Discrimination of lung cancer by serum using fluorescence and principal component analysis

Su Zhang[a], Xiaozhou Li[b*], Tianyue Yang[b] and Siqi Li[c]

[a]Section of Physiology and Biochemistry, Shenyang Sport University, Shenyang, China
[b]School of Science, Shenyang Ligong University, Shenyang, China
[c]Integrated Life Sciences, Kent State University, Kent, US America

_____

ABSTRACT

The technology of laser-induced auto-fluorescence spectroscopy was used on serum for the diagnosis of lung cancer. Serum from 30 lung cancer patients and 20 healthy people were collected and measured. The results have shown that there was significant difference in the fluorescence spectroscopy between those from lung cancer patients and the controls. Then, we use principal component analysis and discriminant analysis to analyze spectra, and got an accuracy of 88% in distinguishing lung cancer patients and healthy people. Our experiment revealed that fluorescence of serum can be an indicator for the diagnosis of lung cancer.

Keywords: fluorescence spectroscopy, serum, lung cancer

_____

INTRODUCTION

According to the estimation of the 21st World Anti-cancer Conference hold in Shenzhen China, there were about 2.6 million incidences of cancer, and 1.8 million mortality each year. The mortality rate of cancer has increased 80% in the past 30 years. Cancer has become the first cause of death in urban and rural China, and lung cancer has been the first killer in China[1]. Lung cancer is the most common primary malignant tumor, most of lung cancer are stemmed from epithelial of bronchial mucosa, so lung bronchogenic carcinoma is another name of lung cancer. Nowadays, researches on lung cancer tissues, cells, serum and saliva using spectroscopy have been reported[2–8], but none has been used clinically.

Serum is one of the most important detect target on the analysis of cancers. There are many tumor markers exited in serum. Nowadays, many analyzing methods are being used on the analysis of serum, such as ELISA, HPLC, and mass spectrometry. However, most of them are time-consuming, expensive and hard to operate. Fluorescence spectroscopy is highly sensitive to the changes of tissues and cells, so it is recognized as an important means of detection[9].

In this study, we use laser induced auto-fluorescence to explore its capability in analyzing serum. A wavelength of 488 nm laser created by Ar ion laser were used to analyze serum from lung cancer patients and healthy people, then the data were analyzed by the principal component analysis (PCA) and linear discriminant analysis (LDA) to find the difference between groups. The results showed that auto-fluorescence of serum can provide some assistance for the prophylaxis, early diagnosis, and treatment for lung cancer.

EXPERIMENTAL SECTION

50 serums were investigated in which 30 were from lung cancer patients, 20 from healthy people, all the patients have been diagnosed lung cancer. Venous blood of volunteers were collected before meals, after 30 min

_____

self-aggregation the blood were centrifuged for 10 min at a speed of 3000 rpm, then the supernatant fluid were sucked as serum for further spectral analysis. All the serum samples were store in 4°C and were measured spectroscopy within one week.

Experimental setup is shown in Figure 1, laser beam produced by Ar ion laser irradiate in the sample cell after the modulation of a chopper. Fluorescence were collected in a double-slit monochrometer through a focusing lens, and then detected by PMT. The fluorescence signals were amplified by a lock-in amplifier and were input into computer by a A/D card.
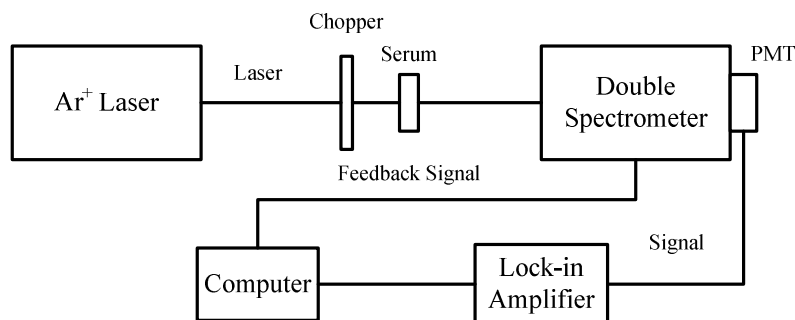

Figure 1. Experimental setup of spectroscopy system

RESULTS AND DISCUSSION

Wavelength of 488.0 nm were used to excite the serum sample, scanning range were 520 - 620 nm, collection interval was 1 Å, time delay was 0 ms.

