Real-time Raman spectroscopy for in vivo skin cancer diagnosis

Harvey Lui, Jianhua Zhao, David McLean, Haishan Zeng

Photomedicine Institute, Department of Dermatology and Skin Science, University of British Columbia and Vancouver Coastal Health Research Institute; Imaging Unit, Integrative Oncology Department, British Columbia Cancer Agency Research Centre
Vancouver, British Columbia, Canada

Corresponding author:
Harvey Lui
Department of Dermatology and Skin Science
University of British Columbia
835 West 10th Avenue
Vancouver, B.C. Canada V5Z 4E8
T 604-875-4747
F 604-873-9919
harvey.lui@ubc.ca

Running title:
Raman spectroscopy of skin cancer

Keywords:
Skin cancer, Raman spectroscopy, melanoma, carcinoma, diagnosis, real-time, in vivo

Word count:
Abstract: 150
Text: 5003
Tables: 3
Figures 3
Supplementary Table for publication: 1
Supplementary Figures for publication: 4

Précis:
A rapid, non-invasive technology is described for accurate diagnosis of skin lesions that might be readily adapted to clinical practice.
ABSTRACT

Raman spectroscopy is a non-invasive 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 one second, including melanomas, basal cell carcinomas, squamous cell carcinomas, actinic keratoses, atypical nevi, melanocytic nevi, blue nevi and seborrheic keratoses. Lesion classification was made using principal component with general discriminant analysis and partial least squares, in three distinct discrimination tasks: skin cancers and precancers from benign skin lesions (ROC=0.879); melanomas from non-melanoma pigmented lesions (ROC=0.823); and melanomas from seborrheic keratoses (ROC=0.898). For sensitivities between 95-99%, the specificities ranged between 15-54%. Our findings establish that real-time Raman spectroscopy can be used to distinguish malignant from benign skin lesions with good diagnostic accuracy comparable to clinical examination and other optical-based methods.
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 4741 pigmented skin lesions evaluated by 468 general practitioners, the biopsy ratio, defined as the number of non-melanoma 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 demonstrated that the accuracy of clinical diagnosis of melanoma by dermatologists varies between 49% to 81%, with approximately one third of melanomas being misdiagnosed as benign lesions (2-6).

Raman spectroscopy is a non-invasive 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 locations 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 (FT) 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 utility of ex vivo Raman spectroscopy is quite limited since suspect lesions must first be biopsied which
necessarily entails an invasive procedure. In a recent *in vivo* study of non-melanoma skin cancer diagnosis using Raman microscopy, Lieber et al measured 21 lesions and their adjacent normal skin (9 BCC, 4 SCC, 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 over 1000 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.

**PATIENTS AND METHODS**

*Patients*

This study was approved by the Clinical Research Ethics Board of the University of British Columbia (Protocol C96-0499). Patients over 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 1022 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: (1) malignancies and premalignancies that require treatment: malignant melanoma (MM), squamous cell carcinoma (SCC), basal cell carcinoma (BCC), and actinic keratosis (AK); and (2) benign conditions that can visually mimic skin cancer: seborrheic keratosis (SK), atypical nevi (AN), melanocytic nevi [junctional (JN), compound (CN), and intradermal (IN)] and blue nevi (BN)]. 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 (1) clinical evaluation by one of two experienced dermatologists (HL or DM), and/or (2) 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 dataset thus consisted of 518 validated lesions from 453 subjects (224 male, 229 female), aged 18-94 (median 61).
Of these, 313 (60%) lesions underwent subsequent treatment including 44 MM, 47 SCC, 109 BCC, and 32 AK.

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

**Instrumentation**

The integrated real-time Raman system was developed in-house and is schematically shown in Figure 1 (16, 17). The hardware comprises a diode laser, a fiber and fiber bundle delivery system, a hand-held Raman probe, a spectrograph, a 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 500-1800 cm\(^{-1}\) with a resolution of 8 cm\(^{-1}\).

To take a Raman measurement the handheld 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 three times at different locations within the lesion, particularly for MM (34%) and SCC (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 did 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 wavenumber (abscissa) for any three 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 demonstrated a systematic change that was wavenumber-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 1055 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. Based on this analysis of Raman measurement reproducibility, the skin spectra were analyzed according to the full acquired spectrum (500-1800 cm\(^{-1}\)), and also by separately considering only lower (500-1055 cm\(^{-1}\)) and higher frequency (1055-1800 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” (AUCs) according to the respective spectral range being analyzed.

