Light-Induced Autofluorescence Spectroscopy for Detection of Nasopharyngeal Carcinoma *in Vivo*

HANPENG CHANG, JIANAN Y. QU,* POWING YUEN, JONATHAN SHAM, DORA KWONG, and WILLIAM I. WEI

Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, P. R. China (H.C., J.Y.Q.); Division of Otorhinolaryngology, Queen Mary Hospital, The University of Hong Kong, Hong Kong, P.R. China (P.Y., D.K., W.I.W.); and Department of Clinical Oncology, The University of Hong Kong, Hong Kong, P.R. China (J.S.)

Autofluorescence spectral signals were measured *in vivo* from 85 nasopharyngeal carcinoma lesions and 131 normal tissue sites of 59 subjects during routine nasal endoscopy. Diagnostic algorithms based on principal component analysis and the ratio of the spectral signals between multiple-wavelength bands were developed for classifying the autofluorescence spectra. Performances of the algorithms were evaluated using the cross-validation method. The principal component analysis based algorithms using information from the entire fluorescence spectrum can differentiate nasopharyngeal carcinoma lesions from normal tissue with 95% sensitivity and 93% specificity. With 94% sensitivity, the specificities of multiple-wavelength ratio algorithms are about 83%. The results demonstrate that light-induced autofluorescence endoscopy with principal component analysis algorithms can provide accurate diagnostic information for the detection of nasopharyngeal carcinoma *in vivo* and may be potentially used in clinical practice combined with the routine white-light endoscopy procedure.

Index Headings: **Fluorescence spectroscopy; Endoscopy; Principal component analysis.**

INTRODUCTION

Nasopharyngeal carcinoma (NPC), most common among Southeast Asians, may occur at any age. When detected in its early stages, successful treatment is possible. However, the currently available detection method, white-light endoscopy, does not provide sufficient accuracy in discriminating NPC from normal tissue, especially in the incidence of flat/small lesions and identification of tumor margins in advanced stages of NPC. The screening program identifies the high-risk group of patients with raised serum EBV antibody titer. For diagnosis of the subclinical tumors from the high-risk group, the common practice is random endoscopic biopsy. However, only 5.4% of patients with elevated serum EBV antibody titer had asymptomatic NPC in random biopsies of the nasopharynx. Therefore, there is a need for new technology that can combine with the commonly used white-light endoscopy to provide more diagnostic information for guiding the routine biopsy procedure and improving the accuracy for identifying malignant lesions at the early stage of tumor development.

Light-induced fluorescence (LIF) spectroscopy can remotely sense the biochemical and morphological state of tissue *in vivo*. The difference in chemical composition and structure between normal and diseased tissue is reflected in the measured autofluorescence spectral characteristics. This makes the LIF spectroscopic technique a potential tool for *in vivo* noninvasive diagnosis of diseased tissue. The use of LIF spectroscopy on tissue has been conducted on a variety of organ sites. Comprehensive reviews of the characterization of tissue by LIF spectroscopy have been presented in Refs. 1–3. In a pilot study, we measured the *in vivo* autofluorescence spectra at nasopharyngeal sites in a small group of subjects. It has been demonstrated that *in vivo* autofluorescence spectra of NPC and normal nasopharyngeal tissues are different, and that the difference can be used to discriminate NPC from normal tissue.[4]

In this extended study, we conduct measurements of the *in vivo* autofluorescence at nasopharyngeal sites in a large group of subjects. We explore the potential of statistical multivariate methods to classify the *in vivo* autofluorescence. Specifically, principal component analysis (PCA), an effective method for analyzing the statistical characteristics of spectral data, is used to process the spectra collected from normal tissue and carcinomas *in vivo*. This method allows us to use the information from the full spectral data to build a diagnostic algorithm. The PCA method has been successfully applied to classifying *in vivo* autofluorescence by other groups.[5–9] In general, PCA reduces the dimension of spectral data into a set of informative principal components that account for most of the variance of the spectral data. Here, we first identify the principal component (PC) loadings that carry the diagnostic information and then develop a probability-based algorithm and a simpler threshold-based algorithm to classify *in vivo* autofluorescence spectral data based on the correlation between projection scores of the autofluorescence to the PC loadings and the pathologic state of tissue sites where the autofluorescence was measured.

In a comparison with the results of the pilot study reported in Ref. 4, we evaluate the performance of the algorithms including the fluorescence signals in two-wavelength bands (2-λ) and three-wavelength bands (3-λ) for the detection of nasopharyngeal carcinoma. The 2- and 3-λ algorithms use the diagnostic information in two- and three-wavelength bands instead of from the entire spectrum. Finally, the factors that could affect the performance of the probability-based PCA algorithm and the threshold-based PCA algorithm are discussed.

MATERIALS AND METHODS

LIF Spectroscopy and Tissue Histology. *In vivo* autofluorescence spectra were collected at the nasopharyn-

Received 11 March 2002; accepted 29 May 2002.
* Author to whom correspondence should be sent.

