Exhaled Breath Analysis

for Detecting Interstitial Lung Diseases

Iris G. van der Sar

EXHALED BREATH ANALYSIS FOR DETECTING INTERSTITIAL LUNG DISEASES

Analyse van uitgeademde lucht voor het opsporen van interstitiële longziekten

Iris Gerdina van der Sar

The work described in this thesis was conducted at the Department of Respiratory Medicine of Erasmus Medical Center, Rotterdam (the Netherlands).

Exhaled breath analysis for detecting interstitial lung diseases
ISBN: 978-94-6506-078-1

Unrestricted grants for research described in this thesis have been received from Boehringer Ingelheim & AstraZeneca-Daiichi Sankyo
Financial support for the printing of this thesis was kindly provided by Erasmus University Rotterdam & Department of Respiratory Medicine of Erasmus Medical Center (the Netherlands)

Layout and printing: Ridderprint | www.ridderprint.nl
Cover design and copyright holder: Anton van Es

Exhaled Breath Analysis for Detecting Interstitial Lung Diseases

Analyse van uitgeademde lucht voor het opsporen van interstitiële longziekten

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof.dr. A.L. Bredenoord

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 23 oktober 2024 om 13:00 uur

door

Iris Gerdina van der Sar
geboren te Delft.

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor	Prof.dr. M.S. Wijsenbeek-Lourens
Overige leden	Prof.dr. V. Cottin Prof.dr. P.J. Sterk Dr. H.M. Janssens
Co-Promotor	Dr. C.C. Moor

TABLE OF CONTENTS

PART I

INTRODUCTION

GENERAL INTRODUCTION

Interstitial lung disease

The field of respiratory medicine comprises many conditions that involve different parts of the respiratory system. Interstitial lung disease (ILD) is a group of rare lung diseases mainly affecting the lung interstitium. ILDs are characterized by interstitial inflammation and/or fibrosis, resulting in reduced gas exchange and generic symptoms like dyspnea, cough, fatigue, and reduced exercise tolerance. A complex interplay between exposures, immune system and genetic susceptibility forms the basis of many ILDs. However, in many cases the underlying cause of an ILD is unknown, these ILDs are considered idiopathic (e.g. idiopathic pulmonary fibrosis, idiopathic non-specific interstitial pneumonia, cryptogenic organizing pneumonia). Other categories of ILDs include autoimmune-related disease, exposure-related disease, cystic disease and sarcoidosis [1, 2]. Sarcoidosis is a multi-organ disease and encompasses a heterogeneous population in terms of number and types of organs involved, severity of symptoms, and response to treatment. As roughly 90% of patients suffer from pulmonary involvement of sarcoidosis, the diagnosis is classified within the group of ILD [3]. Both between and within specific ILD diagnosis groups, there is much heterogeneity amongst patients in terms of etiology, symptom severity and progression, extent of affected lung tissue, response to and side effects of treatment, and overall prognosis.

Diagnosis

Interstitial abnormalities can be detected on a high resolution chest computed tomography (CT) scan, but a CT scan alone is not conclusive for the diagnosis of a specific ILD. Defined patterns like usual interstitial pneumonia have a high predictive value for idiopathic pulmonary fibrosis, but inter-observer variation exists and a usual interstitial pneumonia pattern can fit other fibrotic ILDs, such as auto-immune and exposure related ILDs, or even sarcoidosis [4]. Therefore, integrating CT patterns with clinical features and additional test results is key for a conclusive diagnosis [1]. This usually requires multiple additional tests like blood samples, lung function tests, and regularly a bronchoscopy with bronchoaveolar lavage and/or tissue biopsy. Bronchoscopies are invasive, bothersome for patients and contain risk on mild to severe complications [5]. Moreover, collected material for histopathologic evaluation do not necessarily confirm a diagnosis. After testing, consensus diagnosis and treatment strategy are discussed in a multidisciplinary team (MDT) with experienced radiologists, pulmonologists, pathologists, and immunologists, among others. MDT discussion is considered the gold standard for ILD diagnosis but may not always lead to a conclusive diagnosis or conclude on a working diagnosis with a low level of certainty. Besides, consensus diagnosis of an individual patient can vary between different MDTs despite clinical guidelines [6].

Altogether, the rarity, generic symptoms, heterogeneous presentation, and lack of a single conclusive test lead to diagnostic delays in patients with suspected ILD. A definite diagnosis may take up to two years from start of first symptoms [5]. Diagnostic delay prolongs time to start of adequate treatment and is associated with worse patient outcomes [7].

Treatment

Determining the optimal treatment strategy for ILD is often challenging. Preferred treatment depends on diagnosis and various other patient-related and disease-related factors. Options for ILDs include immunosuppressive and anti-fibrotic drugs. Anti-fibrotic agents are currently available for patients with idiopathic pulmonary fibrosis and other ILDs that manifest progressive pulmonary fibrosis [8]. These drugs have improved outcomes for patients [9-12]. Nevertheless, means for prospective prediction of treatment response for an individual patient are lacking. Moreover, despite therapy, ILDs are often progressive and ultimately many patients die from respiratory failure.

Biomarkers

A single non-invasive test to quickly and accurately establish a specific ILD or phenotype is highly warranted to limit diagnostic and treatment delays for patients with a suspected ILD. Therefore, the quest for a single accurate test or biomarker to diagnose ILDs has started decades ago and is still ongoing. A biomarker, or biological marker, is a measurable trait that can indicate a (patho)physiologic process or treatment response [13]. Biomarker sources used in general medical practice include blood serum (e.g. to diagnose and monitor diabetes), stool (e.g. to screen for colorectal cancer), sweat (e.g. to diagnose cystic fibrosis), breath (e.g. to monitor asthma) and radiologic imaging (e.g. to monitor treatment response in cancer).

