Raman biophysical markers in skin cancer diagnosis

Xu Feng
Austin J. Moy
Hieu T. M. Nguyen
Yao Zhang
Jason Zhang
Matthew C. Fox
Katherine R. Sebastian
Jason S. Reichenberg
Mia K. Markey
James W. Tunnell

Raman biophysical markers in skin cancer diagnosis

Xu Feng,[a] Austin J. Moy,[a] Hieu T. M. Nguyen,[a] Yao Zhang,[a] Jason Zhang,[a] Matthew C. Fox,[b]
Katherine R. Sebastian,[b] Jason S. Reichenberg,[b] Mia K. Markey,[a] and James W. Tunnell[a,*]
[a]University of Texas at Austin, Department of Biomedical Engineering, Austin, Texas, Unites States
[b]University of Texas at Austin, Dell Medical School, Department of Medicine, Austin, Texas, United States

Abstract. Raman spectroscopy (RS) has demonstrated great potential for *in vivo* cancer screening; however, the biophysical changes that occur for specific diagnoses remain unclear. We recently developed an inverse biophysical skin cancer model to address this issue. Here, we presented the first demonstration of *in vivo* melanoma and nonmelanoma skin cancer (NMSC) detection based on this model. We fit the model to our previous clinical dataset and extracted the concentration of eight Raman active components in 100 lesions in 65 patients diagnosed with malignant melanoma (MM), dysplastic nevi (DN), basal cell carcinoma, squamous cell carcinoma, and actinic keratosis. We then used logistic regression and leave-one-lesion-out cross validation to determine the diagnostically relevant model components. Our results showed that the biophysical model captures the diagnostic power of the previously used statistical classification model while also providing the skin's biophysical composition. In addition, collagen and triolein were the most relevant biomarkers to represent the spectral variances between MM and DN, and between NMSC and normal tissue. Our work demonstrates the ability of RS to reveal the biophysical basis for accurate diagnosis of different skin cancers, which may eventually lead to a reduction in the number of unnecessary excisional skin biopsies performed. © *2018 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.23.5.057002]

Keywords: Raman spectroscopy; skin cancer; diagnosis; biophysical marker; optical sensing.

Paper 180081RR received Feb. 6, 2018; accepted for publication Apr. 23, 2018; published online May 11, 2018.

1 Introduction

Raman spectroscopy (RS) has emerged as a powerful tool for clinical diagnosis of skin malignancies. RS offers a number of advantages compared with gold standard biopsy: it has high endogenous molecular specificity, it is minimally invasive, and it does not require sample preparation.[1] The development of Raman optical fiber probes has greatly promoted its application in real time *in vivo* skin cancer screening.

Previous studies have shown that RS is highly sensitive in differentiating malignant melanoma (MM, the deadliest version of skin cancer) from benign pigmented lesion (PL, frequently confused in the clinic with MM).[2–4] Our group has demonstrated that MM (12 lesions) can be discriminated from PL (17 lesions) with 100% sensitivity and specificity using a Raman probe-based system[5] and principal components analysis (PCA) with a logistic regression classifier.[4] Schleusener et al.[6] discriminated MM (23 lesions) and PL (33 lesions) with a balanced accuracy of 91% using partial least squares discriminant analysis (PLS-DA). Lui et al.[3] discriminated MM (44 lesions) from PL (286 lesions) with 90% to 99% sensitivity and 15% to 68% specificity using PCA with generalized discriminant analysis (PCA-GDA). A follow-up independent validation study from the same group showed consistent results for discriminating MM (53 lesions) from PL (336 lesions).[7] Their research later led to the commercial launch of a clinical skin cancer detection device (Verisante Aura) in Canada.[3]

RS also has been used to detect nonmelanoma skin cancer (NMSC), mainly basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and actinic keratosis (AK, a precancerous state). Lieber et al.[8] developed a portable confocal Raman system with a handheld probe and achieved 100% sensitivity and 91% specificity in discriminating BCC, SCC, and inflamed scar tissues from normal tissues (21 versus 21). The spectral differences were extracted through the maximum representation and discrimination feature statistical method. Silveira et al.[9] discriminated BCC, SCC, and AK from nontumorous tissue (44 versus 55) with ~91.9% accuracy using a dispersive Raman system and PLS-DA. Schleusener et al.[6] discriminated BCC from normal skin (35 versus 104) and SCC from normal skin (22 versus 104) based on PLS-DA with a balanced accuracy of 73% and 85%, respectively. Lui et al.[3] distinguished skin cancer and AK from benign lesions with 90% to 99% sensitivity and 24% to 66% specificity based on PLS.

Despite these successes, these studies have employed statistical classifiers, sometimes called "black box" methods, to describe the spectral differences between pathologies. The challenge with these statistical algorithms lies in interpreting the biophysical basis for their discriminant ability. That is, they do not provide insights into the most relevant cancer biomarkers that RS relies on to make an accurate diagnosis. Therefore, we aim to determine the biophysical basis of skin cancer detection based on RS. This may enable the pathologist to interpret the spectral data in a familiar manner (such as a thickening epidermis, the change in collagen and lipid content, etc.) and guide a dermatologist in determining the most appropriate treatment.

Our group recently proposed a Raman biophysical model, an inverse model that derived the skin's biochemical makeup from its Raman spectrum.[10] The model described the Raman spectra

*Address all correspondence to: James W. Tunnell, E-mail: jtunnell@mail.
utexas.edu

