

A Review of Technological Progression from Radiomics to Breathomics for Early Detection of Lung Cancer

¹Funmilayo S. Moninuola, ¹Oluwadamilola Oshin, ^{1,2,3}Emmanuel Adetiba, ¹Anthony A. Atayero, ⁴Ademola Adeyeye, ¹Victoria Oguntosin, ⁵Olushola O. James, ⁶Anthony A. Adegoke, ^{7,8}Obiseye O. Obiyemi, ⁷Surendra Thakur and ⁹Abdultaofeek Abayomi

¹Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria

²Covenant Applied Informatics and Communication Africa Center of Excellence (CApIC-ACE), Covenant University, Ota, Nigeria

³HRA, Institute for Systems Science, Durban University of Technology, Durban, P.O. Box 1334, Durban, South Africa

⁴Division of Oncology, Department of Surgery, Afe Babalola University, Ado-Ekiti, Nigeria

⁵Department of Chemistry, Kwara State University, P.M.B. 1530, Malete, Kwara State, Nigeria

⁶Department of Microbiology, University of Uyo, PMB 1017, Akwa-Ibom State, Uyo, Nigeria

⁷KZN e-Skills CoLab, Durban University of Technology, Durban, South Africa

⁸Department of Electrical and Electronic Engineering, Osun State University, Osogbo, Osun State, Nigeria

⁹Department of Information and Communication Technology, Mangosuthu University of Technology, P.O. Box 12363 Jacobs, 4026 Durban, South Africa

Article history

Received: 03-05-2021

Revised: 18-08-2021

Accepted: 25-08-2021

Corresponding Author:

Emmanuel Adetiba
Department of Electrical and
Information Engineering,
Covenant University, Ota,
Nigeria

Email:
emmanuel.adetiba@covenantuniversity.edu.ng

Abstract: Lung Cancer (LC) is formed by an uncontrolled growth of abnormal cells in the lungs, which alters the body's metabolism. Early detection of LC is very important for a better survival rate and prognosis in patients and persons at risk of the disease. Traditional approaches to LC detection are less sensitive to cancerous cells at the early stage, are invasive and consume a lot of resources in terms of time and cost. Exhaled breath is a good source of Volatile Organic Compound (VOC) biomarkers originating from the airways and tissues. Breath analysis is an ideal approach for the detection of metabolites relating to cancer cells in the lungs. The electronic nose (E-nose) approach for LC detection is apt for detecting VOC biomarkers found in the bronchi. E-nose is non-invasive, inexpensive and easy to use. This review discusses the progression of LC diagnosis from radiomics to breathomics (E-nose). Studies on the potential use of E-nose as an easy and non-invasive tool to analyze VOCs in the exhaled breath of lung cancer patients to characterize suspected lung nodules are also presented. Furthermore, the current gaps in respect of the application of E-nose for early detection of LC are unearthed.

Keywords: Lung Cancer, Early Detection, Radiomics, Breathomics, Electronic Nose

Introduction

Lung Cancer (LC) also referred to as carcinoma of the lung or bronchogenic carcinoma, is an aggressive lung tumor caused by unrestrained growth of epithelial cells in the tissues of the lung, usually in the bronchi and the airways. This development may spread to a distant site from the lungs and produce metastatic lesions in the brain, bone, liver or adrenal glands. Often times, lung cancer patients are symptomatic of cough, hemoptysis, dyspnea, chest pain, non-resolving pneumonia, metastatic disease like skeletal pain or neurological issues (Hu *et al.*, 2016). Due to the asymptomatic character of this cancer, most cases are not diagnosed till the illness has advanced significantly and therapeutic treatment is not an option anymore.

The International Agency for Research on Cancer (IARC) and World Health Organization (WHO) estimate that lung cancer is the world's most prevalent cancer fatality and the 6th leading cause of mortality among all forms of cancer in Africa (Sanni *et al.*, 2018). WHO also predicted that Chronic Obstructive Pulmonary Disease (COPD) will be liable for every third mortality in the globe by 2030. COPD was also proposed as a contributing factor for lung cancer (Welniak *et al.*, 2015). In addition, when localized, only 16% of cases of lung cancer are found, while 22 and 57% are detected at regional and distant stages, respectively. In spite of the understanding that timely diagnosis of lung cancer will have a significant effect on survival rates, there has not been an appropriate diagnostic process that offers proof of early detection as is

available for other types of cancer such as breast and colorectal (Herbst *et al.*, 2018; Lanni *et al.*, 2018).

Lung carcinoma is traditionally classified histologically into 2 key forms: Small Cell LC (SCLC) and Non-Small Cell LC (NSCLC). Mutations involving EGFR, KRAS and TP53 have also been used to sub-classify lung cancers on a molecular level. SCLC, which mostly begins in the wider airways as well as the main and secondary bronchi, accounts for around 10-15% of lung cancers. It is the most severe form of lung cancer and develops faster than NSCLC, frequently metastasizing at the early onset of the disease to other areas of the organ. At the onset of symptoms, most SCLC cases have widespread metastasis. SCLC's 5-year survival rate (6%) is lower than that of NSCLC (21%) (Linning *et al.*, 2019).

The most prevalent form of lung cancer is NSCLC, grouped on the basis of therapeutic outlook and the cancer cells' gene expression. Roughly 85% of lung cancer cases fall into this category. Adenocarcinoma (40%), squamous cell carcinoma (25-30%) and large cell carcinoma (10-15%) are the three major NSCLC subtypes.

Many patients diagnosed with NSCLC are at a complicated phase in which it is important to envisage the probability of survival and therapeutic options. Regrettably, during diagnosis, 70% of NSCLC tumors cannot be resected, which is important for a clinical analysis that will disclose molecular details and thus the outcome of the prognosis. Knowing the prognosis helps to assess if seeking a certain procedure is useful (Linning *et al.*, 2019; Vavala and Novello, 2017).

If LC is identified early, at least 50% of the patients will survive and be free of relapse tumors. The significant causes of lung cancer are firsthand smoking of tobacco and secondhand smoking in nonsmokers. Other factors include asbestos contamination, automotive and factory air pollution, as well as arsenic, chromium, nickel, aromatic hydrocarbons and ether contaminations (Hecht, 2012).

Recognition of LC in its early stages can be characterized on the grounds of lead time, length of time and biases of tumor selection. If it is discovered at an indolent stage, the survival rate will improve significantly. Chemo and radiotherapy treatments are usually used for SCLC, whereas NSCLC is usually treated with surgery. Research shows that use of imaging techniques such as Chest Radiograph (CRG), Computed Tomography (CT), Low-Dose CT (LDCT), Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) are standard approaches for detecting lung cancer.

Nonetheless, all these procedures are expensive and protracted. Also, because of their invasiveness, they are unresponsive to carcinogenic cells in their initial phase. Biopsy and autoantibodies are other common ways of identifying lung cancer, but they are also expensive, time-consuming and require expert physicians. A non-invasive breath test will have significant promise as

means a of detecting lung carcinoma (Antoniou *et al.*, 2019). A diagnostic test that can identify lung cancer early before it has spread is effective in minimizing lung cancer death rates. The earlier the detection, the higher the chances of successful treatment.

In this review, we focus on the evolution of existing approaches for early lung cancer detection. The resurgence of interest in early identification of LC and the medical imagery used have raised innumerable issues about this international epidemic. Each of the existing methods for LC detection has particular advantages and weaknesses, hence there is continued focus on developing better approaches. The rest of this review is divided into two sections: The trends in LC diagnosis and detection from customary methods to the most recent improvements in exhaled breath analysis are presented, while the latter section presents the research opportunities and the future direction in this area of research.

Literature Review

Conventional Imaging Approaches

The traditional methods of identification and characterization for LC include: CRG, CT, LDCT, MRI and PET. The purpose of these techniques is to locate lesions in subjects who are at high risk of LC. While they are common strategies with a wide pool of professional users, they are impaired by demerits, which include a high false rate, radiation exposure, high cost and poor sensitivity for the identification of cancer cells at early stages (Kovalchik *et al.*, 2013; Kumar and Latte, 2019). A key advancement in the traditional method for identifying and characterizing LC is the use of advanced algorithms to interpret clinical images called Radiomics. Radiomics improves the precision of LC nodule detection and enables doctors and radiologists to precisely analyze radiographic scans.

Chest Radiograph

A chest radiograph is the traditional screening method for the detection of lesions and pulmonary nodules in patients with LC. It uses X-rays to generate lung images, but is associated with the issues of radiation exposure dose and poor image quality. It has also been reported in the literature that this approach has low sensitivity and is not helpful in reducing LC mortality (Khobragade *et al.*, 2016).