In ILD, several sources of biomarkers and novel clinical tests have been investigated over the years. A vast amount of serum, bronchoaveolar lavage fluid and genetic biomarkers have been associated with ILD diagnosis and prognosis in patients with pulmonary fibrosis [14]. However, the lack of robust validation studies hamper the translation to clinical implementation. Only a few biomarkers are routinely used to determine the potential underlying etiology of an ILD. For instance, serological testing may support a connective tissue disease diagnosis, like anti Scl-70 antibodies positivity for systemic sclerosis [15]. Bronchoaveolar lavage lymphocyte level is indicative of hypersensitivity pneumonitis when performed in the right clinical context [16].

Exhaled breath

Exhaled breath is a novel potential biomarker source for ILD. Compounds of exhaled breath range from very small gaseous molecules (e.g. nitric oxide) to relatively large particles (e.g. microbes) as schematically shown in **Figure 1**. For monitoring a person's health status using exhaled breath, volatile organic compounds (VOCs) are of particular interest.

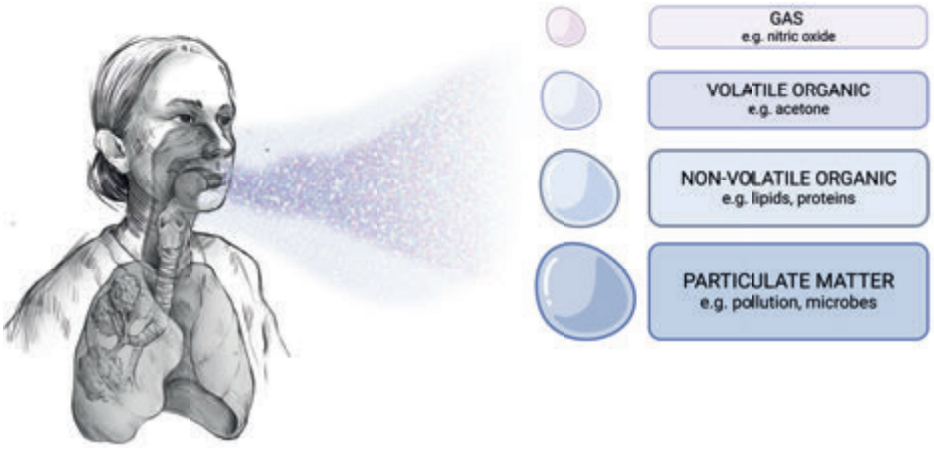


Figure 1: Compound classes of human exhaled breath.
Created with Biorender

VOCs originate from systemic physiologic and pathogenic processes in the human body, and diffuse via the blood stream to the alveoli. VOCs can also be produced locally by the lung cells and microbiome. Lastly, exhaled VOCs can result from inhaling or absorbing exogenous VOCs from various sources, such as from cigarette smoke or micro-organisms. Exogenous VOCs may, among others, be inhaled and exhaled without tissue interaction or interact with human tissue and alter VOC production [17]. However, the exact origin of most exhaled and disease-related VOCs remains unknown.

The total of VOCs that a person exhales, is called a breath or VOC profile, breathprint, breath fingerprint, or the volatilome. The field of VOC analysis is called breathomics, because of similarities with other omics studies in output and processing of data. In omics studies, full profiles of a specific biospecimen (e.g. genes, genomics; proteins, proteomics) or a clinical data source (e.g. radiologic imaging, radiomics) are investigated. To process the large amount of generated data, machine learning algorithms are typically used following a datamining approach, i.e. looking for valuable information without a specific pre-defined target [18, 19]. Subsequently, data can be used for designing artificial intelligence-based models to allow translation to clinical practice.

Various methods for VOC analysis exist; however, none of the methods has reached clinical application yet in any field of medicine. To date, most evidence for ILD is collected using gas chromatography-mass spectrometry (GC-MS) and electronic nose (eNose) sensor technology. eNose seem to have the greatest potential as point-of-care tool seeing its quick noninvasive nature. Moreover, compared to other currently used diagnostic tests as chest CT scans, biopsy and blood investigations it is relatively cheap, quick, generates results in real-time, does not expose patients to radiation or complication risks, and is accessible for health care settings across the world. In contrast to eNose technology, GC-MS allows identification of specific VOCs present in a breath sample and might therefore be an appropriate method for unravelling disease-specific biological mechanisms. GC-MS is less feasible as clinical test as analyses are time-consuming, need expensive devices and require highly-trained personnel.

The main aim of this thesis is to investigate exhaled breath analysis for ILD, with a focus on the potential of eNose technology to improve future diagnostic trajectory.

Thesis outline and aims

Part I is the general introduction of this thesis and contains three chapters. **Chapter 1** introduces eNose technology and includes an update of the current evidence of eNose research in lung disease. **Chapter 2** elaborates on different types of exhaled breath analysis and the available evidence in ILD. **Chapter 3** describes patient-reported diagnostic delay, highlighting the need for fast and accurate diagnostic testing.

Part II focuses on the use of exhaled breath analysis by an eNose to diagnose ILD. **Chapter 4** reports the accuracy of eNose technology in differentiating patients with ILD from other common respiratory diseases. **Chapter 5** shows whether an eNose can be used in international cohorts to distinguish fibrotic ILDs. **Chapter 6** evaluates whether sarcoidosis has a distinct eNose breath profile, and if this differs in patients with and without pulmonary involvement. **Chapter 7** explores a method for designing a robust diagnostic model using eNose data for pulmonary sarcoidosis.

