Turkish Journal of Electrical Engineering and Computer Sciences

Volume 31 | Number 3 Article 5

5-1-2023

Breath analysis for detection of lung cancer with hybrid sensor-based electronic nose

ÜMİT ÖZSANDIKCIOĞLU

AYTEN ATASOY

Follow this and additional works at: https://journals.tubitak.gov.tr/elektrik

 Part of the Computer Engineering Commons, Computer Sciences Commons, and the Electrical and Computer Engineering Commons

Recommended Citation

ÖZSANDIKCIOĞLU, ÜMİT and ATASOY, AYTEN (2023) "Breath analysis for detection of lung cancer with hybrid sensor-based electronic nose," *Turkish Journal of Electrical Engineering and Computer Sciences*: Vol. 31: No. 3, Article 5. https://doi.org/10.55730/1300-0632.4001
Available at: https://journals.tubitak.gov.tr/elektrik/vol31/iss3/5

This Article is brought to you for free and open access by TÜBİTAK Academic Journals. It has been accepted for inclusion in Turkish Journal of Electrical Engineering and Computer Sciences by an authorized editor of TÜBİTAK Academic Journals. For more information, please contact academic.publications@tubitak.gov.tr.

Turkish Journal of Electrical Engineering & Computer Sciences
http://journals.tubitak.gov.tr/elektrik/
Research Article
Turk J Elec Eng & Comp Sci
(2023) 31: 550 – 565
© TÜBİTAK
doi:10.55730/1300-0632.4001

Breath analysis for detection of lung cancer with hybrid sensor-based electronic nose

Ümit ÖZSANDIKCIOĞLU*, Ayten ATASOY
Department of Electrical and Electronics, Faculty of Engineering, Karadeniz Technical University, Trabzon, Turkey

Received: 08.09.2022 • **Accepted/Published Online:** 13.03.2023 • **Final Version:** 28.05.2023

Abstract: Lung cancer has the highest death rates among all types of cancer worldwide. Detection of lung cancer in its early stages significantly increases the survival rate. In this study, the aim is to improve the lung cancer detection performance of electronic noses (e-noses) with breath analysis by using two different types of gas sensor-based e-nose. The developed e-nose system consists of 14 quartz crystal microbalance (QCM) sensors and 8 metal oxide semiconductor (MOS) sensors. Breath samples were collected from a total of 100 volunteers, including 60 patients with lung cancer, 20 healthy nonsmokers, and 20 healthy smokers, and were classified using decision tree (DT), support vector machine (SVM), k-nearest neighbour (kNN), and random forest (RF) algorithms. Principal component analysis (PCA) and linear discriminant analysis (LDA) algorithms were used for dimension reduction. A classification accuracy of 86.34% and 75.48% was obtained using MOS and QCM sensor data, respectively. The overall results have shown that combining the sensor data increases the accuracy to 88.54%. Additionally, it can be indicated that the PCA and LDA algorithms have a positive effect on the performance. By using PCA and LDA algorithms, the accuracy increased up to 92.67% and 95.36%, respectively.

Key words: Biomedical signal processing, breath analysis, electronic nose, lung cancer diagnosis

1. Introduction

Motivation: Lung cancer is the most important cancer type that causes death all over the world [1]. According to World Health Organization (WHO), the number of people who lose their lives annually due to lung cancer is higher than kidney, bowel, and prostate cancer [2]. For the individuals with lung cancer to be healthy again, the cancer should be detected as early as possible, and patients should be treated immediately. If the lung cancer is detected in advanced stages, the five-year survival rate is around 10%, while this rate rises to 80% if it is detected in early stages [3]. This information reveals the importance of detecting lung cancer at early stages. Since lung cancer has few symptoms in the early stages, medical imaging techniques are the most widely used methods to find cancer-related tumors [4]. For early detection of lung cancer, individuals included in the high-risk group were examined together with many medical methods. These methods are generally salivary cytology, biomarker tumors in the body, examination of the protein structure of the blood, magnetic resonance combinations such as chest tomography and nuclear magnetic resonance (NMR), chest radiography (or chest X-ray), and low-dose computed tomography (CT). Since these methods, which have been used for years, have limited diagnostic features, studies to increase their performance are being carried out. Chest radiography

*Correspondence: umitozsandikci@ktu.edu.tr

This work is licensed under a Creative Commons Attribution 4.0 International License.

(CXR) is one of these methods that have failed to provide a significant decrease in the mortality rate of lung cancer patients in the past few decades [5, 6]. The sensitivity of this method is low and it also produces high false negative values. In addition, the high radiation intensity in the CXR method limits its usage in this area. Mazzane et al. stated that the computer-assisted tomography (CT) method can fix some of the deficiencies of the traditional CXR method, but there are still aspects that need improvement [7]. The CT method produces a large number of false negatives and misses a lot of first-level lung cancer disease states. In addition, because CT detects lesions smaller than 1 mm in diameter, it produces a significant amount of false positive results [8]. To prevent this negative situation, a verification process is performed using positron emission scanning (PET) [9]. However, the high costs of the CT and PET methods limits their use in early detection of lung cancer. Since imaging methods have not achieved much success in detecting lung cancer in its early stages and reducing the death rate, researchers have tended to work on breath analysis.

Proposed approach: Pauling et al. investigated the content of human breath in their study in 1970 and identified more than 200 chemical compounds in human breath using gas chromatography (GC) [10]. After this pioneering work, scientists have continued researching breath analysis. In the past few decades, the content of human breath has been determined in general with the help of medical electronic devices [11]. As a result of the experiments, it was determined that there are basically 3 types of chemicals in the content of human breath: inorganic compounds such as nitrogen monoxide (NO), carbon dioxide (CO_2) and carbon monoxide (CO); volatile organic compounds (VOCs) such as isoprene, ethane, pentane, and acetone; and nonvolatile compounds such as cytokines, isoprostanes, and nitrogen [12]. VOCs form as a result of the functioning of human metabolism and circulate around the body through the bloodstream. These compounds pass tissues according to certain biological rules and diffuse from the bloodstream into the lung air space across the alveolar wall. Then, VOCs are released from the blood with all metabolites prior to exhalation [13]. It has been determined that the number of VOCs in human breath is around 3000 [11]. The types of these VOCs include saturated and unsaturated hydrocarbons, oxygen-containing compounds, sulphur-containing compounds and nitrogen-containing compounds [12]. Saturated hydrocarbons such as ethane and pentane come into existence when the w3 and w6 fatty acids react with reactive oxygen species in the cell membrane [14]. Isoprene is an unsaturated hydrocarbon and a significant portion of isoprene forms as a result of cholesterol synthesis. In addition, a small amount of isoprane originates from the bacteria existing in the body [15]. Oxygen-containing compounds such as acetone, acetaldehyde, methanol and ethanol occur as a result of decarbonization of acetoacetate. Acetaldehyde occurs with the oxidation of ethanol produced in the body. The source of ethanol is bacteria which live in the intestines [16, 17]. Sulphur-containing compounds, such as dimethyl sulfide and methyl and ethyl mercaptans, generally occur as a result of irregularities in the excretory system. For this reason, the number of sulphur-containing compounds increases in the breath of individuals with kidney disease [18]. Nitrogen-containing compounds in human breath are ammonia and di/trimethyl amine. These VOCs are found especially in the breath of individuals with renal disorder [19]. Since VOCs circulate inside the human body and pass a lot of tissues, any disorder in organs causes some changes in the levels of VOCs [20]. Chen et al. summarized the exhaled breath analysis studies in the literature and their study gives important information about breath composition versus some common diseases [21]. It is determined in the literature that the chemicals found in lung disease cases are carbon-based compounds [22], alpha-olefin, styrene, and dodecane-4-methyl [23]. In cases with airway inflammation disease, it is reported that adenosine [24], H_2O_2 [25], NO [26], 8-isoprostaglandin [27] and periosteum [28] are encountered. Studies regarding gastrointestinal disorders have shown that 12

