

Research Article

Raman Spectroscopy in Colorectal Cancer Diagnostics: Comparison of PCA-LDA and PLS-DA Models

Wenjing Liu,^{1,2} Zhaotian Sun,³ Jinyu Chen,⁴ and Chuanbo Jing³

¹State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³No. 4 Hospital, Jinan 250031, China

⁴National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

Correspondence should be addressed to Chuanbo Jing; chbjing@eyou.com

Received 13 August 2016; Revised 25 October 2016; Accepted 8 November 2016

Academic Editor: Feride Severcan

Copyright © 2016 Wenjing Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Raman spectra of human colorectal tissue samples were employed to diagnose colorectal cancer. High-quality Raman spectra were acquired from normal and cancerous colorectal tissues from 81 patients. Subtle Raman variations, such as for peaks at 1134 cm^{-1} (protein, C-C/C-N stretching) and 1297 cm^{-1} (lipid, C-H₂ twisting), were observed between normal and cancerous colorectal tissues. The average peak intensity at 1134 and 1297 cm^{-1} was increased from approximately 235 and 72 in the normal group, respectively, to 315 and 273 in the cancer group. The variations of Raman spectra reflected the changes of cell molecules during canceration. The multivariate statistical methods of principal component analysis-linear discriminant analysis (PCA-LDA) and partial least-squares-discriminant analysis (PLS-DA), together with leave-one-patient-out cross-validation, were employed to build the discrimination model. PCA-LDA was used to evaluate the capability of this approach for classifying colorectal cancer, resulting in a diagnostic accuracy of 79.2%. Further PLS-DA modeling yielded a diagnostic accuracy of 84.3% for colorectal cancer detection. Thus, the PLS-DA model is preferable between the two to discriminate cancerous from normal tissues. Our results demonstrate that Raman spectroscopy can be used with an optimized multivariate data analysis model as a sensitive diagnostic alternative to identify pathological changes in the colon at the molecular level.

1. Introduction

Colorectal cancer has high morbidity and mortality rates and is the third most commonly diagnosed cancer as well as the third leading cause of cancer death for both males and females in the United States [1]. Accurately detecting this cancer is a crucial and foremost step toward improving the survival rate of patients with colorectal cancer. Currently, colonoscopy and histopathology are standard screening and diagnostic techniques for colorectal tissues. Though colonoscopic screening has significantly increased the survival rate of patients with colorectal cancer, it remains a challenge to distinguish adenomas and early adenocarcinomas from benign hyperplastic polyps using colonoscopy [2]. This difficulty is due mainly to

the fact that conventional white light reflectance colonoscopy deeply relies on subjective visual assessment of colorectal polyps [3]. The gold standard for cancer diagnostics is histopathology, which is based on the visual investigation of tissue biopsies. A pathologist can diagnose a sample using specific staining, for example, with hematoxylin and eosin, to highlight the focus. Disadvantages of histopathology include time-consuming sample preparation and the subjectivity of pathologists [4]. It is of great necessity to develop an objective and sensitive technique that can assist clinicians in the differential diagnosis of benign and malignant cysts.

Raman spectroscopy, a vibrational analysis technique, is gaining popularity in cancer diagnostics. This technology investigates molecular vibrations that can be used for

functional group identification and compositional analysis. Extensive research has demonstrated that Raman spectroscopy can support gold-standard techniques and substantially improve clinical diagnostics [5–9]. Raman measurements on biopsies can help pathologists identify the tumor margins in a fast and precise way.

Raman spectroscopy has been employed to study human colorectal tissues *in vivo* or *ex vivo* to collect spectral information for cancer diagnosis. Since biochemical changes only lead to subtle changes in the Raman spectra, statistical methods are necessary to extract diagnostic information [10, 11]. Typical statistical methods can be categorized into supervised and unsupervised approaches. The unsupervised approach relies only on the Raman spectra to make a decision, whereas the supervised method uses additional information acquired by the gold-standard method. Principal component analysis (PCA), a frequently used unsupervised approach, reduces the number of variables and assesses the data as a first step. Following PCA, a supervised approach such as linear discriminant analysis (LDA), which takes advantage of PCA and the histopathological results, can classify tissues or cells [12–14]. A study on mice with colon cancer showed that the PCA-LDA model can correctly discriminate tumors from healthy tissues with an accuracy of 86.8% [15].

