The Performance of MelaFind

A Prospective Multicenter Study

Gary Monheit, MD; Armand B. Cognetta, MD; Laura Ferris, MD, PhD; Harold Rabinovitz, MD; Kenneth Gross, MD; Mary Martini, MD; James M. Grichnik, MD, PhD; Martin Mihm, MD; Victor G. Prieto, MD, PhD; Paul Googe, MD; Roy King, MD; Alicia Toledano, ScD; Nikolai Kabelev, BCSc; Maciej Wojton, MS; Dina Gutkowicz-Krusin, PhD

Objective: To demonstrate the safety and effectiveness of MelaFind, a noninvasive and objective computer-vision system designed to aid in detection of early pigmented cutaneous melanoma.

Design: A prospective, multicenter, blinded study. The diagnostic performance of MelaFind and of study clinicians was evaluated using the histologic reference standard. Standard images and patient information for a subset of 50 randomly selected lesions (25 melanomas) were used in a reader study of 39 independent dermatologists to estimate clinicians’ biopsy sensitivity to melanoma.

Setting: Three academic and 4 community practices in the United States with expertise in management of pigmented skin lesions.

Patients: A total of 1383 patients with 1831 lesions enrolled from January 2007 to July 2008; 1632 lesions (including 127 melanomas—45% in situ—with median Breslow thickness of invasive lesions, 0.36 mm) were eligible and evaluable for the study end points.

Main Outcome Measures: Sensitivity of MelaFind; specificities and biopsy ratios for MelaFind and the study investigators; and biopsy sensitivities of independent dermatologists in the reader study.

Results: The measured sensitivity of MelaFind was 98.4% (125 of 127 melanomas) with a 95% lower confidence bound at 95.6% and a biopsy ratio of 10.8:1; the average biopsy sensitivity of dermatologists was 78% in the reader study. Including borderline lesions (high-grade dysplastic nevi, atypical melanocytic proliferations, or hyperplasias), MelaFind’s sensitivity was 98.3% (172 of 175), with a biopsy ratio of 7.6:1. On lesions biopsied mostly to rule out melanoma, MelaFind’s average specificity (9.9%) was superior to that of clinicians (3.7%) (P = .02).

Conclusion: MelaFind is a safe and effective tool to assist in the evaluation of pigmented skin lesions.

Trial Registration: clinicaltrials.gov Identifier: NCT00434057


In 2009, there were an estimated 68 700 new cases of invasive melanoma and over 53 000 new cases of melanoma in situ in the United States.¹ The actual number might be significantly higher because melanoma incidence is considered to be underreported by 30% to 40% in cancer registries.² The number of deaths due to melanoma was estimated at 8650 in 2009.¹ Melanoma is virtually 100% curable if detected when it is confined to the epidermis (melanoma in situ). Thin melanomas, with a Breslow thickness of 1 mm or thinner, have a 94% rate of survival after 5 years.³ However, once melanoma has advanced and metastasized to other parts of the body, it is difficult to treat. The survival rate for patients with stage IV melanoma is less than 15% for 5 years, with most patients dying within 6 to 10 months.⁴ Therefore, early detection and prompt treatment are essential to improve the prognosis for patients with melanoma. The challenge is that early melanoma may be difficult to differentiate from many benign simulants.

CME available online at www.jamaarchivescme.com

This multicenter prospective trial was designed to establish the safety and effectiveness of MelaFind (MELA Sciences Inc, Irvington, New York) as an aid in evaluating pigmented lesions (PLs) that have 1 or more clinical or historical characteristics of melanoma. MelaFind is a noninvasive, fully automatic, computer-vision diagnostic system designed as an aid to detection of early melanoma and developed to identify PLs that should be con-
sidered for biopsy to rule out melanoma. This study evaluates the performance of MelaFind using sensitivity and specificity as metrics and comparing the specificity of MelaFind to that of the study investigators.

METHODS

PROSPECTIVE INCLUSION AND EXCLUSION CRITERIA

Patients with at least 1 PL scheduled for biopsy in toto were invited to participate in the trial. Exclusion criteria were as follows: (1) failure to give informed consent; (2) known allergy to isopropyl alcohol; (3) diameter of the PL less than 2 mm or greater than 22 mm; (4) anatomic site of PL not accessible to the device; (5) lesion previously biopsied, excised, or traumatized; (6) skin not intact (eg, open sores, ulcers, bleeding); (7) lesion within 1 cm of the eye; (8) lesion on palmar, plantar, or mucosal (eg, lips, genitals) surface or under nails; (9) lesion in an area of visible scarring; or (10) lesion containing foreign matter (eg, tattoo ink, splinter, marker).