Part III comprises eNose results for the purpose of ILD screening. Breath profiles of patients who have developed ILD are compared to those at risk. **Chapter 9** compares patients of systemic sclerosis with and without ILD. **Chapter 10** parallels patients diagnosed with drug-induced ILD due to cancer treatment, with those without ILD.

Part IV aims to explore the value of GC-MS breath analysis for ILD. **Chapter 11** investigates VOC differences between several distinct phenotype groups of ILD patients to reveal underlying pathobiology.

Part V contains **Chapter 11**, the general discussion. It discusses the outcomes and challenges of the results presented in is thesis, including the future implications for research and daily ILD care. Lastly, perspectives on artificial intelligence-based medical testing are challenged.

References

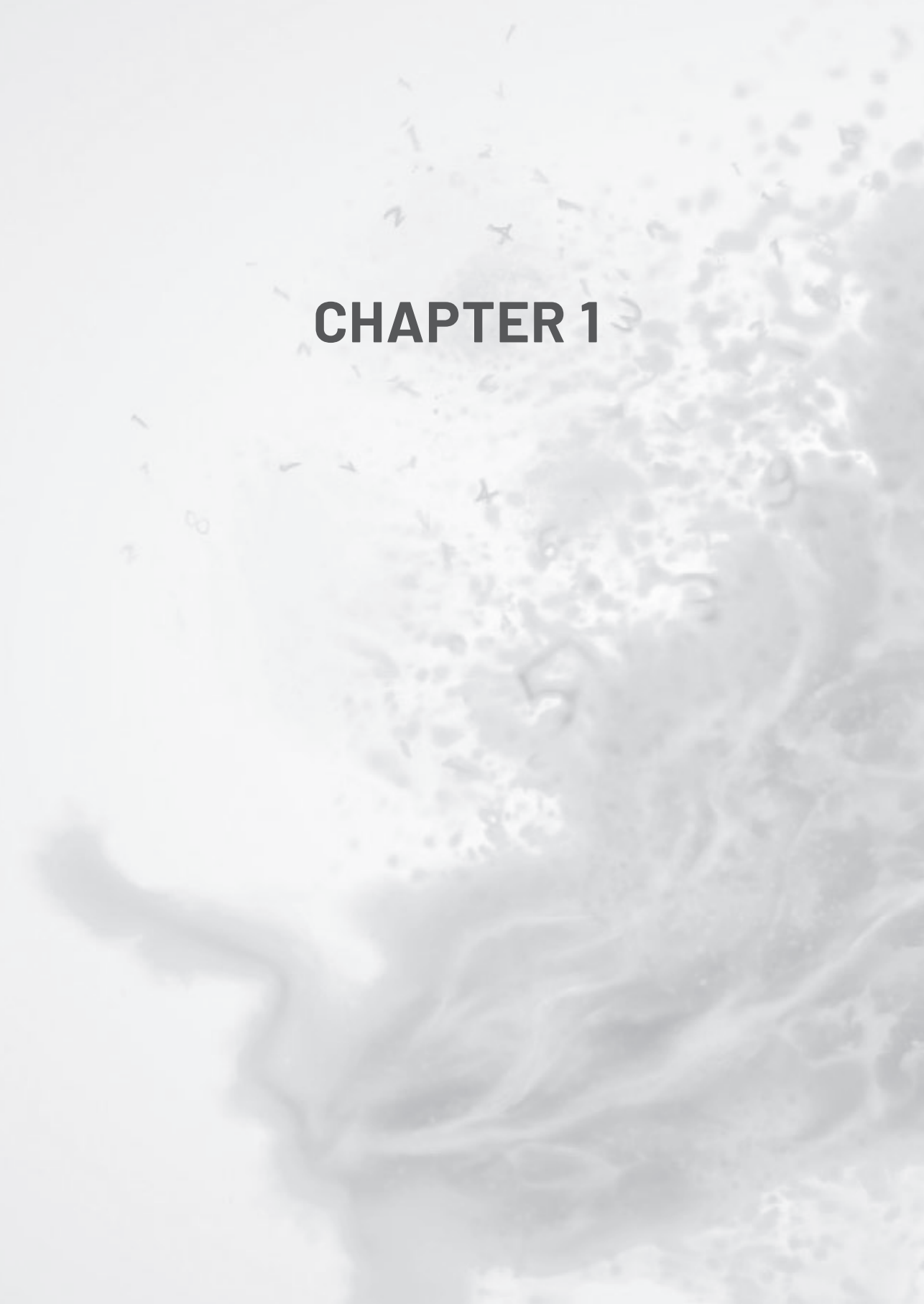
1. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med. 2013 Sep 15;188(6):733-48.
2. Wijsenbeek M, Suzuki A, Maher TM. Interstitial lung diseases. Lancet. 2022 Sep 3;400(10354):769-86.
3. Grunewald J, Grutters JC, Arkema EV, Saketkoo LA, Moller DR, Müller-Quernheim J. Sarcoidosis. Nat Rev Dis Primers. 2019 Jul 4;5(1):45.
4. Walsh SLF, Devaraj A, Enghelmayer JI, Kishi K, Silva RS, Patel N, et al. Role of imaging in progressive-fibrosing interstitial lung diseases. Eur Respir Rev. 2018 Dec 31;27(150).
5. Spagnolo P, Ryerson CJ, Putman R, Oldham J, Salisbury M, Sverzellati N, et al. Early diagnosis of fibrotic interstitial lung disease: challenges and opportunities. Lancet Respir Med. 2021 Sep;9(9):1065-76.
6. Cottin V, Martinez FJ, Smith V, Walsh SLF. Multidisciplinary teams in the clinical care of fibrotic interstitial lung disease: current perspectives. LID - 220003.
7. Hoyer N, Prior TS, Bendstrup E, Shaker SB. Diagnostic delay in IPF impacts progression-free survival, quality of life and hospitalisation rates. BMJ Open Respir Res. 2022 Jul;9(1).
8. Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1;205(9):e18-e47.
9. Nathan SD, Albera C, Bradford WZ, Costabel U, Glaspole I, Glassberg MK, et al. Effect of pirfenidone on mortality: pooled analyses and meta-analyses of clinical trials in idiopathic pulmonary fibrosis. Lancet Respir Med. 2017 Jan;5(1):33-41.
10. Flaherty KR, Wells AU, Cottin V, Devaraj A, Walsh SLF, Inoue Y, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. N Engl J Med. 2019 Oct 31;381(18):1718-27.
11. Kreuter M, Wuyts WA, Wijsenbeek M, Bajwah S, Maher TM, Stowasser S, et al. Health-related quality of life and symptoms in patients with IPF treated with nintedanib: analyses of patient-reported outcomes from the INPULSIS(R) trials. Respir Res. 2020 Jan 30;21(1):36.
12. Wijsenbeek M, Swigris JJ, Inoue Y, Kreuter M, Maher TM, Suda T, et al. Effects of nintedanib on symptoms in patients with progressive pulmonary fibrosis. Eur Respir J. 2024 Feb;63(2).
13. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001 Mar;69(3):89-95.
14. Inoue Y, Kaner RJ, Guiot J, Maher TM, Tomassetti S, Moiseev S, et al. Diagnostic and Prognostic Biomarkers for Chronic Fibrosing Interstitial Lung Diseases With a Progressive Phenotype. Chest. 2020 Aug;158(2):646-59.
15. Crouser ED, Maier LA, Wilson KC, Bonham CA, Morgenthau AS, Patterson KC, et al. Diagnosis and Detection of Sarcoidosis. An Official American Thoracic Society Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2020 2020/04/15;201(8):e26-e51.