The lesion types selected for analysis in this study primarily included common conditions whose clinical behaviour and appearance raise suspicion for skin cancer. The diagnostic performance of in vivo Raman spectroscopy for classifying lesions was tested according to three 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 which require treatment (MM, BCC, SCC, and AK) versus benign conditions (AN, BN, CN, IN, JN, and SK). The other two tasks tested the discrimination of melanoma (all forms, MM) versus benign pigmented skin lesions (AN, BN, CN, IN, JN, and SK), and melanoma (all forms, MM) versus seborrheic keratoses (SK) which can also be confused due to 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 three dichotomous groupings of interest. All multivariate data analyses were implemented within MATLAB (version 2010a, MathWorks, Natick, MA) and STATISTICA (version 6.0, StatSoft, Tulsa, OK) 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 principal component (PC) factors and the PC loadings of the training spectra were calculated. A discrimination model was developed based on the PC factors derived from the training spectra where the
classifications were known \textit{a priori}. The PC factors of the test spectra were then calculated based on the PC loadings of the training spectra and tested against the discrimination model for classification. A posterior probability (PP) of the test lesion for skin cancer was calculated. The PPs 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 paper, we utilized a nonlinear iterative partial least squares algorithm (18). Similar to PC-GDA, all PLS analyses are based on LOO-CV.

The receiver operating characteristic (ROC) curve, i.e. sensitivity versus (1-specificity) was calculated from the PPs derived above, and represents the diagnostic performance. With good discrimination between two groups the ROC curve moves towards the left and top boundaries of the graph, whereas poor discrimination yields a curve that approaches the diagonal line function. The areas under the ROC curves (AUC) were calculated using the trapezoidal rule (22). The significance of the AUCs and comparisons between different AUCs were carried out in standard fashion (23-25). All ROC analyses were based on non-parametric techniques, and were performed separately for the PC-GDA and PLS analyses and for each of the three classification tasks.

\textbf{Skin cancer biopsy ratios according to Raman spectroscopy}

The skin cancer biopsy ratio is defined as the number of negative biopsies that are performed 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 in order 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). In order to compare Raman spectroscopy to other noninvasive diagnostic techniques as well as to 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). In order to test whether the discrimination capability of the Raman spectrometer system for skin cancer may have been influenced by body site, we did an sub-analysis to discriminate malignant melanoma (MM) from other non-melanoma pigmented lesions (AN, BN, CN, IN, JN, SK) using only the lesions located on the head and neck. Insufficient lesion numbers at other body sites did not permit site-specific sub-analyses 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 MM from non-melanoma 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 two 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 (MM) versus biopsied non-melanoma pigmented lesions (AN, BN, CN, IN, JN, SK), and the other involved biopsied skin cancers (MM, SCC, BCC) versus biopsied non-cancerous lesions (AN, BN, CN, IN, JN, SK, AK).

RESULTS

The mean Raman spectra for different skin pathologies in this study are depicted in Figure 2. All spectra were normalized to their respective areas under the curve (AUC) 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 1445 cm\(^{-1}\) with other major Raman bands centered at 855, 936, 1002, 1271, 1302, 1655 and 1745 cm\(^{-1}\). There are no distinctive Raman peaks or bands that could be uniquely assigned to specific skin cancer by visual inspection alone. Statistical methods are thus used to extract the diagnostic information 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 AK) versus benign skin lesions (non-cancer): When Raman spectroscopy is used to distinguish cancerous and precancerous skin lesions, which require treatment (n = 232), from benign skin lesions (n
that can simply be observed, the ROC AUC is 0.879 (95% CI: 0.829-0.929, PC-GDA) and statistically significant (p<0.001). The results are depicted in Figure 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 Figure 3D. At a sensitivity of 90%, the overall specificity is over 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, BR) for sensitivities of 90%, 95%, and 99%.

