Biomedical Optics

SPIEDigitalLibrary.org/jbo

Micro-Raman spectroscopy study of cancerous and normal nasopharyngeal tissues

Yongzeng Li
Jianji Pan
Guannan Chen
Chao Li
Shaojun Lin
Yonghong Shao
Shangyuan Feng
Zufang Huang
Shusen Xie
Haishan Zeng
Rong Chen



Micro-Raman spectroscopy study of cancerous and normal nasopharyngeal tissues

Yongzeng Li, a* Jianji Pan, b* Guannan Chen, a* Chao Li, b Shaojun Lin, b Yonghong Shao, c Shangyuan Feng, a Zufang Huang, a Shusen Xie, a Haishan Zeng, d and Rong Chen a* Fujian Normal University, Key Laboratory of OptoElectronic Science and Technology for Medicine, Ministry of Education and Fujian Provincial

Key Laboratory for Photonics Technology, Fuzhou 350007, China

^bFujian Provincial Tumor Hospital, Fuzhou 350001, China

Abstract. The capabilities of micro-Raman spectroscopy for differentiating normal and malignant nasopharyngeal tissues were evaluated. Raman scattering signals were acquired from 22 normal and 52 malignant nasopharyngeal tissue samples. Distinctive spectral differences in Raman spectra between normal and malignant nasopharyngeal tissues were found, particularly in the spectral ranges of 853, 937, 1094, 1209, 1268, 1290 to 1340, 1579, and 1660 cm⁻¹, which primarily contain signals related to proteins, DNA, and lipids. Compared to normal tissues, the band intensity located at 853, and 937 cm⁻¹ were significantly lower for cancerous tissues (p < 0.05), while the band intensity located at 1094, 1209, 1268, and 1579 cm⁻¹ were significantly higher (p < 0.05). The band intensity located at 1290 to 1340, and 1660 cm⁻¹ were also higher for cancerous tissues; but the differences were not statistically significant (p > 0.05). Principal component analysis (PCA) and linear discriminate analysis (LDA) were employed to generate diagnostic algorithms for classification of Raman spectra of the two nasopharyngeal tissue types. The PCA-LDA algorithms together with leave-one-out, crossvalidation technique yielded diagnostic sensitivity of 92% and specificity of 82%. This work demonstrated that the Raman spectroscopy technique associated with PCA-LDA diagnostic algorithms has potential for improving the diagnosis of nasopharyngeal cancers. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1 .JBO.18.2.027003]

Keywords: Raman micro-spectroscopy; nasopharyngeal carcinoma; principal component analysis; linear discriminate analysis. Paper 12510 received Aug. 9, 2012; revised manuscript received Dec. 19, 2012; accepted for publication Dec. 20, 2012; published online Feb. 1, 2013.

1 Introduction

Nasopharyngeal malignancies remain one of the major causes of cancer-associated death and have high incidence rates in East Asia, particularly in South China. Presently, routine whitelight endoscopy and histological examination are the primary methods for clinical identification of nasopharyngeal carcinoma (NPC). Due to the anatomical position of NPC, it is not easy to diagnose this lesion early. Only about 10% and 20% of the diagnosed lesions are stage I and II diseases. Almost two thirds of the diagnosed NPCs are at an advanced stage.^{2,3} Due to the fact that the therapeutic outcome is highly related to the stage of the disease, early identification of malignant NPC is crucial for improving the survival rate of patients.4

Recent developments of spectral techniques may significantly expand our ability to diagnose this tumor rapidly and accurately. Optical spectroscopic techniques such as fluorescence spectroscopy, 5-7 light scattering spectroscopy, and Raman spectroscopy⁸ have been investigated for the evaluation

Address all correspondence to: Haishan Zeng, Imaging Unit - Integrative Oncology Department, British Columbia Cancer Agency Research Centre, Vancouver, BC, Canada V5Z 1L3. Tel: +1 604 675-8083; Fax: +1 604 675-8099; E-mail: hzeng@bccrc.ca

of malignancies in tissues. Although fluorescence spectroscopy has shown its promising diagnostic potential for in vivo detection of preneoplastic and early neoplastic lesions with highdetection sensitivities, it still suffers from moderate diagnostic specificities owing to the inter-observer dependence and the lack of ability to reveal specific biomolecular information about the tissue. Among the optical approaches currently under investigation for in vivo endoscopic applications, Raman spectroscopy is a very promising technique.