Another commonly used supervised approach, partial least-squares-discriminant analysis (PLS-DA), can provide additional group affinity information by classifying memberships as zeros and ones and thus can maximize the variations between groups of samples. PLS-DA rotates the latent variables (LVs) to achieve maximum group separation [16, 17]. Thus, the LVs consider the diagnostically relevant variations rather than the significant differences in the dataset. PLS-DA model has been employed to analyze colon tissues [3, 18]. In the previous study of Bergholt et al., the PLS-DA model was performed to diagnose the colorectal cancer with an accuracy of 88.8% [3].

Different models can result in different diagnostic performance when employed to analyze the same dataset. Thus, the use of a proper statistical model plays an important role in achieving diagnostic accuracy. The diagnostic performances obtained from different statistical models in terms of sensitivity and specificity were compared to find the optimal model. Based on this optimal model, the suitable diagnostic method was identified in the target tissue system. However, the relevant study on model comparison in the colorectal cancer diagnosis is lacking. This study evaluated and compared two statistical models for colorectal tissue classification. Our aim is to bridge the knowledge gap in identifying the appropriate model for Raman spectroscopy in cancer diagnosis.

Two multivariate statistical methods, PCA-LDA and PLS-DA, were used in combination with leave-one-patient-out cross-validation to establish the discrimination model. This work demonstrates that Raman spectroscopy is a prospective tool in the diagnosis of colorectal cancer during clinical examinations and that the PLS-DA model is superior in detecting spectral differences between normal and cancerous colorectal tissues.

2. Experimental Methods

2.1. Sample Preparation. The formalin-fixed, paraffin-embedded colorectal tissues were retrieved from the Jinan No. 4 Hospital in accordance with the regulations of its ethics committee. The Jinan No. 4 Hospital has approved this study. Normal regions were outside of the tumor areas in the tissue that was obtained during surgery. The paraffin-embedded tissues were sectioned into 10 μm thick sections. Each section was put on a glass microscope slide and stained with hematoxylin and eosin for histopathological diagnosis of the suspected area. The adjacent section was placed on a glass slide without being stained for Raman spectroscopy analysis [12, 19]. The histological analysis was conducted by professional medical doctors who are board certified pathologists.

2.2. Raman Spectroscopy. Raman spectra were acquired with a 10 s integration time in the spectral range of 400–4000 cm^{-1} using a Raman system (Horiba JY HR evolution, France) equipped with an Olympus BXFM open space optical microscope and a charge-coupled device (CCD) detector. A 532 nm laser was focused through a 100x objective ($\text{NA} = 0.9$, $\text{WD} = 0.21$ mm) to excite the samples. The laser power on the sample was about 1.33 mW. A 520.7 cm^{-1} band of silicon wafer was used for calibration. The spectral resolution was about 0.65 cm^{-1} , and the wavenumber accuracy was ± 0.03 cm^{-1} . The normal spectra were acquired from healthy regions outside the tumor areas in tissue.

2.3. Data Processing and Multivariate Data Analysis. A linear baseline correction was applied to the Raman spectra using Labspec6 software (Horiba JY). About 10 spectra were collected for each tissue and then averaged. Mean-centering was carried out prior to multivariate statistical analysis to remove common variance from the colorectal tissue Raman spectra dataset.

PCA-LDA and PLS-DA methods were applied for discriminant analysis. Leave-one-patient-out cross-validation was used to validate and optimize the PLS-DA model. Distinct molecular features of the colorectal tissues were extracted and visualized through loadings and scores. The statistical significance among the PCA/PLS scores for normal and cancerous tissues was calculated using a p value less than 0.05.

The PCA-LDA statistical analysis was performed using in-house written scripts. The PLS-DA statistical analysis was carried out with PLS toolbox (Eigenvector Research, Wenatchee, US). All statistical analyses were carried out in the Matlab programming environment (Mathworks Inc., Natick, US).

