Development of Raman microspectroscopy for automated detection and imaging of basal cell carcinoma

Marta Larraona-Puy
Adrian Ghita
Alina Zoladek
University of Nottingham
School of Physics and Astronomy
University Park
Nottingham, NG7 2RD
United Kingdom

William Perkins
Sandeep Varma
Nottingham University Hospital NHS Trust
Dermatology Department
QMC Campus
Derby Road
Nottingham, NG7 2UH
United Kingdom

Iain H. Leach
Nottingham University Hospital NHS Trust
Histopathology Department
QMC Campus
Derby Road
Nottingham, NG7 2UH
United Kingdom

Alexey A. Koloydenko
University of London
Mathematics Department
Royal Holloway
Egham, TW20 0EX
United Kingdom

Hywel Williams
University Hospital NHS Trust
Centre of Evidence-Based Dermatology
C Floor South Block, Nottingham
QMC Campus
Derby Road
Nottingham, NG7 2UH
United Kingdom

Ioan Notingher
University of Nottingham
School of Physics and Astronomy
University Park
Nottingham, NG7 2RD
United Kingdom

Abstract. We investigate the potential of Raman microspectroscopy (RMS) for automated evaluation of excised skin tissue during Mohs micrographic surgery (MMS). The main aim is to develop an automated method for imaging and diagnosis of basal cell carcinoma (BCC) regions. Selected Raman bands responsible for the largest spectral differences between BCC and normal skin regions and linear discriminant analysis (LDA) are used to build a multivariate supervised classification model. The model is based on 329 Raman spectra measured on skin tissue obtained from 20 patients. BCC is discriminated from healthy tissue with 90±9% sensitivity and 85±9% specificity in a 70% to 30% split cross-validation algorithm. This multivariate model is then applied on tissue sections from new patients to image tumor regions. The RMS images show excellent correlation with the gold standard of histopathology sections, BCC being detected in all positive sections. We demonstrate the potential of RMS as an automated objective method for tumor evaluation during MMS. The replacement of current histopathology during MMS by a "generalization" of the proposed technique may improve the feasibility and efficacy of MMS, leading to a wider use according to clinical need. [DOI: 10.1117/1.3251053]

Keywords: Raman imaging; basal cell carcinoma; Raman microspectroscopy; Mohs micrographic surgery; skin cancer.

Paper 09225R received Jun. 5, 2009; revised manuscript received Aug. 3, 2009; accepted for publication Aug. 11, 2009; published online Oct. 27, 2009.

1 Introduction

Skin cancer is a growing source of concern, not only for being the most common of all types of cancers, but also due to its increasing incidence rate. Each year, there are more new cases of skin cancer than the combined incidence of cancers of breast, prostate, lung, and colon.[1] In the United Kingdom and the United States more than 100,000 and 1,000,000 cases, respectively, are diagnosed annually.[1,2]

About 80% of skin cancer cases worldwide are basal cell carcinomas (BCCs). This type of skin cancer belongs to the keratinocyte or nonmelanoma family, and arises from the epidermis. It commonly occurs in areas exposed to the sun, such as the head or the neck.[3]

Mohs micrographic surgery (MMS) is the most suitable treatment for large, rare, or recurrent BCCs, those growing

Please address all correspondence to: Ioan Notingher, University of Nottingham, School of Physics and Astronomy, University Park, Nottingham, NG7 2RD United Kingdom Tel: 44-0-11595-15172; E-mail: ioan.notingher@nottingham.ac.uk

1083-3668/2009/14(5)/054031/10/$25.00 © 2009 SPIE

into the surrounding skin tissue and in critical areas, e.g., the high-risk zone of the face: nasolabial folds, eyelids, and periauricular areas.[4] MMS was first developed by Frederick Mohs and maximizes the evaluation of the surgical margin by pathologic observation of the histologic slides during surgery. Sequential layers of tissue are removed until the lesion is clear of BCC. If the pathologic evaluation indicates tumor persistence, accurate location is recorded and further tissue removal is performed by the surgeon. This procedure ensures high cure rates and enables maximal conservation of healthy tissue, which can be particularly important on areas such as the face.[5]

It is widely accepted that MMS is the most effective current method for removal of aggressive BCC in terms of compromise between maximum conservation of healthy skin and minimum recurrence rates.[6–8] While 5-yr recurrence rates for BCC treated by MMS are 1.4% for primary tumors and 4% for recurrent tumors,[9] for standard excision, this rate reaches 3.2 to 10% for primary tumors, and more than 17% for recurrent BCCs.[10] It has also been reported that at shorter follow-up periods (18 months), MMS was considerably more advantageous than surgical excision especially for recurrent tumors[7] (recurrence rate 0 versus 3%).[7]

However, in many cases, traditional methods such as surgical excision, cryosurgery, curettage, and electrodessication rather than MMS are applied for high-risk BCC removal, despite its lower effectiveness, based solely on availability and cost considerations.[11–13] Focusing on the United Kingdom as an example of the current worldwide situation, it has been reported that there are fewer MMS centers and specialist surgeons than the number recommended by the medical community according to clinical needs.[14] The main reason for the inequity of service provision is the need of time-consuming and costly procedures to obtain and evaluate tissue sections during MMS, as well as specialized staff, including trained technicians for frozen section preparation. In addition, several studies have reported that even the gold standard of histopathology has interobserver differences.[15] In a study[16] on 48 samples evaluated by 20 pathologists, overall sensitivity was 87% (range, 55 to 100%) and specificity 94% (range, 83 to 100%). Another study of 592 histopathology slides using two pathologists, found interobserver agreement in only 93% of cases.[17] However, the real values for effectiveness in BCC diagnosis during MMS in terms of sensitivity and specificity may be lower because most MMS surgeons are not trained histopathologists. Therefore, an automated, reliable, low-cost method for BCC detection and imaging in MMS excised skin sections, which can be used in a surgery theater environment as an alternative to current histopathology tissue evaluation, would enable a wider use of MMS according to clinical needs. This would be a significant advance in the management of BCCs.

Many optical techniques have been proposed for the detection and imaging of BCC. Methods using fluorescence spectroscopy showed differences between BCC and healthy skin.[18] Bispectral fluorescence imaging, combining skin autofluorescence with d-aminolaevulinic- acid-induced fluorescence was proposed to improve imaging of tumor margins.[19] However, comparisons between this technique and histopathological mapping showed that good correlations were observed in only less than 50% of patients.[19] The low sensitivity for BCC detection is an intrinsic limitation of fluorescence imaging caused by the broadband nature of the fluorescence spectra of biomolecules. Considering the complex composition of the tissue, the fluorescent bands of the biomolecules overlap to produce featureless fluorescent spectra.[20] These wide fluorescence bands hide the subtle spectral details required for reliable discrimination between healthy skin and BCC. In addition, images produced by fluorescence spectroscopy are difficult to quantify as discrimination relies mainly on emission intensity differences. Therefore, fluorescence images were only qualitative and did not provide quantitative information as required for automated imaging.

