Clinical Studies

Cancer Research

Real-time Raman Spectroscopy for *In Vivo* Skin Cancer Diagnosis

Harvey Lui[1,2], Jianhua Zhao[1,2], David McLean[1], and Haishan Zeng[1,2]

Abstract

Raman spectroscopy is a noninvasive optical technique capable of measuring vibrational modes of biomolecules within viable tissues. In this study, we evaluated the application of an integrated real-time system of Raman spectroscopy for *in vivo* skin cancer diagnosis. Benign and malignant skin lesions ($n = 518$) from 453 patients were measured within 1 second each, including melanomas, basal cell carcinomas, squamous cell carcinomas, actinic keratoses, atypical nevi, melanocytic nevi, blue nevi, and seborrheic keratoses. Lesion classification was made using a principal component with general discriminant analysis and partial least-squares in three distinct discrimination tasks: skin cancers and precancers from benign skin lesions [receiver operating characteristic (ROC) = 0.879]; melanomas from nonmelanoma pigmented lesions (ROC = 0.823); and melanomas from seborrheic keratoses (ROC = 0.898). For sensitivities between 95% and 99%, the specificities ranged between 15% and 54%. Our findings establish that real-time Raman spectroscopy can be used to distinguish malignant from benign skin lesions with good diagnostic accuracy comparable with clinical examination and other optical-based methods. *Cancer Res; 72(10); 2491–500. ©2012 AACR.*

Introduction

The clinical diagnosis of skin cancer is based on visual examination followed by biopsy of suspicious lesions. The accuracy of clinicians is highly variable according to the level of formal training and experience. For example, in a retrospective study of 4,741 pigmented skin lesions evaluated by 468 general practitioners, the biopsy ratio, defined as the number of nonmelanoma lesions that underwent biopsy for each confirmed case of melanoma ranged from 58:1 to 21:1 for new versus experienced general practitioners, respectively (1). A number of studies have shown that the accuracy of clinical diagnosis of melanoma by dermatologists varies between 49% and 81%, with approximately one third of melanomas being misdiagnosed as benign lesions (2–6).

Raman spectroscopy is a noninvasive optical method under investigation for cancer diagnosis (7). Arising from the inelastic scattering of light within tissue, Raman signals correlate with the molecular vibrations of various tissue biomolecules. The positions and relative magnitudes of spectral peaks correspond to the vibrational energies associated with specific chemical bonds. Raman spectroscopy is capable of detecting molecular and/or biochemical changes associated with pathology (8). The probability of inelastic Raman scattering is exceedingly low, and as a consequence, long integration times are required to acquire sufficient scattering signals for a single spectrum. For example, a traditional Fourier-transform Raman system requires up to 30 minutes of integration time to acquire one spectrum. Most prior studies involving the skin have been limited to either *ex vivo* samples or a few *in vivo* skin measurements, all requiring relatively long integration times (9–14). The clinical use of *ex vivo* Raman spectroscopy is quite limited, as suspect lesions must first be biopsied, which necessarily entails an invasive procedure. In a recent *in vivo* study of nonmelanoma skin cancer diagnosis using Raman microscopy, Lieber and colleagues measured 21 lesions and their adjacent normal skin [9 basal cell carcinoma (BCC), 4 squamous cell carcinoma (SCC), and 8 inflamed scar] with an integration time of 30 seconds and reported 100% sensitivity and 91% specificity for discriminating lesional from normal skin (15). One significant limitation of previous Raman studies of skin cancer is their small sample sizes.

We have developed a rapid, real-time Raman spectrometer system for *in vivo* skin measurements (16) that substantially reduces spectral acquisition time to less than 1 second. This system was optimized by combining a unique signal binning method and software processing for direct point-of-care use (17). We now have *in vivo* measurements of more than 1,000 cases of skin cancers and other skin diseases and report the results and performance of this system for guiding the clinical evaluation of skin lesions.

Authors' Affiliations: [1]Photomedicine Institute, Department of Dermatology and Skin Science, University of British Columbia and Vancouver Coastal Health Research Institute; and [2]Imaging Unit, Integrative Oncology Department, British Columbia Cancer Agency Research Centre, Vancouver, British Columbia, Canada

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Corresponding Author: Harvey Lui, Department of Dermatology and Skin Science, University of British Columbia, 835 West 10th Avenue, Vancouver, British Columbia, Canada V5Z 4E8. Phone: 604-875-4747; Fax: 604-873-9919; E-mail: harvey.lui@ubc.ca

doi: 10.1158/0008-5472.CAN-11-4061

©2012 American Association for Cancer Research.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

Patients and Methods

Patients

This study was approved by the Clinical Research Ethics Board of the University of British Columbia (Vancouver, BC, Canada; Protocol C96-0499). Patients older than 18 years of age attending the Vancouver General Hospital Skin Care Centre were invited to volunteer in this study if they provided informed consent and had any discrete skin lesions amenable to spectral characterization. Patients with lesions of potential concern for skin cancer as well as those with incidental skin lesions of clinical interest were considered for this study. Subjects underwent spectral measurements of up to 10 separate skin lesions, each with its own diagnosis. Lesions were not considered for inclusion if they were less than 1 mm in lateral dimension (which cannot be measured by the spectrometer), located at a body site that was inaccessible to the spectrometer probe, were infected, or had previously been biopsied, excised, or traumatized.

Between January 2003 and May 2011, Raman spectra were acquired from 1,022 separate benign and malignant skin lesions from 848 patients. The analysis presented here is focused specifically on those diagnostic classes of skin lesions that characteristically give rise to patient and physician concern over skin cancer including: (i) malignancies and premalignancies that require treatment: malignant melanoma, SCC, BCC, and actinic keratosis and (ii) benign conditions that can visually mimic skin cancer: seborrheic keratosis, atypical nevi, melanocytic nevi (junctional, compound, and intradermal), and blue nevi. There were a total of 553 such lesions including 35 that were invalidated for analysis because their spectra had obvious spurious Raman peaks arising from accidental ambient light leakage into the spectrometer system for a few weeks during the spring of 2009. The final diagnosis for each measured lesion was established through (i) clinical evaluation by 1 of 2 experienced dermatologists (H. Lui or D. McLean) and/or (ii) histopathologic analysis if a skin biopsy of the lesion was taken subsequent to the optical Raman measurement. All of the lesions deemed to be cancerous were confirmed by skin biopsy; 31% of the premalignant lesions (i.e., actinic keratoses) and 28% of the benign lesions underwent skin biopsy. Dermoscopy was not used as an aid for establishing the final diagnosis of any of the lesions. The final data set thus consisted of 518 validated lesions from 453 subjects (224 male, 229 female), aged 18 to 94 years (median, 61 years). Of these, 313 (60%) lesions underwent subsequent treatment including 44 malignant melanomas, 47 SCCs, 109 BCCs, and 32 actinic keratoses.

