Near-infrared fiber optic spectroscopy as a novel diagnostic tool for the detection of pancreatic cancer

Venkata R. Kondepati
Johann Zimmermann
Mannheim University of Applied Sciences
Institute for Instrumental Analysis and Bioanalysis
Mannheim, Germany

Michael Keese
Joerg Sturm
Bernd C. Manegold
University of Heidelberg
Faculty for Clinical Medicine
Surgical Clinic
Mannheim, Germany

Juergen Backhaus
Mannheim University of Applied Sciences
Institute for Instrumental Analysis and Bioanalysis
Mannheim, Germany

Abstract. We have investigated the application of near-infrared (NIR) fiber-optic spectroscopy for the diagnosis of pancreatic cancer. Cluster analysis of the Fourier transformed near-infrared (FTNIR) fiber-optic spectra of surgically resected pancreatic tumor tissues allowed discrimination of tumor from normal tissue with high sensitivity and specificity. The sensitivity of the method using spectral information of the CH stretching first overtone region ($5951-5608$ cm^{-1}) was 83.3% with a specificity of 83.3%. Based on the CH stretching second overtone region ($8605-7938$ cm^{-1}) we could achieve a sensitivity of 88.9% and specificity of 72.2%. These findings suggest that NIR spectroscopy offers the potential for minimally invasive in-vivo diagnosis of pancreatic cancer. © *2005 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.2060568]

1 Introduction

Biopsy followed by pathology is the gold standard and the most common approach to diagnose cancer. However, it is a time-consuming method and based on the pathologist's expertise. Though diagnostic procedures for cancer have improved tremendously during the past decades, they are established as highly expensive techniques. Thus, there is the need to develop an accurate, fast, convenient, and inexpensive method for cancer diagnosis. Several research groups have investigated the application of near-infrared (NIR) spectroscopy as a noninvasive and time-saving method for cancer diagnosis.[1–4] Near-infrared spectroscopy is based on overtones and combinations of the molecular vibrations of C–H, N–H, and O–H groups, which are part of all biological molecules. Therefore, the absorption of near infrared radiation ($12000-4000$ cm^{-1}) leads to a complex spectrum, which provides qualitative and quantitative information about the chemical composition of the tissue with its main components of lipids, proteins, carbohydrates, and water. Any alteration in the chemical composition of the tissues can thus be detected by NIR spectroscopy and be evaluated with the help of mathematical nonsupervised methods such as cluster analysis, which is a pattern recognition technique. In the present study, we investigated the qualitative spectral differences between cancerous and its adjacent normal tissues in surgically resected pancreatic cancer specimens using NIR fiber-optic spectroscopy.

2 Materials and Methods

2.1 FTNIR Spectrometer and Fiber Optic Probe

A Vector 22/N Fourier transform near-infrared (FTNIR) spectrometer (Bruker Optics GmbH, Ettlingen, Germany) was used for the NIR measurements of the tissues. The spectrometer was equipped with a fiber-optic port for the remote sampling applications and was controlled by the Bruker Optics User Software (OPUS) of version 4.0 (Bruker Optics GmbH, Ettlingen, Germany). The spectra were recorded, preprocessed, and analyzed using this software. The fiber optic probe (Richard Wolf GmbH, Knittlingen, Germany) used for the measurements was a Y-shaped probe with one arm used for sample illumination and the other for radiation detection. Both arms contained the same number of quartz fibers and possessed a high transmittance within the used wave-number range. The head of the probe had an outer diameter of 2.9 mm and an inner diameter of 2.6 mm. Within the inner diameter, the active surface was of 1.8 mm diameter and the NIR radiation and detection fibers alternate relatively evenly.

2.2 Pancreatic Samples

Patient specimens were collected immediately after surgery from the central operation theatre of the University Hospital, Mannheim, with informed patient consent. The study was approved by the local ethics committee. A total of 18 surgically resected pancreatic tumor specimens, 17 from adenocarcinoma patients (eight males and nine females) and one from a cystadenoma patient (female) were collected. Of the 18 specimens, 15 were derived from Traverso procedure, two from pancreas pancreas tail resection, and one from Whipple resection. The mean age of the patients was 66.5 years with the oldest 84 years and the youngest 40 years.