VOCs, including pentanoic acid [29], 2-propenenitrile, furfural, and 6-methyl-5hepten-2-one [30], are found in breath composition. One of the most common chemical is acetone for metabolic disorders and Shin et al. analysed this chemical by using stannic oxide-based gas sensors [31]. Other studies have been performed regarding metabolic disorders and significant findings were obtained [32, 33]. In another study, Moon et al. investigated kidney diseases, one of the most common disease types, and detected H_2S, NH_3 and NO [34]. Also, studies of breath analysis show that the pentane concentration in breath increases in case of febrile and sepsis diseases [35, 36]. High levels of methyl nitrate and acetone compounds are also found in the breath of individuals with diabetes, which is one of the most common diseases all over the world [37, 38]. Peng et al. researched the exhaled breath of people with different cancer types. They identified a number of VOCs, which are present in $>$ 80% of cancer patients and healthy subjects. The number of these VOCs is 33 for LC, 39 for colon cancer (CC), 54 for breast cancer (BC) and 36 for prostate cancer (PC). They determined that for the most appropriate VOCs there was no overlap in abundance between healthy controls and cancer patients. These VOCs for lung cancer are 1-methyl-4-(1-methylethyl)benzene, toluene, dodecane, 3,3-dimethyl pentane, 2,3,4-trimethyl hexane and 1,1-(1-butenylidene)bis benzene. For colon cancer (CC), the determined VOCs are 1,1-(1-butenylidene)bis benzene, 1,3-dimethyl benzene, 1-iodo nonane, [(1,1-dimethyl ethyl)thio]acetic acid and 4-(4-propylcyclohexyl)-4-cyano[1,1- biphenyl]-4-yl ester benzoic acid. In their study, they determined five VOCs associated with breast cancer (BC): 3,3-dimethyl pentane, 2-amino-5-isopropyl-8-methyl-1-azulenecarbonitrile, 5-(2-methylpropyl)nonane, 2,3,4-trimethyl decane and 6-ethyl-3-octyl ester 2-trifluoromethyl benzoic acid. The last cancer type that they studied is prostate cancer (PC). The VOCs whose levels differ in healthy and cancer patients are toluene, 2-amino-5-isopropyl-8-methyl-1-azulenecarbonitrile, p-xylene and 2,2-dimethyl decane [39]. Jia et al. published a review article on lung cancer disease, and they listed VOCs identified as a lung cancer biomarker in four or more studies. These were propanol, isoprene, acetone, pentane, hexanal, toluene, benzene, ethyl benzene, heptane, 2-butanol, styrene, propanal, pentanal, ethanol, butanol, decane, undecane1-hexane, 1,2,4, trimethyl, propylene heptanal [40]. Poli et al. found aldehydes in the exhaled breath of nonsmall cell lung cancer patients (NSCLC) in their study. They revealed that VOC levels in the breath of lung cancer patients significantly differ from healthy people. The exhaled aldehyde levels can be seen in Table 1. (In Table 1 GM, GSD and pM represent geometric mean, geometric standard deviation and concentration level in 10^{-12} Molar, respectively.)

Breath analysis studies focus on identifying and detecting VOCs in the breath that can be disease biomarkers. Because lung cancer detection by breath analysis is a noninvasive, sensitive, and simple method, this subject has attracted substantial attention from researchers. In the literature, it is seen that gas chromatography–mass spectrometry (GC-MS), selected ion flow tube–mass spectrometry, laser absorption spectrometry and infrared spectroscopy are widely used to determine VOCs in the breath. However, these devices are quite expensive and the performed tests are time consuming. In addition, these methods require predensification of breath samples in order to improve performance of VOC detection.

Hardware approaches: E-noses are one of the devices which are used in breath analysis. These devices were developed with inspiration from the human olfactory system. They can detect odours and VOCs with the help of sensor arrays in their structures. E-noses allow us to distinguish simple or complex odours by using artificial intelligence and pattern recognition subsystems in their structure. These devices mainly consist of three units: a sensor unit, an electronic unit, and a pattern recognition unit. To detect target gases with high accuracy, sensors must be selected very carefully. In the literature, gas sensitive materials can be classified as "electrochemical

Table 1. Exhaled aldehyde levels.GM [GSD] and median values (25–75th percentile) [41].

Exhaled Aldehyde	NSCLC	NSCLC		Controls
		Non-/exsmokers	Smokers	
Propanal (pM)	53.6 [1.5] 52.4 (42.4–72.6)	47.9 [1.4] 49.8 (40.7–61.3)	59.3 [1.6] 66.3 (47.8–81.7)	30.6 [2.8] 24.4 (17.1–46.9)
Butanal (pM)	26.2 [1.8] 26.2 (18.7–41.0)	24.5 [1.7] 23.6 (17.9–33.8)	27.9 [2.0] 28.6 (19.1–46.1)	10.9 [2.4] 10.8 (6.9–18.6)
Pentanal (pM)	19.1 [2.4] 17.7 (12.7–42.6)	16.1 [2.4] 17.1 (12.8–22.3)	22.2 [2.5] 20.3 (12.1–49)	7.6 [2.7] 8.2 (4.4–14.7)
Hexanal (pM)	37.3 [1.9] 38.1 (26.6–57.7)	37.1 [1.6] 38.2 (26.7–54.6)	37.5 [2.2] 35.9 (20.9–66.7)	8.5 [2.7] 10.3 (7.0–13.8)
Heptanal (pM)	13.9 [1.8] 16.1 (9.3–21.3)	15.2 [1.6] 15.4 (10.4–21.3)	12.9 [2.1] 17.0 (8.4–24.0)	6.1 [2.0] 6.9 (3.8–10.1)
Octanal (pM)	23.0 [1.7] 23.6 (17.7–33.2)	25.7 [1.5] 26.9 (19.1–33.5)	20.8 [1.9] 22.4 (16.9–34.2)	10.0 [1.8] 11.6 (7.2–16.2)
Nonanal (pM)	44.0 [1.8] 48.2 (31.6–62.5)	50.9 [1.7] 51.7 (32.4–72.1)	38.5 [1.9] 36.5 (31.1–60.1)	12.7 [1.8] 13.3 (7.2–22.7)

principals" and "other principals". Electrochemical principals contain a metal-oxide semiconductor (MOS), conductive polymer composites (CPCs) and carbon nanomaterials (CNs). Other principals contain acoustic wave sensors (AWS) and catalytic sensors [42]. MOS sensors are the most popular gas sensors for e-nose applications. Their high sensitivity, low cost, stability and compatibility with modern electronic devices make them very attractive for e-nose applications. These sensors are coated with sensing materials such as SnO_2, ZnO, CuO, TiO_2, WO_3, and NiO. The sensing materials react with gases at an appropriate temperature. As a result of this reaction, the conductivity of the sensor changes. The change of sensor conductivity is converted to electrical signals by the electrical unit of e-noses and is then transferred to a computer. To maintain the appropriate temperature for sensors, an external heating voltage is applied to MOS sensors. It can be seen in the literature that MOS sensors are used in both commercial and custom made e-noses [43]. Another commonly used gas sensor for e-noses is quartz crystal microbalance (QCM) sensors, which consist of a quartz disk placed between two gold, parallel electrodes. Quartz is a piezoelectric material and when an alternating current is applied to electrodes a mechanical deviation occurs on the quartz. Also if the frequency of the current is equal to the resonant frequency of a QCM sensor, a standing wave is generated in the resonator. This standing wave is related to the mass of the sensor. QCM sensors are coated with a thin film and this film is sensitive to target gases. When particles of a target gas are absorbed by the thin film, the resonant frequency of the sensor changes. This change is calculated and transferred to the computer by the appropriate electronic circuit. By using this frequency deviation, meaningful information can be obtained about the target gases. It can be seen in the literature that QCM sensors are used in both commercial and specially made e-noses [44].