16. Raghu G, Remy-Jardin M, Ryerson CJ, Myers JL, Kreuter M, Vasakova M, et al. Diagnosis of Hypersensitivity Pneumonitis in Adults. An Official ATS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2020 Aug 1;202(3):e36-e69.

17. van der Schee MP, Paff T, Brinkman P, van Aalderen WMC, Haarman EG, Sterk PJ. Breathomics in Lung Disease. Chest. 2015 2015/01/01/;147(1):224-31.

18. Newton CA, Herzog EL. Molecular Markers and the Promise of Precision Medicine for Interstitial Lung Disease. Clin Chest Med. 2021 Jun;42(2):357-64.

19. Barnes H, Humphries SM, George PM, Assayag D, Glaspole I, Mackintosh JA, et al. Machine learning in radiology: the new frontier in interstitial lung diseases. The Lancet Digital Health. 2023 2023/01/01/;5(1):e41-e50.



CHAPTER 1

The smell of lung disease: a review of the current status of electronic nose technology

Respir Res. 2021 Sep 17;22(1):246.

Iris G. van der Sar[#], Nynke Wijbenga[#], Gizal Nakshbandi, Joachim G.J.V. Aerts, Olivier C. Manintveld, Marlies S. Wijsenbeek, Merel E. Hellemons[#], Catharina C. Moor[#]

Erasmus Medical Center, Rotterdam (the Netherlands);
[#]Authors share first/last authorship

Abstract

There is a need for timely, accurate diagnosis, and personalised management in lung diseases. Exhaled breath reflects inflammatory and metabolic processes in the human body, especially in the lungs. The analysis of exhaled breath using electronic nose (eNose) technology has gained increasing attention in the past years. This technique has great potential to be used in clinical practice as a real-time non-invasive diagnostic tool, and for monitoring disease course and therapeutic effects. To date, multiple eNoses have been developed and evaluated in clinical studies across a wide spectrum of lung diseases, mainly for diagnostic purposes. Heterogeneity in study design, analysis techniques, and differences between eNose devices currently hamper generalization and comparison of study results. Moreover, many pilot studies have been performed, while validation and implementation studies are scarce. These studies are needed before implementation in clinical practice can be realised. This review summarises the technical aspects of available eNose devices and the available evidence for clinical application of eNose technology in different lung diseases. Furthermore, recommendations for future research to pave the way for clinical implementation of eNose technology are provided.

Keywords

Electronic nose, breath analysis, respiratory medicine, personalised medicine, machine learning, sensor technology

Background

The field of pulmonary medicine has rapidly evolved the last decades, with increasing knowledge about pathophysiology and aetiology leading to better targeted treatment strategies. Nevertheless, many chronic lung diseases have non-specific, often overlapping symptoms, which delays the diagnostic process and timely start of adequate treatment. Moreover, even specific disease entities can be very heterogeneous with varying phenotypes, and thus disease courses and optimal treatment strategies vary per patient. Accurate, non-invasive, real-time diagnostic tools and biomarkers to predict disease course and response to therapy are currently lacking in most lung diseases, but are indispensable to achieve a personalised approach for individual patients.

An emerging tool that has the potential to meet this need is an electronic nose (eNose). This device 'smells' exhaled breath for clinical diagnostics, a concept probably as old as the field of medicine itself. Exhaled breath contains thousands of molecules, also known as volatile organic compounds (VOCs). These VOCs can be divided into compounds derived from the environment (exogenous VOCs) and compounds that are the result of biological processes in the body (endogenous VOCs). Endogenous VOCs can be associated with normal physiology, but also with pathophysiological inflammatory or metabolic activity [1, 2]. Identification of individual VOCs using techniques as gas chromatography or mass spectrometry is a specific but time-consuming exercise. An eNose can be used in real-time to recognise patterns of VOCs and has therefore potential as point-of-care tool in clinical practice.

