Front. Optoelectron. 2015, 8(2): 183–186
DOI 10.1007/s12200-015-0498-7

RESEARCH ARTICLE

Diagnostics of bronchopulmonary diseases through Mahalanobis distance-based absorption spectral analysis of exhaled air

A. A. BULANOVA (✉)[1], E. B. BUKREEVA[1], Yu. V. KISTENEV[1,2], O. Yu. NIKIFOROVA[3]

1 Siberian State Medical University, Tomsk 634050, Russia
2 Tomsk State University, Tomsk 634050, Russia
3 V.E. Zuev Institute of Atmospheric Optics SB RAS, Tomsk 634055, Russia

Abstract Accurate diagnosis of different bronchopulmonary diseases is important in clinical practice. This study involved 20 healthy volunteers and 77 patients with bronchopulmonary diseases, including chronic obstructive pulmonary disease (COPD), bronchial asthma, pulmonary tuberculosis, and community-acquired pneumonia. The absorption spectrum of exhaled air samples was recorded on an intra-cavity photo-acoustic gas analyzer (ILPA-1, Special Technologies, Ltd., Russia) with photo-acoustic detectors and CO_2 laser with a tuning range from 9.2 to 10.8 μm. In conclusion, analysis of the Mahalanobis distance-based absorption spectral profiles of breath air from bronchopulmonary patients and healthy volunteers allows the formulation of a preliminary diagnosis.

Keywords bronchopulmonary diseases, exhaled air, Mahalanobis distance, laser photo-acoustic spectroscopy, CO_2 laser

1 Introduction

Accurate diagnosis of different bronchopulmonary diseases is important in clinical practice, particularly when conventional examination methods are inapplicable for certain patients, such as children, pregnant women, and extremely severe patients [1].

Many studies explored the relationship between the composition of exhaled air and various diseases. The exhaled air of patients with chronic obstructive pulmonary disease (COPD) differs from that of healthy individuals in terms of the composition of volatile organic compounds (VOCs) [2,3]. Van Berkel et al. [3] selected six types of VOCs that exhibit 92% sensitivity toward COPD diagnostics. Phillips et al. [4] created a diagnostic model based on VOC profile analysis to distinguish COPD patients from healthy individuals and achieved 64% diagnostic accuracy. The diagnostic accuracy of COPD patients improved up to 74% after smokers were excluded from the patients with COPD. Moreover, the amounts of hydrogen cyanide and isoprene in exhaled air were analyzed to identify patients with community-acquired pneumonia [5]. Patients with pulmonary tuberculosis were also determined on the basis of the absorption spectral profiles of exhaled air [6].

VOCs are not strictly specific for any disease, thereby complicating the practical application of VOC measurement for diagnostics. Bronchial asthma significantly increases the amount of exhaled NO and moderately augments the amount of CO, whereas COPD slightly increases NO and markedly increases CO [7]. Thus, using the set of VOCs or the direct absorption spectral profile of breath samples as a fingerprint of bronchopulmonary diseases is convenient. In the latter case, various distant metrics should be used to estimate the "difference" between spectra. Distance metrics with good quality can identify important features and discriminate relevant and irrelevant features. Specifying which pairs of data points are similar or dissimilar is important in biomedical analysis. The Mahalanobis distance is a measure between two data points in the space defined by relevant features by assigning corresponding weights to the features of data points.

This study aims to compare the absorption spectral profiles of breath air from patients with various bronchopulmonary diseases and from healthy volunteers on the basis of Mahalanobis distance.

Received December 9, 2014; accepted March 10, 2015

E-mail: anjuta107@gmail.com

The absorption spectra of breath air from patients with various bronchopulmonary diseases and from healthy volunteers are visually similar. We used the Mahalanobis distance to quantify the similarity or difference in the absorption spectra of breath air from patients with various bronchopulmonary diseases and from healthy volunteers. Previous similar studies [6] used this method to separate patients with pulmonary tuberculosis from healthy volunteers and patients with other diseases. In the present study, the same equipment was used, but breath air samples were obtained from a different group of patients.

2 Methods

The study involved 20 healthy volunteers and 77 patients with bronchopulmonary diseases, including COPD, bronchial asthma, pulmonary tuberculosis, and community-acquired pneumonia. Details regarding participants are presented in Table 1.