IR microspectroscopy studies showed that the discrimination accuracy between normal epidermis, BCC, squamous cell carcinoma, and melanocytic lesions was[21] 93%. The distinction between normal epidermis and BCC was based on the higher content of nucleic acids in the BCC regions. Despite the high chemical specificity and discrimination accuracy, this technique imposes severe limitations regarding sampling depth. Since IR spectroscopy is based on the absorption of IR radiation by tissue, this technique requires tissue sections thinner than 10 μm to avoid complete absorption of the incident radiation.[21,22] Therefore, IR techniques still require preparation of thin frozen sections, which represents an intrinsic limitation toward the development of an automated quantitative imaging during MMS, particularly on tissue blocks.

The technique proposed in this paper for creating images of skin tissue excised during MMS relies on Raman microspectroscopy (RMS). In RMS, the Raman signal of different micrometric regions within a sample is collected to produce an image based on the biochemical composition of the sample. Raman spectra are "chemical fingerprints" of the molecular constituents of the sample and are based on the Raman effect.

For the last 2 decades, RMS has been recognized as a powerful optical technique for biomedical applications.[23] Compared to fluorescence spectroscopy, vibrational spectra of biomolecules are characterized by molecule-specific narrow peaks, which are sensitive to molecular structure, conformation, and interactions. This high chemical specifity constitutes a major advantage of RMS, making it able to detect slight chemical changes in biological samples.[24] RMS achieves diffraction-limited lateral resolution in the micrometer range, which makes it an appropriate tool to imaging cells[25] and tissues.[26]

RMS is a suitable technique for cancer diagnosis because of its high sensitivity to molecular and structural changes associated with cancer, such as an increased nucleus-to-cytoplasm ratio, disordered chromatin, higher metabolic activity, and changes in lipid and protein levels.[27] The potential of RMS for detection and diagnosis of human cancers, both *in vivo* and *in vitro*, has been demonstrated for a large number of cancer types, including skin,[28–34] breast,[35] esophagus,[36] lung,[37] cervix,[38] and prostate.[39]

Early studies on skin using RMS presented Raman spectroscopy as a useful tool in dermatological diagnosis, comparing Raman spectra from normal, healthy human stratum corneum with other skin tissues, such as callus tissue or hyperkeratotic psoriatic plaques.[37] The capability of Raman spectroscopy to detect biochemical alterations in skin tissue caused by BCC was first demonstrated by Gniadecka et al.[28]

Several protein and lipid alterations characteristics of BCC tissue, such as the alterations of the amide bands, attributed to the conformational changes of proteins (essentially, changes in collagen), were reported.[28,33] Further experiments showed 97% sensitivity and 98% specificity on BCC detection were realized using principal component analysis (PCA) for dimension reduction along with a neural network classifier for spectral clustering.[34] A more recent work also demonstrated the ability of RMS combined with fiber optics for skin tumor *in vivo* diagnosis.[32]

However, apart from tumor detection, MMS requires high-spatial-resolution imaging of BCC regions in tissue blocks and sections. Quantitative Raman spectroscopic images can be built by representing the intensity of a certain spectral peak, score, or weight obtained with a multivariate spectral analysis method for each individual location in the 2-D region where Raman spectra were acquired. Raman spectroscopic measurements do not require sample preparation, e.g., dying the tissue; they are free of variations due to changes in the molecular composition and structure of the sample, which may be caused by preparation protocols. Consequently, RMS is an objective and quantitative method that can be used continuously with the same level of accuracy, making it ideally suitable for automatic implementation and biochemical imaging. Since RMS does not rely on light absorption, tissue, thickness, and water tissue these have little effect on the measurements. Therefore, this technique could be used on both tissue sections and excised tissue blocks.

Many studies have applied this technique to tumor discrimination, to create spectral maps of tissue sections containing cancerous cells. Images of brain tumors,[40] gastrointestinal (GI) tract,[41] lymph,[42] lung,[26] esophagus,[43] and skin[30] cancer have been created. These studies employed unsupervised methods for imaging, such as the intensities of the scores of a selected number of principal components or *k*-means clustering. However, these unsupervised methods for creating Raman images have an important disadvantage when applied to detection and imaging of tumors: for building a specific image of a tissue sample only information present in that particular tissue is used. Therefore, images obtained by these methods do not provide an automated objective diagnosis, but require additional expert information. Supervised methods based on measurements of a large number of tissue specimens have the potential to overcome this limitation. The prior information gathered from a large number of samples and patients can be used to produce pseudocolor images based on biochemical contrast as well as providing an objective diagnosis on new tissue sections obtained during surgery.

In this study, a supervised classification method was developed to investigate the ability of RMS to detect and image BCC in skin tissue excised during MMS and skin surgery. A spectral database using 329 tissue regions from 20 randomly chosen patients was developed. The spectra were divided into three classes—BCC, dermis, or epidermis—according to histopathology diagnosis. Once the classification accuracy was established, the model was applied on tissue specimens obtained from new patients for imaging tumor regions.

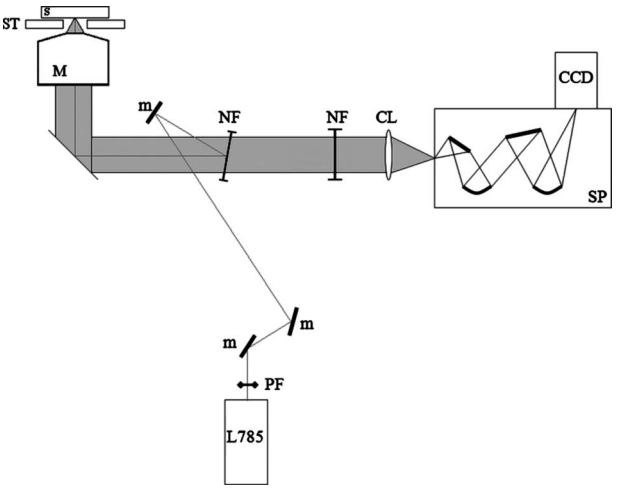

Fig. 1 Schematic diagram of the Raman micro-spectrometer: L785, laser; PF, plasmaline filter; m, mirror; M, microscope; ST, stage; s, tissue sample; NF, notch filter; CL, collecting lens; SP, Czerny-Turner spectrograph; CCD, detector.

2 Methods and Materials

2.1 *Skin Tissue Samples*

Skin tissue sections were obtained from the Nottingham University Hospitals National Health Service (NHS) Trust. Consent was obtained from the patients and ethical approval was granted from Nottingham Research Ethics Committee. Tissue sections were cut from blocks removed during MMS and standard BCC excision into 20-μm sections for RMS investigations. After the RMS measurements, the analysed sections were stained using conventional hematoxylin and eosin (H&E) staining. Diagnosis was given by a consultant histopathologist.

2.2 *Raman Spectroscopy*

A Raman microspectrometer was built based on an inverted microscope (IX71, Olympus, Essex, United Kingdom) equipped with an automated *XYZ* translation stage (H117, Prior Ltd., Cambridge, United Kingdom), deep-depletion back-illuminated charge coupled device (CCD) detector (DU401A-BR-DD, Andor Ltd., Belfast, United Kingdom) and spectrograph (SR-303i, Andor Ltd., Belfast, United Kingdom) and a 785-nm continuous wave GaAs diode laser (XTRA, Toptica Photonics, Munich, Germany). The laser power was set to 50 mW at the sample to avoid sample damage and ensure repeatable measurements. The objective lens of the microscope has a numerical aperture of 0.75 and a magnification of $\times$50. A microscopy camera (2-1C Infinity, Lumenera, Ottawa, Canada) was used to record images of the tissue sections. Figure 1 presents a schematic description of the instrument. The wave-number (vibrational frequency) axis was calibrated using Raman standard samples (ASTM E 1840), such as naphthalene and 1,4 bis (2-methylstyryl benzene) (Sigma, United Kingdom). The wave-number accuracy was found to be ± 0.5 cm^{-1}.