3. Results and Discussion

3.1. Raman Spectra of Colorectal Tissues. Figure 1(a) shows the averaged Raman spectra of normal ($n = 78$) and cancerous ($n = 81$) colorectal tissues, and the peak assignments are listed in Table S1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2016/1603609> [20–26].

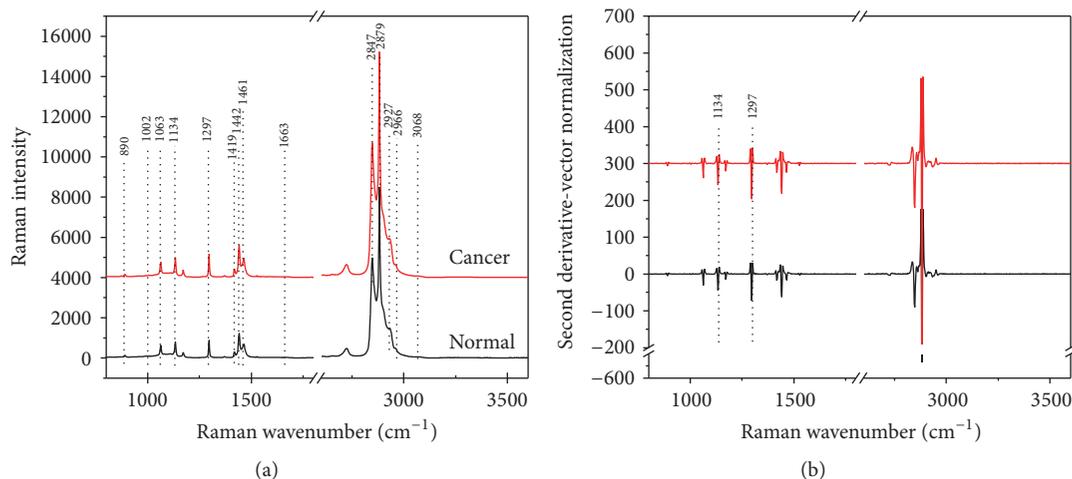


FIGURE 1: (a) Averaged Raman spectra for normal ($n = 78$) and cancerous ($n = 81$) colorectal tissues. (b) Second derivative-vector normalized spectra of Raman spectra from normal and cancerous tissues. The broken interval (//) indicates the region of $1800\text{--}2600\text{ cm}^{-1}$, which does not contain tissue-related biochemical information.

General Raman-active tissue components were comparable among colorectal tissues, and subtle variations, even though highly molecule-specific, were observed including peak position and intensity. Prominent Raman bands were observed for normal and cancerous colorectal tissue at about 1063 (lipids/collagen), 1134 (fatty acids and proteins), 1174 (L-tryptophan), 1297 (lipids and phospholipids), 1414 (lipids), 1442 (fatty acids and triglycerides), 1461 (lipids/proteins), 2847 (fatty acids and triglycerides), 2879 (lipids), and 2927 cm^{-1} (proteins and lipids). The difference between normal and cancerous tissues reflects the molecular changes in the tissue associated with the dysplastic progression (Figure 1(b)). For instance, the peak intensities at 1134 and 1297 cm^{-1} increased significantly in cancerous tissue, relative to the normal tissue, suggesting a higher amount of lipid material compared with the normal tissue (Figure S1). But the subtle variations were also hard to differentiate two types of tissues. This motivated further studies of PCA-LDA and PLS-DA to analyze the suitability of each in colorectal cancer diagnosis.

3.2. PCA-LDA Analysis of Raman Spectra. To reduce the dimension and complexity of the biological dataset, we performed PCA-LDA on normal and cancerous colorectal tissues in the spectral range of $400\text{--}4000\text{ cm}^{-1}$. PCA modeling is able to extract most fundamental features, resolving highly specific biomolecular information. Figure 2 shows that the first two PC components accounted for 98% (PC1, 82.7%; LV2, 15.3%) of the total Raman variations around the major Raman peak positions. These two PC components alone contributed to the most characteristic vibrational frequencies. They are dominated by the vibrational features of fatty acids, lipids, proteins, and nucleic acids from the colorectal tissues (Figure 3). The PC1 loading contained Raman peaks for fatty acids (1134 cm^{-1}); proteins (1174 cm^{-1} from L-tryptophan and 1461 cm^{-1} from C-H wagging); and lipids and fatty acids (1442 cm^{-1} from CH_2 or CH_3 deformations, 2847 cm^{-1} from