Address all correspondence to Juergen Backhaus, Institute for Instrumental Analysis and Bioanalysis, Mannheim University of Applied Sciences, Windeckstrasse 110, 68163 Mannheim, Germany. Tel: +49 621 292 6486; Fax: +49 621 292 6420; E-mail: j.backhaus@fh-mannheim.de
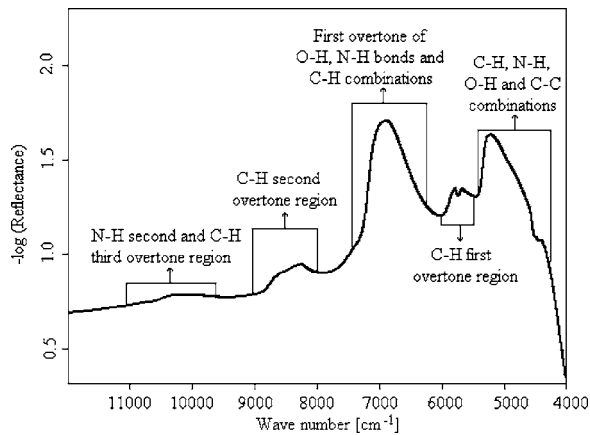
Fig. 1 Assignment of bands of various chemical substructures in the NIR spectrum of the tissue.

Fig. 3 Population mean and standard deviation spectra of tumor and normal tissue measurements.

2.3 NIR Measurements

The spectra of the tissue samples were measured at room temperature before the microscopic histological assessment was provided. White Spectralon (Bruker Optics GmbH, Ettlingen, Germany) was used as a reflectance standard. The spectra were collected with the co-addition of 32 scans (measurement time 70 s) in diffuse reflectance mode from 12000 to 4000 cm^{-1} with a spectral resolution of 8 cm^{-1}, applying Blackman-Harris three-term apodization and a zero filling factor of 2. The resected tissue kept in a plastic box was placed on a laboratory jack and the fiber optic probe was attached to a clamp to make its tip contact the tissue area by carefully lifting the stage. In all patients the spectra were collected from the tumor region and from the resection edges. The assignment of the spectra to normal and cancer was based on the histological examination of the tumor and resected margins of the tissue by a pathologist. In 17 samples three different regions were selected from cancer and from the resected border of the tissue; three spectra were measured from each region for a total of nine spectra each from cancerou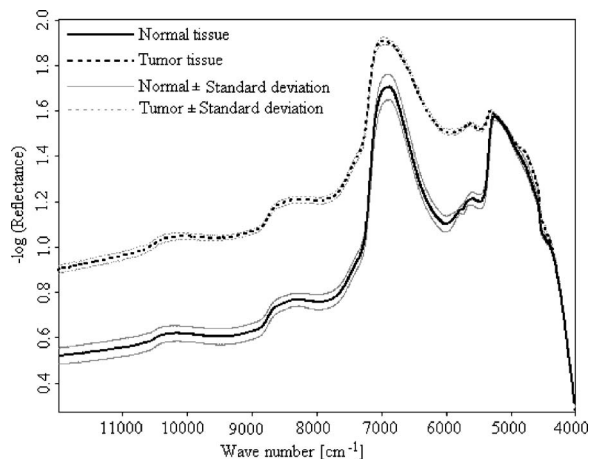s and normal tissue. In one sample only six spectra were collected from both cancerous and the resected border of the tissue from two different regions due to the small size of the tissue.

Fig. 2 Mean and standard deviation spectra of the tumor and normal tissue measurements of one patient.