2.3 *Building the RMS Database for BCC Discrimination*

2.3.1 *Data acquisition*

First, an adjacent H&E-stained skin section was placed on the microscope and the regions of interest (BCC, epidermis, or dermis) were identified. The corresponding unstained skin section, which had been deposited on a MgF_2 window, was placed on the microscope and the position coordinates of the measured regions were recorded. After RMS measurements, the tissue section was returned to the pathology laboratory, H&E stained, and then placed on the Raman microscope for retrospective acquisition of images to be used for diagnosis by a consultant histopathologist. Once identified, the measured regions were classified into three groups: BCC, epidermis, and dermis. The precision of retrospective location was determined to be less than 5 μm based on two marks engraved on each slide.

To account for tissue heterogeneity, each spectrum used for the multivariate analysis represented the average of 100 spectra measured at 5-μm intervals over a 50- $\times$50-μm region. The integration time for each position was 1 s. A total of 329 mean Raman spectra were measured from tissue specimens from 20 patients: 127 BCCs (nodular and morphoeic), 92 epidermis, and 110 dermis.

2.3.2 *Data analysis*

Prior to analysis, the contribution of the microscope objective was subtracted. All spectra were baseline corrected using a sixth-order polynomial and normalized to zero mean and unity standard deviation. Finally, data were smoothed using the Savitsky-Golay algorithm (five points, second-order polynomial).

Data analysis was performed using linear discriminant analysis (LDA). LDA is a statistical technique for data classification that tries to provide the maximum class separability and draw a decision region between given classes.[44] This technique has been widely used in oncological research to classify spectral data from healthy and malignant tissue.[22,30,43,45] In our study, the LDA model was built using the area of several selected Raman bands. The selection criterion was to maximize the differences among the spectra representative of each class. The peak areas were calculated after a local linear baseline was subtracted in the regions of interest. This approach enables direct control over the selection of spectral features that could be assigned to specific tissue biochemicals and avoided measurement artefacts, such as baseline variations or spectral shifts. Thus, we selected a small number of features by directly optimizing their discriminatory power. Other automated methods for dimensionality reduction have been reported,[30,43] such as PCA. However, the use of selected Raman bands in our case produced more reliable results and was therefore preferred. Moreover, automated methods may not be optimal as a feature selection approach when the ultimate goal is classification because class-related information would not be directly taken into account by such methods.[46]

Our LDA-based classifier enables automated diagnosis of a measured tissue region as BCC, epidermis, or dermis. We implemented our three-class classifier as a composition of two two-class classifiers. The classification process can be flexibly tuned to a desirable regime by controlling either specificity or sensitivity while optimizing the other measure. To emulate a realistic scenario of BCC detection, we currently target highly sensitive regimes (90% or higher). As our current data size is modest, we used cross-validation (CV) to assess the performance of our method.[47] A k-fold CV splits the data set into k equally sized parts: $k-1$ parts are used to produce a model, and the last part is used to test it.[47] The different CV methods can also be labeled with the percentages corresponding to the splitting, e.g., the fivefold CV can also be refered to as the 80% to 20% CV. The case in which k is equal to the total number of elements in the data set is called leave-one-out CV (LOOCV). In this method, all spectra except one are used to build a model and then to classify the left out spectrum. This method is repeated so that each spectrum is predicted once.[47] As CV estimation might depend on the splitting ratio, we obtained our CV estimates with three different settings, namely, 70% to 30%, and the more commonly used fivefold CV and LOOCV schemes.

To express the diagnosis accuracy of our model, the classification results of the CV procedures were expressed in terms of its sensitivity and specifity. While the sensitivity is related to how precise the technique is in detecting true BCC, the specifity shows how accurately it identifies healthy tissue. The errors in sensitivity and specificity of a k-fold CV were obtained by repeating the algorithm for a large number of k-fold partitions. Note that in medical practice, it may be beneficial to provide an end-user with the option of adjusting sensitivity (or specificity) and, in effect, operating in several regimes. In this scenario, it is imperative for the regime to be controlled with high degree of confidence. The statistical models employed in our classifiers enable us to achieve this (cf. Sec. 3.2).

2.4 *Spectral Imaging of BCC*

After the LDA model was built, the ability of RMS to detect and image BCC was tested on a set of six skin sections obtained from three new patients (no samples from these patients were included into the LDA classification model). Raman spectra from a selected region were acquired at 5-μm intervals with 2-s integration time at each position (total acquisition time was 3 to 5 h, depending on the size of the sample). Each spectrum was smoothed using the Savitsky-Golay algorithm (five points, second-order polynomial). Raman spectra of the MgF_2 substrate were detected using a threshold filter in the 1370 to 1500-cm^{-1} spectral range and eliminated from the classification model.

Spectra were binned over 10 or 15 μm to account for tissue heterogeneity, i.e., each new spectra was the average of the four or nine adjacent spectra to ensure that the acquired spectral data were representative of the tissue class. Thus, the spatial resolution of the biochemical images achieved was 10 or 15 μm, respectively. The LDA model was then applied to predict the class of each spectrum as BCC, epidermis, or dermis. An image was constructed based on the LDA model classification. An alternative method combining unsupervised k-means clustering followed by LDA discrimination was also proposed (k-means-LDA). The aim of this method was to observe any loss in spectral images caused by the data compres-

sion methods required when large data sets are used. The *k*-means clustering is a classification technique that groups the Raman spectra of the data set into a fixed number of clusters *k* by minimizing the intragroup and maximizing intergroup differences. First, the initial data set is separated into an arbitrary number of *k* clusters. Second, the vector of centroids of every cluster is calculated. The centroid is defined as the average expression value for all the samples in that cluster. Then, the distance between every sample and every centroid is calculated, each sample being reassigned to the closest cluster. The process is repeated iteratively until it converges.[48] To produce qualitative false-color images, the centroid Raman spectra corresponding to each group were obtained by *k*-means clustering. The LDA model was then applied on the centroid spectra to classify these image groups into BCC, epidermis, or dermis, thus providing an automated objective diagnosis. Note that *k*-means clustering is not part of the classifier training phase, but it is used only to apply the final classifier to the new data from the unseen tissue sections.