0003-7028/02/5610-1361$2.00/0

geal tissue site through a standard nasoendoscope. The details of instrumentation for the measurement of LIF signals *in vivo* have been described in a previous paper.[4] Briefly, the light source used as the excitation was a mercury arc lamp filtered with a band pass filter with a bandwidth from 390–450 nm. The excitation light was delivered to the tissue through the illumination channel of the endoscope. The autofluorescence was mainly excited by the strong spectral peaks of mercury at 404.66 and 435.84 nm. The mercury emission lines at 404.66 and 435.84 nm share the same upper initial energy level, $^1S_0$. Therefore, the relative intensities between two strong spectral peaks are determined by their relative transition probabilities and independent of the operation conditions of the light source. The major fluorophores excited by the mercury lines at 404.66 and 435.84 nm are collagen and FAD.[1,3] The fluorescence from the other major fluorophore, NADH, which is related to local metabolic processes, is not excited because the absorption coefficients of NADH at the mercury peaks are negligible.[1,3]

The fluorescence and reflection signals from the illuminated tissue surface were imaged by the endoscope and separated by a dichoic mirror with a cut-on wavelength at 470 nm. Seven optical fibers were evenly distributed in the image plane of the endoscope to conduct the fluorescence signals to an imaging spectrograph for spectral analysis and recording. The spectral response of the system was calibrated. The autofluorescence spectra were recorded in the wavelength range from 470–680 nm. The dominant absorbers in this wavelength range are hemoglobin and oxyhemoglobin.

The LIF spectroscopy measurement at the nasopharyngeal tissue site was conducted in the Department of Otorhinolaryngology and the Department of Clinical Oncology at Queen Mary Hospital, The University of Hong Kong. Fifty-nine subjects were included in this study. The *in vivo* autofluorescence spectra were collected from 85 carcinoma tissue sites and 131 normal tissue sites. The typical *in vivo* autofluorescence for normal tissue and carcinoma are shown in Fig. 1. The biopsy specimens were taken from the tissue sites where the autofluorescence spectra were collected. Histologic examinations on biopsies were then performed by pathologists.

Diagnostic Algorithm Development. A variety of multivariate statistical analysis methods have been used in tissue spectroscopy. A comparison between different methods on the classification of autofluorescence spectra was made in Ref. 5. It was found that PCA-based algorithms produced the highest diagnostic accuracy. The commonly used multivariate linear regression methods fail to perform better than PCA when the spectral data set is not large enough, because the former method requires the sample number in the data set to be greater than the spectral dimension and that there be no colinearities in the data set.[5,8] In this work, we use the PCA method for processing the collected autofluorescence spectra and develop the diagnostic algorithms based on principal components scores for classification of autofluorescence spectra.

Principal components analysis involves a mathematical procedure that transforms a number of correlated or uncorrelated variables into a smaller number of uncorrelated variables called principal components (PCs).[10] The first

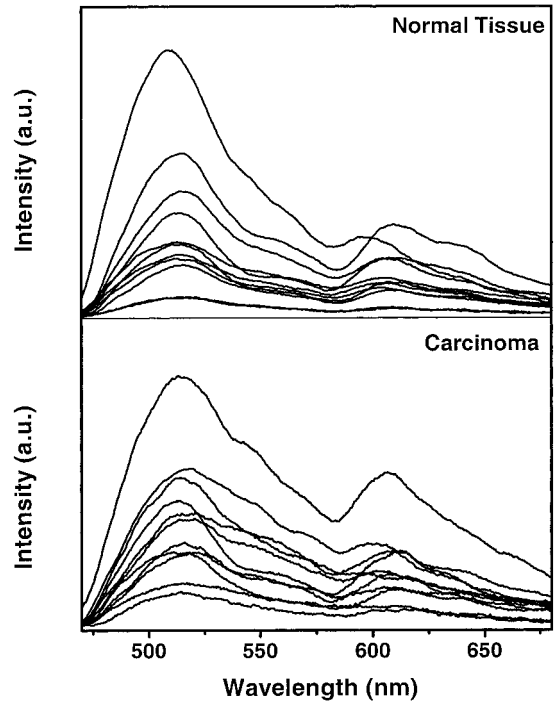

FIG. 1. Typical raw autofluorescence spectra directly collected from normal tissue sites and carcinoma tissue sites *in vivo*.

PC accounts for as much of the variability in the data set as possible, and each succeeding component accounts for as much of the remaining variability as possible. In the application of the PCA method for processing autofluorescence spectral data, the original spectral variables (wavelength) are transformed into a set of PC spectra in the space of the original spectral variables (wavelength). The PCs are arranged in the order of their contribution to the variance of the entire spectral data set. Each spectrum is thus a combination of the PC loading spectra.

In general, a few PCs can explain most of the informative variations of an autofluorescence spectrum because spectral characteristics of tissue are determined by a limited number of biological fluorophores and absorbers. Therefore, PCA processing can dimensionally reduce the variables of the original spectral data into a small number of informative PCs that fully describe the variations in the spectral data within the limitations of noise. The projection of *in vivo* autofluorescence signals onto the informative PCs may then be related to the tissue pathology and used to create a diagnostic algorithm for classification of the autofluorescence signal.

