

## CLINICAL STUDY

# Novel chemometrics-assisted spectroscopic methods for diagnosis and monitoring of invasive ductal carcinoma in breast tissue

Albayrak M<sup>1</sup>, Senol O<sup>2</sup>, Demirkaya-Miloglu F<sup>2</sup>, Calik M<sup>3</sup>, Kadioglu Y<sup>2</sup>

Department of Medical Laboratory Techniques, Health Services Vocational Training School, Ataturk University, Erzurum, Turkey. [m\\_albayrak25@hotmail.com](mailto:m_albayrak25@hotmail.com)

**ABSTRACT**

**OBJECTIVES:** Early diagnosis of breast cancer is extremely important because it is the most common female cancer and a leading cause of cancer death in adult women. In this study, it is aimed to create Raman mapping with developed chemometrics-assisted Raman and FT-IR spectroscopy methods for the diagnosis of invasive ductal carcinoma (IDC) in breast tissue samples.

**METHODS:** Samples were deparaffinized and 20-micron layers of each tissue were located on a coverslip. Mapping of both healthy and cancerous tissues were performed by exposing them to Raman laser at 532 and 758 nm while excitation was recorded at wavenumbers in range of 100–4,000 cm<sup>-1</sup>. Orthogonal partial least square (OPLS) algorithm was applied to evaluate obtained Raman spectra. Latent variable was selected to explain the whole model.

**RESULTS:** Healthy and IDC tissues were accurately and precisely clustered with Raman mapping and obtained results were compared to those obtained by means of histopathology and FT-IR methods. It is claimed that the proposed method has a great potential in clustering and separating IDC tissues from the healthy ones.

**CONCLUSION:** This novel, rapid, precise, easy and objective diagnosis method may be an alternative to conventional diagnostic methods for IDC in breast tissue (Fig. 5, Ref. 22). Text in PDF [www.elis.sk](http://www.elis.sk).

**KEY WORDS:** chemometrics, FT-IR spectroscopy, invasive ductal carcinoma, Raman spectroscopy, qualitative analysis.

**Introduction**

Breast cancer is the most commonly seen cancer type, affecting especially women at any age in our country and elsewhere around the world. Histological and biological features, prognosis and clinical responses of disease show differences depending on the type of breast cancer. Invasive ductal carcinoma (IDC) is the most frequent breast cancer, and about 80 % of these cases are diagnosed with IDC. The incidence of breast cancer is around 43 per 100,000 individuals in accordance with the statistical report of the health ministry in Turkey (1). Therefore, as with other types of cancer, breast cancer is a serious disease which has to be diagnosed early in order to improve breast cancer outcomes and survival. Current diagnostic methods of breast cancer include breast self-examination, clinical examination, and then mammography.

Mammography quantitatively determines the density changes in breast tissue and as such is considered to be a very sensitive method for detecting breast cancer at its early stage. Often however, the mammography findings are not specific. Breast biopsy follows in order to determine the features of the lesion. Histopathologic evaluation can take months because this method has a long experimental procedure and needs expertise to interpret the results. At the end of this process, 70–90 % of detected lesions are found to be benign (2) and thus the patient unnecessarily suffers both physical and psychological depression. This is the reason why scientists begin to research alternative methods to evaluate the breast tissues of patients.

The term of medical photonics is described as use of optical technology and instruments for the diagnosis and therapeutic and basic science application in medicine. Nowadays, medical photonic methods such as FTIR and Raman spectroscopy assisted with chemometrics became popular for the diagnosis of breast cancer as well as for other types of cancer. Chemical analysis can be performed with these non-invasive analytical methods by providing spectral differences observed as a function of tissue physiology or pathology (3). The obtained data can be analyzed with chemometric methods while various statistical methods and diagnostic algorithms are used in the process of confirming or disproving the cancer diagnosis. An important advantage of this

<sup>1</sup>Department of Medical Laboratory Techniques, Health Services Vocational Training School, Ataturk University, Erzurum, Turkey, <sup>2</sup>Department of Analytical Chemistry, Faculty of Pharmacy, Ataturk University, Erzurum, Turkey, and <sup>3</sup>Department of Pathology, Faculty of Medicine, Firat University, Elazig, Turkey

**Address for correspondence:** M. Albayrak, Department of Medical Laboratory Techniques, Health Services Vocational Training School, Ataturk University, 25240, Erzurum, Turkey.

diagnostic method lies in automation, which allows for objective and real-time diagnosis of pathologies.