Raman spectroscopy is a nondestructive, inelastic light scattering technique in which the scattered photon is shifted to another wavelength with respect to the incident excitation light, depending on the specific vibrational modes of molecules in tissue and cells. Thus, Raman spectroscopy reveals specific biochemical information and biomolecular structures of tissue, providing the unique opportunity to distinguish between different pathological tissue types at the molecular level. Over the last two decades, studies using Raman spectroscopy have been conducted on samples from a variety of organs, including brain, breast, colon, cervix, gastric tissue, heart, liver, lung, lymph system, prostate, skin and thyroid, etc. 9-16 The results of all these studies demonstrated that normal and malignant tissues can be differentiated on the order of 80% to 100% accuracy with the use of various statistical analyses.

cShenzhen University, Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, Shenzhen 518060,

^dBritish Columbia Cancer Agency Research Centre, Imaging Unit—Integrative Oncology Department, 675 West 10th Avenue, Vancouver, BC, V5Z 1L3, Canada

^{*}These authors contributed equally to this work.

Very few reports have focused on Raman spectroscopy of NPC. In 2003, we reported a preliminary study of excised NPC and normal tissues from 6 patients using a fiber optic Raman system.8 That time our system was not optimized and it covers only a narrow spectral window of 950 to 1650 cm⁻¹. The signal to noise ratio (S/N) of the spectra was not very good. To the best of our knowledge, there are no other publications on Raman spectroscopy of nasopharyngeal tissues. This motivated us to conduct a more systematic study of NPC and normal nasopharyngeal tissues. This time we used a commercial micro-Raman system covering a broad spectral windows and with better spectral resolution and S/N. Furthermore, the micro-Raman system gets more of the tissue intrinsic Raman spectral properties independent of the tissue optics modifications to the spectra due to the confocal arrangement and the micron size tissue volume being measured. The previous fiber optic based Raman system⁸ interrogated a millimeter size tissue volume. Due to strong scattering of tissue to light, the Raman spectra are largely affected by tissue optics. 17 The fractional contributions of different tissue components to the total spectra vary due to tissue optics modifications. 17 This will result in significant spectral shape differences between the spectra obtained from the fiber optic system and the micro-Raman system. In this study, intrinsic (micro-) Raman spectral results from 74 excised tissue samples (52 NPC and 22 normal) will be discussed.

2 Materials and Methods

2.1 Subjects and Protocol

Seventy-four patients from the Fujian Provincial Tumor Hospital were recruited for this study with institutional ethical approval and informed consent. Spectra were obtained from 22 normal and 52 cancerous nasopharyngeal biopsies (50 nonkeratinizing undifferentiated and two differentiated carcinoma). The mean ages and standard deviations are 43 ± 13 for the normal group and 48 ± 13 for the cancer group. A recent study on the Raman spectral properties of healthy and diseased oral mucosa tissues demonstrated that "the aging related changes do not seem to have any bearing on classification of normal from abnormal conditions" (Ref. 18). We expect that the slight mismatch on ages (five years difference on mean values) between the normal and cancer groups in our study will not affect the diagnostic performances. 41 cancer samples are from male and 11 from female. 14 normal samples are from male and 8 from female. Fresh tissue samples were stored in a -80°C refrigerator until spectral measurements. The samples were placed on pure aluminum plate and thawed at room temperature immediately before experimental measurements.

2.2 Raman System

The targets were directly placed on a pure aluminum plate and all the Raman spectra of the tissue samples were recorded using a Renishaw inVia micro-Raman system with a $50 \times$ objective. A 785 nm diode laser was used for excitation and Raman spectra were recorded from 800 to 1800 cm⁻¹ with an integration time of 30 s. The excitation laser power focused on the targets was lower than 30 mW. There was no notable tissue damage during the measurements.

