The Diagnostic Performance of Expert Dermoscopists vs a Computer-Vision System on Small-Diameter Melanomas

Robert J. Friedman, MD; Dina Gutkowicz-Krusin, PhD; Michele J. Farber; Melanie Warycha, MD; Lori Schneider-Kels, MPH; Nicole Papastathis, BA; Martin C. Mihm Jr, MD; Paul Googe, MD; Roy King, MD; Victor G. Prieto, MD; Alfred W. Kopf, MS; David Polsky, MD, PhD; Harold Rabinovitz, MD; Margaret Oliviero, ARNP; Armand Cognetta, MD; Darrell S. Rigel, MD; Ashfaq Marghoob, MD; Jason Rivers, MD, FRCPC; Robert Johr, MD; Jane M. Grant-Kels, MD; Hensin Tsao, MD, PhD

Objective: To evaluate the performance of dermoscopists in diagnosing small pigmented skin lesions (diameter ≤6 mm) compared with an automatic multispectral computer-vision system.

Design: Blinded comparison study.

Setting: Dermatologic hospital-based clinics and private practice offices.

Patients: From a computerized skin imaging database of 990 small (≤6-mm) pigmented skin lesions, all 49 melanomas from 49 patients were included in this study. Fifty randomly selected nonmelanomas from 46 patients served as a control.

Main Outcome Measures: Ten dermoscopists independently examined dermoscopic images of 99 pigmented skin lesions and decided whether they identified the lesions as melanoma and whether they would recommend biopsy to rule out melanoma. Diagnostic and biopsy sensitivity and specificity were computed and then compared with the results of the computer-vision system.

Results: Dermoscopists were able to correctly identify small melanomas with an average diagnostic sensitivity of 39% and a specificity of 82% and recommended small melanomas for biopsy with a sensitivity of 71% and specificity of 49%, with only fair interobserver agreement (κ=0.31 for diagnosis and 0.34 for biopsy). In comparison, in recommending biopsy to rule out melanoma, the computer-vision system achieved 98% sensitivity and 44% specificity.

Conclusions: Differentiation of small melanomas from small benign pigmented lesions challenges even expert physicians. Computer-vision systems can facilitate early detection of small melanomas and may limit the number of biopsies to rule out melanoma performed on benign lesions.

Arch Dermatol. 2008;144(4):476-482

Detection of early malignant melanoma (in situ and thin lesions) is one of the most effective ways of preventing mortality from this disease. Prognosis for patients with melanoma is dependent on early detection, as evidenced by the 10-year survival rates as high as 99.5% that have been reported for thin melanomas smaller than 0.76-mm thick in the New York University melanoma database; these rates markedly decrease to 48% for lesions larger than 3 mm in thickness. The effectiveness of this strategy is further confirmed because the reported marked reduction in mortality from melanoma, from 60% for those patients with melanoma diagnosed in 1960 to approximately 11% in 2005, is mainly due to early detection of thinner lesions followed by appropriate treatment.

The incidence of melanoma in the general population is increasing in the United States and worldwide. Several reports have also indicated the presence of small melanomas, defined as those with diameters of 6 mm or smaller. The mere presence of melanomas 6 mm or smaller warrants efforts to facilitate their identification because many of these small melanomas may appear benign by clinical criteria and are therefore more difficult to diagnose. In light of findings that smaller melanomas tend to be less deeply invasive than melanomas larger than 6 mm and, as such, gen-

See also pages 469 and 533

Author Affiliations are listed at the end of this article.
erally have a more favorable prognosis.\textsuperscript{10,11} Efforts that focus on early detection of small melanomas could potentially reduce mortality from this disease.

Supplementing clinical examination with techniques such as dermoscopy is an effective strategy for improving the sensitivity of dermoscopy-trained dermatologists in melanoma diagnosis.\textsuperscript{12} Although dermoscopy may be more effective than clinical examination in diagnosing thin malignant melanomas,\textsuperscript{13,14} a study by Carli et al\textsuperscript{15} demonstrated that dermoscopy did not improve diagnostic performance in identifying small (\(<6\text{-mm}\)) melanomas. This finding indicates that additional diagnostic tools may be necessary to diagnose smaller melanomas more effectively.