Samples of exhaled air were collected in the morning (08:00 am to 09:00 am) on an empty stomach before taking inhaled drugs and after three to five times of mouth rinsing with boiled water. The smokers refrained from smoking for a minimum of 6 h before sampling. Exhaled air was collected into a preliminary sterilized glass tube with dense cotton-gauze tubes. Each participant exhaled in a relaxed manner for one to two times into the tube with lips tightly clasped. Then, the tube was tightly closed with sterile cotton-gauze plugs.

The absorption spectrum of exhaled air samples was recorded on an intra-cavity photo-acoustic gas analyzer (ILPA-1, Special Technologies, Ltd., Russia) with a photo-acoustic detector and CO_2 laser with a tuning range from 9.2 to 10.8 μm [8]. Three samples of exhaled air were obtained from every participant. The absorption spectrum of each sample was recorded for five times to reduce random error.

The measured spectra for participants between the test group S (patients with bronchopulmonary diseases) and the reference group S_0 (healthy volunteers) were compared on the basis of the Mahalanobis distance. The feature vectors of the participants from groups S and S_0 were $\boldsymbol{y}_j$, $j = \overline{1,N_S}$ and $\boldsymbol{x}_i$, $i = \overline{1,N_{S_0}}$, correspondingly. N_S and N_{S_0} are the total quantities of feature vectors that correspond to all participants in the group. Thus, the specific average square of Mahalanobis distance can be defined as

$$I_{S_0}(\boldsymbol{y}_j) = \frac{1}{2mN_{S_0}} \sum_{i=1}^{N_{S_0}} d_M^2(\boldsymbol{y}_j,\boldsymbol{x}_i),$$

where $d_M(\boldsymbol{x},\boldsymbol{y}) = \sqrt{(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{y})}$ is the Mahalanobis distance, $\boldsymbol{C}$ is the covariance matrix of the features of participants from group S_0 [9], and m is the dimension of the feature space.

The sets of absorption coefficients of exhaled air from patients with bronchopulmonary diseases and from healthy volunteers were used as feature vectors $\boldsymbol{y}_j$ and $\boldsymbol{x}_i$, correspondently. We marked the specific average square of the Mahalanobis distances of exhaled air absorption spectra in the 10P and 10R spectral bands of CO_2 laser generation for participants as I_1, I_2, correspondingly. We selected the 10P and 10R spectral bands because the measurement error was small in these bands.

3 Results

We used healthy volunteers as the reference group. The values of I_1, I_2 were calculated for every patient with bronchopulmonary disease. On the basis of the calculation results, the values of I_1, I_2 were not subjected to the law of normal distribution. Thus, we used the median and quartile values (25% and 75%, respectively) for analysis (Table 2). To compare the values of I_1, I_2 for various groups, pairwise statistical analysis in terms of Mann–Whitney coefficients was carried out. Statistical significance was considered at $p < 0.05$. Considering that pneumonia and tuberculosis are urgently arising lung diseases, we combined the values of I_1, I_2 for these diseases in the joint group of "urgent lung diseases" (ULD).

The analysis of exhaled air absorption spectrum in the 10P spectral band revealed a significant difference ($p \leqslant 0.01$) in I_1 between the patients with pulmonary diseases and the healthy volunteers. A significant difference ($p = 0.097$) in I_1 was also observed between the COPD patients and the healthy volunteers in the 10R spectral band (Table 2).

Table 1 Information about the groups

group	gender and number		age	total number in the group
	male	female		
healthy volunteers	5	15	26.90±6.96	20
patients with COPD	27	4	61.90±8.14	31
patients with bronchial asthma	3	13	59.30±12.85	16
patients with tuberculosis	8	2	60.0±5.67	10
patients with pneumonia	10	10	41.85±17.60	20

Table 2 Values of I_1, I_2 in the groups

parameter	healthy volunteers		patients with ULD		patients with bronchial asthma		patients with COPD		p value					
	1		2		3		4							
	N	median (25%–75%)	N	median (25%–75%)	N	median (25%–75%)	N	median (25%–75%)	p_{12}	p_{13}	p_{14}	p_{23}	p_{24}	p_{34}
I_1	20	1.11 (0.86–1.32)	30	3.96 (2.59–28.33)	16	3.37 (2.30–6.45)	31	1.56 (1.18–2.26)	0.001	0.001	0.002	1	0.001	0.006
I_2	20	1.03 (0.86–1.38)	30	2.81 (1.86–4.71)	16	2.61 (1.90–4.28)	31	1.26 (1.09–1.82)	0.001	0.001	0.097	1	0.001	0.001