The detailed distribution of the patients and lesions including diagnostic subtypes and location is provided in Table 1 and Supplementary Table S1. For the purposes of this study, each individual lesion was considered an experimental unit for analysis.

Table 1. Summary of patients and lesions evaluated by Raman spectroscopy

	Subjects					Location			
Final lesion diagnosis	Mean age, y (range)	Male	Female	Number of lesions	Number biopsied (%)	Head and neck	Trunk	Upper limb	Lower limb
MM									
LM	69 (51–88)	12	8	20	20 (100)	19	1	0	0
LMM	67 (42–85)	7	1	8	8 (100)	8	0	0	0
SS	60 (22–77)	6	8	14	14 (100)	3	3	7	1
MM other	61 (60–62)	2	0	2	2 (100)	1	1	0	0
BCC									
Superficial	63 (34–86)	10	13	28	28 (100)	10	9	5	4
Nodular	66 (39–94)	34	29	73	73 (100)	52	10	9	2
Pigmented	67 (46–83)	2	4	6	6 (100)	2	4	0	0
Other BCC	68 (60–75)	1	1	2	2 (100)	1	1	0	0
SCC									
In situ	70 (56–88)	12	5	18	18 (100)	7	4	5	2
Invasive	66 (39–94)	16	10	28	28 (100)	16	1	5	6
Other SCC	78	0	1	1	1 (100)	1	0	0	0
Actinic keratosis	66 (43–92)	13	14	32	10 (31.3)	28	0	3	1
Atypical nevus	48 (20–75)	22	26	57	24 (42.1)	3	39	8	7
Junctional nevus	43 (18–70)	12	17	34	4 (11.8)	5	11	15	3
Compound nevus	35 (18–67)	13	15	30	6 (20)	9	8	9	4
Intradermal nevus	50 (28–83)	9	26	38	12 (31.6)	21	8	7	2
Blue nevus	37 (18–66)	4	9	13	4 (30.8)	4	1	6	2
Seborrheic keratosis	64 (25–89)	49	42	114	31 (27.2)	47	47	14	6

Abbreviations: MM, malignant melanoma; MM subtypes: LM, lentigo maligna; LMM, lentigo maligna melanoma; SS, superficial spreading melanoma.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

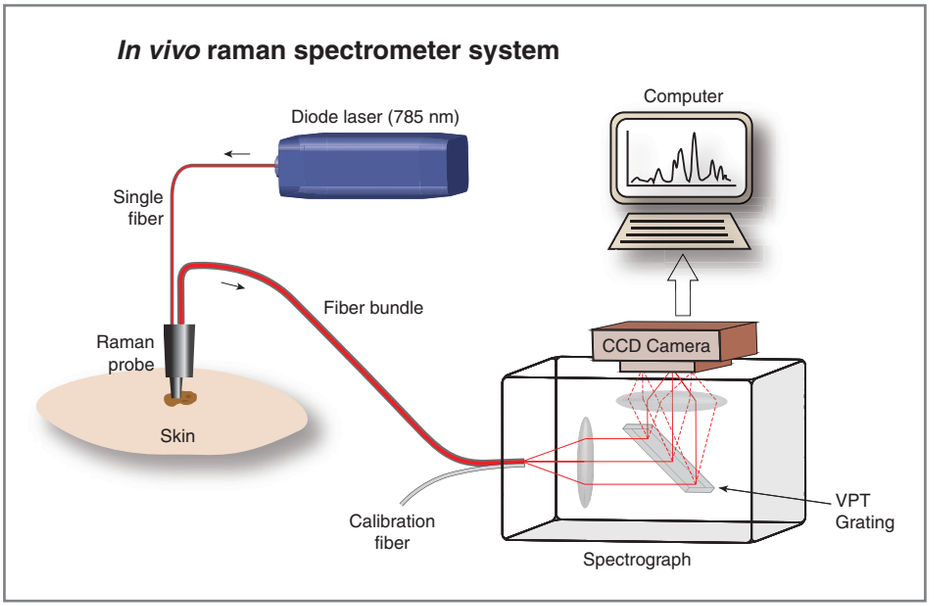

Figure 1. Schematic configuration of the real-time Raman system for skin cancer diagnosis.

Instrumentation

The integrated real-time Raman system was developed in-house and is schematically shown in Fig. 1 (16, 17). The hardware comprises a diode laser, a fiber and fiber bundle delivery system, a hand-held Raman probe, a spectrograph, a charge coupled device (CCD) camera detector, and a computer. A 785-nm laser beam is delivered to the Raman probe through a 200-μm core diameter single fiber and illuminates a 3.5-mm diameter skin area. The raw signal from the skin, which is composed of the Raman scattering signal and tissue autofluorescence, is collected by the probe and transmitted to the spectrometer through a fiber bundle for spectral analysis.

The integrated software contains all calibration procedures and real-time data processing, including intensity calibration and fluorescence background removal according to the Vancouver Raman algorithm with a fifth-order polynomial fitting (17). The effective spectral range of the system is 500 to 1,800 cm^{-1} with a resolution of 8 cm^{-1}.

To take a Raman measurement, the hand-held spectrometer probe was placed in gentle contact with the target skin site without compressing it. Spectral measurements for skin lesions of interest were taken in duplicate by separately measuring each lesion itself and then the normal-appearing surrounding skin from the same anatomic region. The "normal" skin measurement site was usually within 5 cm of the visible border of the corresponding skin lesion. Each spectral measurement was acquired with a 1-second integration time, facilitating measurements from multiple paired sites (i.e., lesion and normal) in a given patient, where applicable. Most of the lesions were measured once, but larger and morphologically inhomogeneous lesions were measured up to 3 times at different locations within the lesion, particularly for malignant melanomas (34%) and SCCs (17%). For these cases, the average of the multiple spectra was used to represent these lesions for statistical analysis.

Statistical analysis

To evaluate the reproducibility of the Raman measurements, we conducted a separate study where repeated spectra were taken from the same sites in triplicate from 15 different skin lesions and 15 different normal skin locations (data not shown). Variability in the Raman frequency shifts by wave number (abscissa) for any 3 consecutive spectra from the same skin site was negligible, confirming the relative stability of spectral peak positions. However, in terms of the Raman signal intensities (ordinate), the variances for the triplicate measurement sets showed a systematic change that was wave number–dependent, with relatively smoother spectra at lower Raman frequencies and increasing fluctuations at higher frequencies. We calculated the successive variances at each Raman frequency and were able to define 1,055 cm^{-1} as a frequency at which the smoother portion of the spectra reached a minimum variance for nearly all the repeat-matched spectral measurements. On the basis of this analysis of Raman measurement reproducibility, the skin spectra were analyzed according to the full acquired spectrum (500–1,800 cm^{-1}) and also by separately considering only lower (500–1,055 cm^{-1}) and higher frequency (1,055–1,800 cm^{-1}) bands. For the full and partial spectral analyses (i.e., full, lower, and higher frequency band analyses), the spectra were first normalized to their integrated spectral areas under the curve (AUC) according to the respective spectral range being analyzed.