The fluorescence intensity of the original spectroscopy varies according to different samples, so normalization was used, and then the spectra were drawn using Origin software (Figure 2). There existed two fluorescence peaks and one wave trough in both of the two groups, but the one from lung cancer patients were smoother.
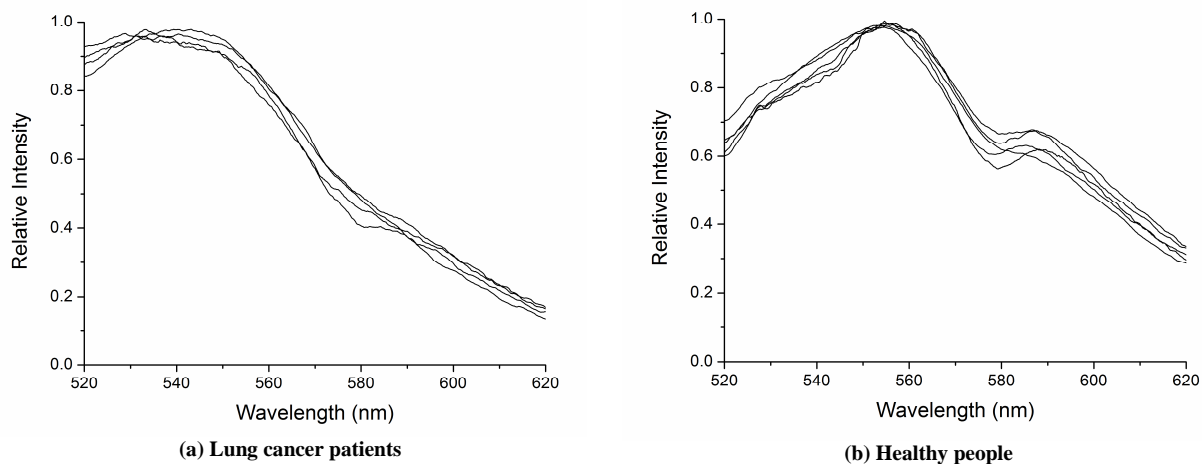

(a) Lung cancer patients   **(b) Healthy people**

Figure 2. Auto-fluorescence spectroscopy of serum from lung cancer patients and healthy people

Table 1. Auto-fluorescence spectral data of 50 serum samples (lung cancer patients and healthy people)

	F1	F2	F3	F4	F5	W_{pos} (nm)	F_{pos} (nm)
Lung1	1.80642	2.01449	1.82609	1.58075	0.87578	578.6	538.6
Lung2	2.16826	2.36252	2.33242	2.24077	1.21067	580.4	534.8
Lung3	2.58663	2.57297	2.45866	2.35083	1.10208	579.8	530.0
……	……	……	……	……	……	……	……
Health20	1.33184	1.72719	1.70642	1.12967	0.97138	576.6	554.8

W_{pos} the the position of wave trough, F_{pos} is the position of the main fluorescence peak

Then we find the peak positions. The main peak from lung cancer patients was around 538 nm, and that of normal people was around 555 nm. Both of the two groups had a wave trough at wavelength of about 578 nm. We use PCA and represent the features of spectroscopy by several principal components (PCs), then discriminate the two groups

through LDA. We define the fluorescence intensity ratio $Fi = I_i/I_t$ (i is the intensity of the ith fluorescence peak) for further comparison (Table 1).

Analyze F_i by PCA using SPSS software, the amount of the eigenvalue of each PC can represent the amount of information the PC carries. The eigenvalue difference between factor 1 and 2, 2 and 3 were bigger, so the first two factors can maintain most of the information. A scattering picture using the PC score was drawn (Figure 4), spot of lung cancer patients and healthy people can be discriminated well.

There were 26 cases under the diagnostic line, and 18 above. We got a diagnostic sensitivity of 86.7%, a specificity of 90%, and an accuracy of 88%.
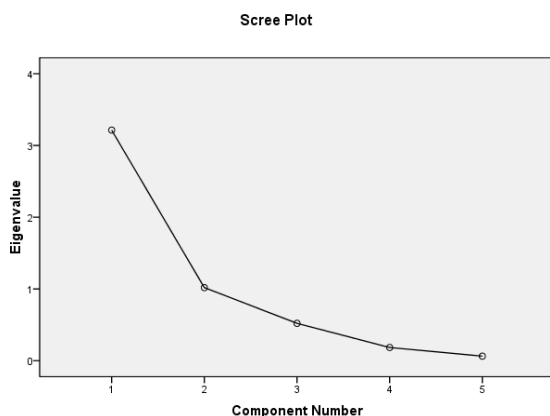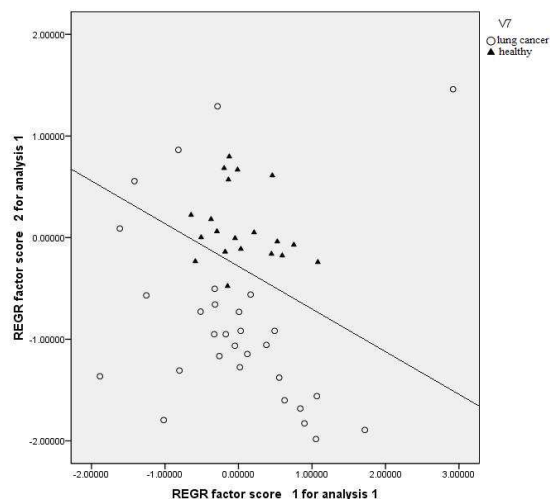