Table 1 Pathological stages of the tumors.

| Patient	Pathology	Tumor	Node	Metastasis	Grade
201	Adenocarcinoma	3	1	0	2–3
202	Adenocarcinoma	3	1	0	2
203	Adenocarcinoma	3	1	0	2
204	Adenocarcinoma	3	1	0	2
205	Adenocarcinoma	3	1	0	2
206	Adenocarcinoma	3	1	0	2
207	Adenocarcinoma	4	1	0	3
208	Adenocarcinoma	3	1	0	2
209	Adenocarcinoma	3	0	0	2
210	Adenocarcinoma	2	0	0	2
211	Adenocarcinoma	4	1	0	3
212	Adenocarcinoma	3	1	0	2
213	Adenocarcinoma	3	1	0	2
214	Adenocarcinoma	3	0	0	2
215	Adenocarcinoma	3	1	0	2
216	Adenocarcinoma	3	1	0	2
218	Adenocarcinoma	3	1	1	2–3
219	Cystadenoma	0	0	0	0
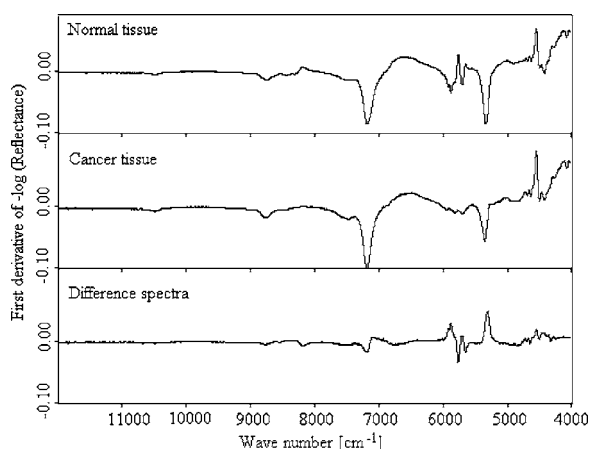
Fig. 4 First derivative and vector normalized mean spectra of pancreatic cancer and normal tissue and their difference spectra.

2.4 Outlier Detection

The spectra from the different tissue categories that were collected from each patient were averaged for a mean spectrum of cancerous and healthy tissue, respectively. The mean and standard deviation for the mean spectrum of cancerous and normal tissues of the patients were calculated and the patient tissue spectra, which were outside the mean ±2 standard deviation, were considered as outliers. No outliers based on the criteria to lie between the mean ±2 standard deviation were detected.

2.5 Data Preprocessing

The spectra of cancerous and normal tissue that were collected from each patient were first derivated using the Savitzky-Golay algorithm with nine smoothing points followed by vector normalization. The preprocessed spectra from cancerous and normal tissue of each patient, respectively, were averaged. These mean spectra were used for further evaluation.
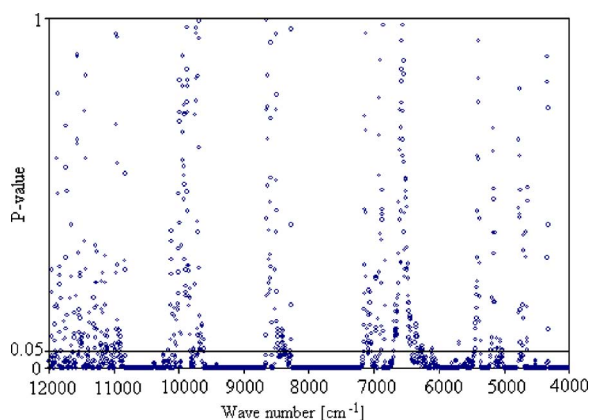