Although generally treated once they are detected, some AK can undergo spontaneous regression with time, and in this respect AKs are therefore not strictly considered to represent skin malignancies. Thus because the diagnosis and treatment of AK are distinct from that of MM, BCC and SCC, we did a further analysis by reclassifying AK with the benign category. In this situation, the AUC of the ROC curve for discriminating skin cancers (MM, BCC, SCC, n = 200) from non-skin cancers (AN, BN, CN, IN, JN, SK, and AK, n = 318) is 0.863 (95% CI: 0.830-0.895, p < 0.001). For a sensitivity of 90%, the overall specificity is over 63%, with a PPV of 60%, NPV of 91% and BR of 0.7:1. Raman spectroscopy appears to detect cancerous skin lesions irrespective of whether AK is considered either benign or malignant.

Melanoma (MM) versus non-melanoma pigmented skin lesions (PIG): When only lesions with pigment are considered, Raman spectroscopy can separate MM (n = 44) from non-melanoma pigmented skin lesions (AN, BN, CN, IN, JN, SK, n = 286) (Figure 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 to range from 5.6:1 to 2.3:1 for sensitivities corresponding to 99% to 90% and specificities from 15% to 68% respectively (Figure 3E, Tables 2 and 3).

*Melanoma (MM) vs seborrheic keratosis (SK):* 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, Figure 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, positive predictive values (PPV), negative predictive values (NPV) and biopsy ratios (BR).

**PLS analysis also demonstrates the discriminative capability of Raman spectroscopy for skin cancer.**

The PLS approach yields results similar to those for PC-GDA in terms of the three classification tasks above (Tables 2 and 3). Furthermore based on the algorithm by Hanley et al (23) there are no significant differences between the PLS and PC-GDA methodologies for the three 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 versus 15-30 principal components).

**Diagnostic performance is spectral band-dependent.**

We did PC-GDA and PLS analyses using the three Raman bands (500 – 1055 cm\(^{-1}\), 1055 – 1800 cm\(^{-1}\) and 500 – 1800 cm\(^{-1}\)) and found that the higher spectral range
from 1055 to 1800 cm\(^{-1}\) was optimal for differentiating melanomas from non-melanoma pigmented lesions, or melanomas from seborrheic keratoses; the full spectral range from 500 to 1800 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 sub-analysis for lesions from only the head and neck in terms of discriminating malignant melanoma (MM, n = 31) from other non-melanoma pigmented lesions (AN, BN, CN, IN, JN, SK, n = 89). The AUC of this ROC curve is 0.789 (95% CI: 0.698 – 0.879, supplemental Figure 1), which is close to the classification results that incorporated all body sites (Figure 3E, n = 330, AUC = 0.823); these two 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 PC-GDA to discriminate MM from non-melanoma 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 (MM, n = 44) versus non-melanoma pigmented lesions (AN, BN, CN, IN, JN, SK, n = 286) is only 0.577 (95% CI: 0.500 – 0.670) when using the difference spectra between corresponding lesions and normal skin (supplemental Figure 2), which is statistically no different from guessing (p = 0.09691). The comparable ROC AUC when lesional spectra alone are used is 0.823 (Figure 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 12-42% 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 based on visual examination by one of the dermatologist investigators. Two separate analyses were carried according to biopsy status, and the overall results remained essentially the same. In one study, we discriminated malignant melanoma (MM, n = 44, all of which were biopsied) from biopsied non-melanoma pigmented lesions (AN, BN, CN, IN, JN, SK, n = 81). The AUC of the ROC curve was 0.833 (95% CI: 0.761 – 0.906, supplemental Figure 3), close to the AUC of all cases regardless of biopsy status (0.823, Figure 3E), and statistically not different (p = 0.8238). In the other study, we tried to discriminate confirmed skin cancers (MM, SCC, BCC, n = 200, all of which were biopsied) from biopsy-verified non-cancerous lesions (AN, BN, CN, IN, JN, SK, AK, 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, supplemental Figure 4), 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 non-melanoma) from normal skin. In our study we aimed to evaluate a more complex and relevant clinical task, namely the discrimination of melanoma and non-melanoma skin cancers from benign skin lesions.