3 Results and Discussion

3.1 *Spectral Database*

Mean Raman spectra of the 329 measured tissue specimens (127 BCCs, 92 epidermis, and 110 dermis) from 20 patients used to construct the multivariate classification model are presented in Fig. 2(a). This figure shows that there exist spectral differences between BCC and epidermis or dermis, in agreement with previous reported works on skin cancer.[28,30,39] The differences between dermis and BCC appear to be mainly due to the presence of collagen I in dermis and not in BCC, as inferred from the computed spectrum difference shown in Fig. 2(c). In addition, the main differences between BCC and epidermis can be explained by the higher amount of DNA in the tumor tissue, as shown in Fig. 2(b). The higher amount of DNA is caused by the smaller amount of cytoplasm and higher density of cells present in the tumor, as can be seen in Fig. 3, where a typical H&E image of measured tissue regions of 50×50 μm diagnosed as BCC, epidermis, and dermis is presented. The H&E technique stains cell nuclei in purple/black (hematoxylin), depending on section thickness and the formulation of hematoxylin used, and most components of the cell cytoplasm in pink/red (eosin).[49] However, the images included in this study showed DNA in dark brown and dermis in pale orange as an effect of the CCD camera calibration. Regions with higher DNA will be darker, as it is the case of cancerous areas. On the contrary, regions with higher collagen will present a paler colour, as it is the case of dermis.

3.2 *LDA Classification Model*

Based on the comparison among the mean Raman spectra of the three classes shown in Fig. 2, six Raman bands were chosen as "fingerprints" to discriminate between healthy skin and tumor regions. The selection criterion was to maximize differences among classes, thus the following band area ratios were chosen:

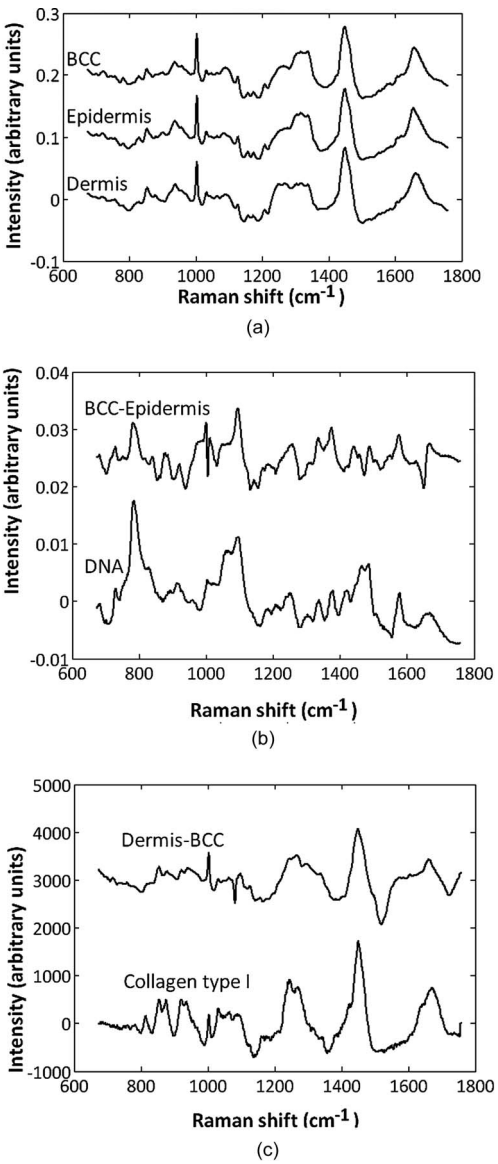

Fig. 2 (a) Mean Raman spectra of 329 tissue specimens—127 BCCs, 92 epidermis, and 110 dermis—from 20 patients used to construct the model. Spectra have been shifted vertically for clarity. Comparison between the Raman spectra of computed differences drmis minus BCC and BCC minus epidermis are shown in (b) and (c) along with Raman spectra of purified Collagen type I and DNA (Sigma, United Kingdom).

$$r_1=\frac{I_{788\ \mathrm{cm}^{-1}}}{I_{1003\ \mathrm{cm}^{-1}}},\quad r_2=\frac{I_{850\ \mathrm{cm}^{-1}}}{I_{1003\ \mathrm{cm}^{-1}}},\quad r_3=\frac{I_{950\ \mathrm{cm}^{-1}}}{I_{1003\ \mathrm{cm}^{-1}}},$$

$$r_4=\frac{I_{1093\ \mathrm{cm}^{-1}}}{I_{1003\ \mathrm{cm}^{-1}}},\quad r_5=\frac{I_{1312\ \mathrm{cm}^{-1}}}{I_{1268\ \mathrm{cm}^{-1}}}.$$

These Raman bands can be assigned to specific vibrations in DNA and collagen type I. The 788- and 1093-cm^{-1} bands correspond to the O-P-O phosphodiester and PO_2 vibrations in DNA, respectively.[50] Regions of healthy dermis are characterized by high Raman bands at 850 and 950 cm^{-1}, which are associated to proline and C-C backbone vibrations in proteins.[51] Significant spectral differences are also observed in

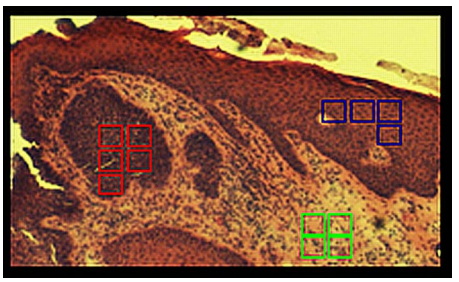

Fig. 3 H&E image of a typical skin tissue sections showing measured regions of 50×50 μm, being represented as empty squares. Color code: blue for epidermis, red for BCC, green for dermis. (Color online only.)

the amide III spectral region (1200 to 1350 cm^{-1}), which is sensitive to the secondary structure of proteins.[50] Most significant differences in this region appeared mainly at 1268 and 1312 cm^{-1}, which have been identified to correspond to molecular vibrations of collagen (Ref 52 and Fig. 2). Intensity of the 1003-cm^{-1} band corresponding to the ring breathing of phenylalanine was chosen as the denominator of the ratio because it showed insignificant differences only between classes.

For building an automated detection and imaging method, a multivariate model based on LDA was developed, in which the ratios of the peak intensities, i.e., r_1, r_2, r_3, r_4, and r_5, were used as input parameters. Two consecutive LDA routines were used.[30] The order of the LDAs takes into account the easiest discrimination among dermis and the other two classes elucidated in Fig. 2(a). The model showed that RMS is able to discriminate nodular and morphoeic BCC from healthy tissue with $90 \pm 9\%$ sensitivity and $85 \pm 9\%$ specificity in a 70% to 30% split CV algorithm. The reported estimates are the means and standard deviations of sensitivity and specificity over randomly chosen partitions. The values obtained for fivefold CV were $89 \pm 11\%$ sensitivity in BCC discrimination and $84 \pm 10\%$ specificity, showing the stability of our estimates. LOOCV also reinforces this statement, achieving 94 and 84% for sensitivity and specificity, respectively.

A typical result for CV is shown in Fig. 4, where all 329 data from the spectral database are classified into three groups, circles for BCC, diamonds for epidermis, and triangles for dermis. Data used for training the model are represented as empty symbols in Fig. 4, while the symbol for data employed for validation of the model in Fig. 4 are color-filled. The employed algorithm consists of two consecutive LDAs. First, dermis is separated from the other two classes, BCC +epidermis, which are considered one only group, by LDA (LDA1). Then, BCC is separated from epidermis by a new LDA (LDA2). The dashed line (boundary 1) and the solid line (boundary 2) represent the 95% target sensitivity discrimination lines of LDA1 and LDA2, respectively. The LDA score plot in Fig. 4 shows that there is a significant clustering of the spectra into three groups corresponding to BCC, epidermis, and dermis. Misclassifications occur mostly between BCC and epidermis, where clusters overlap over the 95% sensitivity boundary. Such overlap is expected considering the great similarities between the mean spectra of BCC and epidermis (see Fig. 2).