Fig. 5 P values of absorbances at wave numbers showing significant differences between pancreatic cancerous and normal preprocessed spectra.
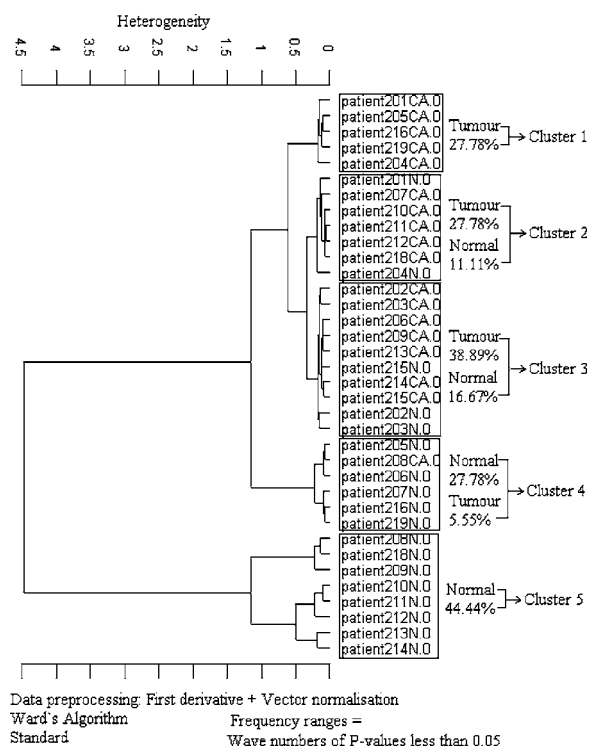

Fig. 6 Hierarchical cluster analysis of the NIR preprocessed spectra of tumor and normal tissue measurements of the wave numbers showing significant differences ($p < 0.05$).

2.6 Calculation of the Statistical Significance

The statistical significance of the differences between preprocessed cancerous and normal tissue spectra was calculated by the double-sided unequal variance T test. The P values were calculated for each wave number and plotted against their corresponding wave numbers. The percentage of wave numbers with $p < 0.05$ was calculated.

2.7 Spectral Evaluation

Spectral evaluation was done by hierarchical cluster analysis using Ward's algorithm as implemented in the OPUS software. In hierarchical cluster analysis the spectra are partitioned into groups or clusters according to the similarity or dissimilarity of the spectra and the result is shown as a dendrogram.[5]
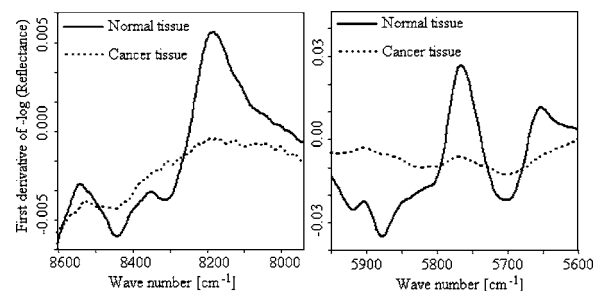