The lesion types selected for analysis in this study primarily included common conditions whose clinical behavior and appearance raise suspicion for skin cancer. The diagnostic performance of *in vivo* Raman spectroscopy for classifying lesions was tested according to 3 tasks or dichotomous groupings based on clinical relevance. In the first task, we considered the ability of Raman spectra to discriminate cancerous and precancerous lesions that require treatment (malignant

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

melanoma, BCC, SCC, and actinic keratosis) versus benign conditions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, and seborrheic keratosis). The other 2 tasks tested the discrimination of melanoma (all forms, malignant melanoma) versus benign pigmented skin lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, and seborrheic keratosis), and melanoma (all forms, malignant melanoma) versus seborrheic keratoses, which can also be confused because of similarities in appearance (2, 3).

Principal component with generalized discriminant analysis (PC-GDA) and partial least-squares (PLS) were each used separately for lesion classification (18), according to the 3 dichotomous groupings of interest. All multivariate data analyses were implemented within MATLAB (version 2010a, MathWorks) and STATISTICA (version 6.0, StatSoft) based on leave-one-out cross-validation (LOO-CV).

For each PC-GDA analysis, successive single lesional spectra were left out for "testing," with the remaining spectra being used for "training." The PC factors and the PC loadings of the training spectra were calculated. A discrimination model was developed on the basis of the PC factors derived from the training spectra where the classifications were known *a priori*. The PC factors of the test spectra were then calculated on the basis of the PC loadings of the training spectra and tested against the discrimination model for classification. A posterior probability of the test lesion for skin cancer was calculated. The posterior probabilities of all lesions were obtained by inputting each lesion as a test according to the LOO-CV protocol.

PLS is typically used for predicting the concentration of components within samples (19, 20). Recently, it was found that PLS could theoretically also be used for discrimination, and in some situations, it was preferred over PC analysis for that purpose (21). In this article, we used a nonlinear iterative PLS algorithm (18). Similar to PC-GDA, all PLS analyses are based on LOO-CV.

The receiver operating characteristic (ROC) curve, that is, sensitivity versus (1 − specificity) was calculated from the posterior probabilities derived above and represents the diagnostic performance. With good discrimination between 2 groups, the ROC curve moves toward the left and top boundaries of the graph, whereas poor discrimination yields a curve that approaches the diagonal line function. The AUCs were calculated using the trapezoidal rule (22). The significance of the AUCs and comparisons between different AUCs were carried out in a standard fashion (23–25). All ROC analyses were based on nonparametric techniques and were conducted separately for the PC-GDA and PLS analyses and for each of the 3 classification tasks.

Skin cancer biopsy ratios according to Raman spectroscopy

The skin cancer biopsy ratio is defined as the number of negative biopsies that are conducted for each true-positive biopsy showing skin cancer. If the decision to biopsy is guided solely by the results of Raman spectroscopy, the corresponding biopsy ratio can be estimated from the ROC analysis by dividing false-positives by true-positives. This ratio depends on the desired sensitivity (more biopsies must be taken to avoid missing any skin cancers) and the accuracy of the diagnostic test (with higher accuracy there will be fewer skin biopsies that are negative for cancer). To compare Raman spectroscopy with other noninvasive diagnostic techniques as well as with clinical diagnosis by visual examination, we calculated skin biopsy ratios at sensitivity levels of 90%, 95%, and 99%, respectively.

Covariate analysis: body site and biopsy status

From previous studies, we have found that the *in vivo* Raman spectra of normal human skin varies according to body site (26). To test whether the discrimination capability of the Raman spectrometer system for skin cancer may have been influenced by body site, we did a subanalysis to discriminate malignant melanoma from other nonmelanoma pigmented lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis) using only the lesions located on the head and neck. Insufficient lesion numbers at other body sites did not permit site-specific subanalyses across all measured lesions. Another means of evaluating the influence of body site on the diagnostic performance of Raman was to analyze the lesions according to the paired measurements of diseased and normal surrounding skin that were taken. This was done by analyzing the difference spectra between lesional and adjacent normal skin by PC-GDA and using this to discriminate malignant melanomas from nonmelanoma pigmented skin lesions. The goal here was to assess whether the difference spectra or the lesional spectra alone have the higher discriminating capability.

In this study, not all of the benign skin lesions were biopsied, and the final diagnosis in those cases was made on clinical grounds by experienced skin oncologists. We therefore also did 2 additional analyses by including only those cases, benign or malignant, where the final diagnosis was established through skin biopsy combined with clinical examination. One of the tests evaluated biopsied malignant melanoma versus biopsied nonmelanoma pigmented lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis), and the other involved biopsied skin cancers (malignant melanoma, SCC, BCC) versus biopsied noncancerous lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, actinic keratosis).

Results

The mean Raman spectra for different skin pathologies in this study are depicted in Fig. 2. All spectra were normalized to their respective AUCs before being averaged in aggregate according to diagnosis. Overall the skin lesions included in this study all appear to share similar major Raman peaks and bands. The strongest Raman peak is located around 1,445 cm^{-1} with other major Raman bands centered at 855, 936, 1,002, 1,271, 1,302, 1,655, and 1,745 cm^{-1}. There are no distinctive Raman peaks or bands that could be uniquely assigned to specific skin cancers by visual inspection alone. Statistical methods are thus used to extract the diagnostic information