The mean Raman spectra of dermis and BCC in Fig. 2 show large differences in some of the selected peaks used in our model, such as those of r_5, i.e., the main collagen peaks. Thus, we could expect that few dermis would misclassified as BCC and vice versa. However, Fig. 4 shows that several dermis spectra are located in the middle of the BCC cluster. Inspection of these spectra showed that they were more similar to the BCC mean spectrum of Fig. 2 than to the mean Raman spectra of the dermis presented also in Fig. 2. Correlation with the H&E images indicated that these spectra corresponded to inflamed dermis regions, which had a higher number of cell nuclei than normal dermis. The variability in Raman spectra of dermis depending on its distance to the tumor has already been reported.[30]

Recall (cf. Sec. 2.3.2) that we are also interested in a reliable control of the classification regime. For example, an end-user might decide in real time to increase sensitivity of BCC detection to 99%. In our implementation, this can indeed be delivered immediately, i.e., without retraining the classification model. Certainly, sensitivity (or specificity) can be controlled with respect to any class, but here we illustrate this using BCC as the class of most interest. Thus, when the target sensitivity is 99%, we observe the following CV estimates of the delivered sensitivity: $98.6 \pm 2.5\%$ (70% to 30%),

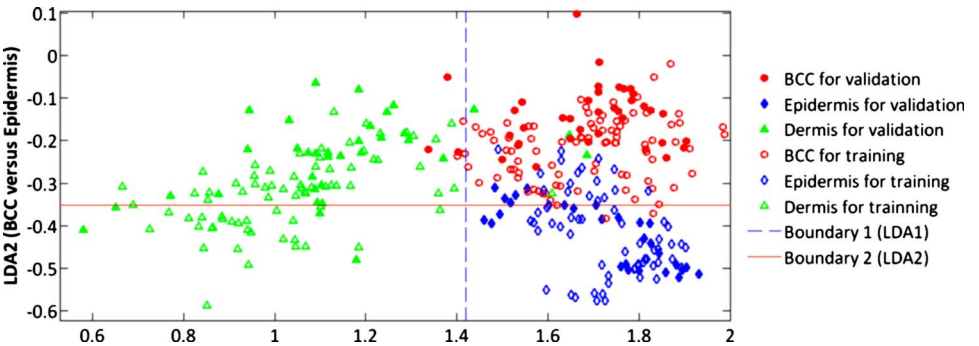

Fig. 4 Classification of 329 Raman spectra from BCC, epidermis, and dermis sampled into three groups by two consecutive LDAs. The dashed line (boundary 1) and the solid line (boundary 2) represent the 95% target sensitivity discrimination lines of the LDA1 and LDA2, respectively. Seventy percent of the data were used for training the model, and are represented in the figure as empty symbols. The other 30% were used for validation of the model, and their symbol in the figure are color-filled symbols. Symbol code: circles for BCC, diamonds for epidermis, and triangle for dermis. (Color online only.)

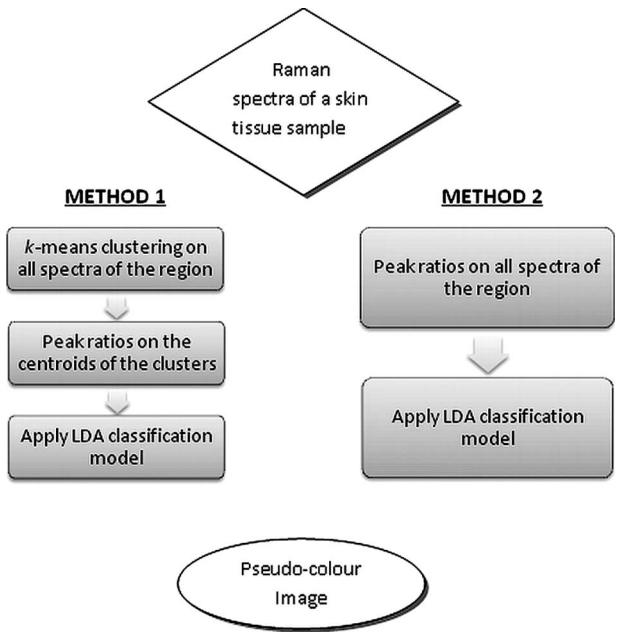

Fig. 5 Schematic diagram of the two different supervised procedures followed to build 2-D biochemical images of skin tissue sections.

98.9 ± 2.6% (80% to 20%), and 98.4 (leave-one-out); also, the corresponding (uncontrolled) specificity estimates obtained on the training subsets agree closely with those on the validation subsets under all three settings, and are around 76%.

3.3 *Raman Spectral Imaging*

The LDA model was applied to create 2-D biochemical images of tissue sections using two different procedures. In the *k*-means method, at least 11 classes were required to ensure discrimination between BCC and epidermis. As *k*-means clustering is an unsupervised method, the presence of any skin-irrelevant element or alteration in the sample was detected and classified first as a new class, prior to BCC and epidermis discrimination. After splitting the spectra into 11 clusters, the Raman band ratios of the centroid spectra corresponding to each cluster were introduced as input in the LDA model. Therefore, each cluster, corresponding to a color in the pseudocolor image, was classified as BCC, epidermis, or dermis. The second method consisted of directly applying the LDA model to the individual Raman spectra measured at each location of the analyzed tissue region to classify each unknown spectra as BCC, epidermis, or dermis. The schemes of both procedures are presented in Fig. 5. Pseudocolor images using both classification methods are presented in Figs. 6–8. These figures contain the automated RMS diagnosis and are presented for comparison with the gold standard histopathology images. Typical RMS images of skin sections containing nodular BCC are shown in Fig. 6. Figures 6(a)–6(c) show the ability of the RMS models to image tumor regions in a tissue section containing only BCC and dermis. Both dermis and tumor were accurately located within the tissue and were correctly classified. Only a few regions of inflamed dermis were misclassified as epidermis. The reason for this misclassification is likely to be due to the lower amount of DNA at those locations. In Figs. 6(d)–6(f), RMS was used to detect nodular BCC in a typical section containing all three regions included in the model: BCC, epidermis, and dermis. The correlation of the spectral images with the H&E image is extremely high. The dermis was correctly identified despite the presence of a large number of cells. The *k*-means-LDA model had a better accuracy in classification of epidermis and fewer misclassifications of BCC as epidermis. Extremely close agreement with H&E–stained images was also obtained with morphoeic BCC