Several papers were reported in literature for the diagnosis of IDC via FT-IR (4–7) and Raman (8–13) spectroscopy. In these studies, subtle changes have been detected in cellular biomolecule content or structure, including nucleic acids, proteins, carbohydrates, and other biomolecules that are closely tied to the carcinogenesis of cells and tissues. Unlike mentioned studies, the aim of the present study is to demonstrate that the Raman mapping has to provide unique chemical and morphological information about IDC tissue status and distinguish cancerous tissue from healthy tissue. In addition to this, reference histopathology and FTIR methods were also performed to compare the prediction power of the proposed method.

## Experiment

### Tissue samples

Invasive ductal carcinoma (IDC) and healthy tissue samples were collected from breasts of 23 women patients (age range of 32–47 years) who underwent transurethral resection or mamotome biopsy at the Department of Breast Surgery in Faculty of Medicine in Ataturk University. The obtained tissue sample sections were deep-frozen and then kept at  $-80^{\circ}\text{C}$  until use. The samples, which were taken from one part of the frozen sections, were prepared for histopathological analysis. The sections were examined by expert breast pathologists and classified as healthy or IDC tissue according to WHO classification (14). The other part of the frozen sections was used for analysis with Raman Spectroscopy.

### Histopathological analysis method

After placing the breast tissue samples from women in 10 % buffered formalin and staining them with hematoxylin and eosin, they were subjected to conventional histological techniques (15). The obtained samples were evaluated by using an electron microscope (Leica CV 5030). These obtained results were accepted as reference.

### Raman spectroscopic measurements and mapping

A WITec Alpha 300R Micro Confocal Raman Spectroscopy including two laser wavelengths (532 and 758 nm) by filtering the Raleigh scattering was used to analyze carcinoma and healthy tissue samples. The grating was 600 g/mm (blaze value of 750 nm). The analyzed spectral range was recorded in the range of  $100\text{--}4000\text{ cm}^{-1}$  to give a resolution of  $0.5\text{ }\mu\text{m}$ . Each Raman spectrum was recorded on the spectrometer using an acquisition period of 0.5 s while the width and height of the pixels of spectrums were set to be 1,024 and 128, respectively. Background spectrum was obtained and removed from sample spectrum automatically. Every line of 20 points were recorded and 50 images were taken for each line which included 20 points

### FTIR spectroscopic measurements

All of the carcinoma and healthy tissue samples were analyzed by FT-IR Bruker Vertex 70V Fourier transform infrared spectro-

scope equipped with attenuated total reflection (ATR). The spectra were obtained in wave number ranging from  $50\text{ cm}^{-1}$  to  $4000\text{ cm}^{-1}$ . Background spectrum was drawn to eliminate the atmospheric effect and removed from sample spectrum automatically.

### Diagnostic algorithm and software

The diagnostic algorithm (healthy or IDC tissue) was constructed by using orthogonal partial least squares algorithm (OPLS). This is aimed at PLS-Toolbox software version 8.1.1 for use with MATLAB R2015a. program was used. OPLS can be used in classification by interpreting output from class differences (17).