Figure 3. Scree plot of eigenvalue of PC1 to PC 5

Figure 4. Scattering plot of PCA score

Discriminant analysis was used on the PCA Score1 and PCA Score2 obtained by PCA, Table 2 was the results.

Table 2. Coefficients of Bayes Fisher's discriminant function

	Type	
	Lung cancer patients	Healthy people
PCA1	.101	-.151
PCA2	-.175	.263
(Constant)	-.714	-.739

Table 3. Results of discrimimant analysis

		Type	Predicted Group Membership		Total
			Lung cancer	Healthy	
Original	Count	Lung cancer	27	3	30
		Healthy	3	17	20
	%	Lung cancer	90.0	10.0	100.0
		Healthy	15.0	85.0	100.0
Cross-validated(a)	Count	Lung cancer	27	3	30
		Healthy	3	17	20
	%	Lung cancer	90.0	10.0	100.0
		Healthy	15.0	85.0	100.0

Bayes discriminant function can be obtained from the coefficients in Table 2,

$$y_1 = 0.101x_1 - 1.75x_2 - 0.714$$

$$y_2 = -0.151x_1 + 0.263x_2 - 0.739$$

In which x_1 is PCA1 score, x_2 is PCA2 score.

Using the PCA score of one serum sample, and calculate the score in different category, then using the discriminant score to determine which category it belong to, it belongs to the category with the highest discriminant score.

Table 3 shows the right prediction number, false prediction number, fault discriminant rate, and results of cross-validation discriminant function of the established discriminant function based on the 50 serum samples. We can see that the results of right categorized and cross-validation of discriminant function are the same: the false diagnostic rate for lung cancer is 10%, the false diagnostic rate for healthy people is 15%, and accuracy is 88%. The results are identical with the one got by previous scattering plot.

CONCLUSION

Serum of lung cancer patients and healthy people were detected auto-fluorescence spectroscopy for the discrimination of the two groups. On the basis of the obtained PCs, using scattering plot and discriminant analysis to analyze the fluorescence of the two groups, an accuracy of 88% was obtained for both the lung cancer patients and healthy people. So, the combination of scattering plot and discriminant analysis can better discriminate the lung cancer group and the control group. This finding may provide some help for the prophylaxis, early diagnosis, and treatment of lung cancer.

REFERENCES

[1] J Langwith. Lung Cancer, Greenhaven Press, New York, **2010**, Vol. Chap. 6.
[2] Y Oshima; H Shinzawa; T Takenaka; C Furihata; H Sato, *J Biomed Opt*, **2010**, 15(1), 017009.
[3] ND Magee; JS Villaumie; ET Marple; M Ennis; JS Elborn; JJ McGarvey, *J Phys Chem B*, **2009**, 113(23), 8137–8141.
[4] A Risch; C Plass, *International Journal of Cancer*, **2008**, 123(1), 1–7.
[5] YK Min; T Yamamoto; E Kohda; T Ito; H Hamaguchi, *Journal of Raman Spectroscopy*, **2005**, 36(1), 73–76.
[6] DM Parkin; F Bray; J Ferlay; P Pisani, *CA Cancer J Clin*, **2005**, 55(2), 74–108.
[7] Z Huang; A McWilliams; H Lui; DI McLean; S Lam; H Zeng, *International journal of cancer*, **2003**, 107(6), 1047–1052.
[8] X Li; T Yang; J Ding, *Spectroscopy and Spectral Analysis*, **2012**, 32(2), 391–393 (in Chinese).
[9] R Kalaivani; V Masilamani; K Sivaji; M Elangovan; V Selvaraj; SG Balamurugan; MS Al-Salhi, *Photomed Laser Surg*, **2008**, 26(3), 251–256.