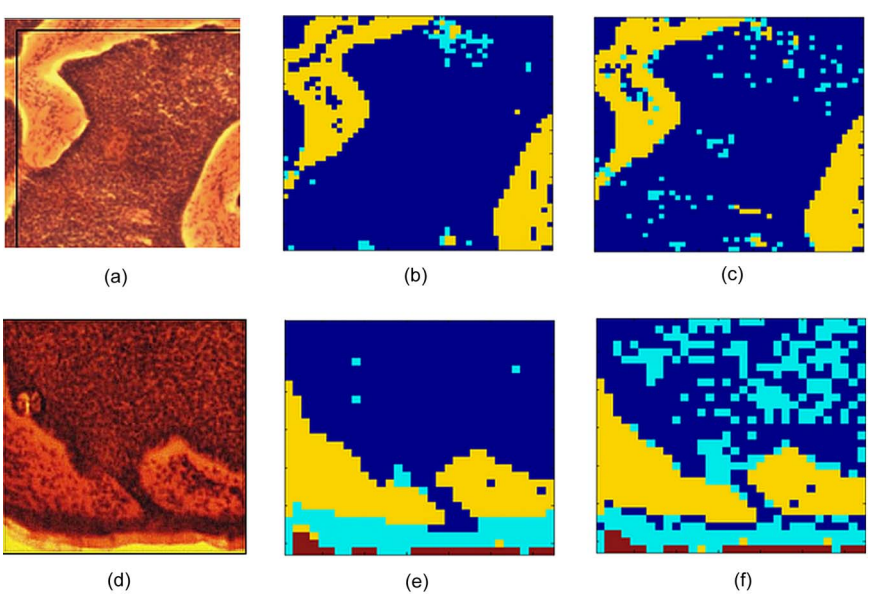

Fig. 6 Imaging nodular BCC in skin sections. Comparison between pseudocolor Raman images (b) and (c) and (e) and (f) produced with the two supervised methods and corresponding hispathology (e) H&E images (a) and (d). Color code: yellow, dermis; light blue, epidermis; and dark blue, BCC. Images (b) and (e) were producd by the *k*-means-LDA method, while (c) and (f) were produced by the direct LDA model. Tissue size: 500 × 500 μm. (Color online only.)

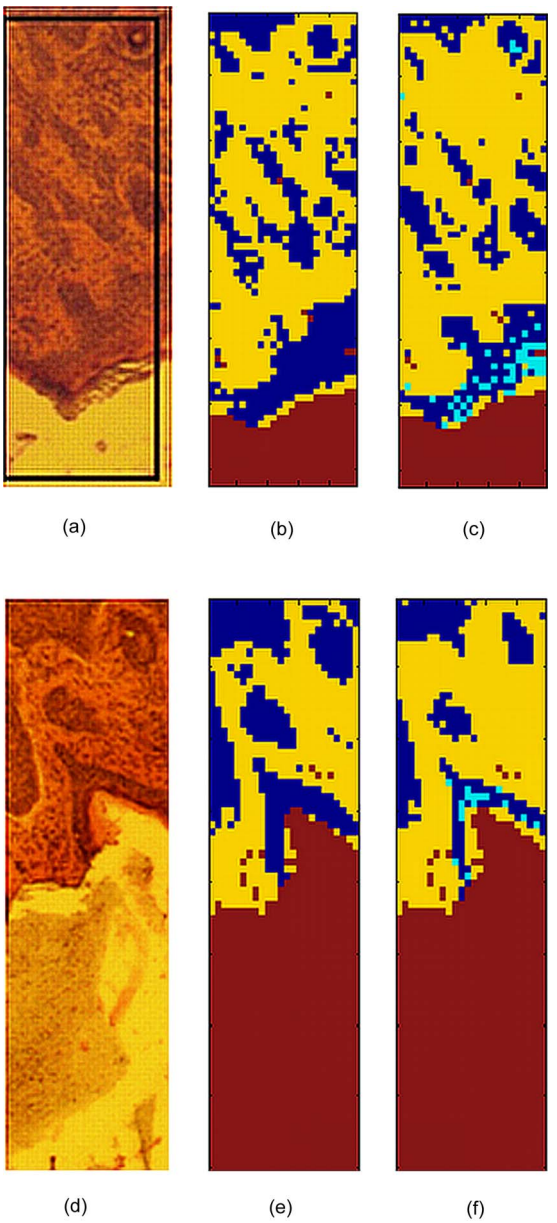

Fig. 7 Imaging morpheic BCC in skin sections. Comparison between pseudocolor Raman images (b) and (c) and (e) and (f) produced with the two supervised methods and corresponding hispathology (e) H&E images (a) and (d). Color code: brown, substrate; yellow, dermis; light blue, epidermis; and dark blue, BCC. Images (b) and (e) were produced by the *k*-means-LDA method, while (c) and (f) are produced by the direct LDA model. Tissue size: (a) 240×720 μm; (d) 240×840 μm. (Color online only.)

in Fig. 7, where BCC regions as small as 30 to 40 μm were detected. In these cases, the *k*-means-LDA method showed a higher number of epidermis misclassifications as BCC than direct application of the LDA model to each individual spectrum. From a clinical point of view, however, it is crucial that both classification methods were able to detect with high accuracy the presence of nodular and morphoeic BCC within the tissue sections, as well as the dermis regions. Nevertheless, due to the overlap of BCC and epidermis clusters shown in Fig. 4 some of the epidermis spectra were misclassified as

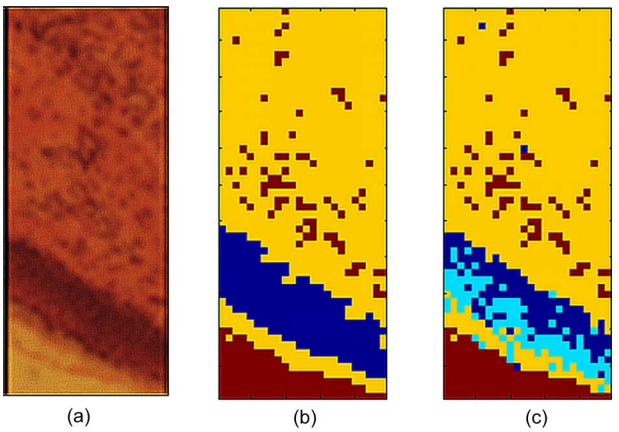

Fig. 8 Imaging of a skin tissue section clear of BCC. Comparison between pseudocolor Raman images (b) and (c) produced with the two supervised methods and corresponding hispathology H&E images (a). Color code: brown, substrate; yellow, dermis; light blue, epidermis; dark blue, BCC. Image (b) was produced by the *k*-means-LDA method and (c) was produced by the direct LDA model. Tissue size: 240×540 μm. (Color online only.)

BCC. However, this misclassification of epidermis as BCC has a lesser clinical significance, because if a region at the edge of the tissue is predicted as BCC but shows no BCC regions within the dermis, it is likely that it is real epidermis that has been misclassified as BCC. Also, areas located deep into the sample predicted as epidermis are more likely to be BCCs or very highly inflamed dermis than epidermis, unless they belong to hair follicles.

Finally, the technique was applied to skin tissue sections excised during MMS and declared clear of BCC by histopathology (Fig. 8). Both methods where able to detect dermis with high accuracy and no BCC regions were predicted within the dermis. Some epidermis regions were misclassified as BCC, particularly by the *k*-means-LDA method, due to the higher spectral similarities between BCC and epidermis.

Therefore, due to the current lower specificity, we envisage that initially RMS may be used to image all tissue layers removed during MMS and only the sections declared clear of BCC or where BCC is detected at the edge of a tissue section would be evaluated by the surgeon using frozen sections to check that no BCC was missed. As the prediction accuracy and specificity would increase in time by inclusion of a large number of tissue specimens, the RMS may be used as an automated imaging method for BCC, eliminating the need of frozen section preparation and histopathology observation. In addition, the provision of automated objective diagnosis may also reduce interobserver variations during the histopathology evaluation of skin sections. These potential changes in surgery practice have the potential to improve the MMS efficiency, enabling more BCC patients to benefit from the best treatment available.