Raman spectra of the skin are data-rich and complex, and we have demonstrated 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 1055-1800 cm⁻¹ is preferred whereas for distinguishing skin cancers from benign lesions overall, the full spectrum from 500-1800 cm⁻¹ is preferred. The spectrometer system used in this study did not extend beyond the 500-1800 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 500-1800 cm⁻¹ range because this is regarded as the finger-print region wherein a denser cluster of Raman peaks can be found (9). It is conceivable that the three 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 (Figure 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 datasets 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 pre-
cancers was associated with a sensitivity of 63.9% for BCC, 41.1% for SCC, and 33.8%
for MM. The positive predictive values for their study were 72.7% for BCC, 49.4% for
SCC and 33.3% for MM, respectively (27). Our results demonstrated that real-time
Raman spectroscopy differentiated skin cancer and pre-cancers from benign skin lesions
with an overall area under the ROC curve of 0.879 (95% CI 0.829-0.929), sensitivity of
90%, and positive predictive value 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 1250 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 multi-spectral imaging (28, 30-38). Westerhoff et al found that the accuracy of melanoma diagnosis could be improved from 63% to 76% with the aid of surface microscopy (30). Robinson et al reported a biopsy ratio of 47:1 to rule out melanoma with digital epiluminescence microscopy (31). Monheit et al found the estimated biopsy ratio to range from 7.6:1 to 10.8:1 for identifying melanoma with or without borderline lesions using a multi-spectral-based method (i.e. MelaFind) (28). Banky et al 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 et al investigated 120 lesions (including 39 MM) 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 et al studied 237 pigmented lesions using multispectral imaging (67 MM and 170 non-MM), and found the AUC of the ROC curve is to be 0.779, with a sensitivity of 80% and specificity of 51% (34). Moncrieff et al studied 348 pigmented lesions (52 MM) 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 2430 lesions (382 MM), 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 favourable in comparison to 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, i.e. 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 (Table 2 and 3). For all three 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 50-100% (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 one 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.

CONCLUSIONS

We have studied skin cancers and a range of benign skin diseases using real-time \textit{in vivo} Raman spectroscopy with an integration times of less than 1 second per lesion. Multivariate PC-GDA and PLS analyses demonstrate that Raman spectroscopy can distinguish (1) malignant and premalignant lesions from benign disorders, (2) melanomas from benign pigmented skin lesions, and (3) melanomas from seborrheic keratoses.
Disclosure of Potential Conflicts of Interest

H. Lui, H. Zeng, D. I. McLean, J. Zhao, and the BC Cancer Agency hold patents for Raman spectroscopy that are licensed to Verisante Technology Inc.

Authors’ Contributions

H. Lui, H. Zeng and D.I. McLean conceived the study, supervised the work and directed the data analysis. H. Lui and D.I. McLean recruited patients and confirmed the lesion diagnoses. H. Zeng and J. Zhao implemented the real-time Raman spectroscopy system and software development. All authors were involved in developing the statistical models and analysis. H. Lui, H. Zeng and J. Zhao wrote the manuscript, which was revised by all authors.

Acknowledgments

The authors gratefully acknowledge the contributions of our collaborators and research team members: Zhiwei Huang, Iltefat Hamzavi, Abdulmajeed Alajlan, Hana Alkhayat, Ahmad Al Robaei, Sunil Kalia, Bernardita Ortiz-Policario, 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.
REFERENCES


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

<table>
<thead>
<tr>
<th>Final lesion diagnosis</th>
<th>Subjects</th>
<th>Number of lesions</th>
<th>Number biopsied (%)</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean age (range)</td>
<td>Male</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM</td>
<td>69 (51-88)</td>
<td>12</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>LMM</td>
<td>67 (42-85)</td>
<td>7</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>SS</td>
<td>60 (22-77)</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>MM other</td>
<td>61 (60-62)</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>BCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superficial</td>
<td>63 (34-86)</td>
<td>10</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td>Nodular</td>
<td>66 (39-94)</td>
<td>34</td>
<td>29</td>
<td>73</td>
</tr>
<tr>
<td>Pigmented</td>
<td>67 (46-83)</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Other BCC</td>
<td>68 (60-75)</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>SCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In situ</td>
<td>70 (56-88)</td>
<td>12</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>Invasive</td>
<td>66 (39-94)</td>
<td>16</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>Other SCC</td>
<td>78</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Actinic keratosis</td>
<td>66 (43-92)</td>
<td>13</td>
<td>14</td>
<td>32</td>
</tr>
<tr>
<td>Atypical nevus</td>
<td>48 (20-75)</td>
<td>22</td>
<td>26</td>
<td>57</td>
</tr>
<tr>
<td>Junctional nevus</td>
<td>43 (18-70)</td>
<td>12</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>Compound nevus</td>
<td>35 (18-67)</td>
<td>13</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Intradermal nevus</td>
<td>50 (28-83)</td>
<td>9</td>
<td>26</td>
<td>38</td>
</tr>
<tr>
<td>Blue nevus</td>
<td>37 (18-66)</td>
<td>4</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Seborrhoeic keratosis</td>
<td>64 (25-89)</td>
<td>49</td>
<td>42</td>
<td>114</td>
</tr>
</tbody>
</table>