from *in vivo* human skin as a linear combination of eight Raman active skin constituents extracted from skin *in situ*, including collagen, elastin, keratin, triolein, ceramide, nucleus, melanin, and water. We have validated the model using previous *in vivo* human skin cancer screening data[10] and identified distinct biophysical changes between pathologies. However, we have not evaluated the diagnostic potential of those biophysical parameters in discriminating skin cancers. We also have not identified the important biophysical features used as diagnostic tools.
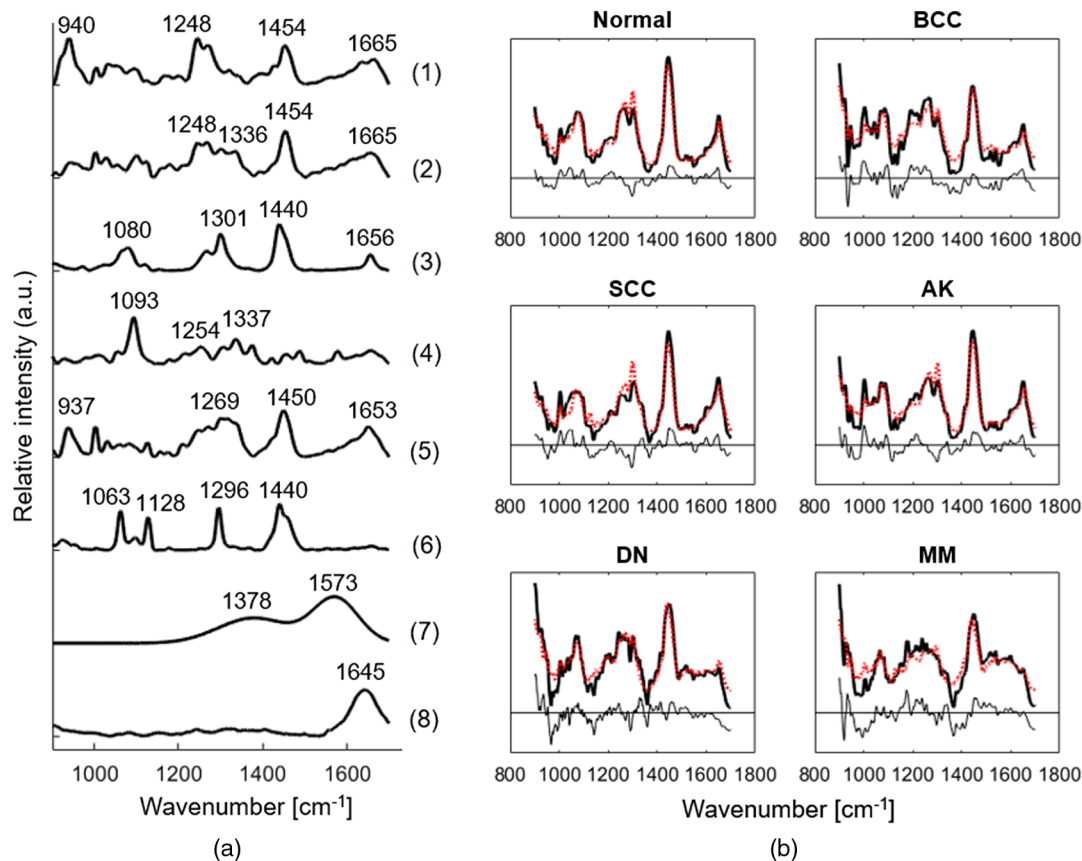
Here, we present a preliminary study of *in vivo* diagnosis of melanoma and NMSC on the biophysical basis. We demonstrated that the biophysical model captures the diagnostic power of the previously used statistical classification model while also providing the skin's biophysical composition. Our work demonstrates the ability of RS in sensing the biochemical composition of skin cancers, thus allowing for better interpretation of the diagnostic results from a pathological basis.

2 Materials and Methods

2.1 Clinical Instrument and Dataset

The clinical skin cancer screening study[4] was conducted using a Raman optical fiber probe[11] integrated in an optical fiber probe-based system.[5] An 830-nm wavelength excitation was used to minimize tissue autofluorescence. Collected signals entered a spectrograph and were imaged onto a camera. Integration time for each measurement was 3 s. Spectral resolution of the probe-based system is around 10 cm^{-1}. This study was approved by the Institutional Review Board at the University of Texas at Austin and the University of Texas MD Anderson Cancer Center (trial registration ID: NCT 00476905). Informed consents were acquired from all patients prior to the study.

In vivo Raman spectra were obtained from 65 patients diagnosed with BCC, SCC, AK, dysplastic nevi (DN, a dysplastic

Table 1 Summary of clinical data.

Lesion type	# patients	# lesions (# spectra)	# adjacent normal tissues (# spectra)
MM	10	12 (33)	11 (23)
DN	11	17 (37)	17 (33)
BCC	14	19 (39)	19 (38)
SCC	20	38 (81)	38 (76)
AK	10	14 (30)	14 (28)
Total	65	100 (220)	99 (198)

Fig. 1 (a) Eight model components: (1) collagen, (2) elastin, (3) triolein, (4) nucleus, (5) keratin, (6) ceramide, (7) melanin, and (8) water. Peak positions of the main Raman bands are labeled. (b) Fitting results for the average Raman spectra of normal tissue, BCC, SCC, AK, DN, and MM. Black solid lines: average tissue spectra. Red dotted lines: model fits. Residuals are also plotted on the bottom. Images are adapted from Ref. 10.

form of PL), and MM. Details of the clinical data are provided in Table 1. In total, there are 100 lesions and 99 adjacent normal tissues because one normal tissue was shared between two lesions. Fourteen out of 38 SCC lesions containing both SCC and AK were grouped into SCC. Multiple spectra were taken from each lesion by moving the probe to different locations to sample as much of the lesion as possible. Multiple spectra were also taken from the normal skin adjacent to each individual lesion. Although not verified by histopathology, normal skin was visually verified to be normal by an experienced dermatologist or physician assistant.

2.2 Data Preprocessing

Spectra underwent wavenumber calibration, dark noise removal, cosmic ray removal, and smoothing, followed by a fifth-order polynomial fitting[12] to remove tissue fluorescence background. Spectral data were spectral response calibrated using a tungsten halogen lamp (LS-1-CAL, Ocean Optics). Spectral band between 800 and 900 cm^{-1} was excluded due to a strong broad fiber background peak around 800 cm^{-1}. A sharp room light peak at 1100 cm^{-1} was removed from five spectra from one MM patient.

2.3 Diagnostic Algorithms

2.3.1 Classification tasks

We used four classification tasks in this study: (1) MM versus DN, (2) MM, DN versus normal (norm), (3) NMSC (BCC, SCC, and AK) versus norm, and (4) SCC, BCC versus AK. Diagnostic algorithms were implemented within MATLAB (version R2015a, MathWorks).

We chose these four classification tasks not only to be consistent with our previous study[4] but also based on their clinical significance. Task (1) is significant, because it directly affects the decision of a clinician to remove the lesion or continue to observe when facing a pigmented lesion of concern. Task (4) is significant, because while a BCC or SCC will require surgical excision, it is often sufficient to treat an AK with cryotherapy or a topical chemotherapeutic agent. Both tasks (1) and (4) are highly related to reducing the number of unnecessary excisional skin biopsies. Although tasks (2) and (3) are not currently clinically actionable, they are very relevant to the perspective of tumor margin detection. We used normal skin as a placeholder for these other diagnoses, with the hope that in the future we can perform the analysis on enough benign lesions to allow the device to distinguish these benign issues from cancer.

2.3.2 Receiver operating characteristic

An ROC curve was used to determine a model's performance in discriminating between two groups. An ROC curve is a graphical representation of the trade-off between sensitivity and specificity. Sensitivity is the ability of the model to correctly identify the positive group, whereas specificity is the ability of the model to correctly identify the negative group. For good discrimination, the ROC curve is predominately in the left and top boundaries of the graph, whereas for poor discrimination, the ROC curve approaches the diagonal line drawn from the bottom-left to the top-right of the plot. ROC curves were calculated separately for PCA and biophysical model, and for each of the four classification tasks.