Computed Tomography

CT is an imaging screening tool that uses sophisticated X-ray technology to detect the condition of the lungs and provides data on tumor features such as dimensions, classification and growth. In addition to staging of the tumor, the 3D CT image presents examination of the chest wall, diaphragm and

mediastinum invasion. According to Latifi *et al.* (2015), LC control is substantially influenced by 4D CT by enabling more accurate targeting of the radiation administered. The key disadvantages of CT remain risky exposure to radiation and the detection of indolent tumors that could lead to over-diagnosis and, in most cases, incur additional cost, anxiety and morbidity for patients (Kovalchik *et al.*, 2013; Ohno *et al.*, 2018; Patz *et al.*, 2014).

Low-Dose Computed Tomography

Low-Dose Computed Tomography (LDCT) was introduced for lung imaging in order to overcome the limitations of radiation in CT. LDCT acquires single-slice spiral Computed Tomography (CT) using Multidetector row CT (MDCT) scanners. Studies show that LDCT spots more lung nodules and cancer cells than a chest radiograph. Up to 90% of new LC cases were detected using LDCT. It uses an iterative image approach, which is more suitable for the detection of LC in high-risk subjects. However, LDCT is limited in reducing LC mortality as it produces high false-positive rates and the problem of excessive exposure to radiation still persists (Journey *et al.*, 2015; Marshall *et al.*, 2017; NLSTR, 2013).

Positron Emission Tomography/Computed Tomography

Positron Emission Tomography/Computed Tomography (PET/CT) is a fusion of two oncological visualization procedures, which deliver CT structural data and PET anabolic information to envision cancer cells in 3D or bi-dimensional slices, either individually or fused. F18-Fluorodeoxyglucose (F18-FDG) is the most widely used radiotracer, which enables the identification of tumor cells and glucose-consuming metastases similar to majority of malignant pulmonary lesions (Chicklore *et al.*, 2013; de Guevara Hernández, 2015). PET provides better responsiveness and enhanced precision for LC identification compared to CT. In cases with developed NSCLC for stereotactic radiotherapy, PET plays a significant function in the identification of subjects and specification of treatment sites (Mac Manus *et al.*, 2013). The main restrictions of this method are erroneous false-positive results, imprecise structural details, and time-consuming nature (Montilla-Soler *et al.*, 2018).

Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) was developed to overcome the drawbacks of PET/CT in the investigation of lung malignancies and other diseases. It is a non-ionization radiation tool that works with the principle of magnetic fields and radio waves. It is interfaced with a computer to produce comprehensive images of the structures within the chest and determine the size, extent and degree of LC spread to adjacent structures (Hochegger *et al.*, 2011). Due to reduced

proton intensity and rapid signal degradation of the lung tissue structure, lung MRI has engineering hurdles. In addition, the quality of lung imaging in MRI depends on the patient's ability to conform to the breath hold instructions (Hochegger *et al.*, 2011; Biederer *et al.*, 2012).

Magnetic Induction Tomography

Magnetic Induction Tomography (MIT) is another type of imagery procedure. MIT is a tomography technique that uses a contactless method for mapping the electrical conductivity of tissue and also imaging the electromagnetic properties of an object using the eddy current effect. It can be used to produce passive electrical 3D images, which include neuroimaging applications, radiology cryosurgery tracking and metal flow modeling in metal working procedures. Intramuscular fat and water-bearing fat-free tissue can be classified by MIT (Han *et al.*, 2016). However, MIT imaging techniques have several issues, including high computation imaging algorithms, hypothetical phantoms of the thoracic cavity, challenging clinical hardware systems and spatial resolution (Xiao *et al.*, 2018).

Computer Aided Diagnosis (CAD) of Lung Cancer using Radiomics

Radiomics is the advancement of high-resolution extraction processes and techniques for quantifiable functions that transform images from traditional imagery methods into databases that can be extracted to facilitate clinical decision-making (Gillies *et al.*, 2016). To improve the precision of neoplasm nodule detection, the CAD approach was created to enable doctors and radiologists to precisely analyze radiographic scans. The traditional computer-aided testing and identification systems focused on medical imaging for LC include; CRG, CT, LDCT, MRI and PET. Such techniques are designed to distinguish tumors in subjects that are particularly prone to LC (Orozco *et al.*, 2012). Therefore, before making final decisions, radiotherapists utilize the device data to ascertain disease state, since the machine is deemed useful for the identification of possible nodules. The key phases used in radiomics are segmentation, feature extraction, classification and nodule detection. Initial radiographs of individuals are collected from the LC testing database of radiographs with illustrated marked-up lesions such as the Lung Image Database Consortium (LIDC) (Demir and Yılmaz Çamurcu, 2015; Junior *et al.*, 2018).

Although radiomics approaches are well-identified methods with a huge network of knowledgeable operators, demerits exist. When analyzing ordinary computed tomography, doctors and radiographers find it tough to distinguish cancerous nodules because the cross-section is complicated. This involves extra efforts in the identification of the LC by radiologists, and hence there is a high risk of error.

Table 1: Some of the CAD using radomics approaches in the literature

S/N	Dataset	Segmentation/Feature Extraction Technique	Classifier	Sensitivity (%)	Reference
1	LIDC	Toboggan Based Growing Automatic TBGA	RNN	97.93	Sankar and George (2021)
2	LIDC	Lesion segmentation	Multiple classifiers voting	100	Saba (2019)
3	LIDC	HoG, PCA, Texture and geometry feature.	K-NN, Naive Bayes, SVM and Ada Boost	98.3	Naqi <i>et al.</i> (2019)
4	DermIS	GLCM, LBP, Color Features	SVM	97	Khan <i>et al.</i> (2019)
5	LIDC-IDRI	Frangi filter	4-channel CNN	80.06	Jiang <i>et al.</i> (2017)
6	LIDC	Watershed segmentation	SVM	92	Makaju <i>et al.</i> (2018)
7	LIDC-IDRI	NIL	Multi crop convolution neural network MC-CNN	77	Shen <i>et al.</i> (2017)
8	LIDC/IDRI	Watershed	SVM, Naive Bayes and FLD	94.4	Firmino <i>et al.</i> (2016)
9	LIDC/IDRI	Clustering	Multi view Convolutional Network	90.1	Setio <i>et al.</i> (2016)
10	LIDC, ELCAP	Discrete wavelet transforms	SVM	90.9	Orozco <i>et al.</i> (2015)
11	LIDC/IDRI	Rule based	SVM	93.6	Demir and Yılmaz Çamurcu (2015)
12	LIDC/IDRI	Structural	LDA, gradient boosting (GB 10)	82	Jacobs <i>et al.</i> (2016)
13	Private	Selective filter based (vessel)	Rule based classification (RBC)	90	Jia <i>et al.</i> (2014)
14	Private	Cylindrical filter and thresholding based	SVM	90	Teramoto <i>et al.</i> (2014)
15	LIDC	3D block based on Entropy Analysis	SVM	95.28	Choi and Choi (2013)
16	JSRT	Clustering watershed segmentation	SVM	85	Chen and Suzuki (2012)
17	LIDC, ANODE09	Stochastic, binarization and active contour model	SVM	89	Keshani <i>et al.</i> (2013)
18	ELCAP	Frequency domain analysis	SVM	96.15	Orozco <i>et al.</i> (2012)

Histogram Of Oriented Gradient (HOG), Gray Level Co-Existing Texture Matrix (GLCM), Local Binary Pattern (LBP), Lung Image Database Consortium (LIDC), Image Database Resource Initiative (IDRI), Early LC Action Program (ELCAP), Japanese Society of Radiological Technology (JSRT), Automatic Nodule Detection 2009 (ANODE09)

Some research works based on this approach as reported in the literature are briefly highlighted in Table 1.

Proteomics and Genomics

Proteomics is the systematic analysis of all the sets of proteins present in cells, chromosomes, or tissues. This technique models how the architecture of proteins is modified by disease. The arrangement of proteins varies from a robust cell to a carcinomatous cell as protein varies or changes occur, which may influence the shape and functionality of the cell (Broodman *et al.*, 2017). Proteomics as well as Genomics are used to sync up gene and protein research demonstrated in a specific subset of genes, cells or tissues and to discover potential molecular targets as well. A concurrent assessment of gene and protein tumor molecular expression uses microarray technology. LC related proteins are produced by LC tissues, which in turn appear to develop high-affinity auto-antibodies to these proteins from the immune system.

Blood biomarkers include specific tumor-associated antigen proteins and auto-antibodies that may be observable 1 to 3 years prior to clinical diagnosis (Jain, 2016; Zhong *et al.*, 2006). Prior studies also used algorithmic solutions for binary classifications of LC biomarkers. Carcino Embryonic Antigen (CEA), CYFRA 21-1 (cytokeratine 19 fragment), Neuron-Specific Enolase (NSE), Progastrin Releasing Peptide (ProGRP) and Squamous Cell Carcinoma Antigen (SCCA) are some commonly recognized and clinically used LC protein biomarkers, according to Broodman *et al.* (2017). Other LC protein biomarkers include Antigen-125 (CA-125), Human

Epididymis Protein 4 (HE4) and Surfactant Form of Protein B (Pro-SFTPb) (Sin *et al.*, 2013; Taguchi *et al.*, 2013).