The aim of this paper is to review the current clinical evidence on eNose technology in lung disease, regarding diagnosis, monitoring of disease course and therapy evaluation. In addition, technical aspects and available eNose devices are discussed.

eNose Technology

In the time of Hippocrates, it was already acknowledged that exhaled breath can provide information about health conditions [3]. For instance, a sweet acetone breath odour indicates diabetes, a fishy smell suggests liver disease, and wounds with smell of grapes point towards pseudomonas infections [4]. Initial breath analysis studies were performed using gas chromatography or mass spectrometry. Throughout the last decades, more techniques were developed for breath analysis, for example ion mobility spectrometry, selected ion flow tube mass spectrometry and laser spectrometry [5]. Although these techniques became more advanced during the years and are very precise in identifying individual VOCs, they are very complex, laborious and thus not suitable as a real-time clinical practice tool.

Exhaled breath analysis by use of eNose technology is recently gaining increasing attention. An eNose is defined as "an instrument which comprises of an array of electronic-chemical sensors with partial specificity and an appropriate pattern recognition system, capable of recognising simple or complex odours" [6]. Sensors are used in eNoses to generate a singular response pattern. The sensors can generally be divided into three categories: electrical, gravimetric, and optical sensors. Each type responds to analytes (i.e. VOCs) in a specific way, and all types have a high sensitivity. Each sensor has advantages and disadvantages, without one type being superior in general. Electrical sensors consist of an electronic circuit connected to sensory materials. Upon binding with specific analytes, an electrical response is provided [7-10]. Consequently, a variation in electrical property of the sensor surface can be detected. Electrical sensors are low-cost, but are sensitive to temperature changes and have a limited sensor life [11]. Gravimetric (or mass sensitive) sensors label analytes based on changes in mass, amplitude, frequency, phase, shape, size, or position. Gravimetric sensors contain a complex circuitry and are sensitive to humidity and temperature [11]. Finally, optical sensors detect a change in colour, light intensity or emission spectra upon analyte binding. Optical sensors are insensitive to environmental changes, but are the most technically complex sensor-array systems and are not portable due to breakable optics and components. Due to the high complexity, they are more expensive than the other sensor types [11]. For each type of sensor, a more in depth explanation can be found in the **Additional file 1**.

Detection and recognition of odours by an eNose is similar to the functioning of the mammalian olfactory system (**Figure 1**). First, an odour is detected (by olfactory receptors in a human nose or eNose sensors), which sends off various signals (to the cortex or software). Then, these signals are pooled together and processed into a pattern. This pattern can be recognised as a particular smell (e.g. a flower) [12]. As a result, an eNose can differentiate between diseases by analysing and comparing the smelled 'breathprints' (i.e. VOC patterns) with those previously learned. The devices are hand-held, patient friendly, easy-to-use and feasible as point-of-care test.

Analysis methods

To analyse eNose breathprints, pattern recognition by machine learning is most commonly used. A machine learning model uses algorithms which automatically improve due to experience with previously presented data. These models are in general established using a five step process: data collection, data preparation, model building, model evaluation, and model improvement. Machine learning is categorised into unsupervised, supervised, and reinforcement learning [13]. In supervised learning, the algorithms are trained with labelled data input, the desired output is thus known. On the

contrary, unsupervised learning allows the algorithm to recognize patterns in the data, and group data without providing labels. Lastly, reinforcement learning encompasses the training of the machine learning models to generate decision sequences. The latter is not used in the eNose studies reviewed in this paper.

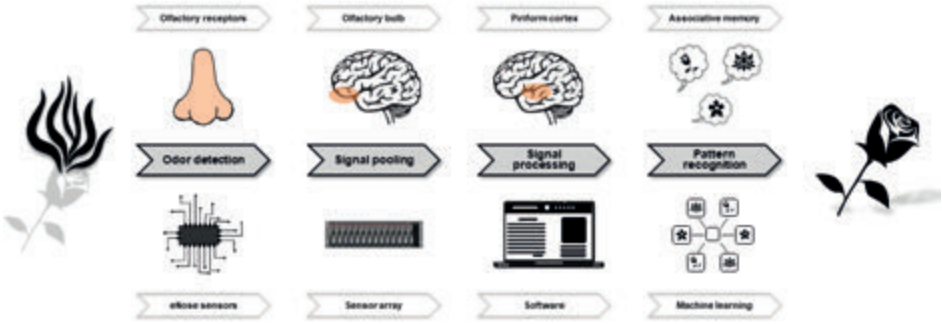


Figure 1: Schematic comparison of eNose technology and the olfactory system [12].

Several machine learning models have been proposed as appropriate algorithms for modelling complex nonlinear relationships in medical research data, such as breathprints. These models include, amongst others, artificial neural networks (mimicking the structure of animal brains to model functions), ensemble neural networks (many neural networks working together to solve a problem), and support vector machines (creating a hyperplane which allows the modelling of highly complex relationships) [14, 15]. A comparison between eNose studies show that SVM algorithm is most frequently used (10 out of 17 studies in 2019) [15]. Possibly, this is due to the fact that this is the easiest model to use for researchers new to machine learning. Another factor can be the existence of many programming languages with well-supported libraries for SVM algorithms. SVM also possesses a high accuracy, is not very prone to overfitting, and is not overly influenced by noisy data [15]. Nonetheless, there is no consensus about the optimal model for breathprint analysis.

Available eNoses

Various eNose devices have been developed and studied in different lung diseases. **Table 1** provides an overview of the specifications of devices used in studies reviewed in this paper. The choice of an eNoses device may, among others, depend on the measurement setting. For example for the BIONOTE, Cyranose 320, PEN3, and Tor Vergata eNoses the exhaled breath is captured into sample bags or cartridges which makes it possible to collect on-site and store samples for later analyses. In other

settings, it could be preferable that the eNose is easily portable, like the Aeonose. The SpiroNose is the only eNose that is capable of adjusting for disturbances from ambient air using its external sensors.