2.3 Data Processing and Analysis

A fifth order polynomial was fitted to the background tissue autofluorescence for each set of raw data.¹⁹ The polynomial was then subtracted from the measured spectrum to obtain the Raman signal. Each Raman spectrum was then normalized to the integrated area under the curve to correct variations in absolute spectral intensity and enable comparison of spectral shapes. To test the capability of tissue Raman spectra for differentiating cancer from normal tissue, principal component analysis (PCA) combined with linear discriminate analysis (LDA) was performed on the measured spectra. PCA is a statistical technique for simplifying complex data sets and determining the key variables in a multidimensional data set that can best explain the differences in the observations. To reduce the dimension of the spectral data, PCA is usually employed to extract a set of orthogonal principal components (PCs) that account for the maximum variance in the dataset for further diagnosis and characterization. In this study, the SPSS software package (SPSS Inc., Chicago) was used for PCA analysis. Subsequently, the PC scores were used as input for LDA to generate diagnostic algorithms. The performance of the PCA-LDA diagnostic algorithm was validated in an unbiased manner using leaveone-out, cross-validation methodology. In the validation procedure, one tissue sample was left out and the PCA-LDA modeling was redeveloped using the remaining Raman spectra. The redeveloped PCA-LDA diagnostic algorithm was then used to classify the withheld Raman spectra. This process was repeated iteratively until all withheld Raman spectra were classified.

3 Results

3.1 *Mean Spectra of Normal and Cancerous Nasopharyngeal Tissues*

Figure 1(a) shows the mean Raman spectra of normal and malignant tumor tissues in the nasopharynx which were measured under the same instrumentation set-up. Prominent Raman bands are observed in both normal and nasopharyngeal tissues at the following peak positions: 853 cm⁻¹ [ν (C–C) of proline], 937 cm⁻¹ [ν (C–C) in α -helix conformation of proline and valine], 1004 cm^{-1} [$\nu_s(\text{C--C})$ breathing of phenylalanine], 1094, 1209, 1268, 1304 cm⁻¹ [δ (CH₂) of lipids and proteins], 1340 cm⁻¹ (CH₃CH₂ wagging mode of proteins and nucleic acids), 1450 cm⁻¹ [δ (CH₂) of phospholipids and collagen], 1660 cm⁻¹ [ν (C=C) of lipid], etc. The normalized intensities of Raman peaks at 1094, 1209, 1268, 1290-1340, 1579, and 1660 cm⁻¹ are more intense for NPC tissue than for normal tissue, while Raman bands at 853 and 937 cm⁻¹ are greater in normal nasopharyngeal samples. Independent-sample t-test showed that the differences are significant for the 853, 937, 1094, 1209, 1268, and 1579 cm⁻¹ bands (p < 0.05), while the differences for the 1290 to 1340 and 1660 cm⁻¹ bands are not statistically significant (p > 0.05). Figure 1(b) shows the mean Raman intensity for the two different tissue groups along with corresponding standard deviations at the major Raman bands mentioned above. The six bands (853, 937, 1094, 1209, 1268, and 1579 cm⁻¹) with significantly (p < 0.05) mean intensity differences between the cancer group and the normal group are labeled with elipitical circles in Fig. 1(b). The band intensities located at 853 and 937 cm⁻¹ were significantly lower for cancerous tissues (p < 0.05), while the band intensities located at 1094, 1209, 1268, and 1579 cm⁻¹ were significantly higher