Table 2 offers a list of some of the recognized LC biomarkers using proteomics and genomics. For instance, earlier work by Adetiba and Olugbara (2015a) revealed that combinations of neural network ensembles with histograms of oriented gradients and affine transformations can be applied for LC detection and classification. In addition, the renowned method of analyzing the Protein Pathway Array (PPA) has been established by the authors in (Li *et al.*, 2013; Liu *et al.*, 2014; Pass *et al.*, 2013; Wang *et al.*, 2011) to classify significant but low sample proteins and phosphoproteins in LC. Despite the specificity of the proteomics and genomics methods, the available data for evaluating the output of these biomarkers for subsequent LC cases and control discrimination is not adequately available (Guida *et al.*, 2018).

Breathomics

Breath analysis, which is also known as breathomics, is a tool that is avant-garde for prompt detection of lung carcinoma. According to Boots *et al.* (2012), the aim of breathomics research is to identify trends in Volatile Organic Compounds (VOCs) that define pathological metabolic activities in human beings. Breath science began approximately 50 years ago when Pauling used the process of gas chromatography to show that around two hundred and fifty gases make up human breath (Pauling *et al.*, 1971). Nevertheless, it is still in its growing stage, especially with the current advances in the application of artificial intelligence methodologies to analyze the signature of VOC structures in breaths (Boots *et al.*, 2012).

In clinical analysis and environmental investigation, breath scrutiny can be used because it is non-invasive. Application of breath analysis has been explored across different spectrum of diseases. For instance, lung gangrene and necrotizing pneumonia patients' breath smells like a sewer, and a stench of mute or fishy reek can suggest liver sickness. Furthermore, the breath smells of untreated diabetic patients are often characterized as "rotten apples" because of the prevalence of acetone in their breath, while failing kidneys can be accompanied by a urine-like smell. Thus, it has been established in the literature that analyzing the breath prints of a LC patient offers several potential advantages (as listed below) than other conventional medical tests (Ahmed *et al.*, 2017; Shehada *et al.*, 2016; Mertin, 2011):

- i. For real-time diagnosis and tracking, breath analysis has great potential
- ii. In a gas matrix (breath), the calculation of gas-phase analytes is simplified than in more diverse genetic matrices
- iii. Expiratory-breath research is uniquely informative, completely non-invasive and an adjunct or precursor to proteogenomics techniques
- iv. Unlike other specimens, breath can be collected easily as often as necessary
- v. With no unnecessary side effects, it is simple to acquire and painless for the subject

Breath is a mixture of inorganic gas, inert gas and a tiny percentage of the VOC concentration. When breathing, several molecules are emitted into the air in the range of parts per million (ppm) to parts per trillion (ppt) by volume. The VOCs in breath are produced by endogenous or exogenic processes, and their quantitative and qualitative characterization in individuals is not the same.

O'Neill *et al.* (1988) claimed that 28 "fingerprints", which include lipid-peroxidation, hexane, aniline methylpentane and o-toluidine cancer, are present in the respiratory tract of LC victims (Gordon *et al.*, 1985; Sponring *et al.*, 2010). The distinction in the breath of LC and healthier participants is also confirmed by indicating alkanes and benzene compounds among the 22 VOCs for effective identification of LC (Phillips *et al.*, 2003). To date, no less than 3000 VOCs in the breath of healthy volunteers have been reported and a few of the VOCs are Isoprene, Ethane, Pentane, Acetone and etc.

In clinical prognosis, VOCs arise as a result of metabolic reactions and are informative (Phillips *et al.*, 2003; Mochalski *et al.*, 2013; Rocco, 2018; Phillips *et al.*, 1999; Horváth *et al.*, 2005). Because LC cells have specific metabolic features, the chemistry of VOCs in the expiratory breath parallels cell metabolism within the body. These metabolic residual products flow through the bloodstream and are transferred to the lungs, where they are breathed out via the circulation system. This means that alterations in the metabolic processes of the body result in distinctive prints or

signatures of VOC. Researchers have also indicated that the unique metabolism of the disease can be found as breath "fingerprints" showing the presence of LC from the study of exhaled breath of LC patients (Sponring *et al.*, 2010; An *et al.*, 2010; Jordan *et al.*, 2010).

Chemical Analysis of Volatile Organic Compound

There are several procedures explained in the literature to sample, investigate and anatomize exhaled VOCs. The Gas Chromatography (GC) technique is the main standard and is mostly used for chemical analysis of VOCs. Proton Transfer Reaction Mass Spectrometry (PTR-MS), Selected Ion Flow Tube Mass Spectrometry (SIFT-MS), Ion Mobility Spectrometry (IMS), Thermal Desorption Gas Chromatography Mass Spectrometry (TD-GC-MS), Solid Phase Microextraction Mass Spectrometry (SPME-MS), Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) and FT-ICR-MS are other methods used to investigate VOCs in LC.

Breath sampling in the chemical analysis approach is an important stage in the procedure. Initially, the exhaled breath is collected and preserved briefly. Each VOC is evaluated by GC following a desorption process, which is typically followed by mass spectrometry (GC-MS). Depending on the chemical characteristics, the different VOCs are first isolated, sequentially ionized, and classified by their mass-to-charge (m/z) ratio.

The most common sampling method for breath collection is to use Tedlar sampling bags because they are inexpensive, can easily be manipulated, and are reusable. They are produced from inert resources to prevent both diffusion and reactions between the compounds and the bag. Other sampling methods found in the literature include, Mylar Bags, Cold Trap Systems, Bio VOCs and Gas Tight Syringes (Callol-Sanchez *et al.*, 2017; Nakhleh *et al.*, 2017; Oguma *et al.*, 2017; Rudnicka *et al.*, 2011). Sampled VOCs are usually processed using different statistical methods. A few of the research papers for the analytical chemistry of VOCs in the respiratory tract are shown in Table 3.

However, these methods of LC detection are not readily available in health care facilities. This is due to the level of sophistication of the equipment required. They are bulky, expensive and only available in specialized laboratories. Also, these procedures require in-depth preparation of the sample, rare skill-sets to interpret the results and are time-consuming. Furthermore, breath VOC analysis at part per billion by volume (ppbv) to part per trillion by volume (pptv) involving GC-MS based methods requires pre-concentration on commencement of the procedure. This can enhance indications of some VOCs and potentially miss other VOCs (Schmidt and Podmore, 2015; Van de Goor *et al.*, 2018). Even though these techniques are effective for early identification of LC, they are not compact, have no flexibility of use, and, as such, cannot be used in hospitals or homes (at the Point of Care (POC)).

Table 2: Some LC detection studies using proteomics and genomics

S/N	Biomarker	Tumor class	Percentage of Sensitivity	Reference
1	DNA methylation	54 NSCLC adjacent normal tissues	85.2	Liu <i>et al.</i> (2018)
2	DNA methylation	138 formalin-fixed -116 LC and 22 controls	100	Walter <i>et al.</i> (2018)
3	miRNA biomarkers	Exosomes from BAL 4 Tissue samples	N/A	Kim <i>et al.</i> (2018)
4	Proteomic	13 AC vs. 15 controls -BAL -139 LC and 49 controls -43 SCLC and 96 NSCLC	95	Kim <i>et al.</i> (2014)
5	Genomic	NSCLC	99.47	Adetiba and Olugbara (2015a)
6	Genomic	NSCLC	95.9	Adetiba and Olugbara (2015b)
7	Proteomic	- Tissue – 13 NSCLC 6 AC and 7 SCC	N/A	Nan <i>et al.</i> (2017)
8	Proteomic	- BAL - 90 suspected LC prospective followed for two years - Tissue	N/A	Carvalho <i>et al.</i> (2017)
9	Liquid biopsy	-108 patients with malignant nodules and -113 patients with benign lung nodules	75	Shen <i>et al.</i> (2011)
10	DNA methylation	Plasma and sputum -150 NSCLC - 60 Control	Sputum – 98 Plasma – 93	Hulbert <i>et al.</i> (2017)
11	Proteomic	- Tissue - 14 AC vs. adjacent normal tissue	N/A	Hsu <i>et al.</i> (2016)
12	miRNA	- Sputum - 30 NSCLC patients vs. 30 controls - SCC vs. AC	90	Bagheri <i>et al.</i> (2019)

Deoxyribonucleic acid (DNA), microribonucleic Acid (miRNA), Bronchoalveolar Lavage (BAL), Not Available (NA) , Non-Small Cell LC (NSCLC), Adenocarcinoma (AC), Squamous Cell Carcinoma (SCC)