Notes: N is the number of participant in the group;
p_{12} is the p value using the Mann–Whitney test in comparing patients with ULD and healthy volunteers;
p_{13} is the p value using the Mann–Whitney test in comparing patients with bronchial asthma and healthy volunteers;
p_{14} is the p value using the Mann–Whitney test in comparing patients with COPD and healthy volunteers;
p_{23} is the p value using the Mann–Whitney test in comparing patients with bronchial asthma and ULD volunteers;
p_{24} is the p value using the Mann–Whitney test in comparing patients with COPD and ULD volunteers;
p_{34} is the p value using the Mann–Whitney test in comparing patients with COPD and bronchial asthma volunteers.

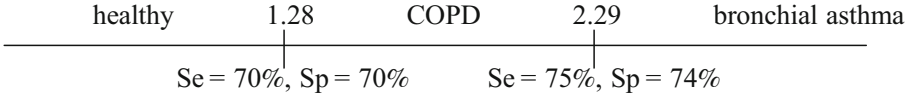
Table 3 Diagnostic intervals of I_1, sensitivity, and specificity of the method

pairwise classification	threshold value of I_1	target disease	sensitivity (Se)/%	specificity (Sp)/%
healthy volunteers – patients with ULD	$\geqslant 1.74$	ULD	90	90
healthy volunteers – patients with bronchial asthma	$\geqslant 1.66$	bronchial asthma	90	81
healthy volunteers – patients with COPD	$\geqslant 1.28$	COPD	70	70
patients with ULD – patients with COPD	$\leqslant 2.45$	COPD	80	80
patients with bronchial asthma – patients with COPD	$\leqslant 2.29$	COPD	75	74

As shown in Table 2, I_1 significantly differed between the patients with bronchial asthma and COPD ($p < 0.001$), as well as between the patients with ULD and COPD ($p = 0.006$). Similarly, I_2 significantly differed between the healthy and asthmatic patients, healthy participants and ULD patients, ULD and COPD patients, and asthmatic and COPD patients (all at $p < 0.001$). The values of I_1, I_2 did not allow distinguishing patients with bronchial asthma from those with ULD.

Furthermore, we calculated the threshold values of I_1 and estimated the sensitivity and specificity of the differential diagnostics within the group of bronchopulmonary diseases using receiver operating characteristic analysis (Table 3).

The results shown in Table 3 can be presented in the form of diagnostic rules:



Therefore, COPD can be expected in 70% of cases when I_1 is between 1.28 and 2.29, whereas bronchial asthma can be expected in 75% of cases when I_1 value exceeds 2.29. ULD was not included in the current scheme because no results were obtained for ULDs in comparison with bronchial asthma in the 10P and 10R bands.

4 Conclusions

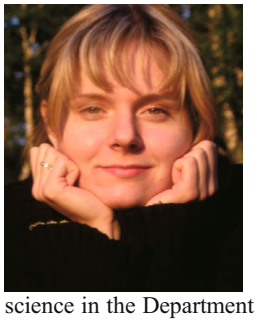
Analysis of the Mahalanobis distance-based absorption spectral profiles of breath air from bronchopulmonary patients and healthy volunteers allows the formulation of a preliminary diagnosis.

Acknowledgements The work was carried out with the partial financial support of the Federal Special Purpose Program contract No 14.578.21.0082 (ID RFMEFI57814X0082).