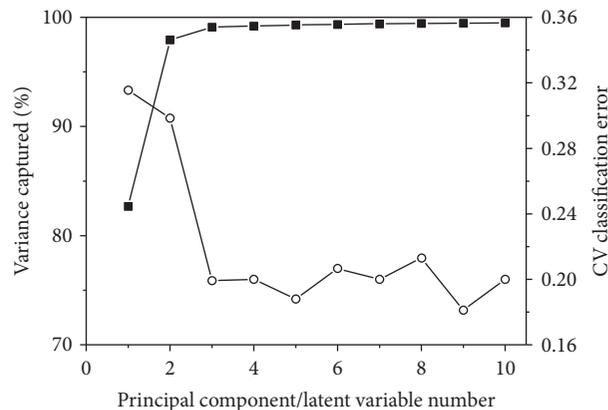


FIGURE 2: Variance captured percent and classification error as function of model complexity (i.e., retained number of PCs and LVs) using PCA-LDA and PLS-DA together with leave-one-patient-out cross-validation.

symmetric CH_2 stretching, and 2879 cm^{-1} from asymmetric CH_2 stretching). The loading on PC2 captured Raman peaks similar to those in PC1 loading, reflecting the main components of Raman spectra. To cross-validate the classification, we used the LDA model with the leave-one-patient-out approach for cross-validation. The PCA-LDA model resulted in a sensitivity of 72.8% and a specificity of 85.9%, which finally yielded a diagnostic accuracy of 79.2%.

3.3. PLS-DA Model for Predicting Cancer. The PLS-DA diagnostic model was performed with leave-one-patient-out cross-validation, to achieve a Raman spectral dataset using an optimum number of components. The optimum number of components was estimated using the local minimum of cross-validation classification error values and was determined to be 3 LVs (Figure 2). Figure 3 shows three significant LV loadings for the Raman spectral dataset. The loading

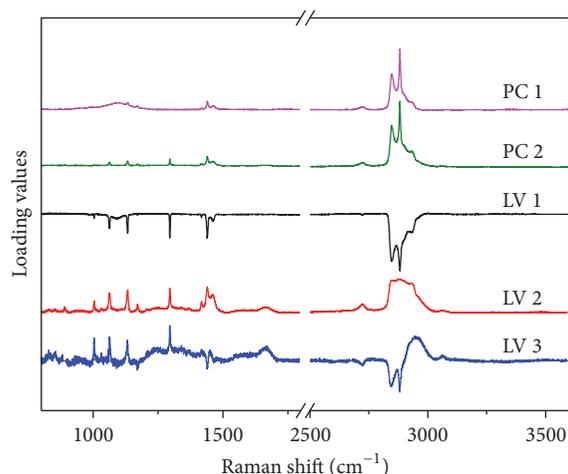


FIGURE 3: Loading plot of PCA components (PC) and PLS components (LV) in the each model calculated from the Raman spectra of colorectal tissues. Each loading is shifted vertically for better visualization. The broken interval (---) indicates the region of $\sim 2000\text{--}2500\text{ cm}^{-1}$, which does not contain much tissue-related biochemical information.

on LV1 contained the following specific Raman peaks from lipids: 1063 cm^{-1} (C-C stretching), 1297 cm^{-1} (CH_2 twisting), 1461 cm^{-1} (C-H wagging), and 2879 cm^{-1} (asymmetric CH_2 stretching). It also contained fatty acids, as evidenced by these peaks: 1442 cm^{-1} (CH_2 or CH_3 deformation), 2847 cm^{-1} (symmetric CH_2 stretching), and 2927 cm^{-1} (symmetric CH_3 stretching). The loading on LV2-captured Raman peaks, aside from lipids and fatty acids, was mainly associated with proteins, as evidenced by the peaks at 1002 cm^{-1} (C-C stretching, phenylalanine), 1174 cm^{-1} (L-Tryptophan), 1663 cm^{-1} (amide I α -helix, C=O stretching), 2927 cm^{-1} (symmetric CH_2 stretching), and 3068 cm^{-1} (nucleic acids/proteins, C-H aromatic vibration). Thus, Raman spectroscopy associated with PLS-DA modeling using 3 LVs provides highly specific signatures of various biomolecules, rendering a sensitivity of 77.7%, a specificity of 91.0%, and collectively a diagnostic accuracy of 84.3% (Table S2).