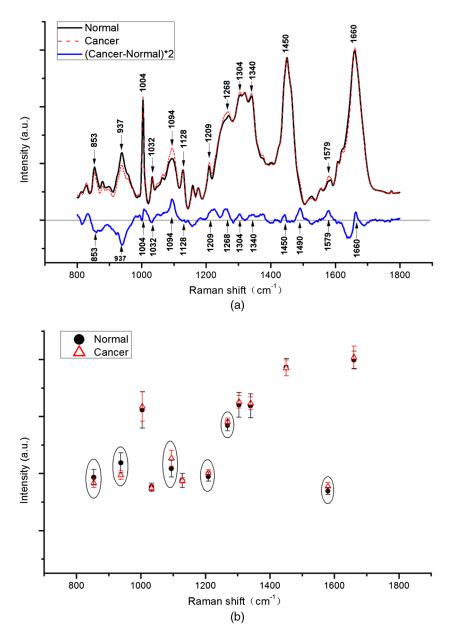


Fig. 1 (a) Mean Raman spectra of 22 normal and 52 cancerous nasopharyngeal tissues. (b) Plot of mean Raman intensity for the two different tissue groups along with corresponding standard deviations at the major Raman bands. The six bands (853, 937, 1094, 1209, 1268, and 1579 cm⁻¹) with significantly (p < 0.05) mean intensity differences are labeled with elipitical circles.

(p < 0.05). These intensity differences indicate that there is a significant increase and decrease in the percentage of distinctive biomolecules relative to the total Raman-active constituents in different tissue types, suggesting the diagnostic potential of Raman spectroscopy for identification of malignant lesions in the nasopharynx.

3.2 PCA and LDA

To test the capability of tissue Raman spectra for differentiating cancer from normal, PCA combined with LDA was performed on the measured Raman spectra. Independent-sample *t*-test on all the PC scores comparing normal and cancerous groups showed that there were three most diagnostically significant PCs (PC1, PC2, and PC4) for discriminating normal and cancerous groups. To illustrate the use of PC scores for diagnostic classification, a three-dimensional (3-D) scatter plot of using the

three PC scores (PC1, PC2, and PC4) as axes are presented in Fig. 2 for direct comparisons between normal (n=22) and cancer (n=52) groups. The data points are clustered into two distinct groups which further confirmed that substantial changes in spectral profiles in the process of malignant transformation from normal to carcinoma.

In order to incorporate all significant Raman spectral features, LDA was used to generate diagnostic algorithms using the PC scores for the three most significant PCs (PC1, PC2, and PC4). To prevent over-training, the leave-one-out and cross-validation procedures were used. Figure 3 shows the posterior probability of each spectrum belonging to the normal and nasopharyngeal cancer groups as calculated from the LDA model. Using a discrimination threshold of 0.5, the diagnostic sensitivity for detecting nasopharyngeal cancer was 92% and the corresponding diagnostic specificity was 82%.

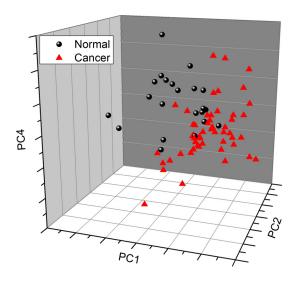


Fig. 2 3D Plot of the first principal component (PC1) versus the second principal component (PC2) and the forth principal component (PC4) for normal group versus nasopharyngeal cancer group.

To further evaluate the performance of the PCA-LDA-based diagnostic algorithm for nasopharyngeal cancer diagnosis, the receiver operating characteristic (ROC) curve was generated from the posterior probability plot in Fig. 3 by varying the threshold level. The results are shown in Fig. 4. The integration area under the ROC curves is 0.968. This further demonstrated that PCA-LDA-based diagnostic algorithms can be used for nasopharyngeal cancer detection with good performance.

4 Discussions

Over the last two decades, studies using Raman spectroscopy have been conducted on samples from a variety of organs and the results in general suggest that Raman spectroscopy has great potential for cancer detection based on the quantitative information about the biochemical differences between normal and neoplastic tissues in terms of proteins, DNA, and lipids. In this work, we investigated the *ex vivo* NIR micro-Raman

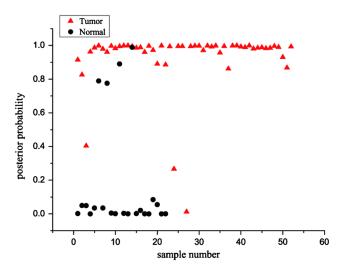


Fig. 3 Scatter plots of the posterior probability of belonging to the normal and tumor categories calculated from the data sets of Raman, using the PCA-LDA-based spectral classification with the leave-one-out, cross-validation method.