Table 1: Characteristics of available eNoses.

	Aeonose	BIONOTE	Cyranose 320	PEN3	SpiroNose	Tor Vergata
Company	The eNose company, Zutphen, the Netherlands	Campus Bio-Medico University, Rome, Italy	Sensigent, California, United States (previously known as: Smith Detections)	Airsense Analytics GmbH, Schwerin, Germany	Breathomix, Leiden, the Netherlands (previously produced by: Comon Invent)	Tor Vergata University, Rome, Italy
Working Principle (i.e. sensors)	Electrical sensors	Gravimetric sensors	Electrical sensors	Electrical sensors	Electrical sensors	Gravimetric sensors
Sensing material	MOS	QCM	Conducting polymer	MOS	MOS	QCM
Array composition	1 array; 3 sensors	1 array; 7 sensors operating at 4 different temperatures	1 array; 32 different polymers	1 array; 10 different sensors	4 exhaled breath and 4 reference arrays; 7 different sensors per array	1 array; 8 sensors
Breath collection	Tidal breathing straight into eNose	Tidal breathing into Pneumopipe cartridge	Exhalation into sample bag	Exhalation into sample bag	Exhalation straight into eNose	Exhalation into sample bag
	5 min tidal breathing (no measurement during first 2 min)	3 min tidal breathing	5 min tidal breaths, deep inhale, exhalation	5 min tidal breathing, deep in- and exhalation	5 tidal breaths, deep inhale, breath hold, slow exhalation	Deep in- and exhalation
Image						
Image source	www.enose.nl	Rocco et al. 2016 [16]	www.sensigent.com/products/cyranose.html	www.airsense.com/sites/default/files/flyer_pen.pdf	www.breathomix.com	Tor Vergata University

An overview of specifications of eNose devices used in studies reviewed in this paper. eNose prototypes are not included. BIONOTE = biosensor-based multisensorial system for mimicking nose tongue and eyes, eNose = electric nose, MOS = metal oxide semiconductor, PEN = portable electronic nose, QCM = quartz crystal microbalance. Images are used with approval of the eNose companies.

The stage of development towards a clinically implemented tool differs substantially per device and disease. Before clinical implementation, each specific eNose has to be tested as a proof of concept and consecutively in substantial cohorts for each specific disease. Subsequently, data validation and clinical implementation needs to be assessed in real-life cohorts. To give more insights in the stage of development for each eNose per lung disease, we divided studies in five different stages: 1) proof of concept study; 2) cohort size of diseased participants less than fifty; 3) cohort size of diseased participants equal or more than fifty; 4) study cohort with an external validation cohort; 5) evaluation of clinical implementation. An overview of the progress per eNose and disease is visualised in **Figure 2**. To the best of our knowledge, none of the devices are currently used in clinical pulmonology practice.

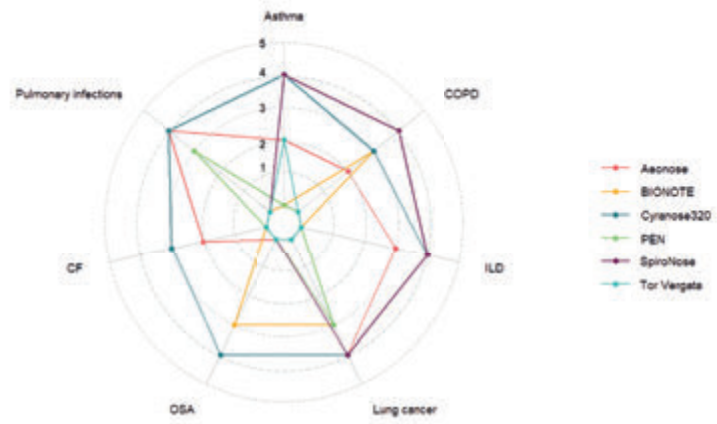


Figure 2: Radar plot of development stages per eNose and disease.
Studies were divided into five different stages: 1) proof of concept study; 2) cohort size of diseased participants less than fifty; 3) cohort size of diseased participants equal or more than fifty; 4) study cohort with an external validation cohort; 5) evaluation of clinical implementation. The highest stage reached for each eNose per lung disease is displayed. eNose prototypes are not included. BIONOTE = biosensor-based multisensorial system for mimicking nose tongue and eyes, CF = cystic fibrosis, COPD = chronic obstructive pulmonary disease, ILD = interstitial lung disease, OSA = obstructive sleep apnoea, PEN = portable electronic nose.

Current clinical application

On 21 October 2020, a systematic literature search was performed in the databases Embase, Medline (Ovid) and Cochrane Central. Search terms and selection criteria are described in the **Additional file 2**. **Table 2** provides an overview of design and results of all studies in this review.

Table 2: Literature overview eNose technology in lung disease.