3.4. Comparison of PCA-LDA and PLS-DA. Figure 4 shows the box chart of significant PC and LV scores to visualize different degrees of diagnostic efficiency. The PCA scores show the classification comparisons between normal and cancerous tissues through PC1 and PC2. Compared with PCs and the other LVs, LV2 ($p = 1.16e - 7$) shows the greatest efficacy in distinguishing colorectal cancer. Analysis of the LV2 scores showed that increased protein and nuclear contents occurred during the neoplastic progression in the colon tissue, indicating the elevated number of cells associated with cancerous development. Channelling this increased biomolecule biosynthesis is absolutely required for tumorigenic transformation [27]. The cancerization could induce lipid and nucleic acid changes, which are reflected in the Raman spectra and further analysis. Human cancer cells express high levels of lipogenic enzymes to meet the great

demand for lipid synthesis [28]. Meanwhile, changes in the levels of nucleic acids were associated with tumor burden and malignant progression [29]. Receiver operating characteristic (ROC) curves (Figure 5) were also generated from spectral datasets to further evaluate the separation. The integrated area under the ROC curve of the PLS-DA model was 0.856, while the integrated area of the PCA-LDA model was 0.696, substantiating the efficiency of using the Raman technique with PLS-DA for diagnosing cancerous colorectal tissues.

Table S2 shows that the PLS-DA model combined with Raman spectroscopy had better diagnostic performance, compared with the PCA-LDA model. Previous studies reported that the PLS-DA model provided a diagnostic sensitivity of 90.9% and specificity of 83.3% for differentiating adenomas from hyperplastic polyps [18]. The PCA-LDA model could distinguish cancer from normal colon tissues in mice with diagnostic accuracy of 86.8% [15]. The diagnostic results from the literature indicated that the PLS-DA model resulted in better diagnostic accuracy than the PCA-LDA model, even in different tissue systems.

In the PLS-DA model, the diagnostic specificity (91.0%) in our study was higher than that (83.3%) in previous study, while the sensitivity (77.7%) from our study was lower than that (90.9%) from previous study. Compared with previous studies, the diagnostic accuracy (79.2%) from the PCA-LDA model in our study was lower than that (86.8%) in the mice colon tissue study. The different statistics may be attributed to the difference in Raman spectra collection and sample preparation. The diagnostic results of our study and previous studies can be acceptable for the clinical application.

PCA is a classical technique for dimensionality reduction. This method identifies several principal directions with a high variance, especially for high-dimensional data X with small number of samples and large number of features, such as Raman spectra. By projecting the original data of X onto these directions, much of the information of X will be maintained by just a small number of these new projected variables (i.e., LVs) [30]. However, PCA only involves one set of data. On the other hand, PLS realizes dimensionality reduction by considering the relations between two data blocks (X and Y) across the same samples. PLS maximizes the covariance between X and Y , which balances the requirement to explain as much variance as possible by considering the correlated relationships between X and Y [31]. Thus, the three LVs identified using PLS captured not only the high variance in the spectral dataset, but also the relationship between the spectral dataset and sample class. In this experiment, the accuracy of PLS-DA modeling (84.3%) was much higher than that of PCA-LDA (79.2%). Thus, out of the two, the PLS-DA model is preferable for discriminating cancerous from normal tissues following the strategies in this study.