Fig. 7 Spectral differences in the mean spectra of cancerous and normal tissues in CH stretching second (8605–7938 cm^{-1}) and first overtone (5951–5608 cm^{-1}) regions.
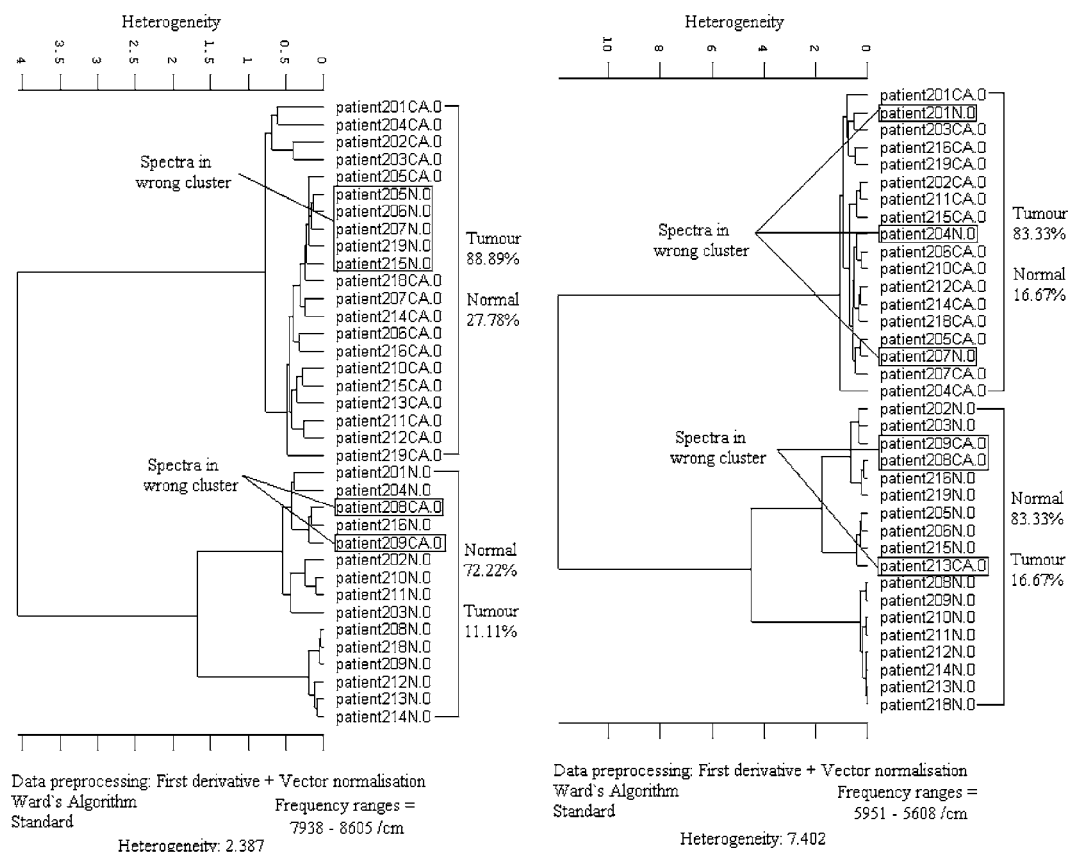
Fig. 8 Hierarchical cluster analysis of the NIR preprocessed spectra of tumor and normal measurements in CH stretching first ($5951-5608$ cm^{-1}) and second overtone ($8605-7938$ cm^{-1}) regions showing dendrograms with separate clusters for tumor and normal tissue.

3 Results and Discussion

The difference in tissue composition between pancreatic cancer and normal tissues has been extensively analyzed using chemical, histochemical, biochemical, and immunohistochemical studies. Altered composition of isoferritin, glycosaminoglycans, glycoproteins, and glycolipids has been described in pancreatic adenomas and carcinomas.[6–10] These differences in the tissue composition will contribute to the spectral variance between cancer and normal tissues. The near-infrared spectrum of the pancreatic tissue and the assignment of the peaks to various chemical substructures[11] in the spectrum are shown in Fig. 1. The spectrum is a mixture of the spectral signatures of many tissue components including water, lipids, proteins, and carbohydrates. The drop of the $-\log$ (reflectance) around 4000 cm^{-1} is due to the low transmittance close to the interval end as shown in Fig. 1.

The mean of tumor and normal tissue measurements from one sample and their standard deviations are shown in Fig. 2. The variance is essentially constant across the spectral range used. This suggests that the spectra are highly reproducible. The mean spectra from tumor and normal tissue measurements of all the samples and their standard deviations are shown in Fig. 3. The relatively high variation between the samples could be due to different tumor stages of the tissues (Table 1) and differences in the tissue thickness of the samples.[12]

Since derivative techniques are normally used to remove or suppress constant background signals and to enhance the visual resolution of the spectra, we have used the first derivative spectra for the evaluation with cluster analysis. However, as derivation amplifies the spectral noise, we smoothed the data using the Savitzky-Golay algorithm. Further, to reduce the differences between each single measurement of the same sample, the spectra were vector normalized after derivation and smoothing. To emphasize the differences between the preprocessed cancer and normal spectra, the mean normal spectrum was subtracted from the mean cancer spectrum. The preprocessed mean spectrum of cancer and normal tissue and their differences are shown in Fig. 4.

To assess whether significant differences existed between preprocessed normal and cancerous tissue spectra, paired t tests were applied for the absorbance values at each wave number and the resulting p values were plotted. Several intervals of the resulting plot contained statistically significant p values (Fig. 5). Of the total wave numbers, 76.2% were significantly different with $p < 0.05$. The result from a hierarchical cluster analysis using the wave numbers showing significant difference is presented in Fig. 6. Five clusters can be found with a majority of cancerous tissue spectra in the first, second, and third clusters and the majority of the normal tissue spectra in the fourth and fifth clusters.