4 Conclusion

We have shown that RMS using supervised classification models can be used for detection and imaging of BCC in skin tissue sections excised during MMS. Therefore, this technique

may represent a feasible alternative toward the development of automated tumor imaging during surgery.

The LDA model was developed using 329 Raman spectra from 20 patients, including 127 BCCs, 92 epidermis, and 110 dermis. Selected Raman bands corresponding to nucleic acids and collagen type I were computed and used as input into the multivariate model. BCC was discriminated from healthy tissue with $90 \pm 9\%$ sensitivity and $85 \pm 9\%$ specificity in a 70% to 30% split cross-validation algorithm.

The model was utilized to build 2-D biochemical images of unknown skin tissue samples excised during MMS. The images were obtained using two supervised methods. The first method applied *k*-means clustering to the whole spectral database, and then ratios of selected Raman bands for each centroid were computed and introduced into the LDA classification model. The second method directly applied the LDA model to compute the peak ratios of the whole spectral database. Both analysis methods provided quantitative images by applying a classification model on new tissue samples. The pseudocolor images produced revealed the presence/absence of tumor without histopathologist intervention and determined accurately its location within the sample.

This study has demonstrated the potential of RMS for automated diagnosis of nodular and morphemic BCC. Further studies are required to establish the ability of this technique to detect more types of BCC (superficial, cystic, and infiltrative BCC) and to identify other skin structures that can be confused with BCC, such as hair follicles: inflammation, especially lymphocytic: trichoblastoma; and trichoepithelioma. In addition, the potential of RMS to image BCC regions in tissue blocks remains to be demonstrated.

Acknowledgments

The authors would like to acknowledge the financial support of the UK National Institute for Health Research (NEAT FSG004), the University of Nottingham Hospitals Charity, and the Engineering and Physical Sciences Research Council (Bridging the Gaps Grant EP/E018580/1).

References

1. Cancer Research UK, revised May 6, 2009, ⟨http://info.cancerresearchuk.org/cancerstats/types/skin/⟩ (accessed May 2009).
2. American Cancer Society, *Cancer Facts & Figures 2008*, revised 2008, ⟨http://www.cancer.org/downloads/STT/2008CAFFfinalsecured.pdf⟩ (accessed May 2009).
3. American Cancer Society, "What are basal and squamous cell skin cancers?" revised July 30, 2008, ⟨http://www.cancer.org/docroot/CRI/content/CRI_2_2_1x_What_Is_Nonmelanoma_Skin_Cancer_51.asp?sitearea⟩ (accessed Nov. 15, 2008).
4. Cancer Research UK, "Skin cancer," "Treating skin cancer," "Surgery," updated May 15, 2007, ⟨http://www.cancerhelp.uk/help/default.asp?page=4326#mohs⟩ (accessed Nov. 15, 2008).
5. J. E. Lane and D. E. Kent, "Surgical margins in the treatment of nonmelanoma skin cancer and mohs micrographic surgery," *Curr. Surg.* **62**, 518–526 (2005).
6. T. W. McGovern, "Mohs surgery: the informed view," *Arch. Dermatol.* **135**, 1255–1259 (1999).
7. N. Smeets, G. Krekels, J. Ostertag, B. Essers, C. Dirksen, F. Nieman, and H. Neumann, "Surgical excision Mohs' micrographic surgery for basal-cell carcinoma of the face: randomised controlled trial," *Lancet* **364**(9447), 1766–1772 (2004).
8. N. R. Telfer, G. B. Colver, and W. Bowers, "Guidelines for the management of basal cell carcinoma," *Br. J. Dermatol.* **141**(3), 415–423 (1999).
9. I. Leibovitch, S. C. Huilgol, D. Selva, S. Richards, and R. Paver, "Basal cell carcinoma treated with Mohs surgery in Australia 11. Outcome at 5-year follow-up," *J. Am. Acad. Dermatol.* **53**, 452–457 (2005).
10. R. Rowe, R. Carroll, and C. Day, "Long-term recurrence rates in previously untreated (primary) basal cell carcinoma: implications for patient follow-up," *J. Dermatol. Surg. Oncol.* **15**(3), 315–328 (1989).
11. T. L. Bialy, J. Whalen, E. Veledar, D. Lafreniere, J. Spiro, T. Chartier, and S. C. Chen, "Mohs micrographic surgery vs traditional surgical excision: a cost comparison analysis," *Arch. Dermatol.* **140**, 736–742 (2004).
12. B. A. B. Essers, C. D. Dirksen, F. H. M. Nieman, N. W. J. Smeets, G. A. M. Krekels, M. H. Prins, and H. A. M. Neumann, "Cost-effectiveness of Mohs micrographic surgery vs surgical excision for basal cell carcinoma of the face," *Arch. Dermatol.* **142**, 187–194 (2006).
13. F. Bath-Hextall, J. Bong, W. Perkins, and H. Williams, "Interventions for basal cell carcinoma of the skin: systematic review," *Br. Med. J.* **329**, 705–709 (2004).
14. National Institute for Health snd Clinical Excellence 2006, "Improving outcomes for people with skin tumours including melanoma," issued Feb. 2006, ⟨http://www.nice.org.uk/guidance/CSGSTIM⟩ (accessed May 2009).
15. M. Mogensen and G. B. E. Jemec, "Diagnosis of nonmelanoma skin cancer/keratinocyte carcinoma: a review of diagnostic accuracy of nonmelanoma skin cancer diagnostic tests and technologies," *Dermatol. Surg.* **33**, 1158–1174 (2007).
16. L. Brochez, E. Verhaeghe, E. Grosshans, E. Haneke, G. Piérard, D. Ruiter, and J.-M. Naeyaert, "Inter-observer variation in the histopathological diagnosis of clinically suspicious pigmented skin lesions," *J. Pathol.* **196**, 459–466 (2002).
17. M. J. Trotter and A. K. Bruecks, "Interpretation of skin biopsies by general pathologists: diagnostic discrepancy rate measured by blinded review," *Arch. Pathol. Lab Med.* **127**, 1489–1492 (2003).
18. L. Brancaleon, A. J. Durkin, J. H. Tu, G. Menaker, J. D. Fallon, and N. Kollias, "*In vivo* fluorescence spectroscopy of nonmelanoma skin cancer," *Photochem. Photobiol.* **73**, 178–183 (2001).
19. B. Stenquist, M. B. Ericson, L. Molne, A. Rosen, O. Larko, and A. M. Wennberg, "Bispectral fluorescence imaging of aggressive basal cell carcinoma combined with histopathological mapping: a preliminary study indicating a possible adjunct to Mohs micrographic surgery," *Br. J. Dermatol.* **154**, 305–309 (2006).
20. R. Na, I. M. Stender, and H. C. Wulf, "Can autofluorescence demarcate basal cell carcinoma from normal skin? a comparison with protoporphyrin IX fluorescence," *Acta Derm Venereol* **81**(4), 246–249 (2001).
21. L. M. McIntosh, M. Jackson, H. H. Mantsch, M. F. Stranc, D. Pilavdzic, and A. N. Crowson, "Infrared spectra of basal cell carcinomas are distinct from non-tumor-bearing skin components," *J. Invest. Dermatol.* **112**, 951–956 (1999).
22. L. M. McIntosh, R. Summers, M. Jackson, H. H. Mantsch, J. R. Mansfield, M. Howlett, A. N. Crowson, and J. W. P. Toole, "Towards non-invasive screening of skin lesions by near-infrared spectroscopy," *J. Invest. Dermatol.* **116**, 175–181 (2001).
23. R. Manoharan, Y. Wang, and M. S. Feld, "Histochemical analysis of biological tissues using Raman spectroscopy," *Surf. Interface Anal.* **52**, 215–249 (1996).
24. K. K. Kumar, A. Anand, M. V. P. Chowdary, K. Thakur, J. Kurien, C. M. Krishna, and S. Mathew, "Discrimination of normal and malignant stomach mucosal tissues by Raman spectroscopy: a pilot study," *Vib. Spectrosc.* **44**, 382–387 (2007).
25. I. Notingher and L. L. Hench, "Raman microspectroscopy: a noninvasive tool for studies of individual living cells *in vitro*," *Expert Rev. Med. Devices* **3**(2), 215–234 (2006).
26. C. Krafft, D. Codrich, G. Pelizzo, and V. Sergo, "Raman mapping and FTIR imaging of lung tissue: congenital cystic adenomatoid malformation," *Analyst (Cambridge, U.K.)* **133**, 361–371 (2008).
27. M. Keller, E. M. Kanter, and A. Mahadevan-Jansen, "Raman spectroscopy for cancer diagnosis," *Spectroscopy* **21**(11), 33–41 (2006).
28. M. Gniadecka, H. C. Wulf, O. F. Nielsen, D. H. Christensen, and J. Hercogova, "Distinctive molecular abnormalities in benign and