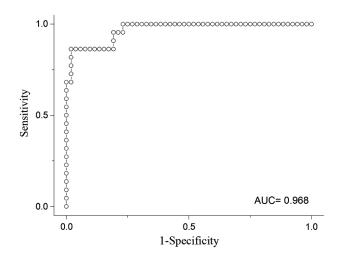


Fig. 4 Receiver operating characteristic (ROC) curve of discrimination results for Raman spectra utilizing the PCA-LDA-based spectral classification with leave-one-spectrum-out, cross-validation method.

spectral properties of normal and malignant lesions in the nasopharynx. Raman spectra of nasopharynx tissue are dominated by many vibrational modes of various biomolecules, such as proteins, lipids, and nucleic acids, which may be altered in quantity or confirmation associated with nasopharyngeal cancer. To better understand the molecular basis, Table 1 lists tentative assignments for the observed Raman bands.^{20–23} Distinctive spectral features and relative intensity differences were observed

Table 1 Tentative assignment of the main peaks from the Raman spectroscopy of human nasopharyngeal tissue.

Peak position/cm ⁻¹	Vibrational mode	Major assignments
853	ν(C C)	Proline ²⁰
937	$ u(\text{CC}) \text{ in } \alpha\text{-helix} $	Proline and valine ²⁰
1004	$\nu_{\rm s}({\rm C-\!C})$ breathing	Phenylalanine ²⁰
1032	$\delta(C-H)$	Phenylalanine ²⁰
1094	<i>ν</i> (C–N)	Lipids, DNA ²¹
1128	<i>ν</i> (C–N)	Proteins ²⁰
1209	$\nu(C-C_6H_5)$)	Tryptophan and Phenylalanine ²⁰
1268	$\delta(=C-H)$, $\delta(N-H)$	Phospholipids, ²² amide III ²¹
1304	$\delta(CH_2)$	Proteins/lipids ^{20,21}
1340	(CH ₃ CH ₂) wagging	Proteins and nucleic acids ^{20,21}
1450	$\delta(CH_2)$	Phospholipids and collagen ^{20,21}
1579	$\delta(C=C)$	Phenylalanine ^{21,23}
1660	<i>ν</i> (C=C)	Lipids ²¹

Note: ν , stretching mode; ν_s , symmetric stretching; δ , bending mode.

between cancer and normal groups, which reflect molecular and cellular changes associated with malignant transformation. For example, the Raman peak intensity at 937 cm⁻¹ due to the $\nu(C-C)$ in α -helix conformation of proline and valine appeared to be more intense for normal tissue than for NPC tissue. The peak at 1094 cm⁻¹, which is probably characteristic of C-N stretching of DNA, was higher for tumor tissue than for normal tissue, indicating that the cancerous tissue may be associated with an increase in the relative amounts of nucleic acids. The band at 1450 cm⁻¹ corresponds to the CH₂ bending mode of collagen has previously been recognized as being of diagnostic significance. The band at 1579 cm⁻¹ was attributed to C=C bending mode of phenylalanine, and the percentage signals were considerably increased in cancer tissue, indicating an increase in the percentage of phenylalanine content relative to the total components in nasopharyngeal cancer tissue. Huang et al.²⁰ also observed an increase of phenylalanine in malignant lung tissue by Raman spectroscopy.

Recently, our group has also carried out surface-enhanced Raman spectroscopy (SERS) studies of blood plasma samples from NPC patients and healthy volunteers.²³ Interestingly, the Raman bands at 1330 and 1579 cm⁻¹ on the SERS spectra showed the same variation trends as observed in this study. This suggests that tissue biochemical changes in the process of malignant transformation from normal to carcinoma seem reflected in the circulating blood. This, from a different perspective, supports the feasibility of using blood plasma SERS spectroscopy for noninvasive cancer detection.