Data analysis approaches: In the literature, significant studies have detected lung cancer with a breath analysis using e-noses. Tirzïte et al. used a Cyranose 320 e-nose device (Sensigent, Baldwin Park, CA, USA). They collected breath samples from 223 healthy and 252 lung cancer patients. They used raw sensor data without preprocessing. In their study, they used logistic regression analysis as a data classifier and they reached

a 91% specificity value for first-stage lung cancer patients [45]. Koronov et al. used an e-nose consisting of MOS sensors in their study, in which 53 healthy volunteers and 65 lung cancer patients participated. They used PCA and logistic regression (LR), as well as the k-NN, random forest (RF), LDA, SVM, and ANN algorithms for data processing and classification. A sensitivity value of 85%–95% and a specificity value of 81.2%–100% were obtained [46]. Krauss and his colleagues used an aeoNose brand e-nose (The eNose Company, Zutphen, The Netherlands) to collect data from 120 LC patients, 33 healthy controls (HC) and 23 chronic obstructive pulmonary disease (COPD) patients. They used the ANN algorithm and obtained 84% sensitivity and 97% specificity values on LC and HC data [47]. Wang and colleagues carried out experiments with an e-nose which included 9 MOS, 3 electrochemical, 1 catalytic combustion and 1 thermal gas sensors. A total of 52 volunteers, including 24 lung cancer patients, 5 volunteers with different respiratory diseases, 10 healthy smokers and 13 nonsmoker healthy volunteers, participated in this study. The dimension of the data was reduced with PCA, LDA, Laplacian eigenmaps (LE), locally linear embedding (LLE) and t-distributed stochastic neighbour embedding (t-SNE) methods. Classification results were obtained as 91.59% accuracy and 91.6% sensitivity by using fuzzy k-NN and SVM classifiers [48]. Kort et al. used the aeoNose commercial e-nose for data collection from 138 subjects with nonsmall cell lung cancer (NSCLC) and 143 controls without NSCLC. In this study, researchers used ANN and multivariate logistic regression algorithms for data classification. They obtained sensitivity values of 94.2%–95.7% and specificity values of 49.0%–59.7% [49]. Chernov et al. collected breath samples from 59 people which included patients with a confirmed diagnosis of respiratory tract cancer. Their sampling device included 14 MOS sensors. To process the data, they used a neural network with perceptron architecture. They obtained a sensitivity value of 90.73% and a specificity value of 61.39% [50]. Huang et al. used a sensor array consisting of carbon nanotubes as a sensing material for lung cancer detection, with 56 volunteers with lung cancer and 188 healthy individuals participating in their study. LDA and SVM algorithms were used for classification. Their results showed that the accuracy values ranged from 85.4% to 92.7% [51]. When studies regarding lung cancer detection using e-nose are examined, it can be clearly seen that most studies include only one type of gas sensor [52]. In our study, MOS and QCM sensors were used together, and thus both conductivity and frequency information were obtained. As seen in the results, using the information from both sensors significantly improves the system performance.

2. Materials and methods

In this study, an e-nose system has been designed to detect lung cancer by analysing human breath samples. The e-nose system used in this study consisted of MOS-type and QCM-type gas sensors. When the literature was investigated, almost all e-noses consist of one type of gas sensor, while usage of a different type gas-sensor based e-nose, especially in breath analysis, is limited. It can be clearly seen in the result section that using different typea of gas sensor data can improve system performance.

2.1. Experimental Setup

A block diagram of the e-nose system that was developed for this study is shown in Figure 1.

The MOS sensors used in this study are commercial-type sensors. To select the MOS sensors, we searched for VOCs which have different concentrations in the breath of lung cancer patients and healthy individuals. MOS sensors used in this study and their target gases are given in Table 2. The second type of sensor used in the experimental system is the QCM sensor. The QCM sensors used in this study were produced by TÜBİTAK

Marmara Research Center (Gebze, Kocaeli, Turkey). The researchers at TÜBİTAK MAM have used AT-cut quartz crystals (Klove Electronics, Westerlo, Belgium). Also the QCM sensors were freed from the remaining solvent in a dry air stream prior to use. The MOS and QCM sensors used in this study can be seen in Figure 2.

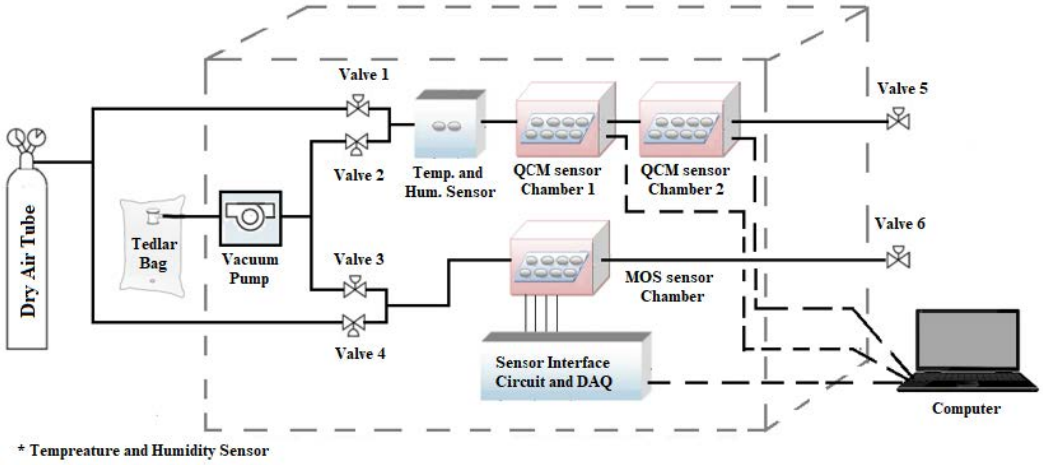

Figure 1. Block diagram of the developed e-nose system.

Table 2. Used sensors and target gases.

Sensors	Target gases
TGS 813	Carbon monoxide, methane, ethanol, propane, isobutane, hydrogen
TGS 816	Carbon monoxide, methane, ethanol, propane, isobutane, hydrogen
TGS 826	Isobutane, hydrogen, ammonia, ethanol
TGS 832	R-12, R134a, ethanol, R-22
TGS 2602	Hydrogen, ammonia, ethanol, hydrogen sulfide, toluene
TGS 2610	Ethanol, hydrogen, methane, isobutane, propane
TGS 2612	Ethanol, methane, isobutane, propane
TGS 2620	Methane, carbon monoxide, isobutane, hydrogen, ethanol

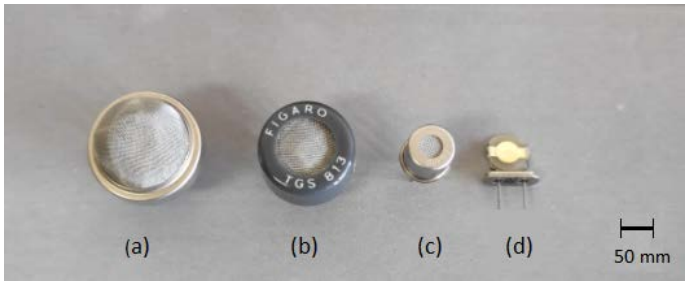

Figure 2. MOS and QCM sensors used in the experimental setup.