By default, the ROC curves were calculated by treating each lesion as an experimental unit. The method is described elsewhere:[13] if one or more spectra from a site were classified as cancer, the site was classified as cancer. If all spectra from a site were classified as normal, the site was classified as normal.

Table 2 Peak positions of the main Raman bands in the Raman active components.

Raman peaks [cm^{-1}]	Band assignments	Components
937	C—C stretching of proline and valine and protein backbone	Keratin
940	C—C stretching of protein backbone	Collagen, elastin
1003	C—C vibration of phenyl ring	Collagen, elastin, keratin
1063	C—C asymmetric skeletal stretching of lipids (trans-conformation);	Ceramide
1080	C—C skeletal stretching in lipids	Triolein
1093	O—P—O symmetric stretching vibration of the DNA backbone	Nucleus
1128	C—C symmetric skeletal stretching	Ceramide
1248	Amide III (β-sheet and random coil conformations)	Collagen, elastin
1254	β sheet/thymine/cytosine (DNA base/DNA and RNA base)	Nucleus
1269	Amide III (α-helix conformation), C—N stretching, N—H in-plane bending	Collagen, elastin, keratin
1301	C—H modes (CH_2 twisting and wagging) of lipids; CH_2/CH_3 bands	Triolein
1336	Amide III, C—N stretching, N—H in-plane bending	Elastin
1337	Adenine, guanine (DNA and RNA base)	Nucleus
1378	Linear stretching of the C—C bonds within the rings	Melanin
1440	CH_2/CH_3 bands	Triolein, ceramide
1450	C—H bending of proteins	Keratin
1454	C—H stretching, C—H asymmetric deformation	Collagen, elastin
1573	In-plane stretching of the aromatic rings	Melanin
1645	O—H bending mode of liquid water	Water
1653	C—O stretching model of amide I	Keratin
1656	C—C lipids	Triolein
1665	C—O amide I vibration	Collagen, elastin

We used this conservative technique to approximate the dermatologist's tendency to err on the side of caution.

2.3.3 Statistical model

The statistical model (PCA) was adopted from our previous publication.[4] For each classification task, we limited the number of principal components (PCs) to 5, because the diagnostic improvements dropped significantly beyond 4.[4] First, we performed PCA for a given classification task and then generated all the possible combinations of 1, 2, 3, 4, or 5 PCs from the first 15 PCs. Next, we selected one combination of PCs and built a logistic regression classifier. Specifically, for each PC-logistic regression analysis, a successive single lesion was left out for testing, with the remaining lesions being used for training. After the posterior probabilities of all lesions were calculated according to the leave-one-lesion-out cross-validation protocol, an ROC curve was then calculated. Using this method, we generated different ROC curves for different combinations of PCs. The combination of PCs that yielded the largest area under the ROC curve (AUC) was selected for subsequent analyses.

2.3.4 Biophysical model

In vivo Raman spectra were fit into the biophysical model with eight primary model components: collagen, elastin, triolein,

Table 3 Comparison of diagnostic performance of the statistical model and biophysical model.

Classification tasks	# lesion	Diagnostically relevant components		ROC AUC	
		Statistical model	Biophysical model	Statistical model	Biophysical model
MM versus DN	12 versus 17	PC 3,4,5,8,9	Collagen, triolein, and melanin	1.00	0.99
[MM, DN] versus norm	29 versus 28	PC 1,6,9	Triolein and melanin	0.89	0.93
[BCC, SCC, AK] versus norm	72 versus 64	PC 3,4,8,9	Collagen, triolein, elastin, nucleus, and ceramide	0.58	0.76
[SCC, BCC] versus AK	68 versus 55	PC 3,6,7,8	Collagen, keratin, and water	0.62	0.65
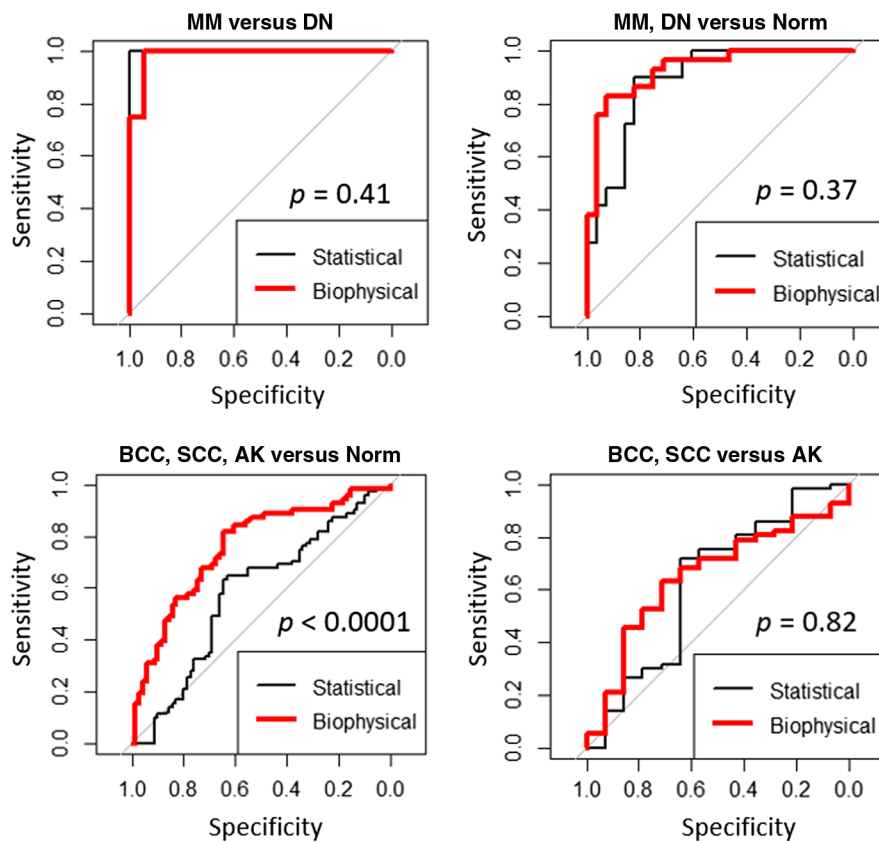
Fig. 2 Comparison of ROC curves between statistical model (thin line) and biophysical model (thick line) for the four classification tasks: MM versus DN, MM, DN versus norm (adjacent normal tissue), BCC, SCC, AK versus norm, and BCC, SCC versus AK. The ROC curves are statistically compared, and the *p* values are labeled. *p* > 0.05 indicates no significant difference between the two curves.

Table 4 Comparison of specificities derived from ROCs according to sensitivities of 95% and 90%.