	Study participants	Outcome measures	Results
Asthma			
Dragonieri 2007 [18]	n=20 asthma • n=10 mild • n=10 severe n=20 HC • n=10 old • n=10 young	Diagnostic accuracy	Mild vs young HC CVV 100%
Fens 2009 [19]	n=20 asthma n=30 COPD n=20 non-smoking HC n=20 smoking HC	Diagnostic accuracy	COPD vs asthma CVA 96%
Lazar 2010 [20]	n=10 asthma • induction of bronchoconstriction with methacholine or saline n=10 controls	Disease course	Bronchoconstriction causes no significant change in breathprint
Montuschi 2010 [21]	n=27 asthma n=24 HC	Diagnostic accuracy	eNose only Acc 87.5%
Fens 2011 [26]	*Training:* [19] n=20 asthma n=20 COPD *Validation*: n=60 asthma • n=21 fixed obstruction • n=39 classic n=40 COPD	Diagnostic accuracy	*Validation*: Classic asthma vs COPD Sens 85%, Spec 90% AUC 0.93 (0.84-1.00) Acc 83%
Van der Schee 2013 [22]	n=25 asthma n=20 HC	Diagnostic accuracy	Before OCS Sens 80.0%, Spec 65.0% AUC 0.766 ±0.14
	n=18 asthma • maintenance ICS, stop ICS (4 weeks) and OCS (2 weeks)	Therapeutic effect	OCS responsive vs not Sens 90.9%, Spec 71.4% AUC 0.883 (±0.16)
	n=25 asthma • maintenance ICS, stop ICS (4 weeks) and OCS (2 weeks) • n=13 Loss of control (LOC)	Disease course	LOC vs no LOC Sens 90.9%, Spec 71.4% AUC 0.814 ± 0.17
Plaza 2015 [28]	n=24 eosinophilic asthma n=10 neutrophilic asthma n=18 paucigranulocytic asthma	Diagnostic accuracy	Neutro vs pauci Sens 94%, Spec 80% AUC 0.88,CVA 89%
Brinkman 2017 [32]	n=22 asthma, induced LOC • maintenance ICS, stop ICS (8 weeks) and restart ICS	Disease course	Baseline vs LOC Acc 95%
Bannier 2019 [23]	n=20 asthma (age >6 years) n=22 HC	Diagnostic accuracy	Sens 74%, Spec 74% AUC 0.79
Brinkman 2019 [29]	n=78 severe asthma • n=51 longitudinal follow-up	Clustering	3 clusters (baseline), acc 93%. Differences: chronic OCS use, percent serum eosinophil and neutrophil count

			eNose	Statistical breathprint analysis
	Severe vs old HC CVV 90%	Mild vs severe CVV 65%	Cyranose 320	PCA; CDA
	COPD vs smoking HC CVA 66%	Non-smoking vs smoking HC Not significant	Cyranose 320	PCA
			Cyranose 320	PCA; mixed model analysis
	eNose + FeNO Acc 95.8%		Tor Vergata	PCA; feed-forward neural network
	Validation: Fixed asthma vs COPD Sens 91%, Spec 90% AUC 0.95 (0.87-1.00) Acc 88%	*Validation*: Fixed vs classic asthma No significant difference	Cyranose 320	PCA; CDA
	After OCS Sens 84.0%, Spec 80% AUC 0.862 ±0.12 Correlation sputum eos - breathprint R = 0.601	Before OCS (FeNO only) AUC 0.738 ±0.15	Cyranose 320	PCA; CDA
	Eosino vs neutro Sens 60%, Spec 79% AUC 0.92, CVA 73%	Eosino vs pauci Sens 55%, Spec 87% AUC 0.79 CVA 74%	Cyranose 320	PCA; CDA
	LOC vs recovery Acc 86%	Correlation sputum eos – breathprint not significant	Cyranose 320	PCA
			Aeonose	ANN
		Follow-up (18 months) n=21 cluster stable n=30 migrated	Cyranose 320, Tor Vergata, Comon Invent	PCA; Ward clustering; Non-hierarchical K-means clustering; PLS-DA; PAM; Topological data analysis

	Study participants		Outcome measures	Results

Cavaleiro Rufo 2019 [34]	n=64 suspected asthma (age 6-18 years)
	• n=45 asthma
	  • n=29 persistent
	  • n=16 intermittent
	• n=19 no asthma
			Diagnostic accuracy	Asthma vs no asthma
				Sens 77.8%, Spec 84.2%
				AUC 0.81 (0.69-0.93)
				Acc 79.7%

Dragonieri 2019 [24]	*Training:* n=14 AAR n=14 rhinitis n=14 HC	*Validation*: n=7 AAR n=7 rhinitis n=7 HC	Diagnostic accuracy	*Training:* AAR vs HC AUC 0.87 (0.70-0.97), CVA 75.0%
Abdel-Aziz 2020 [119]	*Training:* n=486 atopic asthma (age >4 years)	*Validation*: n=169 atopic asthma (age >4 years)	Diagnostic accuracy	*Training:* AUC 0.837-0.990 Sens, spec and acc only visually available
Farraia 2020 [31]	*Training:* n=121 asthma suspected (age >6 years)	*Validation*: n=78 asthma suspected (age >6 years)	Clustering	*Training:* 3 clusters (hierarchic), Differences: food/drink intake 2 h prior to sampling, percentage of asthma diagnosis in group, PEF%, age <12 y
Tenero 2020 [25]	n=28 asthma (age 6-16 years) • n=9 controlled • n=7 partially controlled • n=12 uncontrolled n=10 HC		Diagnostic accuracy	HC+controlled vs. partially+ uncontrolled Sens 79%, Spec 84% AUC 0.85 (0.72 – 0.98)
Chronic obstructive pulmonary disease (COPD)				
Fens 2011 [46]	n=28 GOLD I + II • airway inflammation (sputum eosinophil cationic protein and myeloperoxidase)		Disease course	Correlation eosinophil cationic protein and breathprint r=0.37
Hattesohl 2011 [38]	n=23 COPD (PEB) n=10 COPD (EBC) n=10 HC (EBC, PEB) n=10 AATd (EBC, PEB)		Diagnostic accuracy	COPD vs HC Sens 100%, Spec 100% CVV PEB 67.6% CVV EBC 80.5%
	n=11 AATd COPD (PEB) • augmentation therapy		Therapeutic effect	Before vs 6 d after therapy Sens 100%, Spec 100% CVV 53.3%
Fens 2013 [43]	n=157 COPD		Clustering	4 clusters (acc 97.4%) Differences: airflow limitation, health related QoL, sputum production, dyspnoea, smoking history, co-morbidity, radiologic density, gender
Sibila 2014 [39]	n=10 COPD bacterial colonised n=27 COPD non-colonised n=13 HC		Diagnostic accuracy	Colonised vs non-colonised Sens 82%, Spec 96% AUC 0.922, CVA 89%
Cazzola 2015 [40]	n=27 COPD • n=8 AECOPD ≥2 per year • n=19 AECOPD <2 per year n=7 HC		Diagnostic accuracy	COPD vs HC Sens 96%, Spec 71% CVA 91%