Irrespective of the identification and quantization of the additional molecular species present in malignant nasopharyngeal spectra, the substantial changes in spectral profiles from normal to malignant can be effectively used for discrimination of the two classes of samples. In order to use the full information contained in Raman spectra for the classification, an accurate discriminate algorithm is required. PCA was performed to reduce the large amount of data contained in the measured Raman spectra into a few important principal components. Three most diagnostically significant PCs (PC1, PC2, and PC4) were extracted after independent-sample *t*-test on all the PC scores. Figure 5 shows the plots of these three PC

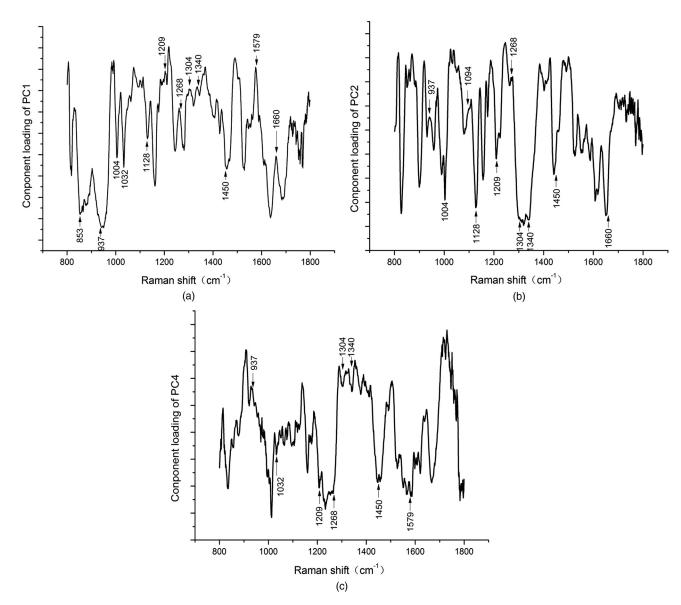


Fig. 5 Plot of the three most diagnostically significant PCs [(a) PC1, (b) PC2, and (c) PC4]. The Raman peaks listed in Table 1 are indicated on the graph.

loadings. The spectral shape of any of the three plots is not dominated by any single molecule in the tissue. Instead it contains many of the 13 Raman peaks listed in Table 1. For example, PC1 contains 12 of the 13 peaks. PC1 also contains other Raman peaks that are not obvious from the measured tissue spectra and thus not listed in Table 1. Similarly, PC2 contains 10 and PC4 contains 8 of the 13 peaks listed in Table 1. PC2 and PC4 also containes other nonlisted peaks. These results suggest multiple chemical origins of the PC loadings. LDA algorithm based on these three PCs shows that malignant nasopharyngeal tissue can be differentiated from normal tissue with a diagnostic sensitivity of 92.3% and specificity of 81.8%. It demonstrates good separations between normal nasopharyngeal tissue and diseased tissue.

5 Conclusions

Micro-Raman spectroscopy was performed on 22 normal and 52 malignant nasopharyngeal tissue samples. Distinctive spectral differences between normal and malignant nasopharyngeal tissue were revealed which primarily contain signals related to proteins, DNA, and lipids The band intensities located at 853 and 937 cm⁻¹ were significantly lower for cancerous tissues (p < 0.05), while the band intensities located at 1094, 1209, 1268, and 1579 cm⁻¹ were significantly higher (p < 0.05). An algorithm derived from PCA-LDA analysis of the spectral data yielded a diagnostic sensitivity of 92.3% and specificity of 81.8%. This suggests great potential for using Raman spectroscopy to improve the diagnosis of nasopharyngeal cancers. It provides solid support for us to develop endoscopic Raman instrumentations for *in vivo* clinical applications.

Acknowledgments

This work was supported by the Program for Changjiang Scholars and Innovative Research Team in University (IRT1115) and the National Natural Science Foundation of China (Nos. 60778046, 11104030, 61210016, and 81101110), the Project of Science Foundation of Ministry of Health and United Fujian Provincial Health and Education Project for Tackling the Key Research (No. WKJ2008-2-046), the Canadian Institutes of Health Research International Scientific Exchange Program, and the Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province.