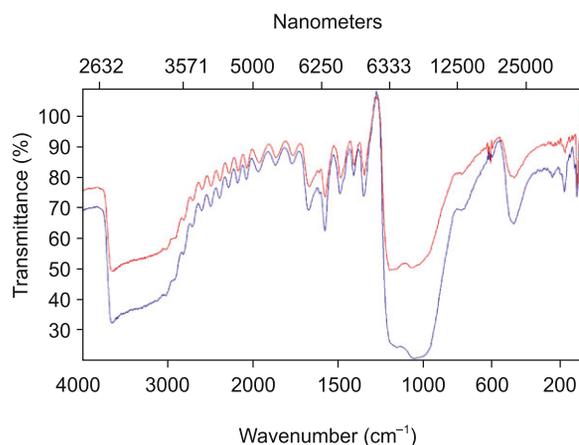
## Results and discussion

### Histopathological evaluations of IDC and healthy tissues

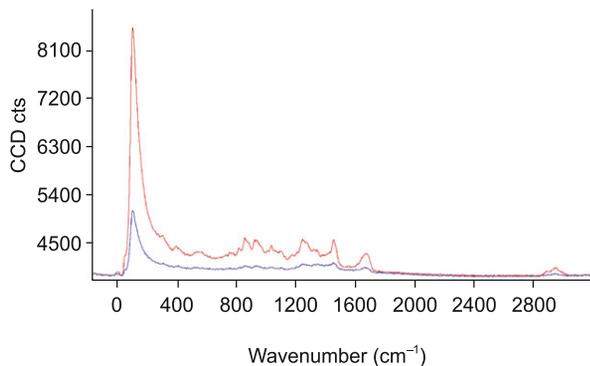
The sections were stained with hematoxylin and eosin and prepared as described in Section 2.2. The total runtime of the histopathological measurement was approximately 1 day. IDC was diagnosed with a basic immunohistochemical profile including the expression of estrogen receptor (ER), HER2 and progesterone receptor. A proportion of 80 % of the samples with IDC cancers were positive for estrogen receptor (ER) and 25 % were positive for HER2. (18). According to this evaluation, out of the total of 23 tissue samples, 12 and 11 were diagnosed as being healthy and positive for IDC, respectively. Histopathological evaluation was accepted to serve as reference for developed chemometrics-assisted Raman spectroscopy.

### FTIR spectral analysis of IDC and healthy tissues

Figure 1 shows the characteristic FTIR spectra of IDC and healthy tissues. According to the spectra, some important biomolecules (DNA, proteins, lipids and fatty acids) were identified similarly to results obtained from literature (6, 5) Bands in range of  $3,080\text{--}3,600\text{ cm}^{-1}$  expressed N-H symmetric stretching and N-H stretching vibrations of protein and nucleic acids. Bands in range of  $2,600\text{--}3,200\text{ cm}^{-1}$  expressed C-H stretching vibrations of lipid,



**Fig. 1. FT-IR spectra of IDC and healthy tissues: differences in the intensity of peaks in the spectra attributed to differences in chemical composition and structure between IDC and healthy tissues.**



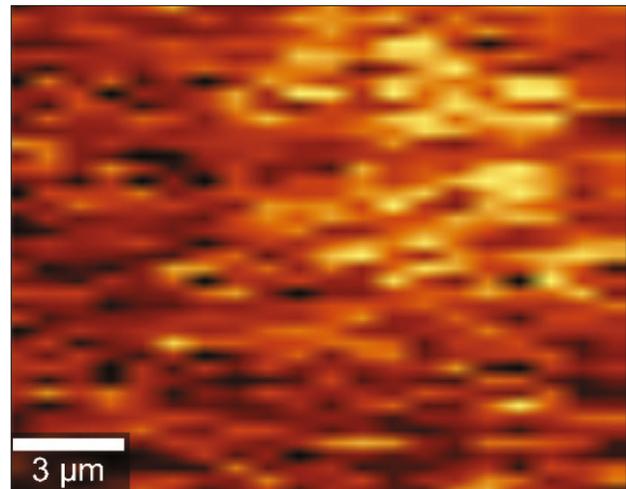
**Fig. 2. Raman spectra of IDC (blue) and healthy (red) tissues: differences in the intensity of peaks in the spectra attributed to alteration of chemical composition and structure in IDC compared to healthy tissues.**