Table 3: Existing Literature on Chemical Analysis of Breath VOCs

Reference	Biomarkers	Detection technique	Breath collection (Sampling) method	Statistical tool
Nakhleh <i>et al.</i> (2017)	Isononane Undecane Nonanal	GC-MS	Mylar bag	Linear regression models
Oguma <i>et al.</i> (2017)	Cyclohexane	GS	Cold Trap System	Wilcoxon's rank test
Callol-Sanchez <i>et al.</i> (2017)	Hexanal Octanal Nonanoic Nonanal Propanoic Heptanal	TD-GC/MS	BioVOC® breath Sampler	Mann-Whitney test
Ligor <i>et al.</i> (2015)	Butane 2-methyl-butane 2-pentanone 4-methyl octane Propane 2,4-dimethyl heptanes Propene	SPME/ GC-MS	Tedlar bags	PCA
Zou <i>et al.</i> (2014)	5-(2-methylpropyl)nonane, 8-hexylpentadecane 2,6,11-trimethyldodecane 2,6-di-tert-butyl-4-methylphenol	SPME/GC-MS	Tedlar bags	Pearson's χ^2 test
Fu <i>et al.</i> (2014)	Hexadecanal 2-Butanone 3-hydroxy-2-butanone 2-hydroxyacetaldehyde 4-hydroxyhexanal	FT-ICR- MS	Tedlar bags	Wilcoxon statistical test
Handa <i>et al.</i> (2014)	Dodecane	MCC/IMS	No storing of samples	Decision Tree
Buszewski <i>et al.</i> (2012)	1-propanol 2-Propanol 2-Pentanone Ethylbenzene ethyl acetate 2-butanone Butanal	GCMS	Tedlar bags	PCA
Wang <i>et al.</i> (2012)	Hexadecanal	SPME/GC-MS	Tedlar bag	Linear DA
Rudnicka <i>et al.</i> (2011)	Acetone Benzene Pentane Styrene Cyclohexane	SPME/GC-TOF-MS	Tedlar bags	Discriminant Factor Analysis (DFA)
Song <i>et al.</i> (2010)	butan-1-ol 3-hydroxybutan-2-one	SPME/GC-MS	Tedlar bag	Wilcoxon statistical test
Fuchs <i>et al.</i> (2010)	Pentanal Nonanal Octanal Hexanal	SPME/GC-MS	Gas tight syringe	Linear Regression
Bajtarevic <i>et al.</i> (2009)	2,3-butanedione Isoprene Acetone methanol 2-Butanone Benzaldehyde 1-propanol	SPME-GCMS	Tedlar bags	Kruskal-Wallis test

GC – gas chromatography; MS-mass spectrometry, GC-MS- gas chromatography– mass spectrometry, TD-GC-MS – thermal desorption-GC-MS, SPME/GC-MS-solid phase microextraction GC-MS, FT-ICR-MS – Fourier transform ion cyclotron resonance MS, MCC/IMS – multi capillary column-ion mobility spectrometry

Electronic Nose Based Volatile Organic Compound Analysis

The E-nose is a system that senses and classifies VOCs in exhaled breath using a sensor array. By using qualitative and pattern-based "breath print" techniques, this technology alleviates the shortcomings of the chemical analysis approach.

The comparison between the human olfactory system and the E-nose is shown in Fig. 1 pictorially. As demonstrated in the top part of Fig. 1, the E-nose consists of a processing unit for sensing, that is, the hardware and the processing unit for automatic pattern recognition. The sensing device consists of an assembly of sensors where, on the basis of chemical properties, each sensor senses a particular biomarker and then transforms it to a digital signal to generate the "breath print". The signature of each biomarker is then used to create a labeled signature database that would be used to teach and build an automated

pattern recognizer. Principal Component Analysis (PCA) is primarily used to obtain main variables while the pattern recognition units employ supervised pattern classification algorithms such as k-Nearest Neighbors (k-NN), Support Vector Machine (SVM) and Artificial Neural Network (ANN) for the classification of datasets (Estakhroyeh *et al.*, 2018). E-nose as a breathomics procedure has phenomenal prospects to promote PoC-centered early detection of lung cancer. This technique is economical, non-invasive, has a quick response time and highly portable. Brief descriptions of some of the current E-nose applications are given in Table 4.

Early diagnosis is supported by E-nose. It is an affordable, fast, portable and non-invasive tool that can detect LC at an early stage at the POC. As shown in Table 5, the practicability of using E-noses for detecting health concerns and detecting diseases in the primary phase has been established by different researchers.

Table 4: Summary of some of the E-nose application fields

Reference	Sensor type	Classifier	Application Area
Adak <i>et al.</i> (2020)	QCM	ANN-ABC	Alcohols detection
Moschini <i>et al.</i> (2018)	MOS	PCA	Meat flavor discrimination
Sarno and Wijaya (2017)	MOS	KNN	Diabetes detection
Zetola <i>et al.</i> (2017)	QMB	KNN	Tuberculosis detection
Långkvist <i>et al.</i> (2013)	ZnO	PCA, SVM	Meat spoilage markers detection
Fedorov <i>et al.</i> (2021)	MQ	LDA	Cooking state of grilled chicken
Tudu <i>et al.</i> (2009)	MOS	Fuzzy	Quality evaluation of black tea

MOS - Metal Oxide Sensor, QMB-quartz microbalance, ZnO-Zinc Oxide

Table 5: Existing Studies on E-nose based Analysis of Breath VOCs

Author	Sensor type	Classifier	Sample size		Model Accuracy (%)	Sensitivity (%)	Specificity (%)	Remark
			Lung Cancer	Control				
Chang <i>et al.</i> (2018)	MOS	Multilayer Perceptron (MPL)	37	48	75	79	72	-The focus was on iso-butane, ethanol and toluene biomarkers. -Only Caucasian samples. -Chemical noise. -Limited smokers.
Van de Goor <i>et al.</i> (2018)	MOS	ANN	60	107	83	83	84	-Early-stage tumor participants were not included.
Wong <i>et al.</i> (2018)	MOS	PCA-KNN	33	12	84	Nil	Nil	-The biomarkers of focus were not mentioned. - System was trained with only shallow learning.
Li <i>et al.</i> (2017)	LDA	Fuzzy	24	13	91.59	91.72	91.58	- C# based software was used to generate the breath pattern. - System was trained off-device with a shallow learning.
Shlomi <i>et al.</i> (2017)	Nanomaterial-based sensor array	DFA	16	30	Nil	75	76	- Small dataset. -System used wild EGFR mutation breath print.
Tirzite <i>et al.</i> (2017)	Cyranose 320 sensor	SVM	45	39	87.3	88.9	66.7	-Feature extraction method was not stated. -The model could not train 31.2% of the control. -Researcher encouraged use of larger number of subjects.
Gasparri <i>et al.</i> (2016)	QMB	PLS-DA	70	76	Nil	58	91	-Biomarkers was not stated. -Higher sensitivity to stage 1 LC but conflicting sensitivity at higher stadiation. -PCA was used for feature extraction.
Shehada <i>et al.</i> (2016)	Silicon nanowire field effect transistor (SiNW FET)	ANN	149	129	77.1	20	94	-The biomarkers were not stated. -SiNW FET was fabricated and was used to check breath print of different disease such as asthma,

Table 5: Continue

Capuano <i>et al.</i> (2015)	QMB	PLS-DA	20	10	76	90	Nil	COPD, gastric cancer and LC - The system exhibited limited sensing to LC -Poor baseline manipulation and signal-to-noise ratio -Short response time and large signal to noise ratio. -Data information were not provided - Sensors have poor adaptability to environment and operate at high temperature. -Catalyst poisoning. -Endoscopic sampling method was used to sample breath -Sensors requires environmental control, baseline drift, low sensitivity to ammonia and carbon dioxide -An array of sensors was fabricated by combining nine different chemo resistors based on the gold nanoparticles and breath sample were simulated. -High power consumption.
Chatterjee <i>et al.</i> (2013)	Quantum resistive sensors	N/A	Nil	Nil	Nil	Nil	Nil	
Mazzone <i>et al.</i> (2012)	Colorimetric sensor array	Logistic prediction model	92	137	81	Nil	Nil	
Santonico <i>et al.</i> (2012)	Metalloporphyrins-coated QCM sensors	PLS-DA	20	10	75	Nil	Nil	
Peng <i>et al.</i> (2009)	Organically Functionalized Au nanoparticle	N/A	40	56	86	Nil	Nil	
Chen <i>et al.</i> (2006)	SAW	ANN	5	5	NIL	NIL	NIL	

-High signal to noise ratio.
 - Incomplete report of result.