Figure 2a–2c represent the TGS826, TGS813 and TGS2620 sensors, respectively. To transfer the dry air and the breath samples in the appropriate direction, six solenoid valves were used. JELPC brand valves (JELPC,

Ningbo, Zhejiang, China) with operating values of 24 V (DC) and 4.8 W were employed. The whole system was cleaned with dry air (21% Oxygen + 79% Nitrogen) before and after the experiments. Breath samples were collected from volunteers by using five-liter volume Tedlar Air Sample Bags (CEL Scientific, Cerritos, CA, USA) and transferred to the sensors with a JH12-65 model vacuum pump. The vacuum pump works on a 12 V (DC) power supply and it has a five-liter/minute flow capacity. Data obtained via the MOS sensors were transferred to the computer using a USB-6218 Multifunction I/O Device (National Instruments, Austin, TX, USA). Data obtained from the QCM sensors were transferred to the computer via a frequency meter circuit which is integrated with the QCM sensor chamber. Experiments with the developed e-nose consist of four stages. In the first stage, the whole system is cleaned with dry air during 130 s and then the valves are closed. After the first cleaning, the breath sample is transferred to the MOS and QCM sensor chambers during 40 s through a vacuum pump. Then, all the valves are closed and the sensors are allowed to respond to the breath samples for 30 s. Finally, the system is cleaned with dry air for 140 s and is ready for the next experiment. The positions of the valves in the e-nose system during an experiment are given in Table 3.

Table 3. Stage of valves during an experiment.

Experiment processes times	State of valves (open-close)					
	V1	V2	V3	V4	V5	V6
Prepurging stage (130 s)	Open	Close	Close	Open	Open	Open
Odor sampling stage (40 s)	Close	Open	Open	Close	Close	Close
Breathing stage (30 s)	Close	Close	Close	Close	Close	Close
Postpurging stage (140 s)	Open	Close	Close	Open	Open	Open

The dry air flow rate was fixed at 10 L/min through a regulated valve of the dry air tube. Thus, the entire system was successfully cleaned before and after the experiment with the same flow rate. By applying constant voltage to the vacuum pump, the breath samples were transferred to the sensor chambers at the same rates. The supply voltage of the vacuum pump was determined as 6 Volts during installation of the e-nose system. At this stage, considering the volumes of the sensor chambers and tedlar bags, the breath samples were delivered to the sensors with a flow rate of 3.6 L/min.

2.2. Collection of breath samples

During this study, a total of 338 experiments were carried out. The breath samples were collected from 40 healthy volunteers and 60 volunteers with lung cancer. LC patients were treated in the Faculty of Medicine, Farabi Hospital, at Karadeniz Technical University. All necessary ethics committee documents were obtained from Karadeniz Technical University before the study and the ethical clearance number is 24237859-517. LC patients participated in this study after their bronchoscopy and biopsy results. The cancer stages of LC patients were determined by the 7th version of the TNM method, which evaluates tumour width (T), regional lymph node metastases (N) and distant metastasis (M), as determined by the American Joint Committee on Cancer (AJCC). The patients did not undergo any surgical operation before breath samples were collected. Patients who were unable to give breath sample were excluded from the study. At least two experiments were performed with the breath samples taken from each volunteer. The healthy volunteers who participated in the study did not have any health complaints and had not been diagnosed with any disease in the last month. In addition, no cases of pregnancy were present in the healthy patients. Half of the healthy volunteers were smokers. Collection

of the breath samples was completely voluntary for both patients and healthy individuals. Information about patients and healthy volunteers is given in Table 4. The breath samples from each volunteer were collected by using tedlar bags and each volunteer used only one tedlar bag.

Table 4. Characterstic of study subjects.

	LC patient (60 Person)	Healthy volunteer (40 Person)
Age (mean/std. deviation)	60.7 / 8	48.2 / 9
Sex (female/male)	13/47	10/30
Smokers/Nonsmokers	0/60	20/20
Exsmokers	45	0

Since tedlar bags do not allow gas leaks from inside or outside, experiments can be carried out without any change in the content of breath samples. The tedlar bags have a polypropylene fitting to collect breath samples and this apparatus provides direction to exhaled breaths. The polypropylene fittings of the tedlar bags can be seen in Figure 3.

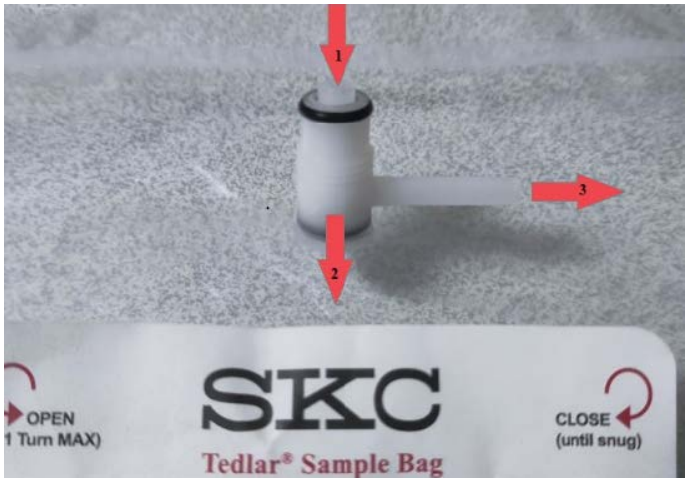

Figure 3. Polypropylene fitting of the tedlar bags.

The breath samples are exhaled along the direction indicated by 1. If it is desired to fill the breath inside the tedlar bag in the direction 2, the pipeline which is indicated by 3 must be closed. If this pipeline is open, breath exhaled along the direction 1 is transferred out of the bag using path 3. To collect alveolar air rather than the upper part of the respiratory system, a timing method was used. In this study, it was assumed that a deep breath takes 6 s. Therefore, as the volunteers exhaled, the pipeline indicated by 3 remained open for 3 s and then closed. Thus, the first part of the breath was discarded and only the breath from the lungs was collected. Volunteers participating in the experiments did not eat anything during the night and gave breath samples while they were hungry. Smoking volunteers also did not smoke for 2 h prior to the test. Lung cancer cell type and cancer stage of the lung cancer patients participating in this study are given in Figure 4 as circular graphics.

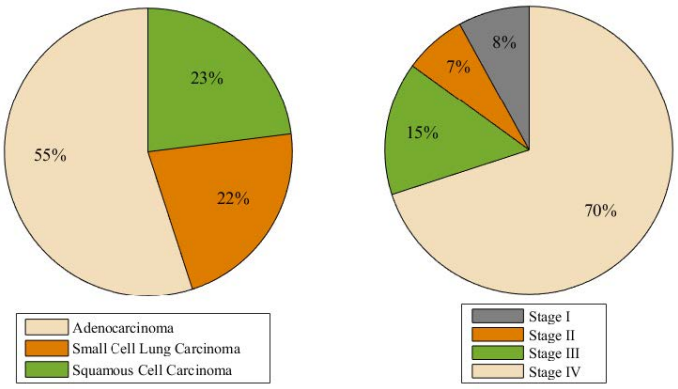

Figure 4. Percentage of patients according to histological type of cancer (left). Percentage of patients according to the cancer stage (right).