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

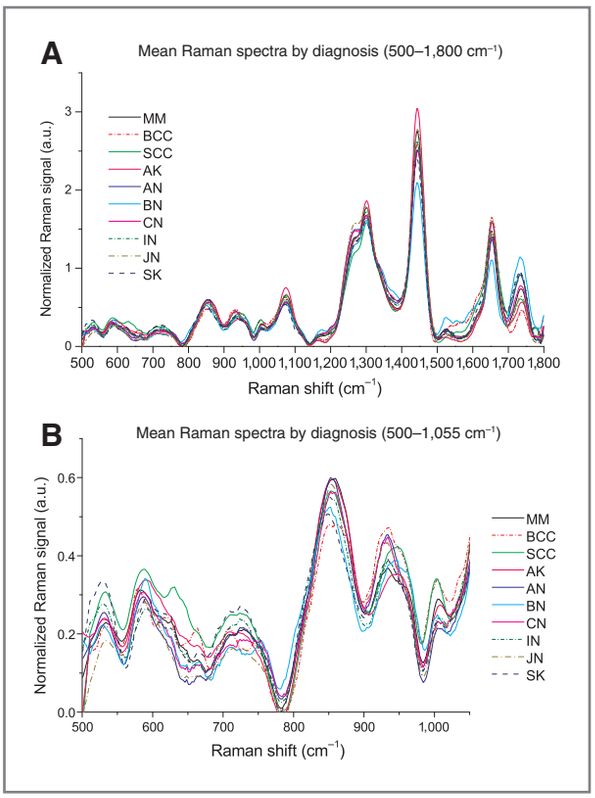

Figure 2. Mean Raman spectra by diagnosis. All lesion spectra were normalized to their AUCs before averaging by diagnosis. Spectral variations were analyzed by PC-GDA and PLS methods for classification purposes. A, full-range spectrum results displayed from 500 to 1,800 cm^{-1}. B, spectral results plotted for lower frequency range only (500–1,055 cm^{-1}) show variability according to lesion diagnosis. Note that the *y*-axis scales for A and B are different. AK, actinic keratosis; AN, atypical nevus, BN, blue nevus; CN, compound nevus; IN, intradermal nevus; JN, junctional nevus; MM, malignant melanoma; SK, seborrheic keratosis.

that is embedded in these seemingly complex and data-rich Raman spectra.

Raman spectroscopy distinguishes skin cancers from benign lesions (classification by PC-GDA)

Cancerous and precancerous skin conditions (cancer plus actinic keratosis) versus benign skin lesions (noncancer). When Raman spectroscopy is used to distinguish cancerous and precancerous skin lesions, which require treatment (n = 232), from benign skin lesions (n = 286) that can simply be observed, the ROC AUC is 0.879 [95% confidence interval (CI), 0.829–0.929, PC-GDA] and statistically significant ($P < 0.001$). The results are depicted in Fig. 3 and Table 2.

Figure 3A shows the posterior probability for each lesion to be classified as a skin cancer or precancer. From the distribution of posterior probabilities, the ROC curve with 95% CIs is generated and shown in Fig. 3D. At a sensitivity of 90%, the overall specificity is more than 64%, with a positive predictive value (PPV) of 67% and negative predictive value (NPV) of 89%. The estimated biopsy ratio is 0.5:1. Table 3 shows the corresponding parameters (specificity, PPV, NPV, and biopsy ratio) for sensitivities of 90%, 95%, and 99%.

Although generally treated once they are detected, some actinic keratoses can undergo spontaneous regression with time, and in this respect, actinic keratoses are therefore not strictly considered to represent skin malignancies. Thus, because the diagnosis and treatment of actinic keratosis are distinct from that of malignant melanomas, BCCs, and SCCs, we did a further analysis by reclassifying actinic keratosis with the benign category. In this situation, the AUC of the ROC curve for discriminating skin cancers (malignant melanoma, BCCs, SCCs, n = 200) from non-skin cancers (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, and actinic keratosis, n = 318) is 0.863 (95% CI, 0.830–0.895; $P < 0.001$). For a sensitivity of 90%, the overall specificity is more than 63%, with a PPV of 60%, NPV of 91%, and biopsy ratio of 0.7:1. Raman spectroscopy appears to detect cancerous skin lesions irrespective of whether actinic keratosis is considered either benign or malignant.

Melanoma (malignant melanoma) versus nonmelanoma pigmented skin lesions. When only lesions with pigment are considered, Raman spectroscopy can separate malignant melanoma (n = 44) from nonmelanoma pigmented skin lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, n = 286; Fig. 3B) with an ROC AUC of 0.823 (95% CI, 0.731–0.915, $P < 0.001$, PC-GDA). Our results showed the biopsy ratio, based on Raman spectroscopy, ranged from 5.6:1 to 2.3:1 for sensitivities corresponding to 99% to 90% and specificities from 15% to 68%, respectively (Fig. 3E and Tables 2 and 3).

Melanoma (malignant melanoma) versus seborrheic keratosis. Figure 3C shows the posterior probabilities for cases of melanoma or seborrheic keratosis to be classified by Raman spectroscopy as melanoma. The AUC for the corresponding ROC curve is 0.898 (95% CI, 0.797–0.999, $P < 0.001$; Fig. 3F). The biopsy ratio ranges from 2.2:1 to 0.9:1, for sensitivities ranging from 99% to 90% and specificities of 25% to 68% (Table 3).

The performance and other diagnostic parameters for real-time Raman spectroscopy are summarized in Tables 2 and 3, including the AUCs of the ROC curves as well as the sensitivities, specificities, PPVs, NPVs, and biopsy ratios.

PLS analysis also shows the discriminative capability of Raman spectroscopy for skin cancer

The PLS approach yields results similar to those for PC-GDA in terms of the 3 classification tasks above (Tables 2 and 3). Furthermore, on the basis of the algorithm by Hanley and McNeil (23), there are no significant differences between the PLS and PC-GDA methodologies for the 3 analyses of interest (P = 0.0644, 0.7494, and 0.1646, respectively). There are fewer factors for the PLS model than for the PC-GDA analysis (4–8 vs. 15–30 PCs).

Diagnostic performance is spectral band–dependent

We conducted PC-GDA and PLS analyses using the 3 Raman bands (500–1,055, 1,055–1,800, and 500–1,800 cm^{-1}) and found that the higher spectral range from 1,055 to 1,800 cm^{-1} was optimal for differentiating melanomas from nonmelanoma