Table 2 Classification results of the hierarchical cluster analysis of NIR preprocessed spectra of pancreatic tumor and normal tissues using data from CH stretching first and second overtone regions.

Spectral region (cm^{-1})	Disease	Number of spectra	Spectra assigned in right cluster	Spectra assigned in wrong cluster	Percentage of spectra assigned correctly	Percentage of spectra assigned incorrectly
8605–7938	Tumor	18	16	2	88.9	11.1
	Normal	18	13	5	72.2	27.8
5951–5608	Tumor	18	15	3	83.3	16.7
	Normal	18	15	3	83.3	16.7

As cancerous tissues mainly differ from normal tissues in glycolipids, glycoproteins, and glycosaminoglycans composition[6–10] and the CH vibrational regions represent these classes,[3,4,13–15] we decided to use the CH stretching bands for the evaluation by a cluster analysis. As a result, the spectra were separated with high sensitivity and specificity using the CH stretching first (5951–5608 cm^{-1}) and second overtone regions (8605–7938 cm^{-1}). The peaks in the cancer tissue spectrum between the wave numbers 8584–8497, 8263–7938, 5795–5729, and 5678–5605 cm^{-1} were reduced in intensity compared with those in the normal tissue spectra. The band intensities between the wave numbers 8605–8585, 8497–8263, 5951–5795, and 5728–5678 cm^{-1} were strongly in relation to those in the normal tissue spectra (Fig. 7). The dendrograms of the cluster analysis using CH first and second overtone regions are shown in Fig. 8 and the classification results are summarized in Table 2.

When evaluating the spectral interval data of 8605–7938 cm^{-1}, the hierarchical cluster analysis classified the spectra into two groups with 88.9% tumor and 27.8% normal spectra in first group and 72.2% normal and 11.1% tumor spectra in the second group. The heterogeneity between the two groups was 2.387. With the evaluation of the spectral region 5951–5608 cm^{-1}, the spectra were classified into two groups with 83.3% tumor and 16.7% normal spectra in the first group and 83.3% normal and 16.7% of the spectra in the second group. The heterogeneity between the two clusters was 7.402. With the evaluation of the CH second overtone region, three spectra from adenocarcinoma and five spectra from normal were misclassified, and with the evaluation of the CH first overtone region three spectra from adenocarcinoma and three spectra from normal were misclassified, respectively. The comparison of the misclassified spectra with the mean spectra of normal tissue and cancer are shown in Figs. 9 and 10. All the normal spectra that were misclassified exhibit spectral features similar to the cancer spectra [Figs. 9(A) and 9(B)] and all the cancer spectra that were misclassified exhibit spectral features similar to the normal spectra [Figs. 10(A) and 10(B)]. The sensitivity of the method in CH first overtone region (5951–5608 cm^{-1}) was 83.3% with a specificity of 83.3%. With the CH second overtone region (8605–7938 cm^{-1}) data we could achieve a sensitivity of 88.9% and specificity of 72.2%. The accuracy of the method using CH first and second overtone region data was 83.3% and 80.5%, respectively.