4. Conclusions

In summary, Raman spectroscopy was applied as a sensitive diagnostic alternative for identifying pathologic changes (e.g., dysplasia) in colon tissue at the molecular level, using an optimized multivariate data analysis model. In a side-by-side comparison of PCA-LDA and PLS-DA with respect

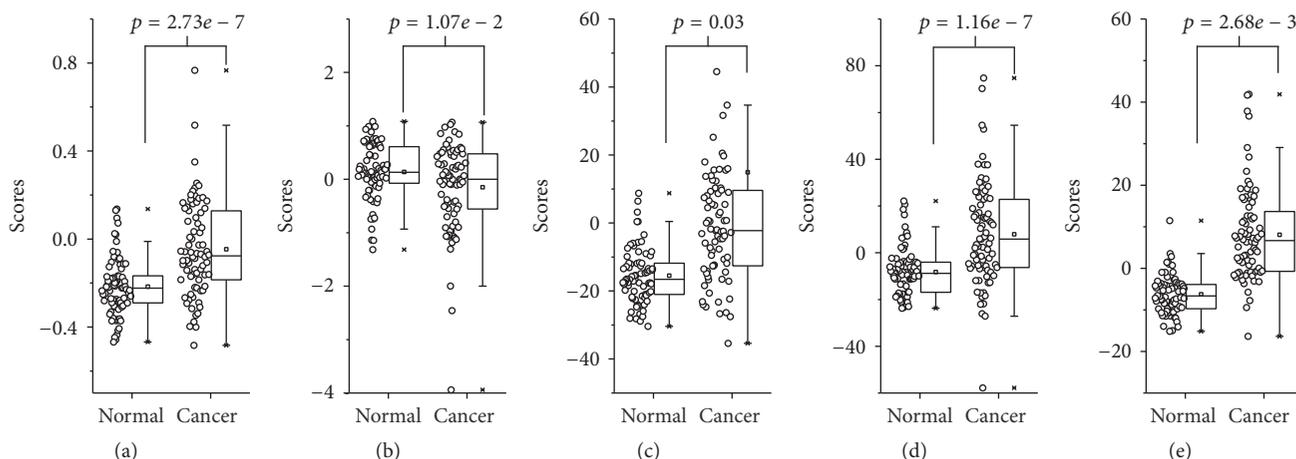


FIGURE 4: Box charts of the two PCA component (PC) scores and three significant PLS component (LV) scores calculated from the Raman dataset for normal and cancerous tissue types: (a) PC1 score, (b) PC2 score, (c) LV1 score, (d) LV2 score, and (e) LV3 score. The line within each box represents the median, while the lower and upper boundaries of the box indicate the first (25th percentile) and the third (75th percentile) quartiles, respectively. Error bars (whiskers) represent the 1.5-fold interquartile range.

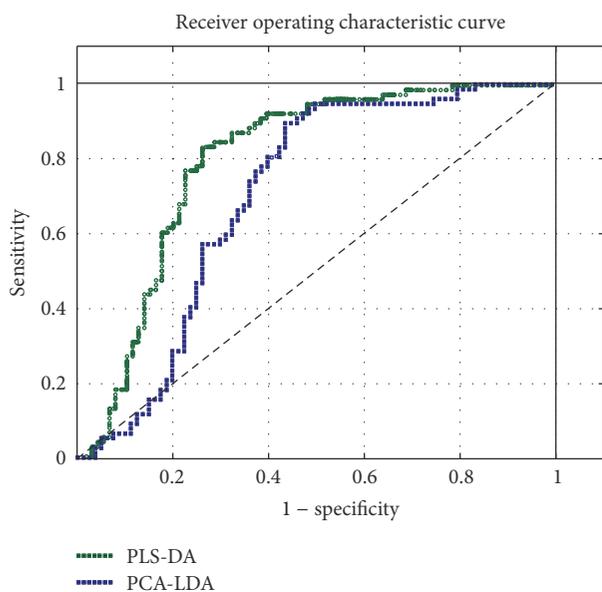


FIGURE 5: Receiver operating characteristic (ROC) curves of discrimination results from Raman spectra for colorectal tissue using Raman spectroscopy and PCA-LDA (blue) and PLS-DA (green) together with the leave-one-patient-out cross-validation method.

to the characterization of molecular profiles (e.g., proteins, lipids, and nucleic acids) of normal and cancerous colorectal tissues, the PLS-DA model was found to be a superior choice. The subtle Raman variations among normal and cancerous colorectal tissues are associated with cancerous tissue transformation. Confocal Raman spectroscopy is a promising technique in the diagnosis and characterization of colorectal cancer.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors acknowledge the financial support of the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB14020201).

