z to JL Rees and RB Fisher.

Funding: CRUK project grant to JLR and RDM C1375/A12060.

The authors declare no conflicts of interest.

REFERENCES

- Balch CM, Soong SJ, Gershenwald JE, Thompson JF, Reintgen DS, Cascinelli N, et al. Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. *J Clin Oncol* 2001; 19: 3622–3634.
- Oliveria SA, Christos PJ, Halpern AC, Fine JA, Barnhill RL, Berwick M. Patient knowledge, awareness, and delay in seeking medical attention for malignant melanoma. *J Clin Epidemiol* 1999; 52: 1111–1116.
- Yee EFT, Hoffman RM, Berwick M. Early diagnosis of melanoma: What do we know? *G Ital Dermatol Venereol* 2007; 142: 55–70.
- Hennrikus D, Girgis A, Redman S, Sanson-Fisher RW. A community study of delay in presenting with signs of melanoma to medical practitioners. *Arch Dermatol* 1991; 127: 356–361.
- Koh HK, Miller DR, Geller AC, Clapp RW, Mercer MB, Lew RA. Who discovers melanoma?: patterns from a population-based survey. *J Am Acad Dermatol* 1992; 26: 914–919.
- Geller AC, Swetter SM, Brooks K, Demierre M-F, Yaroch AL. Screening, early detection, and trends for melanoma: current status (2000–2006) and future directions. *J Am Acad Dermatol* 2007; 57: 555–572.
- Richard MA, Grob JJ, Avril MF, Delaunay M, Gouvenet J, Wolkenstein P, et al. Delays in diagnosis and melanoma prognosis (I): the role of patients. *Int J Cancer* 2000; 89: 271–279.
- Liu W, Hill D, Gibbs AF, Tempany M, Howe C, Borland R, et al. What features do patients notice that help to distinguish between benign pigmented lesions and melanomas?: the ABCD(E) rule versus the seven-point checklist. *Melanoma Res* 2005; 15: 549–554.
- Rees JL. Melanoma: What are the gaps in our knowledge? *PLoS Medicine* 2008; 5: 878–880.
- Weatherhead SC, Lawrence CM. Melanoma screening clinics: are we detecting more melanomas or reassuring the worried well? *Br J Dermatol* 2006; 154: 539–541.
- Rigel DS, Friedman RJ, Kopf AW, Polsky D. ABCDE – an evolving concept in the early detection of melanoma. *Arch Dermatol* 2005; 141: 1032–1034.
- Rigel DS, Russak J, Friedman R. The evolution of melanoma diagnosis: 25 years beyond the ABCDs. *CA Cancer J Clin* 2010; 60: 301–316.
- Girardi S, Gaudy C, Gouvenet J, Teston J, Richard MA, Grob JJ. Superiority of a cognitive education with photographs over ABCD criteria in the education of the general population to the early detection of melanoma: a randomized study. *Int J Cancer* 2006; 118: 2276–2280.
- Gachon J, Beaulieu P, Sei JF, Gouvenet J, Claudel JP, Lemaitre M, et al. First prospective study of the recognition process of melanoma in dermatological practice. *Arch Dermatol* 2005; 141: 434–438.
- Aldridge RB, Zanutto M, Ballerini L, Fisher RB, Rees JL. Novice Identification of Melanoma: Not quite as straightforward as the ABCDs. *Acta Derm Venereol* 2011; 91: 125–130.
- Aldridge RB, Glodzik D, Ballerini L, Fisher RB, Rees JL. Utility of non-rule-based visual matching as a strategy to allow novices to achieve skin lesion diagnosis. *Acta Derm Venereol* 2011; 91: 279–283.
- Laskaris N, Ballerini L, Fisher RB, Aldridge B, Rees J. Fuzzy description of skin lesions. *Proc SPIE 7627, Medical Imaging 2010: Image Perception, Observer Performance, and Technology Assessment, 762717* (February 23, 2010); doi:10.1117/12.845294.
- Robertson K, McIntosh RD, Bradley-Scott C, Macfarlane S, Rees JL. Image training, using random images of melanoma, performs as well as the ABC(D) criteria in enabling novices to distinguish between melanoma and mimics of melanoma. *Acta Derm Venereol* 2014; 94: 265–270.
- Hanrahan PF, Hersey P, Watson AB, Callaghan TM. The effect of an educational brochure on knowledge and early detection of melanoma. *Aust J Public Health* 1995; 19: 270–274.
- Miles F, Meehan JW. Visual discrimination of pigmented skin lesions. *Health Psychol* 1995; 14: 171–177.
- Borland R, Mee V, Meehan JW. Effects of photographs and written descriptors on melanoma detection. *Health Educ Res* 1997; 12: 375–384.
- Gaudy-Marqueste C, Dubois M, Richard MA, Bonnelye G, Grob JJ. Cognitive training with photographs as a new concept in an education campaign for self-detection of melanoma: a pilot study in the community. *J Eur Acad Dermatol Venereol* 2011; 25: 1099–1103.
- Eysenbach G, Kohler Ch. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. *AMIA Annu Symp Proc* 2003: 225–229.
- Sabel MS, Strecher VJ, Schwartz JL, Wang TS, Karimipour DJ, Orringer JS, et al. Patterns of Internet use and impact on patients with melanoma. *J Am Acad Dermatol* 2005; 52: 779–785.
- Zanutto M, Ballerini L, Fisher RB, Aldridge B, Rees J. Visual cues do not improve skin lesion ABC(D) grading. *Proc. Medical Imaging 2011: Image Perception, Observer Performance, and Technology Assessment, SPIE 2011; 7966: 79660U-1-79660U-10*.
- Murphy GL. *The big book of concepts*. Cambridge, Mass: The MIT Press, 2002.