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

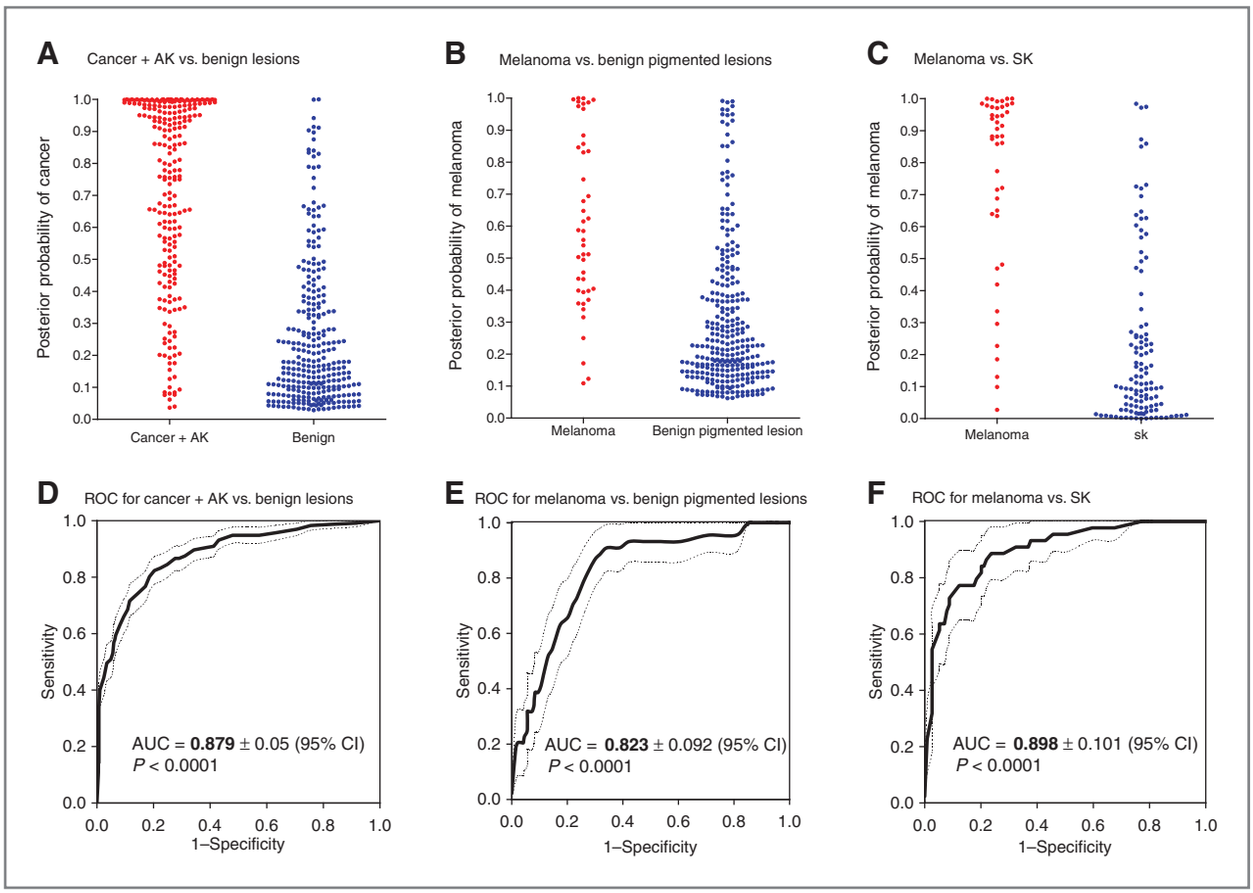

Figure 3. Lesion classification by Raman spectroscopy based on PC-GDA analysis. Posterior probabilities for discriminating (A) skin cancers and precancers (cancer + actinic keratosis, including MM, BCC, SCC, AK, $n = 232$) from benign skin disorders (including atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, $n = 286$; A), melanoma ($n = 44$) from benign pigmented skin diseases (including atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, $n = 286$; B), and melanoma ($n = 44$) from seborrheic keratosis [($n = 114$); C]. D–F, the corresponding ROC curves and 95% CIs are derived from the respective posterior probabilities, and all AUCs are significant ($P < 0.0001$). AK, actinic keratosis; SK, seborrheic keratosis.

pigmented lesions or melanomas from seborrheic keratoses; the full spectral range from 500 to 1,800 cm^{-1} was optimal for separating skin cancers and/or precancers from benign skin lesions.

Classification by Raman spectroscopy is not influenced by lesion location

There were sufficient data to do a subanalysis for lesions from only the head and neck in terms of discriminating malignant melanoma ($n = 31$) from other nonmelanoma pigmented lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, $n = 89$). The AUC of this ROC curve is 0.789 (95% CI, 0.698–0.879; Supplementary Fig. S1), which is close to the classification results that incorporated all body sites (Fig. 3E, $n = 330$, AUC = 0.823); these 2 ROC curves were not statistically different ($P = 0.5493$).

For each skin lesion assessed, we also measured the Raman spectra of the adjacent normal-appearing skin. We used

Table 2. Diagnostic performance of Raman spectroscopy according to optimized wavebands

		ROC AUC	
Diagnosis classification task	**Raman waveband, cm^{-1}**	**PC-GDA (95% CI)**	**PLS (95% CI)**
Skin cancers + actinic keratosis vs. benign lesions	500–1,800	0.879 (0.829–0.929)	0.896 (0.846–0.946)
Melanoma vs. benign pigmented lesions	1,055–1,800	0.823 (0.731–0.915)	0.827 (0.735–0.929)
Melanoma vs. seborrheic keratosis	1,055–1,800	0.898 (0.797–0.999)	0.894 (0.793–0.995)

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

Table 3. Summary of Raman spectroscopy diagnostic parameters derived from ROCs according to various levels of sensitivity

		PC-GDA				PLS			
Diagnosis classification task	**Sensitivity level (95% CI)**	**Specificity (95% CI)**	**PPV[a]**	**NPV[b]**	**Biopsy ratio**	**Specificity (95% CI)**	**PPV**	**NPV**	**Biopsy ratio**
Skin cancers + actinic keratosis vs. benign lesions	0.99 (0.98–1.00)	0.17 (0.13–0.21)	0.49	0.95	1.03:1	0.24 (0.19–0.29)	0.51	0.97	0.95:1
	0.95 (0.92–0.99)	0.41 (0.35–0.48)	0.57	0.91	0.77:1	0.52 (0.48–0.58)	0.62	0.93	0.62:1
	0.90 (0.86–0.94)	0.64 (0.58–0.70)	0.67	0.89	0.49:1	0.66 (0.61–0.71)	0.68	0.89	0.47:1
Melanoma vs. benign pigmented lesions	0.99 (0.96–1.00)	0.15 (0.11–0.19)	0.15	0.99	5.58:1	0.14 (0.10–0.18)	0.15	0.99	5.65:1
	0.95 (0.89–1.00)	0.38 (0.32–0.44)	0.19	0.98	4.24:1	0.44 (0.38–0.50)	0.21	0.98	3.83:1
	0.90 (0.81–0.99)	0.68 (0.63–0.73)	0.30	0.98	2.31:1	0.63 (0.57–0.69)	0.27	0.98	2.67:1
Melanoma vs. seborrheic keratosis	0.99 (0.96–1.00)	0.25 (0.17–0.33)	0.34	0.98	1.96:1	0.46 (0.37–0.55)	0.41	0.99	1.41:1
	0.95 (0.89–1.00)	0.54 (0.45–0.63)	0.44	0.97	1.25:1	0.52 (0.43–0.61)	0.43	0.96	1.31:1
	0.90 (0.81–0.99)	0.68 (0.59–0.77)	0.52	0.95	0.92:1	0.66 (0.57–0.75)	0.51	0.94	0.98:1

[a]PPV is the ratio of true-positives to the total of true-positives and false-positives.
[b]NPV is the ratio of true-negative to the total of true-negatives and false-negatives.