The development of a diagnostic algorithm seeks to make use of the reduced set of PC components to extract the principal differences between normal and carcinoma tissue for diagnostic purposes. The spectral dimension of the raw autofluorescence spectra is determined by the wavelength range and resolution of the recorded spectral data. In this work, the wavelength range of the raw autofluorescence signal was from 470 to 680 nm and the resolution of the spectrograph was 1.0 nm. The dimension of the raw spectral data was therefore 211. To prepare the data for PCA, the raw spectra were smoothed to remove high-frequency random noise using adjacent averaging. Each smoothed spectrum was then normalized

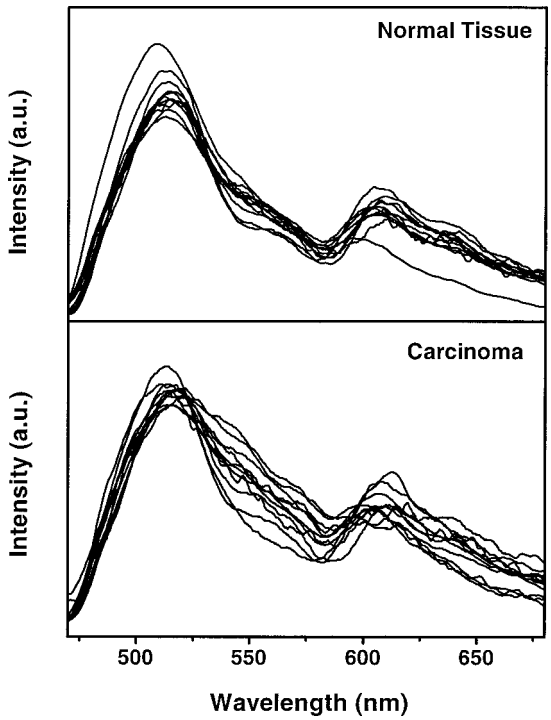

FIG. 2. Preprocessed autofluorescence spectra for principal component analysis.

to its area to eliminate variations between site-to-site and individual-to-individual measurements. The processed spectra of the normal tissue and the carcinoma lesions are displayed in Fig. 2.

Applying PCA to the preprocessed autofluorescence spectra, the original spectra are transformed into a set of PC scores. Since most of PCs account mainly for noise and do not provide diagnostic information, it is necessary to identify a small group of informative PCs to build the algorithms for classification of autofluorescence spectra. The complexity of a diagnostic algorithm will be reduced by including a minimal number of PCs. The procedure for the selection of informative PCs consists of two steps. First, the contribution of each PC to the total variance of spectral data is proportional to its eigenvalue. High-order PCs often account for less than 1% of the total data variation and represent noise. In this work, the PCs that explain more than 1% variance in the spectral data are considered as informative PCs. In the second step, an unpaired students' *t*-test is used to evaluate the difference in projection scores on the informative PCs between the autofluorescence spectra of normal tissue and carcinoma lesions. Only the PCs that have significantly different projection scores for normal tissue and carcinoma lesions will be selected for construction of diagnostic algorithms.

After the PCs with diagnostic information are identified, a probability-based method is used to build the algorithm for classification of the autofluorescence spectra. This method has been successfully applied in the autofluorescence diagnosis of cervical precancer, laryngeal lesions, and oral carcinogenesis by other groups.[6,8,9] Briefly, the classification is based on posterior probability of the given observation being normal or carcinoma tissue. The posterior probability of a given PC score, x_i, being carcinoma, C, can be estimated using Bayes' theorem as follows:

$$P(C|x_i) = \frac{P(x_i|C) \times P(C) \times k}{P(x_i|C) \times P(C) \times k + P(x_i|\bar{C}) \times P(\bar{C}) \times \bar{k}} \quad (1)$$

where $P(x_i|C)$ is the conditional probability that the PC score of a spectral data collected from carcinoma tissue is x_i. $P(C)$ and $P(\bar{C})$ are the prior probability of the tissue being carcinoma and normal tissue, respectively. k is the cost for misclassification of a carcinoma tissue as a normal tissue. The sum of k and $\bar{k}$ is equal to 1. In the calculation of conditional probability, the PC scores are assumed to be normally distributed. The joint multivariate normal distribution for the PCs included in Eq. 1 is then:

$$f(x) = \frac{1}{\sqrt{(2\pi)^p|\Sigma|}}\exp\left[-\frac{1}{2}(x - u)'\Sigma^{-1}(x - u)\right] \quad (2)$$

where u and Σ are the mean and covariance matrix of the multivariate p-dimensional variable, x, respectively. The dimension, p, is determined by the number of PCs used in the algorithm.

The advantage of the probability-based method is that it can use an unlimited number of PCs for the spectral analysis when there is a large volume of samples. It is convenient to use the resulting posterior probability ranging from 0 to 1 for the classification of the autofluorescence spectra. However, it should be pointed out that the performance of the algorithm relies on the accuracy in the estimation of the PC score density function.[11] With an increase of the number of PCs included in the algorithm, the dimension of the score probability density function increases. When the total number of samples is limited, the accuracy of density estimation using Eq. 2 drops. Also, the probability-based algorithm may be corrupted when the score probability density function has a significant deviation from the normal distribution.