References

1. Bukreeva E B, Bulanova A A, Kistenev Y V, Kuzmin D A, Tuzikov S A, Yumov E L. Analysis of the absorption spectra of gas emission of patients with lung cancer and chronic obstructive pulmonary disease by laser optoacoustic spectroscopy. In: Proceedings of SPIE 8699, Saratov Fall Meeting 2012: Optical Technologies in Biophysics and Medicine XIV; and Laser Physics and Photonics XIV. 2013, 86990K
2. Bessa V, Darwiche K, Teschler H, Sommerwerck U, Rabis T, Baumbach J I, Freitag L. Detection of volatile organic compounds

(VOCs) in exhaled breath of patients with chronic obstructive pulmonary disease (COPD) by ion mobility spectrometry. International Journal for Ion Mobility Spectrometry, 2011, 14(1): 7–13
3. Van Berkel J J B N, Dallinga J W, Möller G M, Godschalk R W L, Moonen E J, Wouters E F M, Van Schooten F J. A profile of volatile organic compounds in breath discriminates COPD patients from controls. Respiratory Medicine, 2010, 104(4): 557–563
4. Phillips C O, Syed Y, Parthaláin N M, Zwiggelaar R, Claypole T C, Lewis K E. Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls. Journal of Breath Research, 2012, 6(3): 036003
5. Boshier P R, Mistry V, Cushnir J R, Curtis S, Elkin S, Kon O M, Marczin N, Hanna G B. Analysis of volatile biomarkers within exhaled breath for the diagnosis of pneumonia. Thorax, 2010, 65 (Suppl 4): A58–A59
6. Ageev B G, Kistenjov J V, Nekrasov E V, Nikiforova O J, Nikotin E S, Nikotina G S, Ponomarjov J N, Urazova O I, Filinjuk O V, Fokin V A, Janova G V. Estimate of expired air samples of patients with the pulmonary tuberculosis using laser photoacoustic spectroscopy technique. Bulletin of Siberian Medicine, 2012, 4: 116–120
7. Kharitonov S A, Barnes P J. Exhaled markers of pulmonary disease. American Journal of Respiratory and Critical Care Medicine, 2001, 163(7): 1693–1722
8. Intracavity laser opto-acoustic sensor ILPA-1. Passport. Technical description. Operating Instructions. Special Technologies, Ltd, Russia, Novosibirsk
9. Kistenev Y V, ed. Applications of laser spectroscopy and nonlinear analysis methods for investigation of medical-biological objects. Tomsk: TPU Ed., 2007



A. A. Bulanova is the leading specialist in the pulmonology department of the State Clinical Hospital No. 3 of Tomsk City. She graduated from the Medical Faculty of the Siberian State Medical University (Tomsk, Russia) in 2009 and obtained a post-graduate specialty in pulmonology.

In 2011, she took the first step into science in the Department of Therapy in the Siberian State Medical University, where together with her scientific adviser, Prof. Bukreeva. Her research interests are the early diagnosis of chronic obstructive pulmonary disease (COPD) and other bronchopulmonary diseases.

E. B. Bukreeva graduated from the Medical Faculty of the Siberian State Medical University (Tomsk, Russia) in 1975 and received a Ph.D. degree in medical sciences in 1988. In 2005, she received her Ph.D. degree of medical sciences.

She is currently a professor in the therapy department at the pediatric faculty in the Siberian State Medical University. Her research interests are pneumology and clinical pharmacology. She studies etiology of inflammation in chronic obstructive pulmonary disease (COPD), genetic causes of COPD, and problems of early diagnosis of COPD and other bronchopulmonary diseases.

Yu. V. Kistenev graduated from the Radio-Physical Faculty of Tomsk State University, and received his Ph.D. degree in optics.

His work experience: engineer of laboratory of optical waves of Siberian Physical-Technical Institute of Tomsk State University (1982–1984), post-graduate student of the radio-physics department of Tomsk State University (1984–1987), research assistant at the laboratory of optical waves of Siberian Physical-Technical Institute of Tomsk State University (1987–1993), doctoral student of the radio-physics department of Tomsk State University (1993–1996), senior researcher at the laboratory of atmospheric absorption spectroscopy of The Institute of Atmospheric Optics (1996–1999), head of the physical department of Siberian State Medical University (1999–present (2014–part-time work at present)), Deputy Vice Rector for research of Tomsk State University (2014–present).

His research interests are nonlinear and coherence optics, numerical simulation of optical properties of resonant absorbing and inhomogeneous media, fractal analysis, laser spectroscopy of biological probes.

O. Yu. Nikiforova graduated from Tomsk State University and received her Ph.D. degree in physics. She is currently a senior researcher in the Institute of Atmospheric Optics.The present work is related to the modeling of the absorption spectra of gaseous species, analysis of simulated and measured spectra, solution of inverse problems and, in particular, the use spectroscopic methods for medical diagnostics.