2.3. Data analysis and feature extraction

Before extracting the features from the MOS and QCM sensor data, some data preprocessing was performed. To visualize the data properly and extract useful information, the data were processed with mathematical formulas. In this study, reference correction was applied to both MOS and QCM sensor data by using equations given in (1) and (2), respectively.

$$v_{n,s,r}(t) = v_{n,s}(t) - v_{n,s}(0) \tag{1}$$

$$f_{n,s,r}(t) = fn, s(t) - f_{n,s}(0) \tag{2}$$

For MOS sensor data, $v_{n,s}(t)$ represents raw MOS sensor data, $v_{n,s}(0)$ represents MOS sensor data obtained when the breath sample is transmitted to the sensors and $v_{n,s,r}(t)$ represents reference-corrected sensor data. For QCM sensor data, $fn, s(t)$ shows raw frequency data, $f_{n,s}(0)$ shows the sensors' nominal frequency value and $f_{n,s,r}(t)$ shows reference-corrected sensor data. When the electronic nose studies were reviewed, it was determined that the conductivity information was used rather than the voltage information obtained from the MOS sensors. In this study, the voltage data obtained from the load resistors were converted to conductivity data by using the equation given in (3).

$$G_{n,s,r}(t) = \frac{v_{n,s,r}(t)}{(V_c - v_{n,s,r}(t))R_{LS}} \tag{3}$$

In (3), V_c, R_{LS}, and $G_{n,s,r}(t)$ show the supply voltage, load resistance series with sensors and the conductivity data of the MOS sensors, respectively. The graphs given in Figure 5 show the raw data obtained from the MOS sensors and the reference-corrected conductivity data obtained by using 1 and 3. Figure 6 shows the raw QCM sensor data and the reference-corrected QCM sensor data obtained by using 2. Additionally, in Figure 6 the constant frequency value in the prepurging stage indicates the resonant frequency of the sensors.

A total of 338 experiments were carried out during the study. Two hundred and nineteen 219 of these experiments were conducted with lung cancer patient volunteers and 119 experiments were performed with breath samples collected from healthy volunteers. In this study, maximum, mean, variance, kurtosis and skewness values and the slope of sensor data between specific time intervals were used as features. The feature

matrix sizes were obtained as (338 × 91) and (338 × 7) for the MOS sensor data and the QCM sensor data, respectively. PCA and LDA methods were used to reduce the dimension of feature matrices. As a result of the dimension reduction, dimension-reduced data contain 95% of the original data. The dimension of the feature matrix obtained from the MOS sensor data was reduced from (338 × 91) to (338 × 14) by using the PCA and LDA algorithms. The dimension of the feature matrix obtained from the QCM sensor data was reduced from (338 × 7) to (338 × 5) and to (338 × 6), respectively, by using the PCA and LDA algorithms. The dimension of the feature matrix which was created by combining the feature matrices obtained from the MOS and QCM sensor data was reduced from (338 × 98) to (338 × 19) by using both PCA and LDA algorithms. Using PCA and LDA reduced the dimension of the feature matrices and increased the classification accuracy. The effect of the LDA algorithm on the scatter plot can be seen in Figure 7 In this figure LD1 and LD2 represent the first feature and the second features obtained as a result of LDA, respectively. It can be seen in Figure 7 that using the LDA algorithm can discriminate classes as much as possible. Also, the positive effect of the LDA algorithm on classification accuracy is given in Section 3.

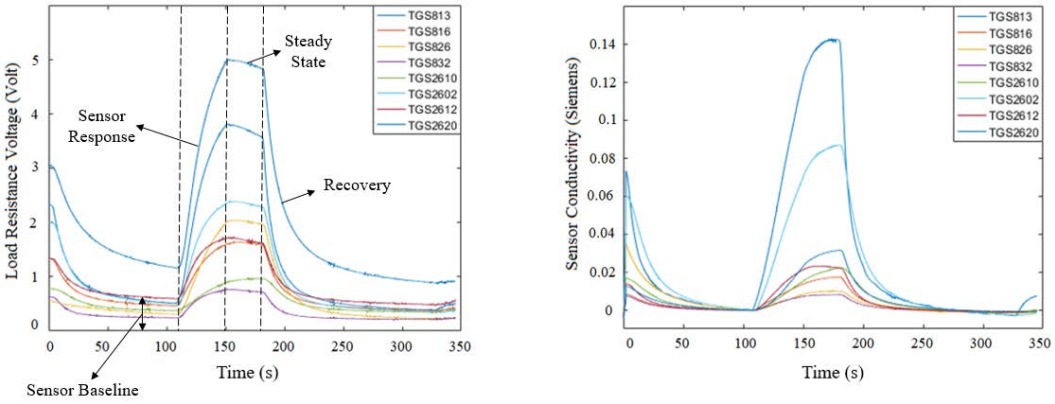

Figure 5. Raw MOS sensor data (left). Preprocessed MOS sensor data (right).

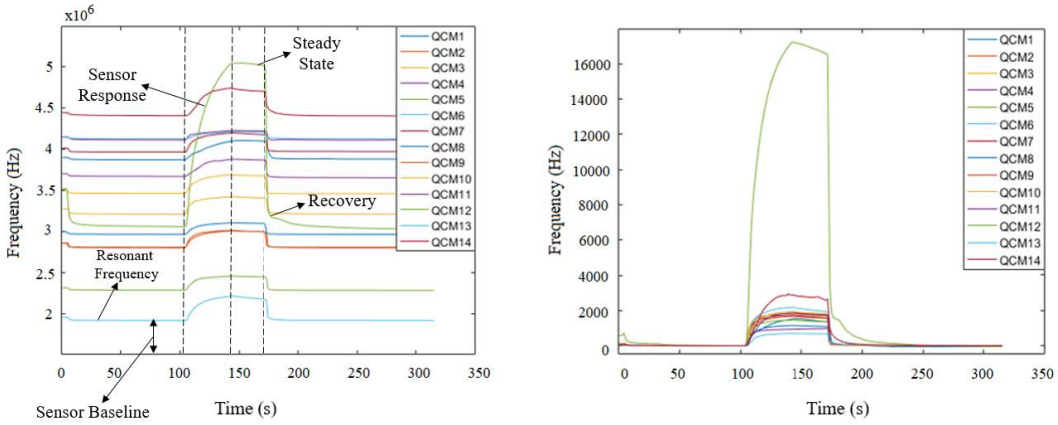

Figure 6. Raw QCM sensor data (left). Preprocessed QCM sensor data (right).

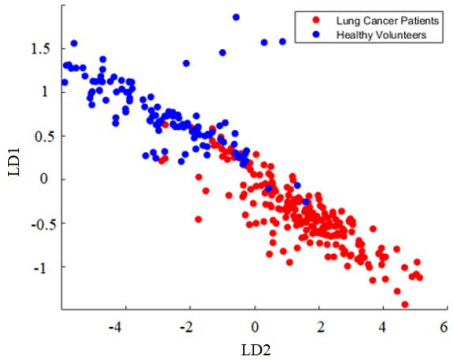

Figure 7. Scatter plot of dimension-reduced feature matrix data.