4 Conclusion

Cluster analysis of the NIR fiber-optic spectra of surgically resected human pancreatic tumor tissues showed high sensitivity and specificity in discriminating the spectra obtained
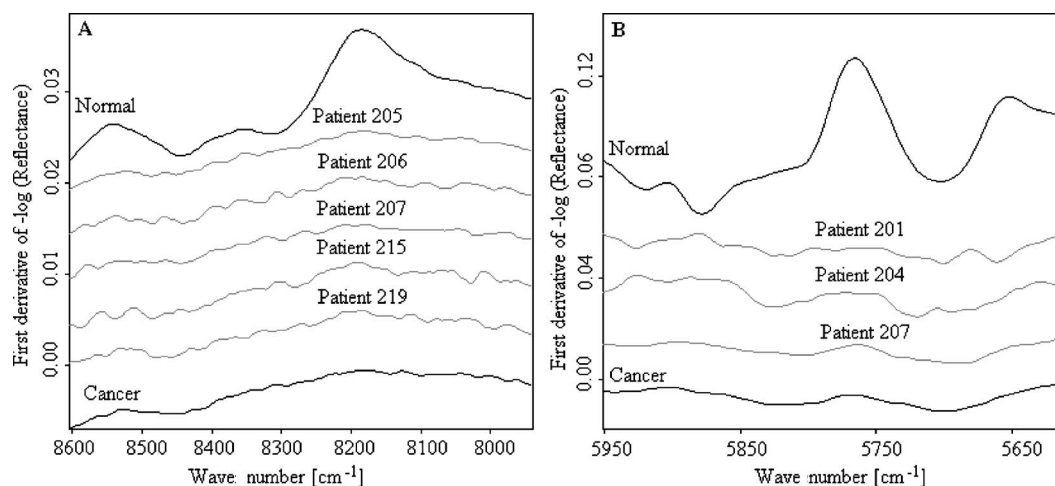

Fig. 9 Comparison of the misclassified normal spectra with the mean spectra of cancer and normal tissue measurements in CH stretching second (A) and first overtone (B) regions. For clarity the individual spectra are displayed by using different offsets.
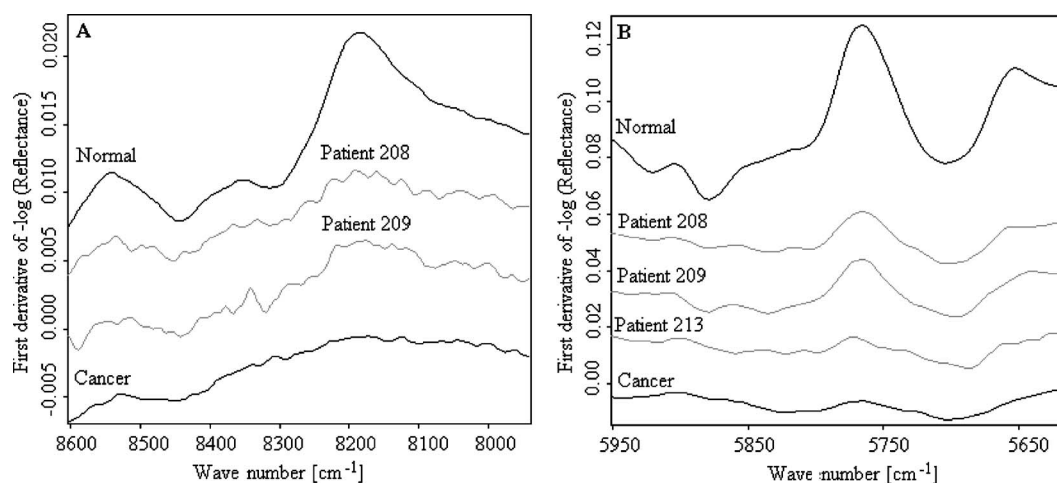
Fig. 10 Comparison of the misclassified cancer spectra with the mean spectra of cancer and normal tissue measurements in CH stretching second (A) and first overtone (B) regions. For clarity the individual spectra are displayed by using different offsets.

from tumor tissues from the spectra measured from normal tissues. With a sample population size of 18 patients, we achieved highest accuracy with a sensitivity of 83.3% and a specificity of 83.3% using data from CH first overtone region of the NIR spectrum, which suggests a potential spectral region to distinguish tumor from normal tissue. The advantage of this method is that it is a very fast, simple, and relatively inexpensive technique. These findings suggest that NIR fiber optic spectroscopy offers the potential for a minimally invasive in-vivo diagnosis of pancreatic tumors.