In a simpler approach to building a PCA-based diagnostic algorithm, scores of spectral data on the PCs with diagnostic information are directly used for classification of the autofluorescence spectra. An exhaustive search can be conducted to find the optimal diagnostic threshold that produces the separation of normal tissue from carcinoma lesions with the highest sensitivity and specificity. When an algorithm includes one PC, the diagnostic threshold is a single point in the PC space. For algorithms with two and three PCs, the thresholds are a line and a plane, respectively. The major advantage of a threshold-based algorithm is that it is independent of the score probability density function. The disadvantage is that the exhaustive search may be time consuming and not accurate in finding the optimal diagnostic threshold when too many PCs are included in the algorithm.

To evaluate the performance of the probability-based algorithm and threshold-based algorithm, a cross-validation test is used. In each round of the validation, one spectrum is held out from the data set that is formed with the preprocessed total of 216 *in vivo* autofluorescence spectra measured from both normal tissue and carcinoma lesions. The remaining spectral data are used as a training set for construction and optimization of the PCA algo-

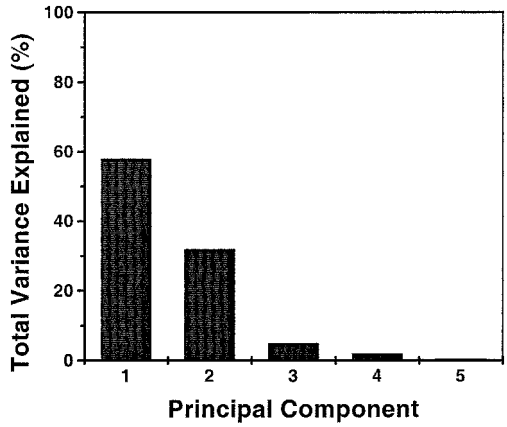

FIG. 3. SCREE plot to illustrate the variation explained by the first five principal components.

rithms. The held-out spectrum is then classified by the optimized algorithm. This process is repeated until all the spectra in the data set are classified. In every round of the cross validation, a new set of PCs and a new algorithm are created using the remaining 215 spectra as the training data. The sensitivity and specificity of a particular algorithm are calculated based on the classification of the held-out spectrum in each round of cross validation.

RESULTS AND DISCUSSION

A MATLAB-based PCA program was used to analyze the spectral data and calculate the PCs. The proportion of variation in the data set explained by the first five PCs are shown in the SCREE plot in Fig. 3. It was found that the first three PCs accounted for over 97% of the total data variance. This result demonstrates again that the significant data variation can be described by a few informative PCs. The higher-order PCs mainly account for random noise and do not contain diagnostic information. The first four principal component loadings obtained from all 216 spectra are shown in Fig. 4. The loading spectra are offset for better illustration.

To identify further the PCs with diagnostic value from the first four PCs, an unpaired students' *t*-test on the hypothesis that there is an equal mean PC score between the autofluorescence spectra collected from normal tissue and from carcinoma lesions was conducted. The small *p*-value for the first two PCs (p-value $\ll 0.001$) shows that there is significant statistical difference in the PC1 and PC2 scores between normal tissue and carcinoma lesions. The *p*-values for PC3 and PC4 are 0.28 and 0.49, respectively. This means that the differences in scores on PC3 and PC4 from the spectra of normal tissue and carcinoma lesions are not statistically significant. Therefore, PC3 and PC4 do not carry the necessary diagnostic information for the separation of carcinoma lesions from normal tissue. For a PCA algorithm, only the first two PCs should be included.

Some spectral characteristics of PC1 and PC2 can be found in Fig. 4. The PC1 loading has a relatively high negative loading value from 500–550 nm and a high opposite (positive) loading value from 600–650 nm. This suggests that the first principal component mainly captures the difference in the spectral data between the regions of 500–550 nm and 600–650 nm. The spectrum of the PC2 loading shows that it captures the difference between a spectral peak around 500 nm and the signal ranging from 525 to 590 nm. The similarity between the PC2 loading and the blood absorption spectrum in the range from 525 to 590 nm suggests that the effect from the variation in blood content in tissue on the total spectral variance is included in PC2.