MM: malignant melanoma, MM Subtypes (LM: lentigo maligna; LMM: lentigo maligna melanoma; SS: superficial spreading melanoma); BCC: basal cell carcinoma; SCC: squamous cell carcinoma.
Table 2. Diagnostic performance of Raman spectroscopy according to optimized wavebands.

<table>
<thead>
<tr>
<th>Diagnosis classification task</th>
<th>Raman waveband (cm⁻¹)</th>
<th>ROC a Area Under Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Principal component analysis (GDA b) (95% CI)</td>
</tr>
<tr>
<td>Skin cancers + Actinic keratosis vs Benign lesions</td>
<td>500-1800</td>
<td>0.879 (0.829-0.929)</td>
</tr>
<tr>
<td>Melanoma vs Benign pigmented lesions</td>
<td>1055-1800</td>
<td>0.823 (0.731-0.915)</td>
</tr>
<tr>
<td>Melanoma vs Seborrheic keratosis</td>
<td>1055-1800</td>
<td>0.898 (0.797-0.999)</td>
</tr>
</tbody>
</table>

a ROC: receiver operating characteristic  b GDA: general discriminant analysis
Table 3. Summary of Raman spectroscopy diagnostic parameters derived from ROCs according to various levels of sensitivity.

<table>
<thead>
<tr>
<th>Diagnosis classification task</th>
<th>Sensitivity level (95% CI)</th>
<th>Principal component analysis (GDA&lt;sup&gt;a&lt;/sup&gt;)</th>
<th>Partial least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Specificity (95% CI) PPV&lt;sup&gt;b&lt;/sup&gt; NPV&lt;sup&gt;c&lt;/sup&gt; Biopsy ratio</td>
<td>Specificity (95% CI) PPV NPV Biopsy ratio</td>
</tr>
<tr>
<td>Skin cancers + Actinic keratosis vs Benign lesions</td>
<td>0.99 (0.98-1.00)</td>
<td>0.17 (0.13-0.21) 0.49 0.95 1.03:1</td>
<td>0.24 (0.19-0.29) 0.51 0.97 0.95:1</td>
</tr>
<tr>
<td></td>
<td>0.95 (0.92-0.99)</td>
<td>0.41 (0.35-0.48) 0.57 0.91 0.77:1</td>
<td>0.52 (0.48-0.58) 0.62 0.93 0.62:1</td>
</tr>
<tr>
<td></td>
<td>0.90 (0.86-0.94)</td>
<td>0.64 (0.58-0.70) 0.67 0.89 0.49:1</td>
<td>0.66 (0.61-0.71) 0.68 0.89 0.47:1</td>
</tr>
<tr>
<td>Melanoma vs Benign pigmented lesions</td>
<td>0.99 (0.96-1.00)</td>
<td>0.15 (0.11-0.19) 0.15 0.99 5.58:1</td>
<td>0.14 (0.10-0.18) 0.15 0.99 5.65:1</td>
</tr>
<tr>
<td></td>
<td>0.95 (0.89-1.00)</td>
<td>0.38 (0.32-0.44) 0.19 0.98 4.24:1</td>
<td>0.44 (0.38-0.50) 0.21 0.98 3.83:1</td>
</tr>
<tr>
<td></td>
<td>0.90 (0.81-0.99)</td>
<td>0.68 (0.63-0.73) 0.30 0.98 2.31:1</td>
<td>0.63 (0.57-0.69) 0.27 0.98 2.67:1</td>
</tr>
<tr>
<td>Melanoma vs Seborrheic keratosis</td>
<td>0.99 (0.96-1.00)</td>
<td>0.25 (0.17-0.33) 0.34 0.98 1.96:1</td>
<td>0.46 (0.37-0.55) 0.41 0.99 1.41:1</td>
</tr>
<tr>
<td></td>
<td>0.95 (0.89-1.00)</td>
<td>0.54 (0.45-0.63) 0.44 0.97 1.25:1</td>
<td>0.52 (0.43-0.61) 0.43 0.96 1.31:1</td>
</tr>
<tr>
<td></td>
<td>0.90 (0.81-0.99)</td>
<td>0.68 (0.59-0.77) 0.52 0.95 0.92:1</td>
<td>0.66 (0.57-0.75) 0.51 0.94 0.98:1</td>
</tr>
</tbody>
</table>