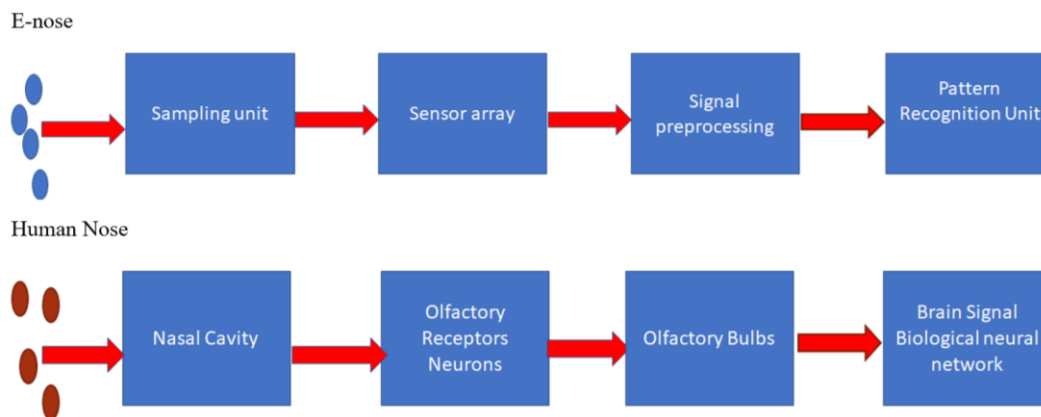


Fig. 1: E-nose versus human olfactory block diagram

As shown in Table 5, the majority of the E-noses developed by researchers are dependent on breath samples from Caucasians, which may not correctly identify black people and other races with LC. In addition, the reported E-noses and inferences were trained off-device. The sample sizes utilized in training the existing E-noses were small populations of high-risk patients. Furthermore, shallow machine learning algorithms were largely used to develop the E-noses in the literature, which could possibly hinder the accuracy, sensitivity and specificity that are sufficient for PoC application.

Research Opportunities and Future Directions

This literature review shows that the metabolic alterations associated with LC can be captured from breath as VOC biomarkers, with clinical significance for early detection of LC. Chemical analysis and E-nose-based approaches are used to identify and analyse these VOCs.

Many improvements have been made on the procedure or sensing material since the inception of this technology to create low-cost, efficient and handheld devices. Nevertheless, the state-of-the-art results on E-nose research have shown that the device is largely application-specific rather than universal. It can detect and discriminate the production profiles of VOC from microbial infections in-situ and other genetic diseases (Phillips *et al.*, 1999; Långkvist *et al.*, 2013).

Furthermore, it has provided a plethora of benefits to a variety of commercial industries, including the agricultural, biomedical, cosmetics, environmental, food, manufacturing, military, pharmaceutical, regulatory and various scientific research fields. However, further multi-disciplinary research efforts on the technology will culminate in advances that will improve its attributes such as uniformity, consistency, sensitivity, specificity and accuracy. The VOC biomarker method for LC detection offers several advantages over customary imaging techniques by enabling early spotting of malignant nodules at the PoC in

a non-invasive and pain-free manner, with the added benefit of low cost. This advantage makes this approach a more probable LC detection method, especially in third world countries.

Conclusion

This study has shown the resurgence of interests in LC early detection and the trends in the application of radiomics through proteomics/genomics to breathomics (E-nose) for LC detection over the years. However, there have been only a few efforts on the deployment of E-nose for early detection of LC at the PoC. With increasing innovations in sensing and detection technologies, machine learning and signal processing, E-nose device can be used as a PoC tool to provide rapid data processing and real-time results. Thus, if VOC biomarkers are leveraged for early detection of LC, there would be drastic reduction in LC patients' mortality rate. A non-invasive, highly specific, sensitive, accurate, cheap, portable and easy to operate intelligent E-nose based on breathomics approach has a great potential for improving the medical outlook of LC.

Acknowledgement

This work was carried out at the Advanced Signal Processing and Machine Intelligence Research (ASPMIR) Group Laboratory, Covenant University, Ota, Nigeria. KZN e-Skills CoLab, Durban University of Technology, Durban, South Africa provided support for the publication of this work.

Author's Contributions

Funmilayo S. Moninuola: Writing of the manuscript

Oluwadamilola Oshin: Contributed to the writing of the manuscript, Critical review and editing of the manuscript.

Emmanuel Adetiba: Designed the research plan and organized the study, contributed to the writing of the manuscript, Editing of the manuscript.

Anthony A. Atayero: Co-designed the research plan and organized the study, Editing of the manuscript.

Ademola Adeyeye: Editing of the medical aspect of the manuscript. critical review of draft article.

Victoria Oguntosin: Contributed to the writing of the manuscript, Critical revision of the article.

Olusola O. James: Editing of the manuscript and critical review of the final draft.

Anthony A. Adegoke: Critical review of the final draft and galley proof.

Obiseye O. Obiyemi: Critical review of final draft of the article and galley proof.

Surendra Thakur: Critical review and revision of final draft of the article. Source for funding of manuscript publication.

Abdultaofoek Abayomi: Editing of the manuscript and critical review of the final draft.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved

References

- Adak, M. F., Lieberzeit, P., Jarujamrus, P., & Yumusak, N. (2020). Classification of alcohols obtained by QCM sensors with different characteristics using ABC based neural network. *Engineering Science and Technology, an International Journal*, 23(3), 463-469. doi.org/10.1016/j.jestch.2019.06.011.
- Adetiba, E., & Olugbara, O. O. (2015a). Improved classification of lung cancer using radial basis function neural network with affine transforms of voss representation. *PloS one*, 10(12), e0143542.
- Adetiba, E., & Olugbara, O. O. (2015b). Lung cancer prediction using neural network ensemble with histogram of oriented gradient genomic features. *The Scientific World Journal*, 2015.
- Ahmed, W. M., Lawal, O., Nijsen, T. M., Goodacre, R., & Fowler, S. J. (2017). Exhaled volatile organic compounds of infection: A systematic review. *ACS Infectious Diseases*, 3(10), 695-710. doi.org/10.1021/acsinfecdis.7b00088.
- An, Z., Chen, Y., Zhang, R., Song, Y., Sun, J., He, J., ... & Abliz, Z. (2010). Integrated ionization approach for RRLC-MS/MS-based metabonomics: finding potential biomarkers for lung cancer. *Journal of proteome research*, 9(8), 4071-4081. doi.org/10.1021/pr100265g.
- Antoniou, S. X., Gaude, E., Ruparel, M., Van Der Schee, M. P., Janes, S. M., Rintoul, R. C., & LuCID research group. (2019). The potential of breath analysis to improve outcome for patients with lung cancer. *Journal of breath research*, 13(3), 034002. doi.org/10.1088/1752-7163/ab0bee.
- Bagheri, A., Khorshid, H. R. K., Tavallaie, M., Mowla, S. J., Sherafatian, M., Rashidi, M., ... & Hosseini, S. M. (2019). A panel of noncoding RNAs in non-small-cell lung cancer. *Journal of cellular biochemistry*, 120(5), 8280-8290.
- Bajtarevic, A., Ager, C., Pienz, M., Klieber, M., Schwarz, K., Ligor, M., ... & Amann, A. (2009). Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC cancer*, 9(1), 1-16.