Although many systems have been devised to facilitate the differentiation between histologically benign and malignant lesions, including dermoscopy and computer-assisted image analysis, this task continues to challenge even the most experienced dermatologist. The clinical ABCD criteria, initially described in 1985, remain an objective succinct algorithm for the early clinical recognition of melanoma.\textsuperscript{1} Recently, E for evolving was added to the ABCD criteria to highlight the element of change as an important diagnostic feature of cutaneous melanoma.\textsuperscript{11} The present study was designed to assess the sensitivity of dermoscopists in diagnosing small melanomas (\(\leq 6\text{-mm}\) diameter) compared with a novel automatic computer-vision system.

### METHODS

#### DATABASE

This study used cases from the digital dermoscopic database acquired by Electro-Optical Sciences Inc for the development and testing of MelaFind (Electro-Optical Sciences Inc, Irvington, New York), a computer-vision system for early detection of melanoma that is undergoing clinical testing. Twenty-six clinical sites in the United States and abroad have contributed to this database.

Only pigmented skin lesions (PSLs) were included in the database. The lesions were scanned using the multispectral computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

### STUDY PROCEDURE

Participants in the study (readers) received a CD-ROM with color dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

Every case in the database consisted of multispectral dermoscopic images created using MelaFind multispectral computer-vision system images and a set of 10 lesions (2 melanomas and 8 low-grade dysplastic nevi) was visually assessed by 3 readers (A.W.K., H.R., and A.C.) on a computer-vision device before excisional or deep shave biopsy in toto. Approximately 80\% of the lesions were biopsied to rule out melanoma, whereas the remaining lesions were biopsied mostly to rule out nonmelanoma skin cancer or because of patient concern. The diameter of eligible PSLs ranged from 2 to 22 mm. Previously biopsied, ulcerated, or bleeding lesions were excluded. Also excluded were lesions on mucosal surfaces and lesions that contained foreign matter (eg, tattoos).

### SELECTION OF LESIONS FOR THE STUDY

Small (\(\leq 6\text{-mm}\)) lesions were selected for the study in July 2005 from the database of 1977 eligible and evaluable PSLs, including 202 malignant melanomas. The lesion diameter was determined automatically by the computer-vision system.\textsuperscript{19} There were 990 (50\%) of the total small lesions, of which 49 were melanomas; thus, 24\% of all malignant melanomas were small. All PSLs were obtained either from the training database (75 small lesions, of which 38 were melanomas and 37 were matched nonmelanomas) or from the blinded set of data (24 small lesions, of which 11 were melanomas and 13 were nonmelanomas). All 49 small malignant melanomas were included in this study. The small nonmelanomas were stratified by patient age (1-30 years, 31-60 years, and 60 years or older), sex (female or male), and lesion location (head and neck, trunk, lower limbs, or upper limbs); 30 nonmelanomas were selected randomly to match the frequency of these characteristics in the melanoma sample. High-grade dysplastic nevi were excluded because no consensus exists on their management. The distribution of lesions in the study is given in \textbf{Table 1}.
A moderately dense lymphohistiocytic infiltrate is also evident in the subjacent dermis (hematoxylin-eosin, original magnification ×60). The histopathologic diagnosis was malignant melanoma in situ.

Table 1. Description of Pigmented Lesions

<table>
<thead>
<tr>
<th>Lesion Type</th>
<th>No. of Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive melanoma</td>
<td>21</td>
</tr>
<tr>
<td>Breslow thickness, median (range), mm</td>
<td>0.32 (0.10-1.40)</td>
</tr>
<tr>
<td>Melanoma in situ</td>
<td>28</td>
</tr>
<tr>
<td>Low-grade dysplastic nevus</td>
<td>32</td>
</tr>
<tr>
<td>Congenital nevus</td>
<td>2</td>
</tr>
<tr>
<td>Blue nevus</td>
<td>1</td>
</tr>
<tr>
<td>Compound nevus</td>
<td>2</td>
</tr>
<tr>
<td>Junctional nevus</td>
<td>2</td>
</tr>
<tr>
<td>Seborrheic keratosis</td>
<td>2</td>
</tr>
<tr>
<td>Hemangioma</td>
<td>1</td>
</tr>
<tr>
<td>Lentigo simplex</td>
<td>3</td>
</tr>
<tr>
<td>Solar lentigo</td>
<td>1</td>
</tr>
<tr>
<td>Lichen planus–like keratosis</td>
<td>1</td>
</tr>
<tr>
<td>Actinic keratosis</td>
<td>1</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Total No. of lesions</td>
<td>99</td>
</tr>
<tr>
<td>Total No. of patients</td>
<td>94</td>
</tr>
</tbody>
</table>

Data are number of lesions unless otherwise indicated.