3. Results

After signal processing, the data were classified using the DT, k-NN, SVM, and RF algorithms. In the k-NN algorithm, the parameter (k) was determined as 3 and the Mahalanobis distance function was used as the distance measurement as it gave the highest accuracy. While classifying with the SVM algorithm, the effects of kernel functions were examined. The last classification algorithm used in this study was the RF algorithm. The most important parameters affecting the performance of the RF algorithm are the number of trees to be developed (N) and the number of variables used in each node to determine the best split (m). In this study, the best results were obtained when using N = 30 and m = 100. A 5-fold crossvalidation method was used to determine the performance of the classification algorithms. Each classification process was repeated 10 times for each classification algorithm. In Table 5, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and the average classification accuracy values for the best performing algorithms can be seen. The specified feature matrices are given by abbreviation according to which sensor data were used and which dimension-reduction algorithm was used. For example, if the feature matrix is obtained from the MOS sensor data and is not reduced, it is represented as "MOS". If the feature matrix is reduced by the PCA algorithm, it is represented as PCA (MOS). The average success rate, maximum success rate, minimum success rate and standard deviation values of the success rate for each classifier are given in Table 6.

Table 5. Classification details for extracted features.

Feature matrix–classification algorithm	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
MOS-RF	92.28	75.74	87.39	83.33	86.34
PCA (MOS) -RF	95.21	88.60	92.11	90.59	92.62
LDA (MOS)-RF	93.41	89.73	94.01	87.60	92.12
QCM-RF	87.16	55.80	78.18	69.47	75.48
PCA (QCM)-RF	94.27	72.57	86.19	86.86	86.64
LDA (QCM)-RF	80.83	67.57	81.94	65.55	74.96
(MOS+QCM)-QSVM	91.83	82.37	90.54	84.34	88.54
PCA (MOS+QCM)-RF	95.83	85.79	92.47	91.07	92.30
LDA (MOS+QCM)-LSVM	95.82	93.90	96.31	91.73	95.36
PCA (MOS) + PCA (QCM)-RF	95.92	89.18	94.17	92.17	93.74
LDA (MOS) + LDA (QCM)-RF	95.07	91.19	94.93	90.75	93.68

Table 6. Classification results for extracted features and classification algorithms.

Classification algorithms							
Types of features		DT	Linear SVM	Quadratic SVM	Cubic SVM	k-NN	RF
	MOS	80.26 82.4–77.5–1.66	82.82 84.6–81.1–1.25	83.76 84.6–83.1–0.65	82.82 87.3–79.6–2.66	80.30 83.1–78.1–1.98	86.34 87.6–85.2–1.17
	PCA(MOS)	91.98 92.6–91.4–0.49	64.62 66.8–63.9–0.84	75.46 76.6–72.5–1.67	90.20 90.8–89.6–0.47	91.52 92.6–91.1–0.62	92.6 93.8–91.1–1.18
	LDA(MOS)	89.02 91.1–87.9–1.23	91.82 92.3–90.5–0.75	90.86 91.7–89.6–0.91	87.24 88.2–85.8–1.22	88.2 89.1–87.3–0.76	92.12 93.2–91.4–0.69
	QCM	70.88 71.9–68.9–1.36	64.14 65.4–63–0.95	74.84 76.9–72.2–1.75	72.5 74.5–70.4–1.55	74.88 77.2–72.8–1.59	75.48 78.4–74.3–1.78
	PCA(QCM)	80.78 82.2–79.00–1.32	64.8 64.8–64.8–0	86.52 88.5–83.4–2.14	83.06 84.6–81.7–1.19	88.30 89.3–87.3–0.83	86.64 88.5–84.6–2.02
	LDA(QCM)	69.44 71.6–68.3–1.29	71.90 73.7–70.4–1.29	74.44 75.1–73.4–0.82	71.60 74.6–68.6–2.40	77.32 79.9–75.1–1.82	74.96 76.1–73.7–0.93
	MOS+QCM	78.52 80.8–76.00–2.17	81.14 82.5–79.9–1.01	88.54 89.9–87.9–0.84	88.28 89.9–86.7–1.41	81.62 82.5–81.1–0.53	86.52 87.9–85.2–1.09
	PCA(MOS+QCM)	90.76 92.3–89.1–1.26	64.60 65–64–0.54	64.60 65–64–0.54	86.40 88.2–84.9–1.25	90.5 91.1–89.6–0.63	92.30 92.9–91.4–0.60
	LDA(MOS+QCM)	93.26 93.8–92.6–0.53	95.36 96.2–94.7–0.57	95.00 95.3–94.7–0.30	93.08 93.8–92.6–0.54	90.04 90.8–88.8–0.77	95.30 96.2–93.8–1.01
	PCA(MOS)+LD(QCM)	91.46 92.6–90.2–1.58	64.60 65–64–0.54	83.18 84.9–82–1.06	89.3 91.1–87–1.69	93.26 94.1–92.6–0.53	93.74 94.4–92.9–0.57
	LD(MOS)+LDA(QCM)	91.22 91.7–90.8–0.34	92.7 93.2–92.3–0.45	92.6 93.5–91.1–0.90	90.62 91.7–89.6–0.75	89.25 87.6–91.4–1.43	93.68 95.9–92.6–1.31

4. Discussion

In this study, an electronic nose system consisting of 14 different QCM sensors and 8 different MOS sensors was developed. Unlike the studies in the literature regarding electronic noses, two different types of sensors were used in this study. Since the information in electronic nose systems is produced using sensors, the number, type and features of these sensors directly affect the performance. Therefore, we aimed to obtain maximum information from breath samples. In this study, the voltage and conductivity information was obtained using MOS sensors, while the frequency deviation information was obtained using QCM sensors. It can be clearly seen from Table 6 that using two different sources of sensor data together has increased the classification accuracy. A total of 100 volunteers, 60 patients with lung cancer, 20 healthy nonsmokers and 20 healthy smokers participated in this study and a total of 338 experiments were performed. In this work, the PCA and LDA algorithms were used for dimension reduction, and for data classification the DT, SVM with different kernel functions, k-NN and RF algorithms were used. The best classification result was obtained as 95.36% by using the SVM algorithm with the LDA (MOS+QCM) feature matrix. When the results are examined, it can be seen that the classification results obtained by using only MOS sensor data are better than the results using only QCM sensor data. Moreover, classification results show that using MOS and QCM sensor data together provide a significant improvement in the classification results. The algorithms used to reduce the size of the feature matrices have also increased the classification success considerably. The PCA method has reduced the size of the feature matrix by eliminating the unrelated features. With this method, the computation time has decreased and the classification success has increased. The LDA algorithm has also increased the classification accuracy up to 95.36% by taking the projection of the feature matrix. Thus, the separation between the classes has been maximum. To see the effect of the LDA algorithm on the classification accuracy, breath samples from healthy nonsmokers (HnS) and

healthy smokers (HS) were also classified. The confusion matrix of this classification, which was carried out with the RF algorithm, is given in Figure 8.

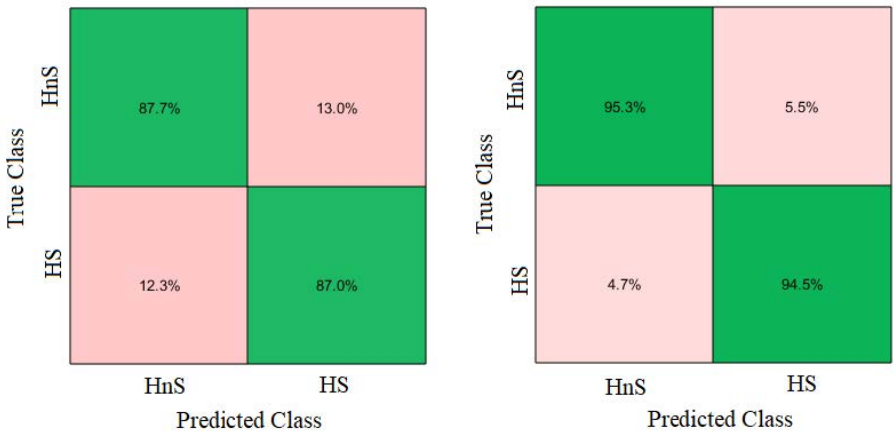

Figure 8. Confusion matrix before LDA (left). Confusion matrix after LDA (right).