Classification tasks	Statistical model		Biophysical model	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
MM versus DN	95	100	95	94
	90	100	90	94
[MM, DN] versus norm	95	64	95	71
	90	71	90	75
[BCC, SCC, AK] versus norm	95	10	95	18
	90	6	90	39
[SCC, BCC] versus AK	95	21	95	11
	90	21	90	21

nucleus, keratin, ceramide, melanin, and water, as shown in Fig. 1. Those components were collected from human skin *in situ* and were averaged over multiple patients.[10] Those components contain both biochemical and structural information. For instance, nucleus refers to the nuclear material in the cell. Collagen and elastin refer to dermal extracellular matrix. Keratin represents epidermal extracellular matrix. Triolein mainly represents subcutaneous fat. Peak positions of the main Raman bands and their physical origin are summarized in Table 2.[10,14–16] The subbands (or subpeaks) were not listed but also played a role in the fitting. The fit coefficients provide the relative concentration of those components and were used as the input variables of the discriminant analysis. Similar to PCA model, for each classification task, we generated all the possible combinations of 1, 2, 3, 4, or 5 components from the eight primary model components and built logistic regression classifiers. We then selected the combination of model components that yielded the largest AUC.

2.4 Comparison of Discriminative Capability Between Statistical and Biophysical Models

Statistical analysis was performed using an open-source package written in R software (version 3.3.3).[17] The AUC of two
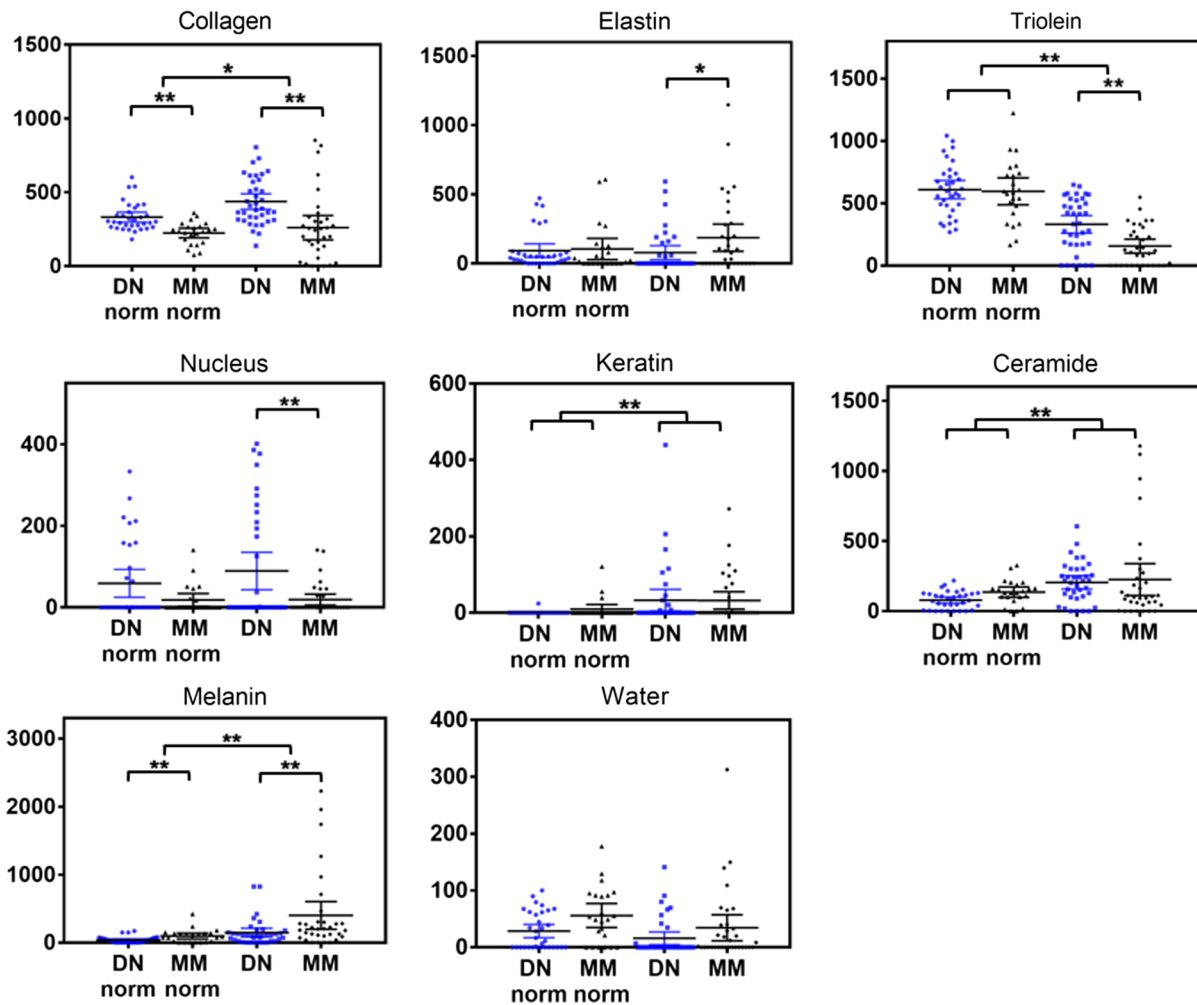
Fig. 3 Fit coefficients of the eight model components computed from the biophysical model. Each point represents a spectrum data. Significance tests are conducted for the fit coefficients of DN norm (the adjacent normal tissue of DN) versus MM norm (the adjacent normal tissue of MM), DN versus MM, and [DN norm and MM norm] versus [DN and MM]. **$p \leq 0.01$, *$p \leq 0.05$.

paired ROC curves was compared using the bootstrap test, with a goal to determine if the biophysical model provides at least equivalent potential for classification compared with the statistical model.

2.5 Interpretation of Biophysical Model Result

The fit coefficients of the eight model components generated by the biophysical model were visualized using scatter plots. Each scatter point represents one spectrum. The error bar generated by the 95% CI is used to represent the variance of the fit coefficient. Unpaired Student's t test was employed, and the corresponding p values were labeled to compare if the fit coefficients have any statistically significant difference between pathologies.
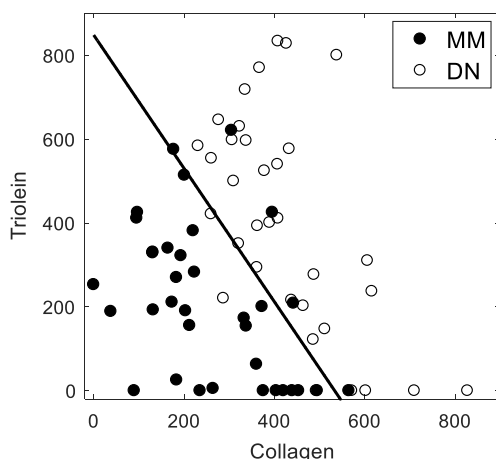
3 Results

3.1 Statistical Model Versus Biophysical Model

In Table 3, the diagnostically relevant model components in statistical and biophysical models are displayed and the AUCs are compared. Figure 2 compares the corresponding ROC curves. The AUCs of the ROC curves of the two models are not statistically distinguishable for the classification tasks of MM versus DN, [MM, DN] versus norm, and [BCC, SCC] versus AK. However, the AUC of the ROC of the biophysical model for [BCC, SCC, and AK] versus Norm is statistically significantly better than the corresponding statistical model ($p < 0.0001$). Table 4 compares the specificities of the two models corresponding to sensitivities of 90% and 95%, respectively.