PC-GDA to discriminate malignant melanoma from nonmelanoma pigmented skin lesions using the difference spectra between the lesions and their adjacent normal sites. The classification results were not as good as those using only the lesional spectra, indicating that the most useful Raman information was embedded in the lesions themselves rather than in any apparent differences between the lesions and normal skin. The AUC of the ROC curve for discriminating malignant melanoma ($n = 44$) versus nonmelanoma pigmented lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, $n = 286$) is only 0.577 (95% CI, 0.500–0.670) when using the difference spectra between corresponding lesions and normal skin (Supplementary Fig. S2), which is statistically no different from guessing ($P = 0.09691$). The comparable ROC AUC when lesional spectra alone are used is 0.823 (Fig. 3E).

Results according to histopathology as a gold standard

The gold standard for skin cancer diagnosis is histopathologic examination of biopsied skin. In this study, not all the benign skin lesions underwent biopsy. As shown in Table 1, all skin cancer cases were confirmed through biopsy and clinicopathologic correlation, whereas noncancerous (i.e., benign) lesions underwent biopsy between 12% and 42% of the time, depending on the diagnosis. For the noncancerous lesions, the reasons for biopsy were either that the appearance of the lesion was difficult to clinically distinguish from skin cancer or the patient gave consent to both Raman measurement and skin biopsy. Lesions deemed to be benign without biopsy were diagnosed on the basis of visual examination by one of the dermatologist investigators. Two separate analyses were conducted according to biopsy status, and the overall results remained essentially the same. In one study, we discriminated malignant melanoma ($n = 44$, all of which were biopsied) from biopsied nonmelanoma pigmented lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, $n = 81$). The AUC of the ROC curve was 0.833 (95% CI, 0.761–0.906; Supplementary Fig. S3), close to the AUC of all cases regardless of biopsy status (0.823; Fig. 3E) and statistically not different ($P = 0.8238$). In the other study, we tried to discriminate confirmed skin cancers (malignant melanoma, SCC, BCC, $n = 200$, all of which were biopsied) from biopsy-verified noncancerous lesions (atypical nevi, blue nevi, compound nevi, intradermal nevi, junctional nevi, seborrheic keratosis, actinic keratosis, $n = 91$). The AUC of the ROC curve based on biopsied lesion spectra was found to be 0.833 (95% CI, 0.783–0.882; Supplementary Fig. S4), close to the results of all cases with/without biopsy (AUC = 0.863; 95% CI, 0.830–0.895) and statistically not different ($P = 0.3206$).

Discussion

One significant advantage of this system over prior Raman technologies is its ability to acquire Raman spectra with reduced integration times of seconds or less. In previous Raman studies, the objectives were primarily to discriminate skin cancers (melanoma or nonmelanoma) from normal skin. In our study, we aimed to evaluate a more complex and relevant clinical task, namely, the discrimination of melanoma and nonmelanoma skin cancers from benign skin lesions.

Raman spectra of the skin are data-rich and complex, and we have shown that relatively conservative statistical techniques can be used to extract the diagnostic information embedded within these signals. Our data also indicate that diagnostically useful information may be contained within certain spectral regions. Specifically for evaluating melanoma, the higher waveband of 1,055 to 1,800 cm^{-1} is preferred whereas for distinguishing skin cancers from benign lesions overall, the full

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

spectrum from 500 to 1,800 cm^{-1} is preferred. The spectrometer system used in this study did not extend beyond the 500 to 1,800 cm^{-1} region and it is thus unknown whether the results would be improved if these were included. Our Raman spectroscopy system was designed for the 500 to 1,800 cm^{-1} range because this is regarded as the fingerprint region wherein a denser cluster of Raman peaks can be found (9). It is conceivable that the 3 major forms of skin cancer as well as their subtypes may themselves be associated with unique Raman characteristics. The sample size did not have sufficient power to analyze the results according to these categories and subtypes.

Although normal and diseased skin share similar Raman peaks, the relative intensities of different Raman peaks vary among skin lesions (Fig. 2), which provides the basis for evaluating skin cancers and other skin diseases. PC-GDA and PLS analyses produced similar results, indicating that both analytic approaches can be used for skin cancer diagnosis. PC-GDA is the most commonly used multivariate approach for analyzing complex data sets such as Raman spectra, and it yields results that tend to be somewhat conservative.

An Australian study found that the clinical diagnosis of skin cancers and precancers was associated with a sensitivity of 63.9% for BCCs, 41.1% for SCCs, and 33.8% for malignant melanomas. The PPVs for their study were 72.7% for BCCs, 49.4% for SCCs, and 33.3% for malignant melanomas, respectively (27). Our results showed that real-time Raman spectroscopy differentiated skin cancer and precancers from benign skin lesions with an overall AUC of the ROC curve of 0.879 (95% CI, 0.829–0.929), sensitivity of 90%, and PPV of 68%. When the sensitivity is set at 95% and 99%, the PPVs are 62% and 51%, respectively.

Many melanomas may appear banal and therefore be overlooked, whereas benign pigmented lesions can sometimes show clinically suspicious features on visual examination and therefore be unnecessarily biopsied. It has been estimated that if all atypical pigmented lesions were to be biopsied to rule out melanoma, the biopsy ratio would be as high as 200:1 (28). A clinical study of 1,250 patients with somewhat conservative results showed that the biopsy ratios ranged from 576:1 for patients without personal history of melanoma to 135:1 for patients with a personal history of melanoma (29). A clinical study by an Australian group found that the biopsy ratios for general practitioners ranged from 82:1 for young patients to 10:1 for older patients. The biopsy ratio appears to be substantially reduced with the use of newer technologies such as dermoscopy, surface microscopy, and multispectral imaging (28, 30–38). Westerhoff and colleagues found that the accuracy of melanoma diagnosis could be improved from 63% to 76% with the aid of surface microscopy (30). Robinson and Nickoloff reported a biopsy ratio of 47:1 to rule out melanoma with digital epiluminescence microscopy (31). Monheit and colleagues found that the estimated biopsy ratio ranged from 7.6:1 to 10.8:1 for identifying melanoma with or without borderline lesions using a multispectral-based method (i.e., MelaFind; ref. 28). Banky and colleagues reported a biopsy ratio as low as 3:1 for patients at high risk of melanoma using a combination of baseline images and dermoscopy, but at the expense of a relatively low sensitivity of only 72% (32). Binder and colleagues investigated 120 lesions (including 39 malignant melanoma) using dermoscopy alone and found that depending on the sample size and selection of lesions, the sensitivity and specificity varied from 93% to 38% and from 84% to 50%, respectively (33). Farina and colleagues studied 237 pigmented lesions using multispectral imaging (67 malignant melanoma and 170 non–malignant melanoma) and found the AUC of the ROC curve to be 0.779, with a sensitivity of 80% and specificity of 51% (34). Moncrieff and colleagues studied 348 pigmented lesions (52 malignant melanomas) using the SIAscope, which is based on narrow-band spectral imaging, and found the sensitivity of 82.7% and specificity of 80.1% (39). Menzies studied 2,430 lesions (382 malignant melanoma), using a dermoscopy-based automated diagnostic SolarScan and found the sensitivity of 85% and specificity of 65% (35). For clinical diagnosis, they found sensitivities and specificities of 90%, 81%, 85%, and 62%; and 59%, 60%, 36%, 63% for experts, dermatologists, trainees, and general practitioners, respectively (35). Overall, our Raman results appear to be favorable in comparison with clinicians and technical diagnostic aids.