protein and DNA. C=O stretching vibrations of proteins and lipid were mostly seen at  $1600\text{ cm}^{-1}$ . Collagen, lipids and protein amide formations were also detected in range of  $1,360\text{--}1,550\text{ cm}^{-1}$ . C-N-H stretching vibrations of protein amide, symmetric C-N-H stretching vibrations of collagen and lipids, and C-N-H stretching vibrations of amide III were seen at  $1,530\text{ cm}^{-1}$ ,  $1,330\text{ cm}^{-1}$  and  $1,260\text{ cm}^{-1}$ , respectively.  $\text{CH}_2$  bending associated with elastin, phospholipids and collagen, as well as symmetric  $\text{CH}_3$  bending associated with elastin, phospholipids and collagen were detected in range of  $1,380\text{--}1,400\text{ cm}^{-1}$ . C-O stretching vibrations of C-OH groups of tyrosine of proteins were confirmed at  $1,170\text{ cm}^{-1}$ . A band of O-P-O vibrations centered at  $1,010\text{--}1,080\text{ cm}^{-1}$  reveal the presence of polysaccharides and carbohydrates. Almost in all regions of FTIR spectrum, both IDC tissue and healthy tissue showed similarly significant absorptions. Nevertheless, there were differences between absolute and relative intensities of the absorption bands of the spectra. This situation confirmed the differences in chemical composition and structure of DNA, proteins and lipids in IDC tissue when compared to healthy tissues.

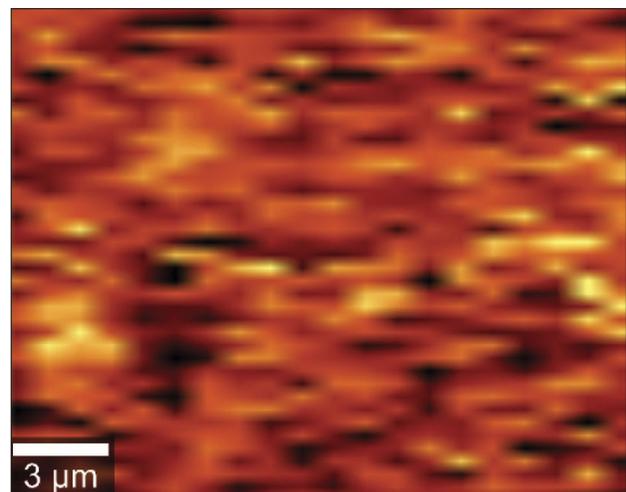
FTIR spectroscopy is a nondestructive technique and in literature it takes place as an alternative to histopathological evaluation for clinical diagnosis of cancerous tissue. In this study, this method was also accepted as a reference method for the developed Raman spectroscopy.

#### *Raman mapping of IDC and healthy tissues*

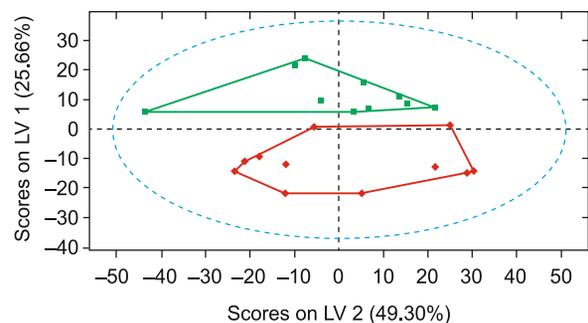
An intercepted Raman spectrum for healthy and cancerous tissue is shown in Figure 2. According to the spectra, several characteristic bands indicating specific biomolecules were detected. At  $854\text{ cm}^{-1}$ , a peak was observed and mostly related to tyrosine molecule in the ring breathing mode while that of  $876\text{ cm}^{-1}$  is characteristic for hydroxyproline also detected in Raman measurements (19). The band in range of  $1,004\text{--}1,033\text{ cm}^{-1}$  is probably a sign of C-C vibrations of phenylalanine (20). Phospholipids were mostly observed at  $1,130\text{ cm}^{-1}$  (21). The tyrosine band is also associated with  $1,390\text{ cm}^{-1}$  (22). Collagen, lipids and amide formations were linked to the range of  $1,100\text{--}1,400\text{ cm}^{-1}$ . This rank was also used