3.2 Biophysical Basis of Classification Results

3.2.1 Malignant melanoma versus dysplastic nevi

The biophysical model reveals the biomarkers responsible for the variances between pathologies. Major bands used for fitting and their physical origin was shown in Table 2 and reported in the literature.[10,14–16] The fit coefficients of the eight model components in DN and MM are shown in Fig. 3. Statistical analysis indicates significant differences in collagen, elastin, triolein, nucleus, and melanin content between MM and DN. Collagen and triolein contributed greatly to the spectral variance between MM and DN. Using the fit coefficients of collagen and

Fig. 4 Scatter plot drawn from triolein and collagen content. The solid logistic regression line separates MM from DN.

triolein, 29 out of 33 MM spectra and 35 out of 37 DN spectra are correctly classified (Fig. 4).

The best result was achieved by employing three components: collagen, triolein, and melanin, resulting in 12 out of 12 MM lesions and 16 out of 17 DN lesions being correctly classified. ROC AUC is 0.99, and specificity is 94% (90% to 95% sensitivity, Table 4).

3.2.2 Pigmented lesions (MM, DN) versus adjacent normal tissue

Figure 3 shows that pigmented lesions and their adjacent normal tissue have significant differences in triolein, collagen, ceramide, keratin, and melanin content. Our results show that triolein and melanin are the most relevant model components to discriminate MM and DN from adjacent normal skin. The ROC AUC is 0.93 (Table 3) for sensitivities from 95% to 90% and specificities of 71% to 75% (Table 4).

3.2.3 Nonmelanoma skin cancers (BCC, SCC) and precancer (AK) versus normal skin

In Fig. 5, significant differences are found between BCC and adjacent normal tissue in collagen, triolein and melanin, and between SCC and adjacent normal tissue in collagen, triolein, keratin, and water. Our results showed that a combination of collagen, triolein, elastin, nucleus, and ceramide was best for discriminating BCC, SCC, and AK from adjacent normal tissue. The ROC AUC is 0.76 (Table 3). For sensitivities from 95% to 90%, specificities range from 18% to 39% (Table 4).

3.2.4 Nonmelanoma skin cancers (BCC, SCC) versus AK

Figure 5 shows significant differences in collagen, nucleus, keratin, and water between SCC and AK, as well as significant differences in keratin and ceramide between BCC and AK. The fit coefficients of collagen, keratin, and water discriminated BCC, SCC from AK with an ROC AUC of 0.65 (2) and specificities range from 11% to 21% for sensitivities corresponding to 95% to 90% (Table 4).

4 Discussion and Conclusions

In our previous work,[4] we demonstrated the capability of RS in detecting skin cancers using a statistical model. Here, we show that a biophysical model can achieve consistent diagnostic performance with the statistical model while simultaneously extracting the relevant biomarkers accounting for the diagnosis.

Our model reveals markedly different biochemical and structural compositions between pathologies. First, the amount of triolein is significantly lower in all skin lesions than surrounding normal skin. Triolein mostly originates from adipose tissue in the subcutaneous layer, with a small contribution from epidermal surface lipids.[18] Triolein has a large Raman scattering cross section, thus contributing greatly to normal skin spectra. The decrease of triolein in skin lesions does not necessarily indicate the actual amount of fat decreases in skin lesions, only that there is a decrease in the triolein sampled by the probe. One possible reason is epithelial thickening associated with dysplastic progression.[19,20] An increased thickness of epidermis would mean that the total volume of tissue sampled would include more epidermis and less adipose tissue, thereby decreasing the amount of Raman emission from deeper skin layers (adipose). Another possible reason for the decrease of triolein in
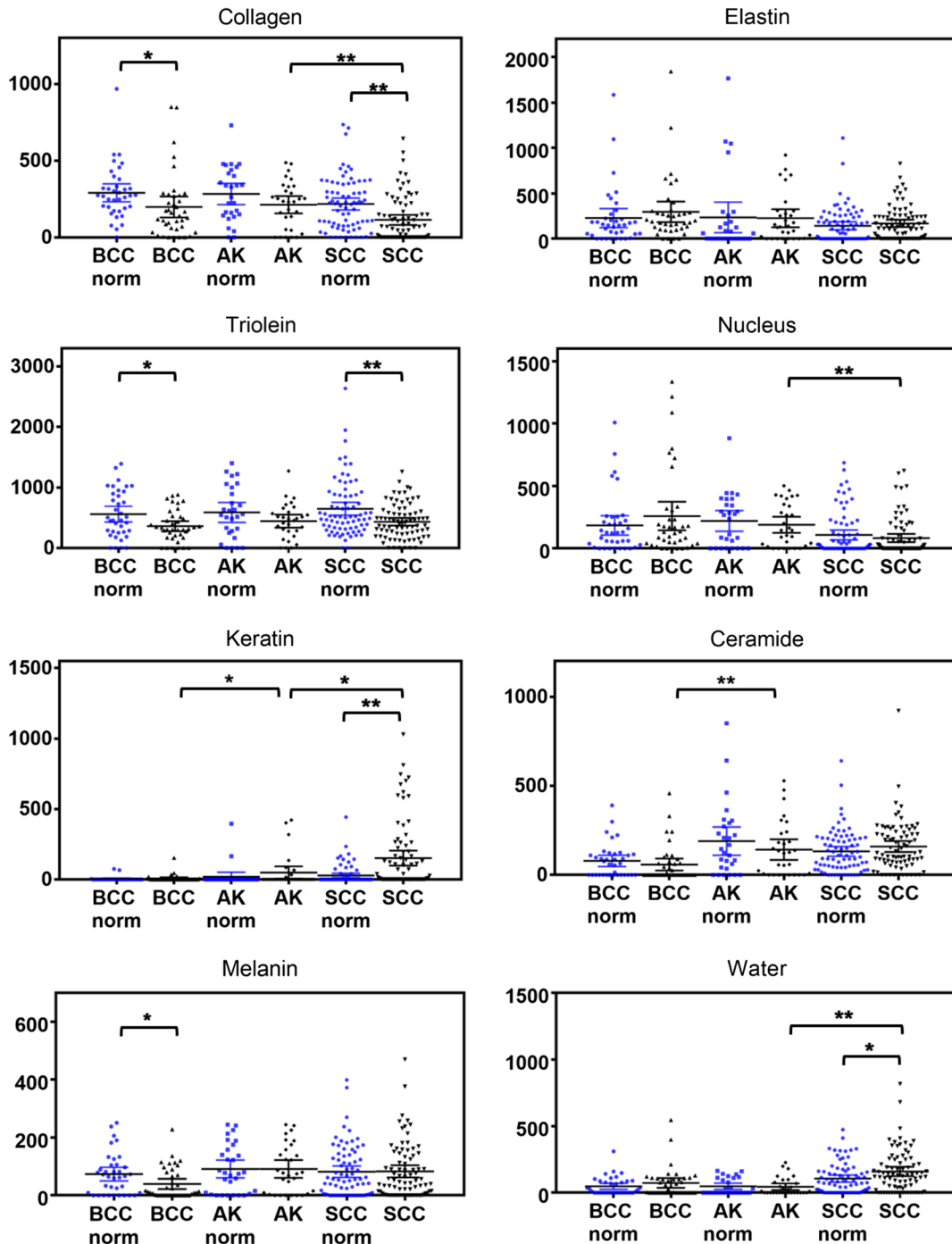
Fig. 5 Fit coefficients of the eight model components computed from the biophysical model. Each point represents a spectrum data. Significance tests are conducted for the fit coefficients of adjacent normal tissue of BCC (BCC norm) versus BCC, the adjacent normal tissue of SCC (SCC norm) versus SCC, SCC versus AK, and BCC versus AK. $**p \leq 0.01$, $*p \leq 0.05$.