<sup>a</sup> GDA: general discriminant analysis.

<sup>b</sup> PPV: positive predictive value, the ratio of true positives to the total of true positives and false positives.

<sup>c</sup> NPV: negative predictive value, the ratio of true negative to the total of true negatives and false negatives.
TABLES

Table 1. Summary of patients and lesions evaluated by Raman spectroscopy. MM: malignant melanoma, MM subtypes (LM: lentigo maligna; LMM: lentigo maligna melanoma; SS: superficial spreading melanoma); BCC: basal cell carcinoma; SCC: squamous cell carcinoma.

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

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

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

Figure 2. Mean Raman spectra by diagnosis. All lesion spectra were normalized to their areas-under-the-curve before averaging by diagnosis. Spectral variations were analyzed by PC-GDA and PLS methods for classification purposes. (A) Full range spectrum results displayed from 500-1800 cm\(^{-1}\). (B) Spectral results plotted for lower frequency range only (500-1055 cm\(^{-1}\)) demonstrate variability according to lesion diagnosis. Note that the y-axis scales for Figures 2(A) and 2(B) are different. MM: malignant melanoma, BCC: basal cell carcinoma, SCC: squamous cell carcinoma, AK: actinic keratosis, AN: atypical nevus, BN: blue nevus, CN: compound nevus, IN: intradermal nevus, JN: junctional nevus, SK: seborrheic keratosis.

Figure 3. Lesion classification by Raman spectroscopy based on PC-GDA analysis. Posterior probabilities for discriminating (A) skin cancers and pre-cancers (Cancer + AK, including MM, BCC, SCC, AK, n = 232) from benign skin disorders (including AN, BN, CN, IN, JN, SK, n = 286), (B) melanoma (n = 44) from benign pigmented skin diseases (including AN, BN, CN, IN, JN, SK, n = 286), and (C) melanoma (n = 44) from seborrheic keratosis (n = 114). (D, E, F) The corresponding ROC curves and 95% confidence intervals (CIs) are derived from the respective posterior probabilities, and all areas under the curve (AUCs) are significant (p <0.0001).
In vivo Raman Spectrometer System

- Diode Laser (785 nm)
- Computer
- Single Fiber
- Fiber Bundle
- Raman Probe
- Skin
- CCD Camera
- Calibration Fiber
- VPT™ Grating Spectrograph
Figure 3

A: Cancer + AK versus Benign lesions

B: Melanoma versus Benign pigmented lesions

C: Melanoma versus Seborrheic keratosis

D: ROC for Cancer + AK versus Benign lesions

E: ROC for Melanoma versus Benign pigmented lesions

F: ROC for Melanoma versus Seborrheic keratosis

AUC = 0.879 ± 0.05 [95% CI]  
\( p < 0.0001 \)

AUC = 0.823 ± 0.092 [95% CI]  
\( p < 0.0001 \)

AUC = 0.898 ± 0.101 [95% CI]  
\( p < 0.0001 \)
Real-time Raman spectroscopy for in vivo skin cancer diagnosis

Harvey Lui, Jianhua Zhao, David I McLean, et al.

Cancer Res  Published OnlineFirst March 20, 2012.

Updated version  Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-11-4061

Supplementary Material  Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2012/03/20/0008-5472.CAN-11-4061.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.