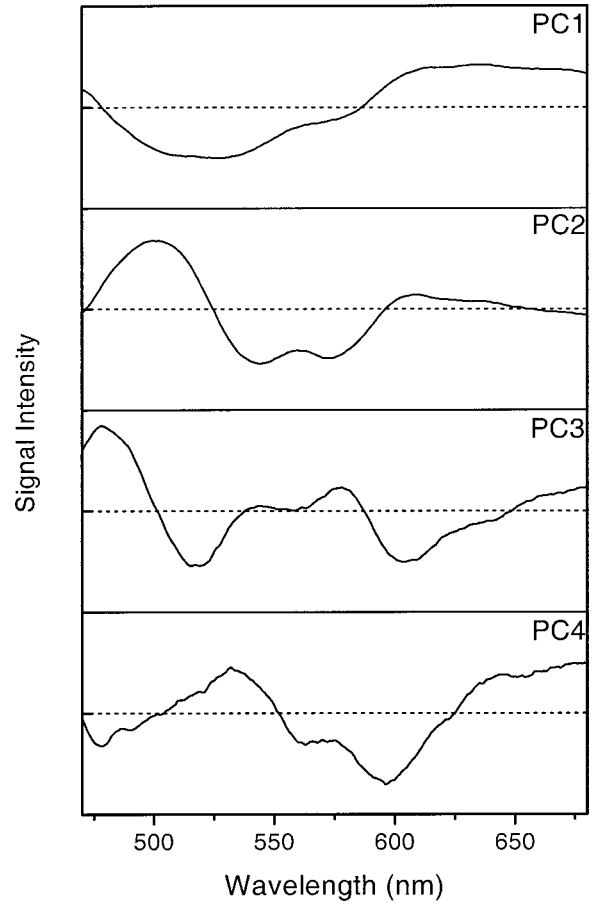

FIG. 4. Spectra of the first four principal component loadings. The dotted lines represent zero lines.

In the calculation of posterior probability for the spectral classification as shown in Eq. 1, the conditional probabilities for the normal tissue group and carcinoma group are determined by fitting a normal probability density function to each PC score. The estimated score distributions of normal tissue and carcinoma lesions for the first four PCs are shown in Fig. 5. The score distributions of normal tissue and carcinoma for PC3 and PC4 almost overlap with each other, which illustrates that PC3 and PC4 do not have diagnostic value. This is consistent with the students' *t*-test results.

The calculated posterior probability using PC1 and PC2 that a given sample belongs to a carcinoma group is shown in Fig. 6. The cost of misclassification was set to 0.6, where the total number of misclassified samples was minimized. Here, a total of 216 spectra were used for the validation. As can be seen, most of the normal tissue and carcinoma lesions are well separated based on the value of the posterior probability that reveals the cer-