The above studies as well as our own have formally assessed specific diagnostic methods in isolation. In the clinical setting, the final diagnosis of any suspect skin lesion is actually rendered by considering all available evidence and data collectively. This is heavily influenced by clinical acumen and experience. Raman spectroscopy should therefore be viewed as a means for assisting the evaluation of suspect skin lesions rather than being a final, definitive arbiter of lesion diagnosis.

Overall, we found evidence to support the use of Raman spectroscopy for guiding skin cancer diagnosis at different levels of clinical interest, that is, malignant/premalignant versus benign, melanomas versus benign pigmented lesions, and melanomas versus seborrheic keratoses, with the ROC AUCs ranging above 0.82 for these tasks (Tables 2 and 3). For all 3 diagnostic tasks, the specificity of the Raman approach is greater than 15% at a sensitivity of 99%, and indeed higher than one study that estimated a 3.7% level of specificity for clinicians (28). Compared with these techniques discussed above, using Raman spectroscopy to guide clinical evaluation may potentially reduce the number of unnecessary biopsies by 50% to 100% (28, 32–35). Raman spectroscopy is complementary to these other noninvasive approaches and has the potential advantage of requiring less extensive user training and expertise. One important limitation of our study is that not all of the lesions deemed to be benign underwent biopsy and histopathologic confirmation. Nevertheless when only biopsy-confirmed lesions were included in the analyses, the overall results remained significant.

Raman scattering within the skin can be measured within 1 second and used to guide the diagnosis of prospective lesions in terms of their propensity for skin cancer. We envision that an algorithm derived from a database of Raman spectra from other lesions would be able to classify a given lesion in less than half a second, making this approach feasible and representing a novel clinical contribution to managing the most common form of malignancy.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

Conclusions

We have studied skin cancers and a range of benign skin diseases using real-time *in vivo* Raman spectroscopy with an integration time of less than 1 second per lesion. Multivariate PC-GDA and PLS analyses show that Raman spectroscopy can distinguish (i) malignant and premalignant lesions from benign disorders, (ii) melanomas from benign pigmented skin lesions, and (iii) melanomas from seborrheic keratoses.

Disclosure of Potential Conflicts of Interest

H. Lui, J. Zhao, D. McLean, H. Zeng, and the BC Cancer Agency hold patents for Raman spectroscopy that are licensed to Verisante Technology Inc. H. Lui has ownership interest in Lumen Health Innovations Inc. and Verisante Technology Inc; and is a consultant/advisory board member for Verisante Technology Inc. J. Zhao has ownership interest and patent ownership licensed to Verisante Technology Inc. D. McLean has ownership interest in Verisante Technology Inc. and Lumen Health Innovations Inc. H. Zeng has received a commercial research grant from Verisante Technology Inc.; ownership interest in Verisante Technology Inc. and Lumen Health Innovations Inc.; and is a consultant/advisory board member for Verisante Technology Inc.

Authors' Contributions

Conception and design: H. Lui, J. Zhao, D. McLean, H. Zeng
Development of methodology: H. Lui, J. Zhao, H. Zeng
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Lui, J. Zhao, D. McLean
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Lui, J. Zhao, H. Zeng
Writing, review, and/or revision of the manuscript: H. Lui, J. Zhao, D. McLean, H. Zeng
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Lui, J. Zhao, H. Zeng
Study supervision: H. Lui, D. McLean, H. Zeng
Directed data analysis: H. Lui, D. McLean, H. Zeng
Confirmed lesion diagnoses: H. Lui, D. McLean
Implemented the real-time Raman spectroscopy system and software development: J. Zhao, H. Zeng

All authors were involved in developing the statistical models and analysis.

Acknowledgments

The authors thank our collaborators and research team members for their contributions: Zhiwei Huang, Iltefat Hamzavi, Abdulmajeed Alajlan, Hana Alkhayat, Ahmad Al Robaee, Sunil Kalia, Bernardita Ortiz-Policarpio, Tsung-Hua Tsai, Soodabeh Zandi, Wei Zhang, Michelle Zeng, Youwen Zhou, Laurence Warshawski, David Zloty, Bryce Cowan, Michael Short, and Hequn Wang. Dr. Michael Schulzer and Edwin Mak provided statistical expertise for the Raman data analysis.

Grant Support

The project was supported by grants from the Canadian Cancer Society (#011031 and #015053), the Canadian Institutes of Health Research (#PP2-111527), the Canadian Dermatology Foundation, the VGH & UBC Hospital Foundation In It for Life Fund, and the BC Hydro Employees Community Service Fund. Funding support for the data analysis was provided in part by Verisante Technology Inc.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 19, 2011; revised March 11, 2012; accepted March 13, 2012; published OnlineFirst March 20, 2012.