malignant skin lesions: studies by Raman spectroscopy," *Photochem. Photobiol.* **66**, 418–423 (1997).
29. T. R. Hata, T. A. Scholz, I. V. Ermakov, R. W. McClane, F. Khachik, W. Gellermann, and L. K. Pershing, "Non-invasive Raman spectroscopic detection of carotenoids in human skin," *J. Invest. Dermatol.* **115**(3) 441–448 (2000).
30. A. Nijssen, T. C. B. Schut, F. Heule, P. J. Caspers, D. P. Hayes, M. H. Neumann, and G. J. Puppels, "Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy," *J. Invest. Dermatol.* **119**, 64–69 (2002).
31. A. Nijssen, S. Koljenovic, T. C. Bakker Schut, P. J. Caspers, and G. J. Puppels, "Towards oncological application of Raman spectroscopy," *J. Biophoton.* **2**(1–2), 29–36 (2009).
32. C. A. Lieber, S. K. Majumder, D. J. Ellis, D. D. Billheimer, and A. Mahadevan-Jansen, "*In vivo* nonmelanoma skin cancer diagnosis using Raman microspectroscopy," *Lasers Surg. Med.* **40**, 461–467 (2008).
33. M. Gniadecka, H. C. Wulf, N. Nymark Mortensen, O. F. Nielsen, and D. H. Christensen, "Diagnosis of basal cell carcinoma by Raman spectroscopy," *J. Raman Spectrosc.* **28**, 125–130 (1997).
34. M. Gniadecka, P. Alshede Philipsen, S. Sigurdsson, S. Wessel, O. F. Nielsen, D. H. Christensen, J. Hercogova, K. Rossen, H. K. Thomsen, R. Gniadecki, L. K. Hansen, and H. C. Wulf, "Melanoma diagnosis by Raman spectroscopy and neural networks: structure alterations in proteins and lipids in intact cancer tissue," *J. Invest. Dermatol.* **122**, 443–449 (2004).
35. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer by using Raman spectroscopy," *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12371–12376 (2005).
36. C. Kendall, N. Shepherd, B. Warren, K. Geboes, H. Barr, and N. Stone, "Raman spectroscopy a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus," *J. Pathol.* **200**, 602–609 (2003).
37. Z. Huang, A. McWilliams, H. Lui, D. I. McLean, S. Lam, and H. Zeng, "Near-infrared Raman spectroscopy for optical diagnosis of lung cancer," *Int. J. Cancer* **107**, 1047–1052 (2003).
38. C. Murali Krishna, N. B. Prathima, R. Malini, B. M. Vadhiraja, R. A. Bhatt, D. J. Fernandes, P. Kushtagi, M. S. Vidyasagar, and V. B. Kartha, "Raman spectroscopy studies for diagnosis of cancers in human uterine cervix," *Vib. Spectrosc.* **41**, 136–141 (2006).
39. P. Crow, N. Stone, C. A. Kendall, J. S. Uff, J. A. M. Farmer, H. Barr, and M. P. J. Wright, "The use of Raman spectroscopy to identify and grade prostatic adenocarcinoma *in vitro*," *Br. J. Cancer* **89**, 106–108 (2003).
40. C. Krafft, S. B. Sobottka, G. Schackert, and R. Salzer, "Raman and infrared spectroscopic mapping of human primary intracranial tumors: a comparative study," *J. Raman Spectrosc.* **37**, 367–375 (2006).
41. G. Shetty, C. Kendall, N. Shepherd, N. Stone, and H. Barr, "Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus," *Br. J. Cancer* **94**, 1460–1464 (2006).
42. M. J. Romeo and M. Diem, "Infrared spectral imaging of lymph nodes: strategies for analysis and artifact reduction," *Vib. Spectrosc.* **38**, 115–119 (2005).
43. J. Hutchings, C. Kendall, B. Smith, N. Shepherd, H. Barr, and N. Stone, "The potential for histological screening using a combination of rapid Raman mapping and principal component analysis," *J. Biophotonics* **2**(1–2), 91–103 (2009).
44. S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis—a brief tutorial," tutorial for Speech Recognition course, March 1998, July 30, 2008, ⟨http://www.isip.msstate.edu/publications/reports/isip_internal/1998/⟩ (accessed Nov. 15, 2008).
45. N. Stone, C. Kendall, N. Shepherd, P. Crow, and H. Barr, "Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers," *J. Raman Spectrosc.* **33**, 564–573 (2002).
46. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., pp. 114–121, John Wiley and Sons, New York (2001).
47. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning Data Minining, Inference, and Prediction*, Springer Series in Statistics, Springer, New York (2001).
48. S. Sasic, D. A. Clark, J. C. Mitchell, and M. J. Snowden, "Raman line mapping as a fast method for analyzing pharmaceutical bead formulations," *Analyst (Cambridge, U.K.)* **130**, 1530–1536 (2005).
49. A. Stevens and J. S. Lowe, *Human Histology*, 2nd ed., p. 6, Mosby, London (1997).
50. A. T. Tu, *Raman Spectroscopy in Biology: Principles and Applications*, John Wiley and Sons, New York (1982).
51. B. G. Frushour and J. L. Koenig, "Raman scattering of collagen, gelatine and elastin," *Biopolymers* **14**(2), 379–391 (1975).
52. Z. Movasaghi, S. Rehman, and I. U. Rehman, "Raman spectroscopy of biological tissues," *Appl. Spectrosc. Rev.* **42**, 493–541 (2007).