When the performances of the classification algorithms used in this study are examined, the superiority of the ensemble learning algorithms comes into prominence. The RF algorithm has given better performance in most cases. The results show that digital signal processing methods have an important place in the success of the implemented systems as well as the used sensors and their types. Obtained features, size reduction algorithms and classification algorithms each significantly affect e-nose performance. This study is the first and pioneering study in Turkey regarding lung cancer detection by using a hybrid sensor-based e-nose. We believe that our study will find a wide application area in the detection of lung cancer in practice.

Acknowledgment

We would like to thank The Scientific and Technological Research Council of Türkiye (TÜBİTAK) and TÜBİTAK Marmara Research Center Materials Sciences Institute for supporting this study within the scope of the project numbered 215E380 and for their contributions to this project.

References

[1] Siegel RL, Miller KD, Jemal A. Colorectal cancer statistics. CA: a cancer journal for clinicians 2020; 70 (3): 145-164 . doi: 10.3322/caac.21601

[2] Luque de Castro MD, Fernández-Peralbo MA. Analytical methods based on exhaled breath for early detection of lung cancer. TrAC Trends in Analytical Chemistry 2012; 38 : 13-20. doi: 10.1016/j.trac.2012.03.018

[3] National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England Journal of Medicine: Research and Review 2011; 365: 395-409. doi: 10.1056/NEJMoa1102873

[4] Marcus PM, Bergstralh EJ, Zweig MH, Harris A, Offord KP et al. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. JNCI: Journal of the National Cancer Institute 2006; 98 (11): 3748-756. doi:10.1093/jnci/djj207

[5] Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. JAMA 2011; 306 (17): 1865-1873. doi:10.1001/jama.2011.1591

[6] Li W, Liu HY, Jia ZR, Qiao PP, Pi XT et al. Advances in the early detection of lung cancer using analysis of volatile organic compounds: from imaging to sensors. Asian Pacific Journal of Cancer Prevention 2014; 15 (11): 4377-4384. doi:10.7314/apjcp.2014.15.11.4377

[7] Mazzone PJ, Obuchowski N, Phillips M, Risius B, Bazerbashi B et al. Lung cancer screening with computer aided detection chest radiography: design and results of a randomized, controlled trial. PLoS One 2013; 8 (3): e59650. doi:10.1371/journal.pone.0059650

[8] Jantus-Lewintre E, Usó M, Sanmartín E, Camps C. Update on biomarkers for the detection of lung cancer. Lung Cancer (Auckl) 2012; 11 (3): 21-29. doi:10.2147/LCTT.S23424

[9] Karki S, Yin Yj, Samanai N, Zhang JX , Pradhan D et al. Breathe analyzer and its importance for the early detection of lung cancer. Sky Journal of Medicine and Medical Sciences 2013; 2 (1): 7-9. doi:10.1016/j.cccn.2004.04.023

[10] Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas–liquid partition chromatography. Proceedings of the National Academy of Sciences of the USA 1971; 68 (10): 2374-2376. doi: 10.1073/pnas.68.10.2374

[11] Belizário JE, Faintuch J, Malpartida MG. Breath Biopsy and Discovery of Exclusive Volatile Organic Compounds for Diagnosis of Infectious Diseases.Frontiers in Cellular and Infection Microbiology 2021; 10: 564194 doi: 10.3389/fcimb.2020.564194

[12] Miekisch W, Schubert JK, Noeldge-Schomburg GF. Diagnostic potential of breath analysis-focus on volatile organic compounds. Clinica Chimica Acta 2004; 347 (1-2): 25-39. doi: 10.1016/j.cccn.2004.04.023

[13] Issitt T, Wiggins L, Veysey M, Sweeney ST, Brackenbury WJ et al. Volatile compounds in human breath: critical review and meta-analysis. Journal of Breath Research 2022; 16 (2). doi: 10.1088/1752-7163/ac5230

[14] Aghdassi E, Wendland BE, Steinhart AH, Wolman SL, Jeejeebhoy K et al. Antioxidant vitamin supplementation in Crohn's disease decreases oxidative stress. A a randomized controlled trial. The American Journal of Gastroenterology 2003; 98 (2): 348-53. doi: 10.1111/j.1572-0241.2003.07226.x

[15] Stone BG, Besse TJ, Duane WC, Evans CD, DeMaster EG. Effect of regulating cholesterol biosynthesis on breath isoprene excretion in men. Lipids 1993; 28 (8): 705-8. doi: 10.1007/BF02535990

[16] Lebovitz HE. Diabetic ketoacidosis. The Lancet 1995; 345 (8952): 767-72. doi: 10.1016/s0140-6736(95)90645-2

[17] Cope K, Risby T, Diehl AM. Increased gastrointestinal ethanol production in obese mice: implications for fatty liver disease pathogenesis. Gastroenterology 2000; 119 (5): 1340-7. doi: 10.1053/gast.2000.19267

[18] Scislowski PW, Pickard K. The regulation of transaminative flux of methionine in rat liver mitochondria. Archives of Biochemistry and Biophysics 1994; 314 (2): 412-6. doi: 10.1006/abbi.1994.1461

[19] Davies S, Spanel P, Smith D. Quantitative analysis of ammonia on the breath of patients in end-stage renal failure. Kidney International 1997; 52 (1): 223-8. doi: 10.1038/ki.1997.324

[20] Wang P, Huang Q, Meng S, Mu T, Liu Z et al. Identification of lung cancer breath biomarkers based on perioperative breathomics testing: A prospective observational study. eClinicalMedicine 2022; 47 (16): 101384. doi: 10.1016/j.eclinm.2022.101384

[21] Chen T, Liu T, Li T, Zhao H, Chen Q. Exhaled breath analysis in disease detection.Clinica Chimica Acta 2021; 515: 61-72. doi: 10.1016/j.cca.2020.12.036

[22] Li M, Biswas S, Nantz MH, Higashi RM, Fu XA. Preconcentration and analysis of trace volatile carbonyl compounds. Analytical Chemistry 2012; 84 (3): 1288-93. doi: 10.1021/ac2021757

[23] Nardi-Agmon I, Abud-Hawa M, Liran O, Gai-Mor N, Ilouze M et al. Exhaled Breath Analysis for Monitoring Response to Treatment in Advanced Lung Cancer. Journal of Thoracic Oncology 2016; 11 (6): 827-37. doi: 10.1016/j.jtho.2016.02.017

[24] Huszár E, Vass G, Vizi E, Csoma Z, Barát E et al. Adenosine in exhaled breath condensate in healthy volunteers and in patients with asthma. The European Respiratory Journal 2002; 20 (6): 1393-8. doi: 10.1183/09031936.02.00005002

[25] Emelyanov A, Fedoseev G, Abulimity A, Rudinski K, Fedoulov A et al. Elevated concentrations of exhaled hydrogen peroxide in asthmatic patients. CHEST Journal 2001; 120 (4): 1136-9. doi: 10.1378/chest.120.4.1136

[26] Kharitonov SA, Yates D, Robbins RA, Logan-Sinclair R, Shinebourne EA et al. Increased nitric oxide in exhaled air of asthmatic patients. The Lancet 1994; 343 (8890): 133-5. doi: 10.1016/s0140-6736(94)90931-8