Figure 1. Dermoscopic images used for patient evaluation. A, Dermoscopic image of pigmented macule located on the leg. B, Machine-generated dermoscopic image of lesion in A. C, Histopathologic evaluation reveals a relatively broad asymmetric proliferation of mostly single and focally nested atypical melanocytes arranged along the dermoepidermal junction and at higher levels of the epidermis. A moderately dense lymphohistiocytic infiltrate is also evident in the subjacent dermis (hematoxylin-eosin, original magnification ×60). The histopathologic diagnosis was malignant melanoma in situ.

EVALUATION OF LESION MANAGEMENT DECISIONS

Diagnostic performance does not provide information about case management by dermatologists. To obtain such information, the readers were asked to answer the following question: “Would you biopsy/excise this lesion?” If the answer was yes, the readers had to specify the reason for biopsy. As with evaluation of diagnostic performance, histologic diagnoses were used as the reference standard to evaluate lesion management decisions. If readers indicated that they would biopsy the lesion because they were sure it was melanoma or to rule out melanoma, then the case was considered true positive (TP) if the histologic diagnosis was melanoma and false positive (FP) otherwise. If the reader would not have biopsied the lesion or would have biopsied the lesion to rule out nonmelanoma skin cancer, the case was considered true negative (TN) if the histologic diagnosis was not melanoma and false negative (FN) otherwise. These data allow determination of biopsy sensitivity and specificity of the readers.

DATA ANALYSIS

The diagnostic performance and lesion management decisions were analyzed by computing, for every reader, sensitivity (TP/[TP + FN]) and specificity (TN/[TN + FP]), as well as 95% confidence intervals, for these quantities. The interobserver variability was assessed using the κ statistic. The average diagnostic sensitivity and specificity and average biopsy sensitivity and specificity were also computed, with 95% confidence intervals determined according to the Obuchowski method, which takes into account correlations among the readers. For the averages, the other metrics of interest were positive predictive value (TP/[TP + FP]), negative predictive value (TN/[TN + FN]), and diagnostic accuracy (TP/[TP + FN + FP]). These variables were compared with the results of the automatic computer-vision system on the same set of small lesions.

RESULTS

Diagnostic sensitivity and specificity for all 10 readers (as well as averages) are displayed in Figure 2 in the format of a receiver operating characteristic plot. The error bars in Figure 2 represent 95% confidence intervals, which were large because of the relatively small sample size for each reader. Despite these large confidence intervals, the interreader variability was even larger; the κ statistic was 0.31, indicating only fair agreement among the readers. On small lesions, the average sensitivity to malignant melanoma was only approximately 40%, but the associated specificity was high (approximately 80%). The median diagnostic sensitivity and specificity were similar, at 43% and 84%, respectively.

The fact that diagnostic sensitivity to melanoma was only approximately 40% does not imply that dermatologists do not treat approximately 60% of small melanomas; it only means that this is not a good measure of lesion management decisions. Information about such decisions can be gained from biopsy sensitivity and specificity (Figure 3). The interreader variability on lesion management decisions was high: biopsy sensitivity ranged
from 37% to 88%, whereas specificity varied from 22% to 80% (κ = 0.34). However, for each reader, biopsy sensitivity was higher than diagnostic sensitivity. This is not surprising because some dermoscopically borderline lesions may be called benign but are nevertheless biopsied to rule out melanoma. Because biopsy sensitivity is higher, it follows that biopsy specificity is lower.

The MelaFind database was randomly divided into the training and blind testing data sets. The biopsy sensitivity and specificity of MelaFind on the small lesions in the training set (38 melanomas and 37 nonmelanomas) was 100% and 46%, respectively. The biopsy sensitivity and specificity of the expert dermoscopist readers on average was 71% and 52%, respectively. On the small lesions in the blind testing set, the biopsy sensitivity of MelaFind was 91% (missed 1 melanoma in situ), with a specificity of 38%.