OUTCOME MEASURES

The primary outcome measures were the sensitivity and specificity of the computer-vision system, MelaFind, and the specificity of clinicians, among PLs with the diagnoses of “melanoma cannot be ruled out” or “not melanoma.” Pigmented lesions with prebiopsy clinical diagnoses of melanoma were excluded from the analysis because such lesions would be biopsied by examining dermatologists regardless of MelaFind results. Melanomas and borderline lesions such as high-grade dysplastic nevi (HGDN) and atypical melanocytic hyperplasias (AMH) or proliferations (AMP) were defined as histologically positive lesions. MelaFind produces a binary output: (1) positive, the lesion should be considered for biopsy to rule out melanoma; and (2) negative, the lesion should be considered for later evaluation. Clinicians were blinded to MelaFind output, and participation in the trial did not affect patient treatment. Thus investigators managed the patient care based on their clinical information, and in the study, the PL was considered positive if the prebiopsy dermatologic diagnosis was melanoma or melanoma precursors or early evolving melanomas and are difficult to differentiate from melanoma in situ. The diagnoses AMP and AMH may be used when dermatopathologists are uncertain about diagnosis and cannot rule out melanoma.

Because concordance of histologic interpretation is limited with early melanomas, histologic slides for each lesion in the trial were evaluated by 2 independent dermatopathologists. In cases of significant discordance, histologic slides were evaluated independently by a third dermatopathologist. When 1 dermatopathologist diagnosed melanoma and 2 others diagnosed a benign lesion, histologic slides were sent again to the dermatopathologist who diagnosed melanoma for a blind review. The final histologic diagnosis was determined following the algorithm detailed in Table 1.

BLINDING AND MASKING

Clinicians did not receive the results of the MelaFind lesion classification algorithm, and information from MelaFind was not used in diagnosing or treating lesions. An independent entity acted as a secure data custodian and verified the integrity of the data received from the clinical sites. In addition, the data custodian analyzed all MelaFind images, using software provided by MELA Sciences Inc. At the conclusion of the accrual phase of the trial, the statistician (A.T.) received the electronic Case Record Forms and lesion classification results from the data custodian, as well as the results of histologic evaluations, and analyzed the performance of MelaFind and of clinicians participating in this trial.

CLINICAL SITES

Seven clinical sites with 23 investigators participated in this trial. Three sites were academic institutions (University of Pittsburgh, Duke University, and Northwestern University), and 4 sites were dermatologic practices highly experienced in managing PLs. All sites were approved to participate in the study by the appropriate institutional review boards.

PILOT STUDY OF BIOPSY SENSITIVITY (READER STUDY)

The reader study investigated the biopsy sensitivity of dermatologists. It used 25 randomly selected melanomas (11 inva-
sive and 14 in situ) and 25 nonmelanomas, matched to melanomas by anatomic location and patient age and sex. Borderline lesions (HGDN, AMP, and AMH) were excluded. The clinical history and images (clinical overview, clinical close-up, and dermoscopy images) were reviewed by the readers. The readers, dermatologists who did not participate in the clinical trial, reported for each lesion whether they would biopsy it to rule out melanoma.

MELAFIND

MelaFind acquires digital multispectral images of a PL in 10 different spectral bands, from blue (430 nm) to near infrared (950 nm). These are contact images that use 91% isopropyl alcohol for refractive index matching. Each image is 1280 × 1024 pixels, with the pixel size in the lesion plane 20 × 20 µm. MelaFind uses automatic image analysis and statistical pat-

tern recognition to help identify lesions to be considered for biopsy to rule out melanoma. The properties of these images as well as image analysis methods have been previously described.15-21

All image analysis and lesion classification algorithms are automatic and were tested prospectively in the clinical trial: (1) calibration algorithms reduce noise and artifacts in the images and determine the diffuse reflectance of the skin; (2) image quality control algorithms detect image problems (eg, overexposure, underexposure, lesion too big, lesion too small, too much hair on the lesion, too many bubbles on the lesion, motion of the handheld imaging device during imaging) and, when appropriate, request the operator to reimage; (3) lesion segmentation algorithm identifies image pixels that belong to the lesion; (4) feature extraction algorithms compute quantitative lesion parameters; (5) lesion classification algorithm differentiates lesions to be considered for biopsy to rule out melanoma (MelaFind positive) from those to be considered for later evaluation (MelaFind negative).