pigmented lesions relative to the adjacent normal skin is that melanin strongly absorbs excitation laser power and therefore reduces the contribution of triolein in Raman signal.

Next, we found that the collagen content is significantly lower in NMSCs than their adjacent normal tissue and AK. For instance, collagen does not change significantly in the progression from normal to AK (benign), but it decreases significantly from AK to SCC (cancer). This trend of decreased collagen in cancer was also observed in previous biophysical models of *ex vivo* human skin fragments,[21] urological tissue,[22] gastric/esophagus tissue,[23] and cervical tissue.[24] This may be partially explained by the thickening of the epithelium as mentioned above. Other reasons may include the release of metalloproteinases by cancerous cells to degrade dermal connective tissue,[25,26] and extracellular-degrading enzymes secreted from fibroblasts that damage the stroma.[27]

Discriminating MM from benign pigmented lesions (especially DN) usually leads to large negative biopsy ratio. Due

to their highly similar appearance, the ratio of negative versus positive biopsies ranges from 22∶1 to 59∶1 for experienced versus new general practitioners.[28] Understanding of the biophysical basis of melanoma skin cancer progression is essential to reduce large negative biopsy ratio and save considerable associated costs and efforts. In our study, we discovered that collagen and triolein are the two most important biomarkers to differentiate MM from DN and NMSCs from normal tissue. Two previous *ex vivo* studies based on Raman biophysical models also showed collagen and triolein (or fat) had important roles in tissue Raman spectra. Bodanese et al.[29] discovered that the amount of collagen and fat extracted from tissue Raman spectra can classify BCC from normal skin with sensitivity and specificity of 95% and 83%, respectively. Haka et al.[30] found that the fit coefficients of collagen and fat can distinguish cancerous breast tissues from normal and benign tissues with 94% sensitivity and 96% specificity, respectively.

Our results show that melanin is an important biomarker for classifying pigmented lesions from adjacent normal tissue, which is as expected because pigmented lesions typically contain more melanin than the surrounding normal skin. However, we also found melanin is not as relevant as collagen and triolein in differentiating MM from DN. In fact, melanomas do not always have more melanin than do benign pigmented lesions. The existence of amelanotic melanoma is a good example—we estimated zero melanin content for the one amelanotic melanoma lesion in our sample. Blue nevi, on the other hand, contain abundant pigment but are not cancer. Thus, more data from amelanotic melanomas is needed to clarify the role that melanin may play in differentiating MM and DN.

We were best able to classify NMSCs from normal skin by employing a model that considered collagen, triolein, elastin, nucleus, and ceramide. To better understand the biophysical changes of each pathology, we examined the lesion-normal pairs for BCC and SCC separately. We found that melanin content is significantly lower in BCC than in adjacent normal skin, likely because the invasion of basal cells takes over the space normally occupied by the melanocytes. Although not statistically significant, the amount of nucleus and elastin is larger in BCC compared with its adjacent normal skin, which may be explained by the proliferation of cancer cells and the enlargement of nuclei. Elastin content is also larger in BCC than adjacent normal, probably because of the existence of solar elastosis.[31] On the contrary, SCC appears to have a higher amount of keratin, ceramide, and water as compared with its adjacent normal skin. The increase of keratin may be attributed to large areas of keratinization in response to malignant epithelial cells.[32] Ceramide indicates abnormal epidermal surface lipid synthesis and thus is a key component to differentiate SCC from normal skin.

AK is the most common precursor lesion of SCC among lightly pigmented individuals. Almost every SCC that arises on sun-damaged skin has evidence of AK in the epidermis, either directly contiguous with or adjacent to the neoplasm.[33] However, AK and SCC have a similar crusted appearance, making it difficult to differentiate by visual examination. We found that the most important components to discriminate SCC from AK are collagen, keratin, and water. AK is confined to foci within the epidermis, whereas SCC may further invade into dermis. Thus, SCC is expected to have a higher amount of keratin than AK. Nucleus content is lower in SCC than AK, likely because the prominent keratinization in SCC occupies the

space of cells. We also observed a higher amount of water content in SCC than AK. High wavenumber Raman will be an ideal tool to study the significance of water in NMSC diagnosis.

An interesting discovery is that the normal tissue adjacent to a DN has significantly more collagen than normal tissue adjacent to a MM (Fig. 2). We were suspicious that the observed difference in collagen could be simply due to aging as the average age of the MM patients ($N = 9$) in our study was 65 years (one patients did not have age information on record), whereas the average age of the DN patients ($N = 11$) was 42 years (Table 5). To control for the effect of aging, we built a generalized linear mixed-effect model[34] using patient age and collagen as fixed effects predictors, and tissue type as the response variable (0 = normal tissue adjacent to DN, 1 = normal tissue adjacent to MM). We also included a random-effects term for intercept grouped by patient to account for patient-specific variations. Our result shows that the p value of collagen is 0.041, indicating the amount of collagen is a significant predictor of tissue type, even after controlling for age. It is plausible that there is more collagen in normal skin adjacent to DN than in that adjacent to MM because melanoma growth is not only associated with malignant growth of cancer cells, but also changes in its stroma microenvironment to support metastasis.[35] Paidi et al.[36] discovered that the use of RS is feasible to detect changes in the stroma of the lung microenvironment in response to primary breast tumors. Sahu et al.[37] found that early malignancy-associated changes in normal contralateral sites of oral cancer may lead to anatomical variability and cause misclassification between contralateral and tumor. Boppart et al.[38] raised the question that molecular surgical margin may be a better way to define tumor boundary than the "gold standard" structural tumor margin. Further studies are needed to study changes in normal stroma in response to dysplastic progression.