References

- W. I. Wei and J. Sham, "Nasopharyngeal carcinoma," *Lancet.* 365(9476), 2041–2054 (2005).
- J. Ma, "Research on treatment for nasopharyngeal carcinoma," J. Sun Yat-sen Univ. (Medical Sciences) 31(2), 179–185 (2010).

- J. Sun and D. Li, "Advances in management of nasopharyngeal carcinoma," Med. Recapitulate 16(21), 3255–3259 (2010).
- J. Guigay, "Advances in nasopharyngeal carcinoma," Curr. Opin. Oncol. 20(3), 264–269 (2008).
- J. Y. Qu et al., "Preliminary study of *in vivo* autofluorescence of nasopharyngeal carcinoma and normal tissue," *Lasers Surg. Med.* 26(5), 432–440 (2000).
- C. Zhou et al., "Design of fluorescence imaging for localizator of early nasopharyngeal carcinoma," *J. Optoelectron. Laser.* 13(6), 639–642 (2002).
- H. Chang et al., "Light induced autofluorescence for detection of nasopharyngeal carcinoma in vivo," Proc. SPIE 4432, 186–195 (2001).
- D. P. Lau et al., "Raman spectroscopy for optical diagnosis in normal and cancerous tissue of the nasopharynx-preliminary findings," *Lasers* Surg. Med. 32(3), 210–214 (2003).
- R. R. Alfano et al., "Human breast tissues studied by IR Fourier transform Raman spectroscopy," *Lasers Life Sci.* 4(1), 23–28 (1991).
- C. J. Frank, R. I. McCreery, and D. C. B. Redd, "Raman spectroscopy of normal and diseased human breast tissues," *Anal. Chem.* 67(5), 777–783 (1995).
- A. Mizuno et al., "Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours," *J. Raman Spectrosc.* 25(1), 25–29 (1994).
- 12. P. J. Caspers et al., "In vitro and in vivo Raman spectroscopy of human skin," *Biospectroscopy* **4**(Suppl. 5), 31–39 (1998).
- H. Lui et al., "Real-time Raman spectroscopy for in vivo skin cancer diagnosis," Cancer Res. 72(10), 2491–2500 (2012).
- H. Yamazaki et al., "The diagnosis of lung cancer using 1064-nm excited near-infrared multichannel Raman spectroscopy," *Radiat. Med.* 21(1), 1–6 (2003).
- U. Utzinger et al., "Near-infrared Raman spectroscopy for in vivo detection of cervical precancers," Appl. Spectrosc. 55(8), 955–959 (2001)
- A. M. K. Enejder et al., "Blood analysis by Raman spectroscopy," *Opt. Lett.* 27(22), 2004–2006 (2002).
- 17. S. Wang et al., "Theoretical study on Raman properties of normal human skin using Monte Carlo method with a multi-layered tissue model," presented at SPIE Photonics West, Abstract#: 8565-22, San Francisco, California (Feb. 2–7 2013).
- A. Sahu et al., "Raman spectroscopy of oral buccal mucosa: a study on age-related physiological changes and tobacco-related pathological changes," *Technol. Cancer Res. Treat.* 11(6), 529–541 (2012).
- J. Zhao et al., "Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy," *Appl. Spectrosc.* 61(11), 1225–1232 (2007).
- Z. Huang et al., "Near-infrared Raman spectroscopy for optical diagnosis of lung cancer," *Int. J. Cancer* 107(6), 1047–1052 (2003).
- Z. Movasaghi, S. Rehman, and I. U. Rehman, "Raman spectroscopy of biological tissues," *Appl. Spectrosc. Rev.* 42(5), 493–541 (2007).
- M. V. P. Chowdary et al., "Biochemical correlation of Raman spectra of normal, benign and malignant breast tissues: a spectral deconvolution study," *Biopolymers* 91(7), 539–546 (2009).
- S. Feng et al., "Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis," *Biosens. Bioelectron.* 25(11), 2414–2419 (2010).