RESULTS

PIGMENTED LESIONS

Between January 2007 and July 2008, 1383 patients with 1831 PLs were enrolled. More female than male patients entered the study, and the median age was 47 years. Most patients (about 98%) were white, consistent with the fact that melanoma is much more frequent among whites.22 There were no significant differences between the group of enrolled patients (1383) and those with evaluable lesions (1257) (Table 2).

Of the 1831 registered lesions, 1 patient with 1 lesion withdrew from the trial. Three lesions were determined to be ineligible by clinicians, and 14 by dermatopathologists (mostly owing to previous scarring that was not identified clinically); 19 lesions were not evaluable because of missing or inadequate histologic slides. One hundred sixty-two lesions were not evaluable owing to unsuccessful imaging attempts: 65 owing to operator errors (eg, too much hair, too many bubbles, lesion not centered in the field of view), 36 owing to MelaFind or standard camera malfunctions, and 61 owing to causes that might have been either

<table>
<thead>
<tr>
<th>Primary Pathologists</th>
<th>Pathologist 3</th>
<th>Rereview by Pathologist</th>
<th>Final Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Diagnosing Melanoma</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Melanoma</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Melanoma</td>
<td>HGDN</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Other</td>
<td>Melanoma</td>
<td>NA</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Other</td>
<td>HGDN</td>
<td>NA</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Other</td>
<td>Other</td>
<td>Melanoma</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Other</td>
<td>Other</td>
<td>HGDN</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Other</td>
<td>Other</td>
<td>Other</td>
</tr>
<tr>
<td>HGDN</td>
<td>HGDN</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>HGDN</td>
<td>Other</td>
<td>HGDN</td>
<td>NA</td>
</tr>
<tr>
<td>HGDN</td>
<td>Other</td>
<td>Other</td>
<td>HGDN</td>
</tr>
<tr>
<td>Other</td>
<td>Other</td>
<td>Other</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: HGDN, high-grade dysplastic nevus (including atypical melanocytic proliferation and atypical melanocytic hyperplasia); NA, not applicable. Other indicates neither melanoma nor HGDN.

aAll pathologic diagnoses are listed left to right in rank order, with the higher grade of malignant diagnosis listed first.
Melanoma cannot be ruled out
Likely melanoma (67%-99% likelihood)
Possible melanoma (34%-66% likelihood)
Unlikely melanoma (1%-33% likelihood)
Not melanoma
Clinical concern only, dermoscopy negative
Nonmelanoma skin cancer
Patient’s concern
Physical discomfort
Cosmetic purposes
Total

<table>
<thead>
<tr>
<th>Dermatologic Diagnosis</th>
<th>Melanoma</th>
<th>HGDN, AMP, or AMH</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melanoma</td>
<td>13</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Likely melanoma (67%-99% likelihood)</td>
<td>30</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td>Possible melanoma (34%-66% likelihood)</td>
<td>44</td>
<td>14</td>
<td>471</td>
</tr>
<tr>
<td>Unlikely melanoma (1%-33% likelihood)</td>
<td>38</td>
<td>27</td>
<td>827</td>
</tr>
<tr>
<td>Not melanoma</td>
<td>1</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>Clinical concern only, dermoscopy negative</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Nonmelanoma skin cancer</td>
<td>1</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td>Patient’s concern</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Physical discomfort</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Cosmetic purposes</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>127</td>
<td>48</td>
<td>1457</td>
</tr>
</tbody>
</table>

Abbreviations: AMH, atypical melanocytic hyperplasia; AMP, atypical melanocytic proliferation; HGDN, high-grade dysplastic nevus.
The secondary end points included all eligible and evaluable lesions with any prebiopsy dermatologic diagnosis. Since no comparisons between MelaFind and study investigators were performed, the data were pooled. These end points are summarized in Table 6. Measured negative predictive values are very high (>98%) owing to the very high sensitivity of MelaFind.

In this multicenter prospective trial, MelaFind achieved very high sensitivity to thin melanomas and borderline lesions (>98%; 95% LCB, >95%). For lesions that were not melanomas and had prebiopsy diagnoses of melanoma, MelaFind had an average specificity of 9.5%, ie, significantly higher than that of investigators (3.7%) (P = .02). However, this trial could not determine the true sensitivity of dermatologists, since melanomas that were not scheduled for biopsy, ie, missed by the examining clinicians, would not be evaluable. The pilot reader study found that dermatologists miss thin melanomas. Thus, even though all lesions in the clinical trial were biopsied by the examining dermatologists, many of the melanomas would not have been biopsied by other dermatologists.