- Biederer, J., Beer, M., Hirsch, W., Wild, J., Fabel, M., Puderbach, M., & Van Beek, E. J. R. (2012). MRI of the lung (2/3). Why... when... how?. Insights into imaging, 3(4), 355-371. doi.org/10.1007/s13244-011-0146-8.
- Boots, A. W., van Berkel, J. J., Dallinga, J. W., Smolinska, A., Wouters, E. F., & van Schooten, F. J. (2012). The versatile use of exhaled volatile organic compounds in human health and disease. Journal of breath research, 6(2), 027108.
- Broodman, I., Lindemans, J., Van Sten, J., Bischoff, R., & Luijck, T. (2017). Serum protein markers for the early detection of lung cancer: A focus on autoantibodies. Journal of proteome research, 16(1), 3-13.
- Buszewski, B., Ligor, T., Jezierski, T., Wenda-Piesik, A., Walczak, M., & Rudnicka, J. (2012). Identification of volatile lung cancer markers by gas chromatography–mass spectrometry: Comparison with discrimination by canines. Analytical and bioanalytical chemistry, 404(1), 141-146.
- Callol-Sanchez, L., Munoz-Lucas, M. A., Gomez-Martin, O., Maldonado-Sanz, J. A., Civera-Tejuca, C., Gutierrez-Ortega, C., ... & Jareno-Esteban, J. (2017). Observation of nonanoic acid and aldehydes in exhaled breath of patients with lung cancer. Journal of breath research, 11(2), 026004.
- Capuano, R., Santonico, M., Pennazza, G., Ghezzi, S., Martinelli, E., Roscioni, C., ... & D'Amico, A. (2015). The lung cancer breath signature: A comparative analysis of exhaled breath and air sampled from inside the lungs. Scientific Reports, 5(1), 1-10.
- Carvalho, A. S., Cuco, C. M., Lavareda, C., Miguel, F., Ventura, M., Almeida, S., ... & Matthiesen, R. (2017). Bronchoalveolar lavage proteomics in patients with suspected lung cancer. Scientific reports, 7(1), 1-13.
- Chang, J. E., Lee, D. S., Ban, S. W., Oh, J., Jung, M. Y., Kim, S. H., ... & Jheon, S. (2018). Analysis of volatile organic compounds in exhaled breath for lung cancer diagnosis using a sensor system. Sensors and actuators B: Chemical, 255, 800-807.
- Chatterjee, S., Castro, M., & Feller, J. F. (2013). An e-nose made of carbon nanotube based quantum resistive sensors for the detection of eighteen polar/nonpolar VOC biomarkers of lung cancer. Journal of Materials Chemistry B, 1(36), 4563-4575.
- Chen, S., & Suzuki, K. (2012). Computerized detection of lung nodules by means of “virtual dual-energy” radiography. IEEE Transactions on Biomedical Engineering, 60(2), 369-378.
- Chen, X., Cao, M., Hao, Y., Li, Y., Wang, P., Ying, K., & Pan, H. (2006, January). A non-invasive detection of lung cancer combined virtual gas sensors array with imaging recognition technique. In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (pp. 5873-5876). IEEE.
- Chicklore, S., Goh, V., Siddique, M., Roy, A., Marsden, P. K., & Cook, G. J. (2013). Quantifying tumour heterogeneity in 18 F-FDG PET/CT imaging by texture analysis. European journal of nuclear medicine and molecular imaging, 40(1), 133-140.
- Choi, W. J., & Choi, T. S. (2013). Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classification approach. Entropy, 15(2), 507-523.
- de Guevara Hernández, D. L. (2015). The role of PET/CT imaging in lung cancer. Journal of Cancer Therapy, 6(08), 690.
- Demir, Ö., & Yılmaz Çamurcu, A. (2015). Computer-aided detection of lung nodules using outer surface features. Bio-medical materials and engineering, 26(s1), S1213-S1222.
- Estakhroyeh, H. R., Rashedi, E., & Mehran, M. (2018). Design and construction of electronic nose for multi-purpose applications by sensor array arrangement using IBGSA. Journal of Intelligent & Robotic Systems, 92(2), 205-221.
- Fedorov, F. S., Yaqin, A., Krasnikov, D. V., Kondrashov, V. A., Ovchinnikov, G., Kostyukevich, Y., ... & Nasibulin, A. G. (2021). Detecting cooking state of grilled chicken by electronic nose and computer vision techniques. Food Chemistry, 345, 128747.
- Firmino, M., Angelo, G., Morais, H., Dantas, M. R., & Valentim, R. (2016). Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. Biomedical engineering online, 15(1), 1-17.
- Fu, X. A., Li, M., Knipp, R. J., Nantz, M. H., & Bousamra, M. (2014). Noninvasive detection of lung cancer using exhaled breath. Cancer medicine, 3(1), 174-181.
- Fuchs, P., Loeseken, C., Schubert, J. K., & Miekisch, W. (2010). Breath gas aldehydes as biomarkers of lung cancer. International Journal of Cancer, 126(11), 2663-2670.
- Gasparri, R., Santonico, M., Valentini, C., Sedda, G., Borri, A., Petrella, F., ... & Spaggiari, L. (2016). Volatile signature for the early diagnosis of lung cancer. Journal of breath research, 10(1), 016007.
- Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: images are more than pictures, they are data. Radiology, 278(2), 563-577.
- Gordon, S. M., Szidon, J. P., Krotoszynski, B. K., Gibbons, R. D., & O'Neill, H. J. (1985). Volatile organic compounds in exhaled air from patients with lung cancer. Clinical chemistry, 31(8), 1278-1282.
- Guida, F., Sun, N., Bantis, L. E., Muller, D. C., Li, P., Taguchi, A., ... & Hanash, S. M. (2018). Assessment of lung cancer risk on the basis of a biomarker panel of circulating proteins. JAMA oncology, 4(10), e182078-e182078.

- Han, M., Cheng, X., & Xue, Y. (2016). Comparison with reconstruction algorithms in magnetic induction tomography. *Physiological measurement*, 37(5), 683.
- Handa, H., Usuba, A., Maddula, S., Baumbach, J. I., Mineshita, M., & Miyazawa, T. (2014). Exhaled breath analysis for lung cancer detection using ion mobility spectrometry. *PloS one*, 9(12), e114555.
- Hecht, S. S. (2012). Lung carcinogenesis by tobacco smoke. *International journal of cancer*, 131(12), 2724-2732. doi.org/10.1002/ijc.27816.
- Herbst, R. S., Morgensztern, D., & Boshoff, C. (2018). The biology and management of non-small cell lung cancer. *Nature*, 553(7689), 446-454. doi.org/10.1038/nature25183.
- Hochegger, B., Marchiori, E., Sedlaczek, O., Irlon, K., Heussel, C. P., Ley, S., ... & Kauczor, H. U. (2011). MRI in lung cancer: A pictorial essay. *The British journal of radiology*, 84(1003), 661-668.
- Horváth, I., Hunt, J., & Barnes, P. J. (2005). Exhaled breath condensate: Methodological recommendations and unresolved questions. *European Respiratory Journal*, 26(3), 523-548.
- Hsu, C. H., Hsu, C. W., Hsueh, C., Wang, C. L., Wu, Y. C., Wu, C. C., ... & Yu, C. J. (2016). Identification and characterization of potential biomarkers by quantitative tissue proteomics of primary lung adenocarcinoma. *Molecular & Cellular Proteomics*, 15(7), 2396-2410.
- Hu, X., Zhao, J., Qian, H., Du, G., Kelly, M., & Yang, H. (2016). Radiological and pathological analysis of LDCT screen detected and surgically resected sub-centimetre lung nodules in 44 asymptomatic patients. *European journal of radiology open*, 3, 223-229. doi.org/10.1016/j.ejro.2016.08.001
- Hulbert, A., Jusue-Torres, I., Stark, A., Chen, C., Rodgers, K., Lee, B., ... & Herman, J. G. (2017). Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. *Clinical Cancer Research*, 23(8), 1998-2005.
- Jacobs, C., van Rikxoort, E. M., Murphy, K., Prokop, M., Schaefer-Prokop, C. M., & van Ginneken, B. (2016). Computer-aided detection of pulmonary nodules: A comparative study using the public LIDC/IDRI database. *European radiology*, 26(7), 2139-2147.
- Jain, K. K. (2016). Role of proteomics in the development of personalized medicine. *Advances in protein chemistry and structural biology*, 102, 41-52.
- Jia, T., Zhang, H., & Meng, H. (2014). A novel lung nodules detection scheme based on vessel segmentation on CT images. *Bio-medical materials and engineering*, 24(6), 3179-3186.
- Jiang, H., Ma, H., Qian, W., Gao, M., & Li, Y. (2017). An automatic detection system of lung nodule based on multigroup patch-based deep learning network. *IEEE journal of biomedical and health informatics*, 22(4), 1227-1237. doi.org/10.1109/JBHI.2017.2725903.
- Jordan, K. W., Adkins, C. B., Su, L., Halpern, E. F., Mark, E. J., Christiani, D. C., & Cheng, L. L. (2010). Comparison of squamous cell carcinoma and adenocarcinoma of the lung by metabolomic analysis of tissue-serum pairs. *Lung cancer*, 68(1), 44-50. doi.org/10.1016/j.lungcan.2009.05.012.
- Journe, N., Rehel, J. L., Le Pointe, H. D., Lee, C., Brisse, H., Chateil, J. F., ... & Bernier, M. O. (2015). Are the studies on cancer risk from CT scans biased by indication? Elements of answer from a large-scale cohort study in France. *British journal of cancer*, 112(1), 185-193. doi.org/10.1038/bjc.2014.526https://www.nature.com/articles/bjc2014526#supplementary-information.
- Junior, J. R. F., Koenigkam-Santos, M., Cipriano, F. E. G., Fabro, A. T., & de Azevedo-Marques, P. M. (2018). Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. *Computer methods and programs in biomedicine*, 159, 23-30. doi.org/10.1016/j.cmpb.2018.02.015.
- Keshani, M., Azimifar, Z., Tajeripour, F., & Boostani, R. (2013). Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system. *Computers in biology and medicine*, 43(4), 287-300.
- Khan, S. A., Nazir, M., Khan, M. A., Saba, T., Javed, K., Rehman, A., ... & Awais, M. (2019). Lungs nodule detection framework from computed tomography images using support vector machine. *Microscopy research and technique*, 82(8), 1256-1266. doi.org/10.1002/jemt.23275.
- Khobragade, S., Tiwari, A., Patil, C. Y., & Narke, V. (2016, July). Automatic detection of major lung diseases using Chest Radiographs and classification by feed-forward artificial neural network. In 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES) (pp. 1-5). IEEE. doi.org/10.1109/ICPEICES.2016.7853683.
- Kim, J. E., Eom, J. S., Kim, W. Y., Jo, E. J., Mok, J., Lee, K., ... & Kim, M. H. (2018). Diagnostic value of microRNAs derived from exosomes in bronchoalveolar lavage fluid of early-stage lung adenocarcinoma: A pilot study. *Thoracic cancer*, 9(8), 911-915.
- Kovalchik, S. A., Tammemagi, M., Berg, C. D., Caporaso, N. E., Riley, T. L., Korch, M., ... & Katki, H. A. (2013). Targeting of low-dose CT screening according to the risk of lung-cancer death. *New England Journal of Medicine*, 369(3), 245-254.
- Kumar, S. P., & Latte, M. V. (2019). Modified and optimized method for segmenting pulmonary parenchyma in CT lung images, based on fractional calculus and natural selection. *Journal of Intelligent Systems*, 28(5), 721-732. doi.org/10.1515/jisys-2017-0028.