One limitation of this study is that it simplifies the model to only eight Raman active components. Although originally we had 15 components, we narrowed down to eight to avoid collinearity issues.[10] We found that including multiple chemically

Table 5 Patient age for MM and DN.

MM patient #	Age	DN patient #	Age
1	70	1	75
2	58	2	69
3	81	3	31
4	—	4	34
5	69	5	29
6	33	6	69
7	68	7	34
8	70	8	28
9	60	9	35
10	78	10	22
		11	34
Average age	**65**		**42**

similar components (e.g., various proteins) would result in fitting results with high variance. However, as there are far more molecules in skin, this method may underestimate the contribution of other molecules to the Raman signal. Another limitation is the limited sample size, which is also the main reason that we used leave-one-out cross validation to compute the ROC AUC. It is worth mentioning that this method comes with the risk of over-optimism. This may be the cause for the discrimination of MM from DN being better than that of (MM, DN) from normal (Fig. 2). Alternative methods include (1) k-fold cross validation (such as $k = 10$) and (2) bootstrapping.[39] The former utilizes 10% of the data as a test set, and the other 90% as the training set. Although it avoids the caveat of using single observation to estimate the model performance in each split of the data, it requires a larger sample size. The latter approach may provide a better estimate of internal validity.[40]

In conclusion, we have demonstrated that the biophysical model has consistent diagnostic capability as our previously published statistical model. By comparing with the statistical model, we have demonstrated that the biophysical model captures the spectral variances between skin pathologies in four distinct classification tasks. More importantly, the biophysical model captures the relevant biophysical changes accounting for the diagnosis. In particular, we found that collagen and triolein were the most important biomarkers in discriminating MM from benign pigmented lesions, and NMSCs and precancers from surround normal skin. Our work demonstrated that RS has great potential in diagnosing skin cancer noninvasively while extracting the skin's biophysical composition. Our future applications involve applying the biophysical Raman model to an ongoing, larger clinical skin cancer screening study and tumor margin detection in Mohs micrographic surgery.

Disclosures

All authors declare no conflict of interests for this paper and have no financial interest in the materials used in the paper.

Acknowledgments

References

1. I. Pence and A. Mahadevan-Jansen, "Clinical instrumentation and applications of Raman spectroscopy," *Chem. Soc. Rev.* **45**(7), 1958–1979 (2016).
2. M. Gniadecka et al., "Melanoma diagnosis by Raman spectroscopy and neural networks: structure alterations in proteins and lipids in intact cancer tissue," *J. Invest. Dermatol.* **122**(2), 443–449 (2004).
3. H. Lui et al., "Real-time Raman spectroscopy for in vivo skin cancer diagnosis," *Cancer Res.* **72**(10), 2491–2500 (2012).
4. L. Lim et al., "Clinical study of noninvasive in vivo melanoma and nonmelanoma skin cancers using multimodal spectral diagnosis," *J. Biomed. Opt.* **19**(11), 117003 (2014).
5. M. Sharma et al., "Design and characterization of a novel multimodal fiber-optic probe and spectroscopy system for skin cancer applications," *Rev. Sci. Instrum.* **85**(8), 083101 (2014).
6. J. Schleusener et al., "In vivo study for the discrimination of cancerous and normal skin using fibre probe-based Raman spectroscopy," *Exp. Dermatol.* **24**(10), 767–772 (2015).
7. J. Zhao et al., "Real-time Raman spectroscopy for automatic in vivo skin cancer detection: an independent validation," *Anal. Bioanal. Chem.* **407**(27), 8373–8379 (2015).
8. C. A. Lieber et al., "In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy," *Lasers Surg. Med.* **40**(7), 461–467 (2008).
9. F. L. Silveira et al., "Discrimination of non-melanoma skin lesions from non-tumor human skin tissues in vivo using Raman spectroscopy and multivariate statistics," *Lasers Surg. Med.* **47**(1), 6–16 (2015).
10. X. Feng et al., "Raman active components of skin cancer," *Biomed. Opt. Express* **8**(6), 2835–2850 (2017).
11. J. T. Motz et al., "Optical fiber probe for biomedical Raman spectroscopy," *Appl. Opt.* **43**(3), 542–554 (2004).
12. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Appl. Spectrosc.* **57**(11), 1363–1367 (2003).
13. N. Rajaram et al., "Pilot clinical study for quantitative spectral diagnosis of non-melanoma skin cancer," *Lasers Surg. Med.* **42**(10), 876–887 (2010).
14. G. Pezzotti et al., "Raman spectroscopy of human skin: looking for a quantitative algorithm to reliably estimate human age," *J. Biomed. Opt.* **20**(6), 065008 (2015).
15. P. Caspers et al., "In vitro and in vivo Raman spectroscopy of human skin," *Biospectroscopy* **4**, S31–S40 (1998).
16. B. Barry, H. Edwards, and A. Williams, "Fourier transform Raman and infrared vibrational study of human skin: assignment of spectral bands," *J. Raman Spectrosc.* **23**(11), 641–645 (1992).
17. X. Robin et al., "pROC: an open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinf.* **12**(1), 77 (2011).
18. A. Pappas, "Epidermal surface lipids," *Dermato-Endocrinol.* **1**(2), 72–76 (2009).
19. S. Duraipandian et al., "Simultaneous fingerprint and high-wavenumber confocal Raman spectroscopy enhances early detection of cervical precancer in vivo," *Anal. Chem.* **84**(14), 5913–5919 (2012).
20. S. K. Teh et al., "Near-infrared Raman spectroscopy for gastric precancer diagnosis," *J. Raman Spectrosc.* **40**(8), 908–914 (2009).
21. L. Silveira et al., "Discriminating model for diagnosis of basal cell carcinoma and melanoma in vitro based on the Raman spectra of selected biochemicals," *J. Biomed. Opt.* **17**(7), 077003 (2012).
22. N. Stone et al., "The use of Raman spectroscopy to provide an estimation of the gross biochemistry associated with urological pathologies," *Anal. Bioanal. Chem.* **387**(5), 1657–1668 (2007).
23. M. S. Bergholt et al., "Characterizing variability in in vivo Raman spectra of different anatomical locations in the upper gastrointestinal tract toward cancer detection," *J. Biomed. Opt.* **16**(3), 037003 (2011).
24. A. Daniel, A. Prakasarao, and S. Ganesan, "Near-infrared Raman spectroscopy for estimating biochemical changes associated with different pathological conditions of cervix," *Spectrochim. Acta, Part A* **190**, 409–416 (2018).
25. S. Zucker et al., "Purification and characterization of a connective-tissue-degrading metalloproteinase from the cytosol of metastatic melanoma cells," *Biochem. J.* **245**(2), 429–437 (1987).
26. T. Quan et al., "Matrix-degrading metalloproteinases in photoaging," *J. Investig. Dermatol. Symp. Proc.* **14**, 20–24 (2009).
27. F. M. Watt and H. Fujiwara, "Cell-extracellular matrix interactions in normal and diseased skin," *Cold Spring Harbor Perspect. Biol.* **3**(4), a005124 (2011).
28. D. R. English, C. Del Mar, and R. C. Burton, "Factors influencing the number needed to excise: excision rates of pigmented lesions by general practitioners," *Med. J. Aust.* **180**(1), 16–19 (2004).
29. B. Bodanese et al., "Discrimination of basal cell carcinoma and melanoma from normal skin biopsies in vitro through Raman spectroscopy and principal component analysis," *Photomed. Laser Surg.* **30**(7), 381–387 (2012).
30. A. S. Haka et al., "Diagnosing breast cancer by using Raman spectroscopy," *Proc. Natl. Acad. Sci. U. S. A.* **102**(35), 12371–12376 (2005).
31. E. C. Naylor, R. E. Watson, and M. J. Sherratt, "Molecular aspects of skin ageing," *Maturitas* **69**(3), 249–256 (2011).