**Fig. 3. Raman image of healthy tissue.**



**Fig. 4. Raman image of cancerous tissue.**



**Fig. 5. Scatter plot of latent variable for healthy (green) and IDC (red) breast tissues regarding to OPLS-DA model.**

for imaging cancerous tissue vs healthy tissue to show the deformation of malignant cells (Figs 3 and 4). This Raman map shows the differences of collagen level in both tissue samples (Fig 3 and 4). The density of darker area in the cancerous tissue sample was quite well observable in Raman image.

All spectrum values were also discussed via chemometric algorithm for cluster analysis. In this calculation, OPLS-DA were preferred to classify each group. Each measurement was recorded and exported into MATLAB 2017 software. Noise and interferences were filtered by pre-processing the data matrix with OSC. OPLS-DA suggested 3 latent variables to explain the method. LV1 vs LV2 score plot is monitored in Figure 5. This graph shows the separation power of the proposed algorithm. Cross validation of the proposed method was performed by venetian blinds w/10 splits and 1 sample per split. Root mean square error of cross validation (RMSECV) values were lower than 0.09 while the regression coefficient of the proposed method is 0.94, which indicates a satisfactory goodness of fit and low bias for the proposed model.

## Conclusion

Latest achievements in biomedical science allow for incorporation of photonic strategies with the aid of chemometrics as new clinical applications in the diagnosis of tumors. The photonic methods have been reported in literature to be producing spectroscopic differences between cancerous tissue and healthy tissue. However, these spectral differences can sometimes fail to detect cancer tissue due to low resolution or differentiated biomolecules not being tied to the tissues. In this study, the mapping based on chemometrics-assisted Raman spectroscopy was used to differentiate IDC from healthy tissue. Therefore, the information was provided on the basis of chemical composition of different structures of material and was confirmed by histopathological evaluations and FTIR spectral analysis. Raman spectroscopic mapping of tissue microarrays can be a candidate for fast analytic methods of early diagnosis of IDC. The other advantages of this method are that it is non-invasive, non-destructive and requires no sample preparation.