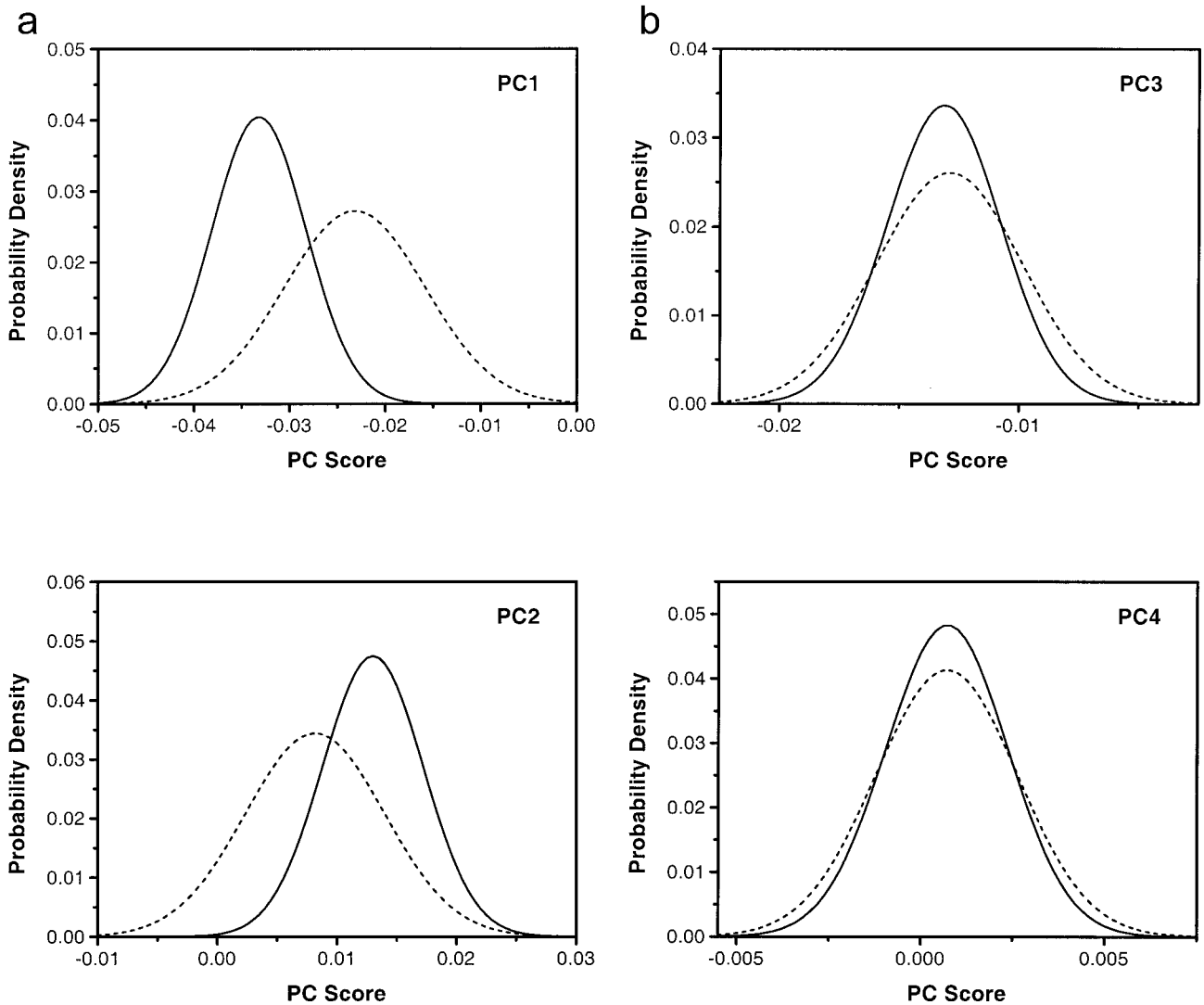

FIG. 5. Estimated score distributions of normal tissue (solid lines) and carcinoma lesions (dashed lines) for the first four PCs.

tainty for a specific sample belonging to the group of normal tissue or carcinoma lesions. Also, we observed that there are a small number of observations that significantly deviate from the majority, which makes it impossible for a diagnostic threshold to be able to separate completely the groups of normal tissue from the carcinoma lesions. In practice, a diagnostic algorithm with the desired sensitivity or specification for classifying the *in vivo* autofluorescence spectra can be obtained by varying the misclassification cost, k. When the total misclassified samples are minimized with setting k to 0.6, the sensitivity and specificity were found to be 94 and 93%, respectively.

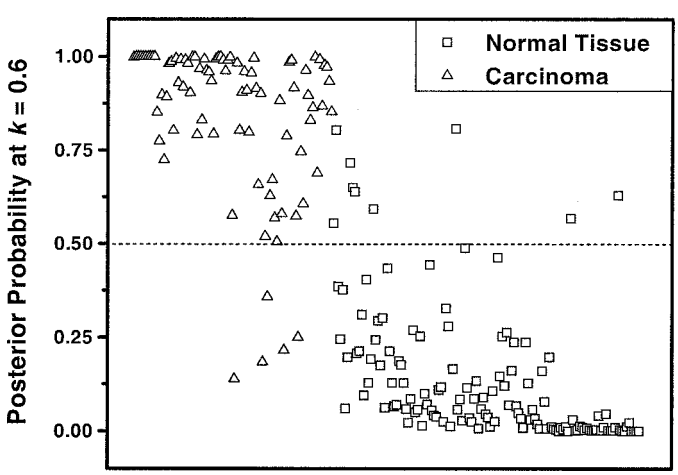

FIG. 6. Posterior probability of a given tissue belonging to the carcinoma group based on PC1 and PC2 scores.

In a simpler approach, we built a threshold-based algorithm with the first two PCs. The scatter plot of the PC1 and PC2 scores is shown in Fig. 7. Again, a total of 216 spectra were used for validation. In the two-dimensional space of PC1 and PC2 scores, a diagnostic threshold line was developed to separate the scores of carcinoma spectra from those of normal tissue and to create a diagnostic algorithm. In detail, the diagnostic line is defined as $Y = aX + b$, where X and Y are the PC1 and PC2 scores. The accuracy for separation of carcinoma from normal tissue is controlled by choosing the parameters of a and b. When a sensitivity is set, an exhaustive search is then performed to determine the values of a and b that maximize the specificity corresponding to

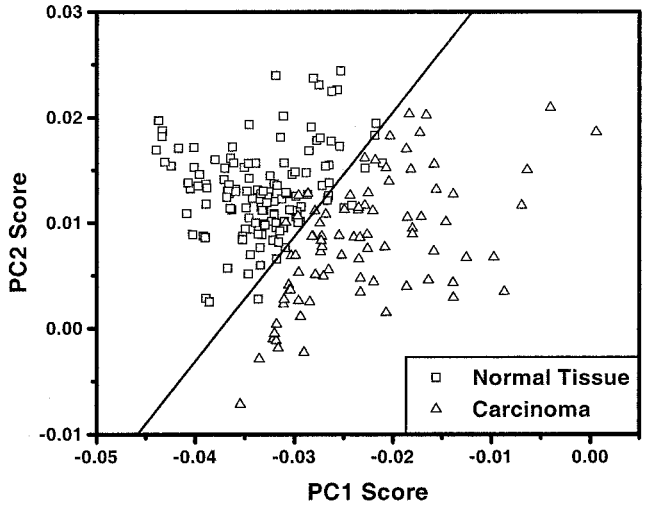

FIG. 7. Scatter plot of PC1 score vs. PC2 score.

a chosen sensitivity. As an example, a diagnostic line classifying the carcinoma spectra at a sensitivity of 95% and a specificity of 93% is depicted in Fig. 7.