- Långkvist, M., Coradeschi, S., Loutfi, A., & Rayappan, J. B. B. (2013). Fast classification of meat spoilage markers using nanostructured ZnO thin films and unsupervised feature learning. *Sensors*, 13(2), 1578-1592.
- Lanni, T. B., Stevens, C., Farah, M., Boyer, A., Davis, J., Welsh, R., ... & Mezwa, D. (2018). Early results from the implementation of a lung cancer screening program. *American journal of clinical oncology*, 41(3), 218-222.
- Latifi, K., Dilling, T. J., Feygelman, V., Moros, E. G., Stevens, C. W., Montilla-Soler, J. L., & Zhang, G. G. (2015). Impact of dose on lung ventilation change calculated from 4D-CT using deformable image registration in lung cancer patients treated with SBRT. *Journal of Radiation Oncology*, 4(3), 265-270.
- Li, W., Liu, H., Xie, D., He, Z., & Pi, X. (2017). Lung cancer screening based on type-different sensor arrays. *Scientific reports*, 7(1), 1-12.
- Li, W., Ye, F., Wang, D., Sun, X., Tong, W., Lian, G., ... & Zhang, D. Y. (2013). Protein predictive signatures for lymph node metastasis of gastric cancer. *International journal of cancer*, 132(8), 1851-1859.
- Ligor, T., Pater, L., & Buszewski, B. (2015). Application of an artificial neural network model for selection of potential lung cancer biomarkers. *Journal of breath research*, 9(2), 027106.
- Linning, E., Lu, L., Li, L., Yang, H., Schwartz, L. H., & Zhao, B. (2019). Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. *Academic radiology*, 26(9), 1245-1252. doi.org/10.1016/j.acra.2018.10.013.
- Liu, F., Zhang, H., Lu, S., Wu, Z., Zhou, L., Cheng, Z., ... & Mao, H. (2018). Quantitative assessment of gene promoter methylation in non-small cell lung cancer using methylation-sensitive high-resolution melting. *Oncology letters*, 15(5), 7639-7648.
- Liu, W., Wu, Y., Wang, L., Gao, L., Wang, Y., Liu, X., ... & Ye, F. (2014). Protein signature for non-small cell lung cancer prognosis. *American journal of cancer research*, 4(3), 256.
- Mac Manus, M. P., Everitt, S., Bayne, M., Ball, D., Plumridge, N., Binns, D., ... & Hicks, R. J. (2013). The use of fused PET/CT images for patient selection and radical radiotherapy target volume definition in patients with non-small cell lung cancer: Results of a prospective study with mature survival data. *Radiotherapy and Oncology*, 106(3), 292-298.
- Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detection using CT scan images. *Procedia Computer Science*, 125, 107-114.
- Marshall, H. M., Zhao, H., Bowman, R. V., Passmore, L. H., McCaul, E. M., Yang, I. A., & Fong, K. M. (2017). The effect of different radiological models on diagnostic accuracy and lung cancer screening performance. *Thorax*, 72(12), 1147-1150.
- Mazzone, P. J., Wang, X. F., Xu, Y., Mekhail, T., Beukemann, M. C., Na, J., ... & Sasidhar, M. (2012). Exhaled breath analysis with a colorimetric sensor array for the identification and characterization of lung cancer. *Journal of Thoracic Oncology*, 7(1), 137-142. doi.org/10.1097/JTO.0b013e318233d80f
- Mertin, I. A. (2011). Monitoring of propofol in breath; pharmacokinetic modeling and design of a control system (Doctoral dissertation, University of Luebeck).
- Mochalski, P., Sponring, A., King, J., Unterkofler, K., Troppmair, J., & Amann, A. (2013). Release and uptake of volatile organic compounds by human hepatocellular carcinoma cells (HepG2) in vitro. *Cancer cell international*, 13(1), 1-9. doi.org/10.1186/1475-2867-13-72
- Montilla-Soler, J. L., Mankanji, R. J., & Barron, B. J. (2018). Oncologic 18F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography: What All Physicians Need to Know. *The American journal of medicine*, 131(4), 357-364. doi.org/10.1016/j.amjmed.2017.12.006
- Moschini, M., Sigolo, S., Brugiapaglia, A., Renna, M., Lussiana, C., & Prandini, A. (2018). Use of electronic nose to discriminate meats from bulls fed diet with or without flaxseed inclusion and subjected to different aging periods. *Italian Journal of Food Science*, 30(3).
- Nakhleh, M. K., Amal, H., Jeries, R., Broza, Y. Y., Aboud, M., Gharra, A., ... & Haick, H. (2017). Diagnosis and classification of 17 diseases from 1404 subjects via pattern analysis of exhaled molecules. *ACS nano*, 11(1), 112-125. doi.org/10.1021/acsnano.6b04930
- Nan, Y., Du, J., Ma, L., Jiang, H., Jin, F., & Yang, S. (2017). Early Candidate Biomarkers of Non-Small Cell Lung Cancer Are Screened and Identified in Premalignant Lung Lesions. *Technology in cancer research & treatment*, 16(1), 66-74.
- Naqi, S. M., Sharif, M., & Lali, I. U. (2019). A 3D nodule candidate detection method supported by hybrid features to reduce false positives in lung nodule detection. *Multimedia Tools and Applications*, 78(18), 26287-26311. doi.org/10.1007/s11042-019-07819-3
- NLSTRT. (2013). Results of initial low-dose computed tomographic screening for lung cancer. *New England Journal of Medicine*, 368(21), 1980-1991.
- Oguma, T., Nagaoka, T., Kurahashi, M., Kobayashi, N., Yamamori, S., Tsuji, C., ... & Asano, K. (2017). Clinical contributions of exhaled volatile organic compounds in the diagnosis of lung cancer. *PloS one*, 12(4), e0174802.
- Ohno, Y., Kauczor, H. U., Hatabu, H., Seo, J. B., van Beek, E. J., & International Workshop for Pulmonary Functional Imaging (IWPFI). (2018). MRI for solitary pulmonary nodule and mass assessment: Current state of the art. *Journal of Magnetic Resonance Imaging*, 47(6), 1437-1458.