## References

1. Yilligi SI. *TC Saglik Bakanligi*. Saglik Arastirmalari Genel Mudurlugu Ankara 2016: 181–209.
2. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002; 225 (1): 165–175.
3. Diem M, Griffiths PR, Chalmers JM. *Vibrational spectroscopy for medical diagnosis*. Wiley Chichester; 2008.
4. Fabian H, Lasch P, Boese M, Haensch W. Mid-IR microspectroscopic imaging of breast tumor tissue sections. *Biopolymers: Orig Res Biomol* 2002; 67 (4–5): 354–357.
5. Fabian H, Thi NAN, Eiden M, Lasch P, Schmitt J, Naumann D. Diagnosing benign and malignant lesions in breast tissue sections by using IR-microspectroscopy. *Biochim Biophys Acta (BBA)-Biomembranes* 2006; 1758 (7): 874–882.
6. Rehman S, Movasaghi Z, Darr JA, Rehman IU. Fourier transform infrared spectroscopic analysis of breast cancer tissues; identifying differences between normal breast, invasive ductal carcinoma, and ductal carcinoma in situ of the breast. *Appl Spectroscopy Rev* 2010; 45 (5): 355–368.
7. Walsh MJ, Holton SE, Kajdacsy-Balla A, Bhargava R. Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology. *Vibrat Spectroscopyc* 2012; 60: 23–28.
8. Rehman S, Movasaghi Z, Tucker AT, Joel SP, Darr JA, Ruban AV et al. Raman spectroscopic analysis of breast cancer tissues: identifying differences between normal, invasive ductal carcinoma and ductal carcinoma in situ of the breast tissue. *J Raman Spectroscopy* 2007; 38 (10): 1345–1351.
9. Han B, Du Y, Fu T, Fan Z, Xu S, Hu C et al. Differences and relationships between normal and atypical ductal hyperplasia, ductal carcinoma in situ, and invasive ductal carcinoma tissues in the breast based on Raman spectroscopy. *Appl Spectroscopy* 2017; 71 (2): 300–307.
10. Chowdary M, Kumar KK, Kurien J, Mathew S, Krishna CM. Discrimination of normal, benign, and malignant breast tissues by Raman spectroscopy. *Biopolymers: Orig Res Biomol* 2006; 83 (5): 556–569.
11. Stone N, Matousek P. Advanced transmission Raman spectroscopy: a promising tool for breast disease diagnosis. *Cancer Res* 2008; 68 (11): 4424–4430.
12. Zhu C, Breslin T. Diagnosing breast cancer by using Raman spectroscopy. *Breast Diseases: A Year Book Quarterly* 2006; 1 (17): 34.
13. Keller MD, Vargis E, Mahadevan-Jansen A, de Matos Granja N, Wilson RH, Mycek M-A et al. Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation. *J Biomed Optics* 2011; 16 (7): 077006.
14. Tavassoli FA, Devilee P. Pathology and genetics of tumours of the breast and female genital organs. *Iarc*; 2003.
15. Senol O, Albayrak M, Demirkaya Miloglu F, Kadioglu Y, Calik M. Application of Photonics in Diagnosis of Papillary Thyroid Carcinoma Tissues through Raman Spectroscopy-Assisted with Chemometrics. *Analyt Lett* 2018; 51 (1–2): 229–235.
16. Boccard J, Rutledge DN. A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion. *Analyt Chim Acta* 2013; 769: 30–39.
17. Schmitz SK, Hasselbach PP, Ebisch B, Klein A, Pipa G, Galuske RA. Application of Parallel Factor Analysis (PARAFAC) to electrophysiological data. *Front Neuroinform* 2015; 8: 84.
18. Badowska-Kozakiewicz AM, Sobol M, Patera J, Kozłowski W. Immunohistochemical evaluation of human epidermal growth factor receptor 2 and estrogen and progesterone receptors in invasive breast cancer in women. *Arch Med Sci* 2013; 9 (3): 466.
19. Nguyen TT, Gobinet C, Feru J, Pasco SB, Manfait M, Piot O. Characterization of Type I and IV Collagens by Raman Microspectroscopy: Identification of Spectral Markers of the Dermo-Epidermal Junction. *Spectroscopy J* 2012; 27 (5–6): 7. doi:10.1155/2012/686183.
20. Azan A, Untereiner V, Gobinet C, Sockalingum GD, Breton M, Piot O et al. Demonstration of the Protein Involvement in Cell Electroporation using Confocal Raman Microspectroscopy. *Scientific Reports* 2017; 7: 40448. doi:10.1038/srep40448.
21. Surovtsev NV, Dzuba SA. Flexibility of phospholipids with saturated and unsaturated chains studied by Raman scattering: the effect of cholesterol on dynamical and phase transitions. *J Chem Phys* 2014; 140 (23): 235103. doi: 10.1063/1.4883237.
22. Madzharova F, Heiner Z, Kneipp J. Surface Enhanced Hyper-Raman Scattering of the Amino Acids Tryptophan, Histidine, Phenylalanine, and Tyrosine. *J Phys Chem C* 2017; 121 (2): 1235–1242. doi: 10.1021/acs.jpcc.6b10905.

Received December 12, 2018.

Accepted January 15, 2019.