References

- [1] R. Siegel, J. Ma, Z. Zou, and A. Jemal, "Cancer statistics, 2014," *CA Cancer Journal for Clinicians*, vol. 64, no. 1, pp. 9–29, 2014.
- [2] M. B. Wallace and R. Kiesslich, "Advances in endoscopic imaging of colorectal neoplasia," *Gastroenterology*, vol. 138, no. 6, pp. 2140–2150, 2010.
- [3] M. S. Bergholt, W. Zheng, K. Lin et al., "Characterizing variability of in vivo Raman spectroscopic properties of different anatomical sites of normal colorectal tissue towards cancer diagnosis at colonoscopy," *Analytical Chemistry*, vol. 87, no. 2, pp. 960–966, 2015.
- [4] K. Eberhardt, C. Stiebing, C. Matthaüs, M. Schmitt, and J. Popp, "Advantages and limitations of Raman spectroscopy for molecular diagnostics: an update," *Expert Review of Molecular Diagnostics*, vol. 15, no. 6, pp. 773–787, 2015.
- [5] M. S. Bergholt, W. Zheng, K. Y. Ho et al., "Fiberoptic confocal raman spectroscopy for real-time in vivo diagnosis of dysplasia in Barrett's esophagus," *Gastroenterology*, vol. 146, no. 1, pp. 27–32, 2014.
- [6] I. Taleb, G. Thiéfin, C. Gobinet et al., "Diagnosis of hepatocellular carcinoma in cirrhotic patients: a proof-of-concept study using serum micro-Raman spectroscopy," *Analyst*, vol. 138, no. 14, pp. 4006–4014, 2013.
- [7] S. Borel, E. A. Prikryl, N. H. Vuong et al., "Discrimination of normal and malignant mouse ovarian surface epithelial cells *in vitro* using Raman microspectroscopy," *Analytical Methods*, vol. 7, no. 22, pp. 9520–9528, 2015.
- [8] O. J. Old, L. M. Fullwood, R. Scott et al., "Vibrational spectroscopy for cancer diagnostics," *Analytical Methods*, vol. 6, no. 12, pp. 3901–3917, 2014.