References

1. S. B. Colak, M. B. van der Mark, G. W. Hooft, J. H. Hoogenraad, E. S. van der Linden, and F. A. Kuijpers, "Clinical optical tomography and NIR spectroscopy for breast cancer detection," *IEEE J. Sel. Top. Quantum Electron.* **5**(4), 1143–1158 (1999).
2. A. E. Cerussi, A. J. Berger, F. Bevilacqua, N. Shah, D. Jakubowski, J. Butler, R. F. Holcombe, and B. J. Tromberg, "Sources of absorption and scattering contrast for near-infrared optical mammography," *Acad. Radiol.* **8**(3), 211–218 (2001).
3. L. M. McIntosh, R. Summers, M. Jackson, H. H. Mantsch, J. R. Mansfield, M. Howlett, A. N. Crowson, and J. W. Toole, "Towards non-invasive screening of skin lesions by near-infrared spectroscopy," *J. Invest. Dermatol.* **116**, 175–181 (2001).
4. N. Hirosawa, Y. Sakamoto, H. Katayama, S. Tonooka, and K. Yano, "Invivo investigation of progressive alterations in rat mammary gland tumors by near infrared spectroscopy," *Anal. Biochem.* **305**, 156–165 (2002).
5. M. J. Adams, "Pattern recognition I—Unsupervised analysis," Chap. 4 in *Chemometrics in Analytical Spectroscopy*, pp. 103–122, Royal Soc. Chem., Cambridge (1995).
6. H. Masuda, T. Ozeki, I. Takazono, and Y. Tanaka, "Composition of glycosaminoglycans in human pancreatic cancer," *Biochem. Med. Metab. Biol.* **41**(3), 193–200 (1989).
7. A. D. Theocharis, M. E. Tsara, N. Papageorgacopoulou, D. D. Karavias, and D. A. Theocharis, "Pancreatic carcinoma is characterized by elevated content of hyaluronan and chondroitin sulphate with altered disaccharide composition," *Biochem. Biophys. Acta.* **1502**(2), 201–206 (2000).
8. S. Linder, E. Castanos-Velez, A. Von Rosen, and P. Biberfeld, "Immunohistochemical expression of extracellular matrix proteins and adhesion molecules in pancreatic carcinoma," *Hepato-Gastroenterology* **48**(41), 1321–1327 (2001).
9. I. Nakazawa, A. Ohneda, and K. Mura, "Biochemical differences between human malignant and benign insulinoma tissues," *Tohoku J. Exp. Med.* **121**(3), 275–280 (1977).
10. H. Hattori, K. Uemura, H. Ishihara, and H. Ogata, "Glycolipid of human pancreatic cancer; the appearance of neolacto-series (type 2 chain) glycolipid and the presence of incompatibile blood group antigen in tumor tissues," *Biochem. Biophys. Acta.* **1125**(1), 21–27 (1992).
11. L. Bokobza, "Origin of near-infrared absorption bands," Chap. 2 in *Near-infrared Spectroscopy*, H. W. Siesler, Y. Ozaki, S. Kawata, and H. M. Heise, Eds., pp. 11–41, Wiley-VCH, Weinheim (2002).
12. H. M. Heise, S. Haiber, M. Licht, C. Moll, and M. Stücker, "In vivo measurements of intact human skin by near- and mid-infrared spectroscopy," *VDI-Berichte 1863*, pp. 107–114, VDI Verlag, Duesseldorf (2004).
13. J. Wang, Y.-J. Geng, B. Guo, T. Klima, B. N. Lal, J. T. Willerson, and W. Casscells, "Near-infrared spectroscopic characterization of human advanced atherosclerotic plaques," *J. Am. Coll. Cardiol.* **39**(8), 1305–1313 (2002).
14. J. M. Canvin, S. Bernatsky, C. A. Hitchon, M. Jackson, M. G. Sowa, J. R. Mansfield, H. H. Eysel, H. H. Mantsch, and H. S. El-Gabalawy, "Infrared spectroscopy: Shedding light on synovitis in patients with rheumatoid arthritis," *Rheumatology* **42**, 76–82 (2003).
15. H. M. Heise, "Applications of near-infrared spectroscopy in medical sciences," Chap. 13 in *Near-infrared Spectroscopy*, H. W. Siesler, Y. Ozaki, S Kawata, and H. M. Heise, Eds., pp. 289–333, Wiley-VCH, Weinheim (2002).