The sensitivity of physicians can be assessed by longitudinal studies, ie, long-term follow-up of the patients to determine whether lesions considered benign later turned out to be clinically suspect for melanoma. In the 9-year longitudinal study at a PL clinic in the United Kingdom by Bataille et al, 221 melanomas (both invasive and in situ) were detected. Melanomas on 14 patients were diagnosed as benign on the first visit and biopsied on the second or third visit; one of these patients died of metastatic melanoma. The biopsy sensitivity to melanoma measured in this study was 93.7% (95% LCB, 90.5%). The median Breslow thickness of melanomas in this study was 0.9 mm. The study by Carli et al26 found biopsy sensitivity to melanoma to be 86.7% (95% LCB, 66.7%) based on a comparison with the cancer registry, but the sample was very small.26 Thus, the 95% LCB on MelaFind’s sensitivity is higher than the biopsy sensitivities of longitudinal studies reported in the literature.

Another method of assessing physician sensitivity is through reader studies. Such studies have been widely used to evaluate diagnostic performance in mammography and colonoscopy.27,28 In dermatology, readers are presented with a series of images of lesions (clinical and/or dermoscopic) and, for each image, may be asked to decide whether a lesion is a melanoma and whether it should be biopsied to rule out melanoma. Assuming acceptable quality of images (clinical and/or dermoscopic) and availability of patient history, and taking into account the fact that most of the training of dermatologists is actually done with such images, reader studies can provide estimates of sensitivity to melanoma. Such studies have been performed in dermatology to evaluate the effectiveness of teledermatology,29,30 to compare the diagnostic performance of dermatologists and primary care physicians,31 to assess the dermoscopic examination of lesions,32 and to determine the diagnostic and biopsy sensitivities of dermatologists.31

The pilot reader study conducted as a part of this clinical trial included a random sample of lesions enrolled in
the trial and provided readers with the clinical overview, a clinical close-up image, contact dermoscopic images, and patient demographic and melanoma risk factor information; all images were checked for quality by an experienced dermatologist. The average biopsy sensitivity of 39 dermatologist readers was 78%, which is similar to a prior reader study by Friedman et al33 for small (<6 mm in diameter) PLs from the MelaFind database. Among 10 expert dermatoscopists reviewing a series of 99 small PLs, the average biopsy sensitivity to melanoma was 71%; the biopsy sensitivity to melanoma in situ was 63%. Interreader variability was very high for lesion management decisions (κ = 0.34), indicating that different experts make different biopsy decisions.

A limitation of the present study is that since only PLs scheduled for biopsy in toto were evaluable in the trial, the benign lesions are not representative of such lesions in the general population. As a result, the specificity reported here cannot be generalized to the general population, for either clinicians or MelaFind. Tsao et al34 estimated that, in the general population in the United States, there are about 70 000 nevi for every invasive melanoma or about 35 000 nevi per invasive or in situ melanoma; the overwhelming majority of these nevi have no clinical or historical characteristics of melanoma. A study by Schäfer et al35 estimated that, in the general population in Germany and found that the average number of atypical nevi per person was 0.074. 36 In the US population of about 300 million, with about 120 000 new cases of invasive and in situ melanomas per year,1 this implies about 200 atypical nevi per melanoma. So, even if all atypical nevi were selected for biopsy to rule out melanoma, the specificity in the general population would be over 99% (ie, about [35 000 − 200]/35 000). Therefore, the fact that the specificities of examining clinicians and MelaFind are rather low in the clinical trial does not mean that the specificities of clinicians and MelaFind would be low in the general population. It is only a reflection of the fact that almost all the lesions in this trial were sufficiently atypical to be selected for biopsy to rule out malignant melanoma.

MelaFind is intended to be used on lesions with 1 or more clinical or historical characteristics of melanoma, ie, atypical lesions. If all atypical lesions were to be biopsied to rule out melanoma, the biopsy ratio (number of false-positive biopsy findings per true-positive biopsy finding) of about 200:1 would be very high. In the trial, the biopsy ratio for MelaFind was 10.8:1 for melanomas and 7.6:1 for melanomas and borderline lesions. The values of the biopsy ratio reported in the literature are highly variable. A study by Cohen et al36 reported biopsy ratios of about 139:1 for patients with a history of melanoma and about 576:1 for patients without personal history of melanoma. Among general practitioners in Australia, the biopsy ratio varied from 82:1 in the youngest patients to 10:1 in the oldest patients.37 For dermatologists, prospective studies reported biopsy ratios of about 8:1 in the general population,38 and from 33:1 to 47:1 among high-risk patients with atypical nevi.39-41 One study by Banky et al42 reported a very low biopsy ratio of about 3:1 in patients at high risk for melanoma using a combination of baseline images and dermoscopy. However, this study also reported that at least 5 of 18 melanomas (4 in situ and 1 invasive at the time of biopsy) were not detected on the first examination. Thus, MelaFind’s biopsy ratio is at the lower end of the values reported in the literature.