32. V. R. Yanofsky, S. E. Mercer, and R. G. Phelps, "Histopathological variants of cutaneous squamous cell carcinoma: a review," *J. Skin Cancer* **2011**, 210813 (2011).
33. W. Fu and C. J. Cockerell, "The actinic (solar) keratosis: a 21st-century perspective," *Arch. Dermatol.* **139**(1), 66–70 (2003).
34. MathWorks, "Generalized linear mixed-effects models," https://www.mathworks.com/help/stats/generalized-linear-mixed-effects-models.html (2017).
35. V. Mundra, W. Li, and R. I. Mahato, "Nanoparticle-mediated drug delivery for treating melanoma," *Nanomedicine* **10**(16), 2613–2633 (2015).
36. S. K. Paidi et al., "Label-free Raman spectroscopy detects stromal adaptations in premetastatic lungs primed by breast cancer," *Cancer Res.* **77**(2), 247–256 (2017).
37. A. Sahu et al., "In vivo subsite classification and diagnosis of oral cancers using Raman spectroscopy," *J. Innov. Opt. Health Sci.* **9**(5), 1650017 (2016).
38. S. A. Boppart et al., "Label-free optical imaging technologies for rapid translation and use during intraoperative surgical and tumor margin assessment," *J. Biomed. Opt.* **23**(2), 1–10 (2017).
39. F. E. Harrell, Jr., *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, Springer, New York (2015).
40. E. W. Steyerberg et al., "Internal validation of predictive models: efficiency of some procedures for logistic regression analysis," *J. Clin. Epidemiol.* **54**(8), 774–781 (2001).

Xu Feng received her BS degree in optical engineering from Wuhan University, and her MS degree from Tsinghua University, Beijing. She is a PhD candidate at the University of Texas at Austin working under professor James W. Tunnell at the Biophotonics Laboratory. Her current research focuses on developing optical imaging and spectroscopy modalities and translating them into the clinic for skin cancer diagnosis and treatment.

Austin J. Moy received his MS and PhD degrees in biomedical engineering from the University of California, Irvine, and his BS degree in engineering physics from Cornell University. He is an application scientist at Trifoil Imaging. He was previously a postdoctoral fellow at the University of Texas at Austin. His research interests include biomedical applications of optical imaging, spectroscopy, and tomography.

Hieu T. M. Nguyen received her BS degree in materials science and engineering from Nanyang Technological University in Singapore, in 2011. She is a PhD student in biomedical engineering at the University of Texas at Austin. Currently, she is working on research areas include tissue optics, photothermal therapy, and immunology.

Yao Zhang received her BS degree in biomedical engineering from Shanghai Jiao Tong University in 2015 and then started her PhD journey in computational biomedical optics. She is a PhD candidate in biomedical engineering at the University of Texas at Austin. Her current research interests include light transport model, machine learning, and data science, especially in the field healthcare.

Jason Zhang received his BS degree in biomedical engineering from the University of Texas at Austin in 2017. He is a graduate student at the Massachusetts Institute of Technology. Currently, he is pursuing a PhD in electrical engineering, working on the development and clinical translation of biomedical optical imaging technologies.

Matthew C. Fox received his MD degree from the University of Texas Medical School at Houston in 2007. He is an assistant professor of medicine and chief of the Division of Dermatology at the University of Texas at Austin. He completed his residency training in dermatology and fellowship in micrographic surgery and dermatologic oncology at the University of Michigan. His clinical interests include the management of skin cancer, including rare and high risk tumors, and Mohs micrographic surgery.

Katherine R. Sebastian received her BA degree in biology from Grinnell College in 1998 and her MSPH degree from the University of Utah in 2004. She is a clinical research coordinator and RN with the Seton Healthcare Family in Austin, Texas. After becoming an RN in 2015, she joined the Seton Healthcare Family as a research coordinator for dermatology.

Jason S. Reichenberg is an associate professor in Dermatology at Dell Medical School at the University of Texas at Austin. He is a graduate of Boston University School of Medicine and completed his residency program at the Mayo Clinic. He is the author of more than 50 journal papers and has coedited a textbook and several book chapters. His current research interests include noninvasive skin imaging and the interplay between the nerves and the skin.

Mia K. Markey is a professor of biomedical engineering at the University of Texas at Austin and adjunct professor of imaging physics at the University of Texas MD Anderson Cancer Center. She is a fellow of both the American Association for the Advancement of Science (AAAS) and American Institute for Medical and Biological Engineering (AIMBE), and is a senior member of both the IEEE and SPIE.

James W. Tunnell is the Roberta Woods Ray Centennial Fellow and associate professor in the Department of Biomedical Engineering at the University of Texas at Austin. He has served as an associate editor for the Annals of Biomedical Engineering and on the program committees for CLEO, OSA, ASLMS, and IEEE-Photonics Society as well as several board positions for OSA and ASLMS. He has a specific research focus on using optical spectroscopy and imaging for disease diagnosis and treatment.