References

1. English DR, Del Mar C, Burton RC. Factors influencing the number needed to excise: excision rates of pigmented lesions by general practitioners. Med J Aust 2004;180:16–9.
2. MacKenzie-Wood AR, Milton GW, de Launey JW. Melanoma: accuracy of clinical diagnosis. Aust J Dermatol 1998;39:31–3.
3. Grin CM, Kopf AW, Welkovich B, Bart RS, Levenstein MJ. Accuracy in the clinical diagnosis of malignant melanoma. Arch Dermatol 1990;126:763–6.
4. Miller M, Ackerman AB. How accurate are dermatologists in the diagnosis of melanoma? Degree of accuracy and implications. Arch Dermatol 1992;128:559–60.
5. Morton CA, Mackie RM. Clinical accuracy of the diagnosis of cutaneous malignant melanoma. Br J Dermatol 1998;138:283–7.
6. Lindelof B, Hedblad MA. Accuracy in the clinical diagnosis and pattern of malignant melanoma at a dermatological clinic. J Dermatol 1994;21:461–4.
7. Zeng H, Zhao J, Short M, McLean DI, Lam S, McWilliams A, et al. Raman spectroscopy for *in vivo* tissue analysis and diagnosis, from instrument development to clinical applications. J Innov Opt Health Sci 2008;1:95–106.
8. Short M, Lui H, McLean DI, Zeng H, Alajlan A, Chen XK. Changes in nuclei and peritumoral collagen within nodular basal cell carcinomas via confocal micro-Raman spectroscopy. J Biomed Opt 2006;11:034004.
9. Barry BW, Edwards HGM, Williams AC. Fourier Transform Raman and infrared vibrational study of human skin: assignment of spectral bands. J Raman Spectrosc 1992;23:641–5.
10. Gniadecka M, Philipsen PA, Sigurdsson S, Wessel S, Nielsen OF, Christensen DH, et al. Melanoma diagnosis by Raman spectroscopy and neural networks: structure alterations in proteins and lipids in intact cancer tissue. J Invest Dermatol 2004;122:443–9.
11. Gniadecka M, Wulf HC, Mortensen NN, Nielsen OF, Christensen DH. Diagnosis of basal cell carcinoma by Raman spectroscopy. J Raman Spectrosc 1997;28:125–9.
12. Nijssen A, Maquelin K, Caspers PJ, Bakker Schut TC, Neumann MHA, Puppels FJ. Discriminating basal cell carcinoma from perilesional skin using high wave-number Raman spectroscopy. J Biomed Opt 2007;12:034004.
13. Nijssen A, Schut TCB, Heule F, Caspers PJ, Hayes DP, Neumann MHA, et al. Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy. J Invest Dermatol 2002;119:64–9.
14. Lieber CA, Majumder SK, Billheimer D, Ellis DL, Mahadevan-Jansen A. Raman microspectroscopy for skin cancer detection *in vitro*. J Biomed Opt 2008;13:024013.
15. Lieber CA, Majumder SK, Ellis DL, Billheimer D, Mahadevan-Jansen A. *In vivo* nonmelanoma skin cancer diagnosis using Raman microspectroscopy. Laser Surg Med 2008;40:461–7.
16. Huang Z, Zeng H, Hamzavi I, McLean DI, Lui H. Rapid near-infrared Raman spectroscopy system for real-time *in vivo* skin measurements. Opt Lett 2001;26:1782–4.
17. Zhao J, Lui H, McLean DI, Zeng H. Integrated real-time Raman system for clinical *in vivo* skin analysis. Skin Res Tech 2008;14:484–92.
18. Hill T, Lewicki P. STATISTICS methods and applications. Tulsa, OK: StatSoft; 2007.
19. Haaland DM, Thomas EV. Partial least squares methods for spectral analyses. 1. relation to other quantitative calibration methods and the extraction of qualitative information. Anal Chem 1988;60:1193–202.
20. Berger AJ, Koo TW, Itzkan I, Horowitz GL, Feld MS. Multicomponent blood analysis by near-infrared Raman spectroscopy. Appl Opt 1999;38:2916–26.
21. Barker M, Rayens W. Partial least squares for discrimination. J Chemometrics 2003;17:166–73.
22. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
23. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839–43.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024

24. Metz CE, Herman BA, Shen J-H. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Stat Med 1998;17:1033–53.
25. Hajian-Tilaki KO, Hanley JA, Joseph L, Collet J-P. A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnostic tests. Med Decis Making 1997;17:94–102.
26. Zhao J, Huang Z, Zeng H, McLean DI, Lui H. Quantitative analysis of skin chemicals using rapid near-infrared Raman spectroscopy. Proc SPIE 2008;6842:684209.
27. Heal CF, Raasch BA, Buettner PG, Weedon D. Accuracy of clinical diagnosis of skin lesions. Br J Dermatol 2008;159:661–8.
28. Monheit G, Cognetta AB, Ferris L, Rabinovitz H, Gross K, Martini N, et al. The performance of MelaFind. Arch Dermatol 2011;147:188–94.
29. Cohen MH, Cohen BJ, Shotkin JD, Morrison PT. Surgical prophylaxis of malignant melanoma. Ann Surg 1991;213:308–14.
30. Westerhoff K, McCarthy WH, Menzies SW. Increase in the sensitivity for melanoma diagnosis by primary care physicians using skin surface microscopy. Br J Dermatol 2000;143:1016–20.
31. Robinson JK, Nickoloff BJ. Digital epiluminescence microscopy monitoring of high-risk patients. Arch Dermatol 2004;140:49–56.
32. Banky JP, Kelly JW, English DR, Yeatman JM, Dowling JP. Incidence of new and changed nevi and melanomas detected using baseline and dermoscopy in patients at high risk for melanoma. Arch Dermatol 2005;241:998–1006.
33. Binder M, Kittler H, Seeber A, Steiner A, Pehamberger H, Wolff K. Epiluminescence microscopy-based classification of pigmented skin lesions using computerised image analysis and an artificial neural network. Melanoma Res 1998;8:261–6.
34. Farina B, Bartoli C, Bono A, Colombo A, Lualdi M, Tragni G, et al. Multispectral imaging approach in the diagnosis of cutaneous melanoma: potentiality and limits. Phys Med Biol 2000;45:1243–54.
35. Menzies SW, Boschof L, Talbot H, Gutenev A, Avramidis M, Wong L, et al. The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma. Arch Dermatol 2005;141:1388–96.
36. Marghoob AA, Swindle LD, Moricz CZM, Negron FAS, Slue B, Halpern AC, et al. Instruments and new technologies for *in vivo* diagnosis of melanoma. J Am Acad Dermatol 2003;49:777–97.
37. Mogensen M, Jemec GBE. Diagnosis of nonmelanoma skin cancer/keratinocyte carcinoma: a review of diagnostic accuracy of nonmelanoma skin cancer diagnostic tests and technologies. Dermatol Surg 2007;33:1158–74.
38. Fuller SR, Bowen GM, Tanner B, Florell SR, Grossman D. Digital dermoscopic monitoring of atypical nevi in patients at risk for melanoma. Dermatol Surg 2007;33:1198–206.
39. Moncrieff M, Cotton S, Claridge E, Hall P. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. Br J Dermatol 2002;146:448–57.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/72/10/2491/2669143/2491.pdf by guest on 16 July 2024