[27] Elsheikh MS, Mohamed NH, Alsharkawy AAA. Improvement of asthma control after laser acupuncture and its impact on exhaled 8-isoprostane as an oxidative biomarker in chronic bronchial asthma. Respiratory Medicine 2019; 156 : 15-19. doi: 10.1016/j.rmed.2019.07.022

[28] Carpagnano GE, Scioscia G, Lacedonia D, Soccio P, Lepore G et al. Looking for Airways Periostin in Severe Asthma: Could It Be Useful for Clustering Type 2 Endotype? Chest Journal 2018; 154 (5): 1083-1090. doi: 10.1016/j.chest.2018.08.1032

[29] Kumar S, Huang J, Abbassi-Ghadi N, Mackenzie HA, Veselkov KA et al. Mass Spectrometric Analysis of Exhaled Breath for the Identification of Volatile Organic Compound Biomarkers in Esophageal and Gastric Adenocarcinoma. Annals of Surgery 2015; 262 (6): 981-90. doi: 10.1097/SLA.0000000000001101

[30] Shehada N, Cancilla JC, Torrecilla JS, Pariente ES, Brönstrup G et al. Silicon Nanowire Sensors Enable Diagnosis of Patients via Exhaled Breath. ACS Nano 2016; 1 0(7): 7047-57. doi: 10.1021/acsnano.6b03127

[31] Shin J, Choi S-J, Lee I, Youn DY, Park CO et al. Thin-Wall Assembled SnO2 Fibers Functionalized by Catalytic Pt Nanoparticles and their Superior Exhaled-Breath-Sensing Properties for the Diagnosis of Diabetes. Advanced Functional Materials 2013; 23: 2357-2367. doi: 10.1002/adfm.201202729

[32] Chen Y, Qin H, Cao Y, Zhang H, Hu J. Acetone Sensing Properties and Mechanism of SnO2 Thick-Films. Sensors (Basel) 2018; 18 (10): 3425. doi: 10.3390/s18103425

[33] Chien PJ, Suzuki T, Tsujii M, Ye M, Toma K et al. Bio-sniffer (gas-phase biosensor) with secondary alcohol dehydrogenase (S-ADH) for determination of isopropanol in exhaled air as a potential volatile biomarker. Biosensors and Bioelectronics 2017; 91: 341-346. doi: 10.1016/j.bios.2016.12.050

[34] Moon HG, Jung Y, Han SD, Shim YS, Shin B et al. Chemiresistive Electronic Nose toward Detection of Biomarkers in Exhaled Breath. ACS Applied Materials and Interfaces 2016; 8 (32): 20969-76. doi: 10.1021/acsami.6b03256

[35] Kokoszka J, Nelson RL, Swedler WI, Skosey J, Abcarian H. Determination of inflammatory bowel disease activity by breath pentane analysis. Diseases Colon and Rectum 1997; 36 (6): 597-601. doi: 10.1007/BF02049868.

[36] Scholpp J, Miekisch JKW, Geiger K. Breath markers and soluble lipid peroxidation markers in critically ill patients. Clinical Chemistry and Laboratory Medicine 2002; 40 (6): 87-94. doi: 10.1515/CCLM.2002.101

[37] Saraoğlu HM, Selvi AO, Taşaltin C. Electronic Nose System Based on Quartz Crystal Microbalance Sensor for Blood Glucose and HbA1c Levels From Exhaled Breath Odor. IEEE Sensors Journal 2013; 13 (11): 4229-4235. doi: 10.1109/JSEN.2013.2265233

[38] Huang H, Zhou J, Chen S, Zeng L, Huang Y. A highly sensitive QCM sensor coated with Ag+-ZSM-5 film for medical diagnosis. Sensors and Actuators B Chemical 2004; 101 (3): 316-321. doi: 10.1016/j.snb.2004.04.001

[39] Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. British Journal of Cancer 2010; 103 (4): 542-51. doi: 10.1038/sj.bjc.6605810

[40] Jia Z, Patra A, Kutty VK, Venkatesan T. Critical Review of Volatile Organic Compound Analysis in Breath and In Vitro Cell Culture for Detection of Lung Cancer. Metabolites 2019; 9 (3): 52.doi: 10.3390/metabo9030052

[41] Poli D, Goldoni M, Corradi M, Acampa O, Carbognani P et al. Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. The Journal of Chromatography B 2010; 878 (27): 2643-51. doi: 10.1016/j.jchromb.2010.01.022

[42] Feng S, Farha F, Li Q, Wan Y, Xu Y et al. Review on Smart Gas Sensing Technology. Sensors (Basel) 2019; 19 (17): 3760. doi: 10.3390/s19173760

[43] Baldini C, Billeci L, Sansone F, Conte R, Domenici C et al. Electronic Nose as a Novel Method for Diagnosing Cancer: A Systematic Review. Biosensors 2020; 10 (8): 84. https://doi.org/10.3390/bios10080084

[44] Behera B, Joshi R, Anil VGK, Bhalerao S, Pandya HJ. Electronic nose: a non-invasive technology for breath analysis of diabetes and lung cancer patients. Journal of Breath Research 2019; 13 (2): 024001. doi: 10.1088/1752-7163/aafc77

[45] Tirzïte M, Bukovskis M, Strazda G, JurkaNand Taivans I. Detection of lung cancer with electronic nose and logistic regression analysis. Journal of Breath Research 2019; 13 (1):016006. doi:10.1088/1752-7163/aae1b8

[46] Kononov A, Korotetsky B, Jahatspanian I, Gubal A, Vasiliev A et al. Online breath analysis using metal oxide semiconductor sensors (electronic nose) for diagnosis of lung cancer. Journal of Breath Research 2019; 14 (1): 016004. doi:10.1088/1752-7163/ab433d

[47] Krauss E, Haberer J, Barreto G, Degen M, Seeger W et al. Recognition of breathprints of lung cancer and chronic obstructive pulmonary disease using the Aeonose® electronic nose. Journal of Breath Research 2020; 14 (4): 046004. doi:10.1088/1752-7163/ab8c50

[48] Wang L, Hongying L, Dandan X, Zichun H, Xititan P. Lung Cancer Screening Based on Type-different Sensor Arrays. Scientific Reports 2017; 7 : 25-39. doi:10.1038/s41598-017-02154-9

[49] Kort S, Brusse-Keizer M, Gerritsen JW, Schouwink H, Citgez E et al. Improving lung cancer diagnosis by combining exhaled-breath data and clinical parameters. ERJ Open Research 2020; 6 (1): 00221-2019. doi:10.1183/23120541

[50] Chernov VI, Choynzonov EL, Kulbakin DE, Obkhodskaya EV, Obkhodskiy AV et al. Cancer Diagnosis by Neural Network Analysis of Data from Semiconductor Sensors. Diagnostics (Basel) 2020; 10 (9): 677. doi:10.1088/1752-7163/ab8c50

[51] Huang CH, Zeng C, Wang YC, Peng HY, Lin CS et al. A Study of Diagnostic Accuracy Using a Chemical Sensor Array and a Machine Learning Technique to Detect Lung Cancer. Sensors (Basel) 2018; 18 (9): 2845. doi:10.3390/s18092845

[52] Kaloumenou M, Skotadis E, Lagopati N, Efstathopoulos E, Tsoukalas D. Breath Analysis: A Promising Tool for Disease Diagnosis-The Role of Sensors. Sensors (Basel) 2022; 22 (3): 1238. doi: 10.3390/s22031238