In a comparison of the performance of PCA-based algorithms with 2-λ and 3-λ algorithms investigated previously in Ref. 4, we studied the characteristics of 216 *in vivo* fluorescence spectra and identified the optimal wavelength bands that separate the carcinoma lesions from normal tissue at maximal accuracy. Briefly, for a 2-λ algorithm, the discrimination between normal and carcinoma tissue was based on the difference in the ratio between the spectral signals at two wavelength bands. An exhaustive search was used to find the optimal wavelength bands that have ratio values with the most significant statistical difference between carcinoma and normal tissue. The statistical difference was evaluated by using an unpaired one-sided students' *t*-test on the ratio values from the groups of carcinoma lesions and normal tissue. The exhaustive search was performed by adjusting the central wavelengths and wavelength bandwidths of the two wavelength bands and searching for a pair of wavelength bands that produced the highest *t*-test value. The optimal wavelength bands for the 2-λ algorithm were found to be 505 ± 20 nm and 605 ± 40 nm. For the 3-λ algorithm, the effect of blood absorption on the spectral characteristics is taken into account because blood has strong absorption in the wavelength region of 530–590 nm. The three-wavelength algorithm was created by forming the dimensionless function:

$$R = \frac{I(C1 \pm B1)}{I(C3 \pm B3)}\left[\frac{I(C1 \pm B1)}{I(C2 \pm B2)}\right]^{j} \qquad (3)$$

where $C1 \pm B1$, $C2 \pm B2$, and $C3 \pm B3$ are the central wavelengths and bandwidths of the three wavelength bands, respectively. j is the parameter to scale the effect of the blood signal in band $C2 \pm B2$ to the R-function. Again, an exhaustive search is used to find optimal wavelength bands that make the R-function discriminate the carcinoma lesions from normal tissue with highest accuracy. The optimal wavelength bands and j-value were found to be 510 ± 20, 550 ± 20, and 610 ± 40 nm, and 0.35, respectively.

TABLE I. Comparison of sensitivity and specificity of PCA algorithms and 2- and 3-λ algorithms for the classification of hold-out samples in cross validation.[a]

Sensitivity set for training data	Sensitivity/specificity (%)			
	PCA-P	PCA-T	2-λ	3-λ
97%	97/78	97/84	97/77	97/76
95%	95/89	95/93	94/83	94/83
93%	93/93	93/95	91/89	91/88

[a] PCA-P: probability-based PCA algorithm; PCA-T; threshold-based PCA algorithm; 2-λ: two-wavelength ratio algorithm; 3-λ: three-wavelength *R*-function algorithm.

The results from the PCA-based algorithms and the 2- and 3-λ algorithms for classification of *in vivo* spectral data are shown in Table I. The sensitivity for classifying the spectral data in the training set by each algorithm in each round of cross-validation were set to be 97, 95, and 93%. The maximal specificity corresponding to each sensitivity for the probability-based PCA method was calculated by varying the cost of misclassification in Eq. 1. For the threshold-based PCA algorithm, the slope and Y-axis interception of the diagnostic line were adjusted to maximize the specificity. For 2- and 3-λ algorithms, the maximal specificities were found by setting the thresholds to separate the carcinoma lesions from the normal tissue based on the values of the 2-λ ratio and the 3-λ *R*-function. The sensitivity and specificity for the prediction set shown in Table I were calculated from the classification of the hold-out sample in 216 rounds of cross validation. It should be noted that the PC spectra calculated from the training set in each round of cross validation are almost identical because only one spectrum was held out from 216 spectra. Over 216 rounds of cross validation for the PCA-based algorithms and 2,3-λ algorithms, it was found that the standard deviations of the specificity at a fixed sensitivity obtained from training sets are less than 0.2%.

As can be seen, the PCA-based algorithms demonstrate obvious advantage in classification accuracy over the 2- and 3-λ algorithms because the PCA methods utilize the information carried across the entire fluorescence spectra instead of a few limited numbers of wavelength bands. We also observe that the accuracy produced by the probability-based algorithm PCA is lower than the simple threshold-based algorithm in all instances. This may be caused by the error in estimation of the PC score density distribution. In the estimation of the PC score distributions and the calculation of the joint distribution of multiple PCs, the probability density function of the PC score was assumed as a normal distribution. In a further investigation on the PC score distribution, we conducted a normality test on the scores of PC1 and PC2 using the Q–Q plot correlation coefficient method.[12] To assess the assumption of normality, the Q–Q plots are commonly used. These plots can be made for the distributions of the sample observations on each variable. In this study, the variables are the scores of PC1 and PC2 from carcinoma lesions and normal tissue.

Briefly, Q–Q plots are the plots of the sample quantiles vs. the quantiles that would be expected if the samples are normally distributed. Let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ rep-

resent n observed samples in ascending order. For a standard normal distribution, the quantile of jth observation, $q_{(j)}$, can be calculated numerically by the following relationship:

$$P[Z \leq q(j)] = \int_{-\infty}^{q(j)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = p_{(j)} = \frac{j - \frac{1}{2}}{n} \quad (4)$$

where $p_{(j)}$ is the probability of getting a value less than or equal to $q_{(j)}$ from a standard normal population. A Q–Q plot can then be made of the ordered data, x, against the normal quantiles, q. When the observations arise from a normal population, the pairs of points $(q_{(j)}, x_{(j)})$ will be approximately linearly related and the pairs of points $(q_{(j)}, x_{(j)})$ lie very nearly along a straight line. The straightness of the Q–Q plot can be measured by calculating the correlation coefficient of the points in the plot. The correlation coefficient for the Q–Q plot is defined by:

$$r_Q = \frac{\sum_{j=1}^{n} (x_{(j)} - \bar{x})(q_{(j)} - \bar{q})}{\sqrt{\sum_{j=1}^{n} (x_{(j)} - \bar{x})^2} \sqrt{\sum_{j=1}^{n} (q_{(j)} - \bar{q})^2}} \quad (5)$$

and the critical points for these Q–Q plot correlation coefficients, which are defined as the hypothesis that normality at significance level is rejected if r_Q falls below the appropriate values, can be used as a test of normality.[12] In the normality test on the PC1 and PC2 scores, it is found that at high significance levels, the normality assumption holds for the PC1 scores of normal tissue and for the PC2 scores of both normal tissue and carcinoma lesions. However, the normality assumption does not hold for the PC1 scores of the carcinoma lesions. This indicates that the PC1 scores for the carcinoma lesions are not normally distributed. Therefore, the assumption of normal distribution for the PC1 scores from the carcinoma tissue may contribute to the misclassification of the fluorescence spectral data. In contrast, the threshold-based algorithm is independent of the distribution of PC scores. This may explain why the diagnostic accuracy of the threshold-based algorithm is higher than that of the probability-based algorithm.

We also noticed that there is no improvement in classification accuracy from the 3-λ algorithm over that from the 2-λ algorithm. The performances of the 2- and 3-λ algorithms developed in this work are poorer in comparison with our previous work in Ref. 4. The optimal wavelength bands are also slightly different from the previous algorithms. This may be due to the difference in the population of the subjects involved in the research. In previous work, all the subjects were patients for follow-up examinations and all of them were diagnosed as having nasopharyngeal cancers at different stages. In this extended study, about 50% of the subjects were new patients and 19 out of the total of 59 subjects enrolled in this study were found not to have nasopharyngeal cancer. In principle, the 2- and 3-λ algorithms developed in the current study should be more reliable and representative because they are based on measurements from both normal subjects and cancer patients from a larger population.

CONCLUSION

We investigated the characteristics of *in vivo* autofluorescence spectra collected from the nasopharynx using the PCA method. The diagnostic algorithms for classification of spectral data were developed based on the PC scores of the fluorescence signals collected from carcinoma lesions and normal tissue. The results of the study reported here demonstrate that PCA-based algorithms can detect nasopharyngeal carcinoma lesions with high sensitivity and specificity. In a comparison with the previously reported two-wavelengths and three-wavelengths algorithms, the PCA method can produce higher diagnostic accuracy because it takes advantage of the diagnostic information carried across the entire spectra instead of from a few wavelength bands. It was found that the threshold-based PCA algorithm performs better than the probability-based PCA algorithm. One of the sources of misclassification in the probability-based algorithm may be from the assumption that PC scores are normally distributed. Overall, our results indicate great promise for autofluorescence-spectroscopy-based detection of small/flat lesions, which would be a big improvement in guiding biopsies and diagnosing small/flat carcinoma lesions in the nasopharynx.

ACKNOWLEDGMENT

The authors acknowledge support from the Hong Kong Research Grants Council, grant HKUST6052/00M.

1. R. Richards-Kortum and E. Sevick-Muraca, Ann. Rev. Phys. Chem. **47,** 555 (1996).
2. S. Andersson-Engels, C. Klinteberg, K. Svanberg, and S. Svanberg, Phys. Med. Biol. **42,** 815 (1997).
3. G. A. Wagnieres, W. M. Star, and B. C. Wilson, Photochem. Photobiol. **68,** 603 (1998).
4. J. Y. Qu, P. W. Yuen, Z. J. Huang, D. Kwong, J. Shan, S. J. Lee, W. K. Ho, and W. I. Wei, Lasers Surg. Med. **26,** 432 (2000).
5. K. M. O'Brien, A. F. Gmitro, G. R. Gindi, M. L. Stetz, F. W. Cutruzzola, L. I. Laifer, and L. L. Deckelbarm, IEEE Trans. Bio. Eng. **36,** 424 (1989).
6. N. Ramanujam, M. F. Mitchell, A. Mahadevan, S. Thomsen, A. Malpica, T. Wright, N. Atkinson, and R. Richards-Kortum, Lasers Surg. Med. **19,** 46 (1996).
7. E. B. Hanlon, I. Itzkan, R. R. Dasari, M. S. Feld, R. J. Ferrante, A. C. McKee, D. Lathi, and N. W. Kowall, Photochem. Photobiol. **70,** 236 (1999).
8. C. Eker, R. Rydell, K. Svanberg, and S. Andersson-Engels, Lasers Surg. Med. **28,** 259 (2001).
9. C. Y. Wang, C. T. Chen, C. P. Chiang, S. T. Young, S. N. Chow, and H. K. Chiang, Photochem. Photobiol. **69,** 471 (1999).
10. J. E. Jackson, *A User's Guide to Principal Components* (John Wiley and Sons, New York, 1991).
11. B. W. Silverman, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, New York, 1986).
12. A. J. Richard and W. W. Dean, *Applied Multivariate Statistical Analysis* (Prentice Hall, Englewood Cliffs, NJ, 1998).