To increase the sample size of the small melanomas for a more robust statistical comparison of expert readers and computer-vision system, the small lesions from the MelaFind training and blinded data sets were combined. This pooling of data for MelaFind would be justified only if its results (sensitivity and specificity) are homogeneous for the 2 sets. Because of high sensitivity, the homogeneity assumption was tested using the Fisher exact test.22 Based on the values of P = .22 for MelaFind sensitivity and P = .75 for MelaFind specificity, the null hypothesis of homogeneity is valid and data were pooled. On average, the expert dermoscopist readers had a biopsy sensitivity of 71%, with a specificity of 49%. The median biopsy sensitivity and specificity of the 10 dermoscopists was 74% and 50%, respectively. The biopsy sensitivity and specificity of MelaFind was 98% (missed 1 melanoma in situ) and 44%, respectively.

Detailed comparison of human and computer vision for the management of small PSLs (biopsy sensitivity and specificity) is presented in Table 2. It clearly demonstrates that for small lesions, the computer-vision system had significantly higher sensitivity than dermoscopists (P < .001), while the difference in specificities was not statistically significant (P = .75). In addition, the computer-vision system had statistically significant (P = .02) higher values of negative predictive value; the differences in positive predictive value (P = .48) and diagnostic accuracy (P = .08) were not statistically significant.

Sensitivity to invasive and in situ melanomas should also be considered separately; the results are given in Table 3. The 95% confidence intervals were large because of the small sample sizes. Nevertheless, it was clear that dermoscopists and the computer-vision system have higher sensitivity to invasive than to in situ melanomas. The data in Table 3 indicate that approximately 19% of small invasive melanomas and approximately 37% of small melanomas in situ may be left unbiopsied, even by expert physicians.

**COMMENT**

Reports vary in their assessment of the prevalence of melanomas 6 mm or smaller among the overall population of cutaneous melanomas.3-7,10-12 In a retrospective review of small-diameter melanomas, Abbasi et al11 concluded that the prevalence of small melanomas (≤ 6 mm) ranged from less than 5% to 14%. The prevalence of small-diameter melanomas in our database, however, was much higher, at 25%, prompting us to evaluate this subset of lesions more closely. Although techniques of lesion recognition, such as dermoscopy, are gaining in popularity, the differentiation of early melanomas from benign PSLs continues to be plagued by uncertainty. This is especially true for the subset of small-diameter melanomas, which frequently display clinical and histologic discord.23 The present study is one of the first, to our knowledge, to comprehensively assess the diagnostic performance of a computer-vision system for identifying these small lesions.
knowledge, to quantify diagnostic accuracy of dermatoscopists in specifically identifying small melanomas. By comparing the diagnostic performance of dermatoscopists with a computer-vision system, the design of this study gives insight into new methods of identifying small melanomas in early stages of development.

In our study, we determined diagnostic sensitivity and specificity and biopsy sensitivity and specificity. This approach uniquely allowed us to differentiate between diagnostic performance and lesion management decisions. Although the average diagnostic sensitivity for all 10 dermatoscopists was only 39%, the average biopsy sensitivity was 71%. A similar disparity was seen between diagnostic sensitivity and biopsy sensitivity of 82% and 58%, respectively, reflecting the fact that lesions suggestive of disease are often biopsied to rule out melanoma. Despite the high biopsy sensitivity of our readers, nearly 30% of melanomas smaller than 6 mm would not have been biopsied. Furthermore, only fair interobserver agreement ($\kappa=0.31$ for diagnosis and 0.34 for biopsy decision) indicates that dermatoscopists differ in their evaluation of small PSLs, emphasizing the challenge in small melanoma diagnosis.