- O'Neill, H. J., Gordon, S. M., O'Neill, M. H., Gibbons, R. D., & Szidon, J. P. (1988). A computerized classification technique for screening for the presence of breath biomarkers in lung cancer. *Clinical chemistry*, 34(8), 1613-1618.
- Orozco, H. M., Villegas, O. O. V., Maynez, L. O., Sánchez, V. G. C., & Domínguez, H. D. J. O. (2012, July). Lung nodule classification in frequency domain using support vector machines. In 2012 11th international conference on information science, signal processing and their applications (ISSPA) (pp. 870-875). IEEE.
- Orozco, H. M., Villegas, O. O. V., Sánchez, V. G. C., Domínguez, H. D. J. O., & Alfaro, M. D. J. N. (2015). Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. *Biomedical engineering online*, 14(1), 1-20.
- Pass, H. I., Beer, D. G., Joseph, S., & Massion, P. (2013). Biomarkers and molecular testing for early detection, diagnosis and therapeutic prediction of lung cancer. *Thoracic surgery clinics*, 23(2), 211-224. doi.org/10.1016/j.thorsurg.2013.01.002
- Patz, E. F., Pinsky, P., Gatsonis, C., Sicks, J. D., Kramer, B. S., Tammemägi, M. C., ... & Aberle, D. R. (2014). Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA internal medicine*, 174(2), 269-274.
- Pauling, L., Robinson, A. B., Teranishi, R., & Cary, P. (1971). Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proceedings of the National Academy of Sciences*, 68(10), 2374-2376.
- Peng, G., Tisch, U., Adams, O., Hakim, M., Shehada, N., Broza, Y. Y., ... & Haick, H. (2009). Diagnosing lung cancer in exhaled breath using gold nanoparticles. *Nature nanotechnology*, 4(10), 669-673. doi.org/10.1038/nnano.2009.235https://www.nature.com/articles/nnano.2009.235#supplementary-information
- Phillips, M., Cataneo, R. N., Cummin, A. R., Gagliardi, A. J., Gleeson, K., Greenberg, J., ... & Rom, W. N. (2003). Detection of lung cancer with volatile markers in the breath. *Chest*, 123(6), 2115-2123.
- Phillips, M., Gleeson, K., Hughes, J. M. B., Greenberg, J., Cataneo, R. N., Baker, L., & McVay, W. P. (1999). Volatile organic compounds in breath as markers of lung cancer: A cross-sectional study. *The Lancet*, 353(9168), 1930-1933.
- Rocco, G. (2018). Every breath you take: The value of the electronic nose (e-nose) technology in the early detection of lung cancer. *The Journal of thoracic and cardiovascular surgery*, 155(6), 2622-2625. doi.org/10.1016/j.jtcvs.2017.12.155
- Rudnicka, J., Kowalkowski, T., Ligor, T., & Buszewski, B. (2011). Determination of volatile organic compounds as biomarkers of lung cancer by SPME-GC-TOF/MS and chemometrics. *Journal of Chromatography B*, 879(30), 3360-3366.
- Saba, T. (2019). Automated lung nodule detection and classification based on multiple classifiers voting. *Microscopy research and technique*, 82(9), 1601-1609. doi.org/10.1002/jemt.23326.
- Sankar, S. P., & George, D. E. (2021). Regression Neural Network segmentation approach with LIDC-IDRI for lung lesion. *Journal of Ambient Intelligence and Humanized Computing*, 1-10. doi.org/10.1007/s12652-020-02069-w
- Sanni, S., Hongoro, C., Ndinda, C., & Wisdom, J. P. (2018). Assessment of the multi-sectoral approach to tobacco control policies in South Africa and Togo. *BMC Public Health*, 18(1), 1-12. doi.org/10.1186/s12889-018-5829-3
- Santonico, M., Lucantoni, G., Pennazza, G., Capuano, R., Galluccio, G., Roscioni, C., ... & D'Amico, A. (2012). In situ detection of lung cancer volatile fingerprints using bronchoscopic air-sampling. *Lung cancer*, 77(1), 46-50.
- Sarno, R., & Wijaya, D. R. (2017, October). Detection of diabetes from gas analysis of human breath using e-Nose. In 2017 11th international conference on Information & Communication Technology and system (ICTS) (pp. 241-246). IEEE. doi.org/10.1109/ICTS.2017.8265677
- Schmidt, K., & Podmore, I. (2015). Current challenges in volatile organic compounds analysis as potential biomarkers of cancer. *Journal of biomarkers*, 2015.
- Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S. J., ... & Van Ginneken, B. (2016). Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging*, 35(5), 1160-1169.
- Shehada, N., Cancilla, J. C., Torrecilla, J. S., Pariente, E. S., Brönstrup, G., Christiansen, S., ... & Haick, H. (2016). Silicon nanowire sensors enable diagnosis of patients via exhaled breath. *ACS nano*, 10(7), 7047-7057.
- Shen, J., Liu, Z., Todd, N. W., Zhang, H., Liao, J., Yu, L., ... & Jiang, F. (2011). Diagnosis of lung cancer in individuals with solitary pulmonary nodules by plasma microRNA biomarkers. *BMC cancer*, 11(1), 1-9.
- Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., ... & Tian, J. (2017). Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognition*, 61, 663-673. doi.org/10.1016/j.patcog.2016.05.029
- Shlomi, D., Abud, M., Liran, O., Bar, J., Gai-Mor, N., Ilouze, M., ... & Peled, N. (2017). Detection of lung cancer and EGFR mutation by electronic nose system. *Journal of Thoracic Oncology*, 12(10), 1544-1551.
- Sin, D. D., Tammemägi, C. M., Lam, S., Barnett, M. J., Duan, X., Tam, A., ... & Taguchi, A. (2013). Pro-surfactant protein B as a biomarker for lung cancer prediction. *Journal of clinical oncology*, 31(36), 4536.

- Song, G., Qin, T., Liu, H., Xu, G. B., Pan, Y. Y., Xiong, F. X., ... & Chen, Z. D. (2010). Quantitative breath analysis of volatile organic compounds of lung cancer patients. *Lung cancer*, 67(2), 227-231.
- Sponring, A., Filipiak, W., Ager, C., Schubert, J., Miekisch, W., Amann, A., & Troppmair, J. (2010). Analysis of volatile organic compounds (VOCs) in the headspace of NCI-H1666 lung cancer cells. *Cancer Biomarkers*, 7(3), 153-161.
- Taguchi, A., Hanash, S., Rundle, A., McKeague, I. W., Tang, D., Darakjy, S., ... & Perera, F. (2013). Circulating pro-surfactant protein B as a risk biomarker for lung cancer. *Cancer Epidemiology and Prevention Biomarkers*, 22(10), 1756-1761.
- Teramoto, A., Fujita, H., Takahashi, K., Yamamuro, O., Tamaki, T., Nishio, M., & Kobayashi, T. (2014). Hybrid method for the detection of pulmonary nodules using positron emission tomography/computed tomography: A preliminary study. *International journal of computer assisted radiology and surgery*, 9(1), 59-69. doi.org/10.1007/s11548-013-0910-y
- Tirzite, M., Bukovskis, M., Strazda, G., Jurka, N., & Taivans, I. (2017). Detection of lung cancer in exhaled breath with an electronic nose using support vector machine analysis. *Journal of breath research*, 11(3), 036009. <http://stacks.iop.org/1752-7163/11/i=3/a=036009>
- Tudu, B., Jana, A., Metla, A., Ghosh, D., Bhattacharyya, N., & Bandyopadhyay, R. (2009). Electronic nose for black tea quality evaluation by an incremental RBF network. *Sensors and Actuators B: Chemical*, 138(1), 90-95.
- Van de Goor, R., van Hooren, M., Dingemans, A. M., Kremer, B., & Kross, K. (2018). Training and validating a portable electronic nose for lung cancer screening. *Journal of Thoracic Oncology*, 13(5), 676-681. doi.org/10.1016/j.jtho.2018.01.024
- Vavala, T., & Novello, S. (2017). SC10. 04 second-line therapy and beyond in squamous cell NSCLC. *Journal of Thoracic Oncology*, 12(1), S101-S102.
- Walter, R. F. H., Rozynek, P., Casjens, S., Werner, R., Mairinger, F. D., Speel, E. J. M., ... & Johnen, G. (2018). Methylation of LIRE1, RARB and RASSF1 function as possible biomarkers for the differential diagnosis of lung cancer. *PloS one*, 13(5), e0195716.
- Wang, D., Ye, F., Sun, Y., Li, W., Liu, H., Jiang, J., ... & Zhang, D. Y. (2011). Protein signatures for classification and prognosis of gastric cancer: A signaling pathway-based approach. *The American journal of pathology*, 179(4), 1657-1666. doi.org/10.1016/j.ajpath.2011.06.010
- Wang, Y., Hu, Y., Wang, D., Yu, K., Wang, L., Zou, Y., ... & Ying, K. (2012). The analysis of volatile organic compounds biomarkers for lung cancer in exhaled breath, tissues and cell lines. *Cancer Biomarkers*, 11(4), 129-137.
- Welniak, T. J., Panzenbeck, A., Koyfman, A., & Foran, M. (2015). Chronic obstructive pulmonary disease: Emergency care in acute exacerbation. *African Journal of Emergency Medicine*, 5(2), 75-84. doi.org/10.1016/j.afjem.2014.08.002
- Wong, D. M., Fang, C. Y., Chen, L. Y., Chiu, C. I., Chou, T. I., Wu, C. C., ... & Tang, K. T. (2018, April). Development of a breath detection method based E-nose system for lung cancer identification. In 2018 IEEE International Conference on Applied System Invention (ICASI) (pp. 1119-1120). IEEE. doi.org/10.1109/ICASI.2018.8394477
- Xiao, Z., Tan, C., & Dong, F. (2018, June). Sensitivity comparison of a cambered magnetic induction tomography for local hemorrhage detection. In 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA) (pp. 1-5). IEEE.
- Zetola, N. M., Modongo, C., Matsiri, O., Tamuhla, T., Mbongwe, B., Matlhagela, K., ... & Di Natale, C. (2017). Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *Journal of Infection*, 74(4), 367-376. doi.org/10.1016/j.jinf.2016.12.006
- Zhong, L., Coe, S. P., Stromberg, A. J., Khattar, N. H., Jett, J. R., & Hirschowitz, E. A. (2006). Profiling tumor-associated antibodies for early detection of non-small cell lung cancer. *Journal of Thoracic Oncology*, 1(6), 513-519.
- Zou, Y., Zhang, X., Chen, X., Hu, Y., Ying, K., & Wang, P. (2014). Optimization of volatile markers of lung cancer to exclude interferences of non-malignant disease. *Cancer Biomarkers*, 14(5), 371-379.