MelaFind is a safe and effective tool to help identify PLs to be considered for biopsy to rule out melanoma. In this trial, MelaFind demonstrated very high sensitivity to early melanomas and borderline lesions, specificity superior to that of clinicians, and a biopsy ratio of about 8:1. Direct comparison of the results of this study with those of other studies is not possible. As pointed out by Menzies et al,43 diagnostic performance depends on the difficulty of lesions included in the study. It would be helpful if the dermatologic community could agree on a standard mix of lesions to be used for future testing of methods developed for early detection of melanoma.

Accepted for Publication: August 11, 2010.
Published Online: October 18, 2010. doi:10.1001/archdermatol.2010.302

Author Affiliations: Total Skin and Beauty Dermatology, Birmingham, Alabama (Dr Monheit); Dermatology Associates of Tallahassee, Tallahassee, Florida (Dr Cogネット); Department of Dermatology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania (Dr Ferris); Skin and Cancer Associates, Plantation, Florida (Dr Rabinovitz); Skin Surgery Medical Group Inc, San Diego, California (Dr Gross); Department of Dermatology, Northwestern University, Chicago, Illinois (Dr Martini); Department of Dermatology, University of Miami Health System, Miami, Florida (Dr Grichnik); Department of Dermatology, Harvard Medical School, Boston, Massachusetts (Dr Mihm); University of Texas M. D. Anderson Cancer Center, Houston (Dr Prieto); Knoxville Dermatopathology Laboratory, Knoxville, Tennessee (Drs Googe and King); Statistics Collaborative, Washington, DC (Dr Toledano); and MELA Sciences, Irvington, New York (Messrs Kabelev and Wojton and Dr Gutkowicz-Krusin).

Correspondence: Dina Gutkowicz-Krusin, PhD, MELA Sciences Inc, 50 S Buckhout St, Ste 1, Irvington, NY 10533 (gutkowicz@melasciences.com).

Author Contributions: Drs Toledano and Gutkowicz-Krusin had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Rabinovitz, Toledano, Kabelev, and Gutkowicz-Krusin. Acquisition of data: Monheit, Cogネット, Ferris, Rabinovitz, Gross, Martini, Grichnik, Mihm, Prieto, and Googe. Analysis and interpretation of data: Gross, Martini, Mihm, Prieto, Googe, King, Toledano, Kabelev, Wojton, and Gutkowicz-Krusin. Drafting of the manuscript: Monheit, Prieto, and Gutkowicz-Krusin. Critical revision of the manuscript for important intellectual content: Monheit, Cogネット, Ferris, Rabinovitz, Gross, Martini, Grichnik, Mihm, Prieto, Googe, King, Toledano, Kabelev, and Wojton. Statistical analysis: Toledano and Gutkowicz-Krusin. Administrative, technical, and material support: Martini, Mihm, Wojton, and Gutkowicz-Krusin. Study supervision: Monheit, Mihm, Prieto, and Kabelev. Provided additional lit-
erature and references: Grichnik. Research and development: Kabelev.

Financial Disclosure: Dr Monheit has served and/or currently serves as consultant and/or clinical investigator for Allergan Corporation (Juvederm), Dermik Laboratories (Sculptra), Genzyme Corporation (Captique, Prevelle), Colbar LifeScience Ltd (now owned by Johnson & Johnson) (Evolence), Contura (Aquamid), Ipsen/Medici (Dysport), Stiefel, Electro-Optic Sciences Inc (MelaFind), Revance, Kythera, Galderma, Mentor, and Mertz. Drs Cognetta, Rabinovitz, and Mihm are members of the MELA Sciences Advisory Committee. Drs Monheit, Cognetta, Ferris, Grichnik, Rabinovitz, Prieto, King, and Toledano serve as consultants to MELA Sciences Inc. Messrs Kabelev and Wojton and Dr Gutkowicz-Krusin are employees of MELA Sciences Inc.

Funding/Support: This study was supported by MELA Sciences Inc.

REFERENCES