Only pigmented lesions scheduled for biopsy were eligible for inclusion in the database. Therefore, in the present study, the true biopsy sensitivity of examining dermatologists could not be determined directly. Thus, the pooled biopsy sensitivity was artificially high and specificity was relatively low. One could, in fact, take an extreme position that the examining physicians had 100% biopsy sensitivity and 0% biopsy specificity on the lesions in the database. The present study determined that the average biopsy sensitivity to small melanomas of the readers was 71%, and the corresponding specificity was 49%. However, only 2 small melanomas in situ were missed by all readers (ie, the combined biopsy sensitivity was 96%). At the same time, the combined biopsy specificity was 6% (only 3 lesions were TN for all readers). The combined biopsy sensitivity and specificity are, therefore, similar to the pooled high sensitivity and low specificity of more than 30 examining dermatologists, who contributed to the database using a variety of approaches to decide on lesion management. Thus, the sensitivity to melanoma can be increased by combining evaluations of lesions by multiple physicians but at the cost of reduced specificity. In practice, this means that, if a patient could consult 10 dermatologists about a single pigmented lesion, the probability that a small melanoma would be biopsied is 96%. However, if a patient is examined by a single dermatologist then the probability that a small melanoma would be biopsied is 71%.

Previously reported sensitivities and specificities for the clinical diagnosis of melanoma are higher than our findings, with sensitivities ranging anywhere from 58% to more than 90% and specificities ranging from 77% to 99%.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Invasive Malignant Melanoma</th>
<th>In Situ Malignant Melanoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human diagnostic sensitivity</td>
<td>48 (36-59)</td>
<td>33 (22-44)</td>
</tr>
<tr>
<td>Human biopsy sensitivity</td>
<td>81 (71-91)</td>
<td>63 (52-74)</td>
</tr>
<tr>
<td>Computer biopsy sensitivity</td>
<td>100 (87-100)</td>
<td>96 (88-100)</td>
</tr>
<tr>
<td>$P$ value for biopsy sensitivity</td>
<td>.11</td>
<td>.02</td>
</tr>
</tbody>
</table>

$^a$Data are given as average percentage (95% confidence interval) unless otherwise indicated.

Table 3. Average Sensitivity to Invasive and In Situ Melanomas

<table>
<thead>
<tr>
<th>Variable</th>
<th>Invasive Malignant Melanoma</th>
<th>In Situ Malignant Melanoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human diagnostic sensitivity</td>
<td>48 (36-59)</td>
<td>33 (22-44)</td>
</tr>
<tr>
<td>Human biopsy sensitivity</td>
<td>81 (71-91)</td>
<td>63 (52-74)</td>
</tr>
<tr>
<td>Computer biopsy sensitivity</td>
<td>100 (87-100)</td>
<td>96 (88-100)</td>
</tr>
<tr>
<td>$P$ value for biopsy sensitivity</td>
<td>.11</td>
<td>.02</td>
</tr>
</tbody>
</table>

$^a$Data are given as average percentage (95% confidence interval) unless otherwise indicated.
they found that the addition of dermoscopy to naked-eye examination did not provide a statistically significant improvement in sensitivity. In contrast, Bono et al reported that the addition of dermoscopy to the clinical assessment of small-diameter melanomas allowed for a higher rate of recognition, from 49% with naked-eye examination to 72% with dermoscopic techniques. Others have also reported improvement in the diagnostic accuracy with the use of dermoscopy, especially in small lesions. The discordant findings among these studies could be attributable to the different levels of expertise of examiners, to different proportions of invasive and in situ melanomas, and to differences in the prevalence of melanoma included in the sample. In a meta-analysis that evaluated the dermoscopic assessment of pigmented lesions, Kittler et al surmised that, as the proportion of melanoma cases that composed a study subset increased, the diagnostic difficulty of the sample increased.

Our results must be examined within the context of the clinical presentation of small melanomas, which often make them more difficult to identify than larger-diameter malignant melanomas. Thomas et al found that the specificity of melanoma diagnosis increases with the number of ABCDE criteria present. Small melanomas, however, do not always display the features typically used to diagnose malignant melanoma. Bergman et al found that small melanomas have a different clinical presentation than large melanomas because smaller lesions do not always exhibit the characteristic ABCD features, and others suggested that malignant melanomas do not display typical clinical and histologic features of melanoma until the lesions are larger than 5 mm. Because morphologic features that help distinguish melanomas from benign lesions may not be visible to the naked eye, it is not surprising that, in light of their clinical presentation, small melanomas are missed much more frequently. The present study found that small melanomas are also difficult to diagnose by dermoscopic evaluation. Given these circumstances, emphasis should, thus, be placed on the E criterion when evaluating small pigmented lesions, which Abbasi et al defined as evolving (ie, change over time in size, shape, color, surface features, or symptoms). In fact, Kamino et al reported that the most frequent reason for excision of small melanomas was a new or changing lesion. Likewise, Helsing and Loeb noted that a change in color was more frequently seen in small melanomas than in larger melanomas.