- [9] A. Sahu, S. Tawde, V. Pai et al., "Raman spectroscopy and cytopathology of oral exfoliated cells for oral cancer diagnosis," *Analytical Methods*, vol. 7, no. 18, pp. 7548–7559, 2015.
- [10] C. Beleites, A. Bonifacio, D. Codrich, C. Krafft, and V. Sergo, "Raman spectroscopy and imaging: promising optical diagnostic tools in pediatrics," *Current Medicinal Chemistry*, vol. 20, no. 17, pp. 2176–2187, 2013.
- [11] K. H. Esbensen and P. Geladi, "Principles of proper validation: use and abuse of re-sampling for validation," *Journal of Chemometrics*, vol. 24, no. 3-4, pp. 168–187, 2010.
- [12] A. Salman, G. Sebbag, S. Argov, S. Mordechai, and R. K. Sahu, "Early detection of colorectal cancer relapse by infrared spectroscopy in 'normal' anastomosis tissue," *Journal of Biomedical Optics*, vol. 20, no. 7, Article ID 075007, 2015.
- [13] A. Synytsya, M. Judexova, D. Hoskovec, M. Miskovicova, and L. Petruzelka, "Raman spectroscopy at different excitation wavelengths (1064, 785 and 532 nm) as a tool for diagnosis of colon cancer," *Journal of Raman Spectroscopy*, vol. 45, no. 10, pp. 903–911, 2014.
- [14] A. Molckovsky, L.-M. Wong Kee Song, M. G. Shim, N. E. Marcon, and B. C. Wilson, "Diagnostic potential of near-infrared Raman spectroscopy in the colon: differentiating adenomatous from hyperplastic polyps," *Gastrointestinal Endoscopy*, vol. 57, no. 3, pp. 396–402, 2003.
- [15] A. Taketani, R. Hariyani, M. Ishigaki, B. B. Andriana, and H. Sato, "Raman endoscopy for the in situ investigation of advancing colorectal tumors in live model mice," *Analyst*, vol. 138, no. 14, pp. 4183–4190, 2013.
- [16] M. Hedegaard, C. Krafft, H. J. Ditzel, L. E. Johansen, S. Hassing, and J. Popp, "Discriminating isogenic cancer cells and identifying altered unsaturated fatty acid content as associated with metastasis status, using K-means clustering and partial least squares-discriminant analysis of raman maps," *Analytical Chemistry*, vol. 82, no. 7, pp. 2797–2802, 2010.
- [17] F. C. De Lucia Jr., J. L. Gottfried, C. A. Munson, and A. W. Miziolek, "Multivariate analysis of standoff laser-induced breakdown spectroscopy spectra for classification of explosive-containing residues," *Applied Optics*, vol. 47, no. 31, pp. G112–G122, 2008.
- [18] M. S. Bergholt, K. Lin, J. Wang et al., "Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy enhances real-time *in vivo* diagnosis of adenomatous polyps during colonoscopy," *Journal of Biophotonics*, vol. 9, no. 4, pp. 333–342, 2016.
- [19] S. Argov, J. Ramesh, A. Salman et al., "Diagnostic potential of Fourier-transform infrared microspectroscopy and advanced computational methods in colon cancer patients," *Journal of Biomedical Optics*, vol. 7, no. 2, pp. 248–254, 2002.
- [20] Q. Matthews, A. Jirasek, J. Lum, X. Duan, and A. G. Brolo, "Variability in raman spectra of single human tumor cells cultured *in vitro*: correlation with cell cycle and culture confluency," *Applied Spectroscopy*, vol. 64, no. 8, pp. 871–887, 2010.
- [21] L. Pantanowitz, J. H. Sinar, W. H. Henricks et al., "Validating whole slide imaging for diagnostic purposes in Pathology: guideline from the College of American pathologists Pathology and Laboratory Quality Center," *Archives of Pathology and Laboratory Medicine*, vol. 137, no. 12, pp. 1710–1722, 2013.
- [22] J. De Gelder, K. De Gussem, P. Vandenabeele, and L. Moens, "Reference database of Raman spectra of biological molecules," *Journal of Raman Spectroscopy*, vol. 38, no. 9, pp. 1133–1147, 2007.
- [23] N. Stone, C. Kendall, N. Shepherd, P. Crow, and H. Barr, "Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers," *Journal of Raman Spectroscopy*, vol. 33, no. 7, pp. 564–573, 2002.
- [24] I. Notingher, "Raman spectroscopy cell-based biosensors," *Sensors*, vol. 7, no. 8, pp. 1343–1358, 2007.
- [25] J. Kneipp, T. Bakker Schut, M. Kliffen, M. Menke-Pluijmers, and G. Puppels, "Characterization of breast duct epithelia: A Raman Spectroscopic Study," *Vibrational Spectroscopy*, vol. 32, no. 1, pp. 67–74, 2003.
- [26] C. Krafft, L. Neudert, T. Simat, and R. Salzer, "Near infrared Raman spectra of human brain lipids," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 61, no. 7, pp. 1529–1535, 2005.
- [27] R. A. Cairns, I. S. Harris, and T. W. Mak, "Regulation of cancer cell metabolism," *Nature Reviews Cancer*, vol. 11, no. 2, pp. 85–95, 2011.
- [28] J. A. Menendez and R. Lupu, "Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis," *Nature Reviews Cancer*, vol. 7, no. 10, pp. 763–777, 2007.
- [29] H. Schwarzenbach, D. S. B. Hoon, and K. Pantel, "Cell-free nucleic acids as biomarkers in cancer patients," *Nature Reviews Cancer*, vol. 11, no. 6, pp. 426–437, 2011.
- [30] E. B. Corrochano, *Handbook of Geometric Computing: Applications in Pattern Recognition, Computer Vision, Neural computing, and Robotics*, Springer Science & Business Media, Berlin, Germany, 2005.
- [31] R. Rosipal and N. Krämer, "Overview and recent advances in partial least squares," in *Subspace, Latent Structure and Feature Selection*, vol. 3940 of *Lecture Notes in Computer Science*, pp. 34–51, Springer, Berlin, Germany, 2006.