In conclusion, the differentiation of small melanomas from small benign pigmented lesions challenges even expert physicians. In comparing human evaluation with a computer-vision system, we found that the computer-vision system recommended biopsy for 98% of small melanomas (1 melanoma in situ missed), whereas dermoscopists, on average, would biopsy only 71% of small melanomas (29% missed, both in situ and invasive). By analyzing features indiscernible to the human eye, automated systems could assist dermatologists in the selection of lesions suggestive of disease for biopsy to rule out melanoma. Not only will the addition of such diagnostic tools limit the number of biopsies necessary to rule out melanoma on clinically suspicious yet histologically benign pigmented lesions, it will also aid in the detection and treatment of melanomas during the curable stages of development.

Accepted for Publication: December 7, 2007.

Author Affiliations: Department of Dermatology, New York University School of Medicine, New York (Drs Friedman, Warycha, Kopf, Polsky, and Rigel and Mss Farber, Schneider-Kels, and Papastathis), Electro-Optical Sciences Inc, Irvington (Dr Gutkowicz-Krusin), and Memorial Sloan-Kettering Cancer Center, New York (Dr Marghoob), New York; Departments of Dermatology (Drs Mihm and Tsao) and Dermatopathology (Dr Mihm), Massachusetts General Hospital, and Harvard Medical School (Dr Mihm), Boston, Massachusetts; Knoxville Dermatopathology Laboratory, Knoxville, Tennessee (Drs Googe and King); The University of Texas M. D. Anderson Cancer Center, Houston (Dr Prieto); Skin and Cancer Associates, Plantation (Dr Rabinovitz and Ms Oliviero), Dermatology Associates of Tallahassee, Tallahassee (Dr Cognetta), and Department of Dermatology, University of Miami School of Medicine, Miami (Dr Johr), Florida; Department of Dermatology, University of British Columbia, and General Hospital and British Columbia Cancer Agency, Vancouver (Dr Rivers); and University of Connecticut Health Center, Farmington (Dr Grant-Kels).

Correspondence: Robert J. Friedman, MD, Department of Dermatology, New York University School of Medicine, 124 E 72nd St, New York, NY 10021.

Author Contributions: Dr Friedman had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Friedman, Gutkowicz-Krusin, Farber, Warycha, Schneider-Kels, and Prieto. Acquisition of data: Gutkowicz-Krusin, Mihm, Googe, King, Prieto, Polsky, Rabinovitz, Oliviero, Cognetta, Rivers, and Tsao. Analysis and interpretation of data: Friedman, Gutkowicz-Krusin, Farber, Warycha, Schneider-Kels, Papastathis, and Prieto. Drafting of the manuscript: Friedman, Gutkowicz-Krusin, Farber, Warycha, Schneider-Kels, Papastathis, and Prieto. Critical revision of the manuscript for important intellectual content: Friedman, Gutkowicz-Krusin, Farber, Warycha, Schneider-Kels, Papastathis, Mihm, Googe, King, Prieto, Kopf, Rabinovitz, Oliviero, Marghoob, Johr, and Grant-Kels. Statistical analysis: Gutkowicz-Krusin. Obtained funding: Friedman. Administrative, technical, and material support: Gutkowicz-Krusin, Farber, Warycha, Schneider-Kels, Papastathis, and Cognetta. Study supervision: Friedman, Kopf, Rigel, Grant-Kels, and Tsao.

Financial Disclosure: Dr Friedman is a shareholder, former director, consultant, and member of the Scientific Advisory Committee for Electro-Optical Sciences Inc (EOS); Dr Gutkowicz-Krusin is an employee of EOS; Mss Farber and Schneider-Kels and Dr Warycha received research support from EOS; Dr Mihm is a member of the Scientific Advisory Committee and the chief dermatopathologist for EOS; Drs Googe and King are dermatopathologists for EOS studies; Dr Prieto is a dermatopathologist for and is involved in a research project with...