		eNose	Statistical breathprint analysis
Persistent vs no asthma Sens 79.7%, Spec 68.6% AUC 0.81 (0.70-0.92) Acc 79.7%	Intermittent vs no asthma Not significant	Cyranose 320	PCA; Hierarchical clustering
Validation: AAR vs HC AUC 0.77 (0.62-0.93), CVA 67.4%	*Validation*: AAR vs rhinitis AUC 0.92 (0.84-1.00) CVA 83.1%	Cyranose 320	PCA; CDA
Validation: AUC 0.18-0.926 Sens, spec and acc only visually available		Cyranose 320, Tor Vergata, Comon Invent, SpiroNose	PLS-DA; adaptive least absolute shrinkage and selection operator; gradient boosting machine
Validation: 3 clusters (hierarchic), differences: food/drink intake 2 h prior to sampling		Cyranose 320	Unsupervised hierarchic clustering; Non-hierarchical K-means clustering; PAM
		Cyranose 320	Penalized logistic regression PCA
Correlation myeloperoxidase and breathprint Not significant	Airway inflammation vs no Sens 50-73%, Spec 77-91% AUC 0.66-0.86	Cyranose 320	PCA
COPD vs AATd Sens 100%, Spec 100% CVV PEB 58.3% CVV EBC 82.0%	HC vs AATd Sens 100%, Spec 100% CVV PEB 62.0% CVV EBC 59.5%	Cyranose 320	LDA
		Cyranose 320	Hierarchical cluster analysis Non-hierarchical K-means clustering
HC vs non-colonised Sens 81%, Spec 86% AUC 0.937, CVA 83%	HC vs colonised Sens 80%, Spec 93% AUC 0.986, CVA 87%	Cyranose 320	PCA; CDA
AECOPD ≥2 vs <2 per y Not significant		Prototype (6 QMB sensors)	PLS-DA

	Study participants	Outcome measures	Results
Shafiek 2015 [41]	n=50 COPD • n=17 sputum PPM growth n=93 AECOPD • n=42 sputum PPM growth n=30 HC	Diagnostic accuracy	COPD vs HC Sens 70-72%, Spec 70-73%
	n=61 AECOPD • during and 2 months after recovery	Disease course	During vs recovery Sens 74%, Spec 67%
Van Geffen 2016 [47]	N=43 AECOPD • n=18 with viral infection • n=22 with bacterial infection	Diagnostic accuracy	With vs without viral infection Sens 83%, Spec 72% AUC 0.74
De Vries 2018 [44]	*Training:* n=321 asthma/COPD *Validation*: n=114 asthma/COPD	Clustering	5 clusters Differences: ethnicity, systemic eosinophilia/ neutrophilia, FeNO, BMI, atopy, exacerbation rate
Finamore 2018 [50]	n=63 COPD • n=32 n6MWD worsened 1 year • n=31 n6MWD stable or improved 1 year	Disease course	n6MWD change predicted by eNose Sens 84%, Spec 88% CVA 86%
Montuschi 2018 [51]	n=14 COPD • maintenance ICS, stop ICS (4 weeks) and restart ICS	Therapeutic effect	Maintenance vs restart ICS Change in 15 of 32 Cyranose sensors; 3 of 8 Tor Vergata sensors
Scarlata 2018 [45]	n=50 COPD • standard inhalation therapy (12 weeks)	Therapeutic effect	Baseline vs after 12 w Significant decline in VOCs
	n=50 COPD	Clustering	3 clusters Differences: BODE index, number of comorbidities, MEF75, KCO, pH/pCO_2 arterial blood
Van Velzen 2019 [48]	N=16 AECOPD • before, during and after recovery	Disease course	Before vs during Sens 79%, Spec 71% CVA 75%
Rodríguez-Aguilar 2020 [42]	n=116 COPD • n=88 smoking, n=28 household air pollution associated • n=64 GOLD I-II, n=52 GOLD III-IV n=178 HC	Diagnostic accuracy	COPD vs HC Sens 100%, Spec 97.8% AUC 0.989 Acc 97.8% (CDA), 100% (SVM)
Cystic fibrosis (CF)			
Paff 2013 [53]	n=25 CF n=25 primary ciliary dyskinesia (PCD) n=23 HC	Diagnostic accuracy	CF vs HC Sens 84%, Spec 65% AUC 0.76
Joensen 2014 [54]	n=64 CF • n=14 pseudomonas infection n=21 PCD n=21 HC	Diagnostic accuracy	CF vs HC Sens 50%, Spec 95% AUC 0.75
De Heer 2016 [55]	n=9 CF colonised A. fumigatus n=18 CF not colonised	Diagnostic accuracy	Sens 78%, Spec 94% AUC 0.80-0.89, CVA 88.9%
Bannier 2019 [23]	n=13 CF (age >6 years) n=22 HC	Diagnostic accuracy	Sens 85%, Spec 77% AUC 0.87

		eNose	Statistical breathprint analysis
COPD vs AECOPD no PPM Sens 89%, Spec 48% (with PPM not significant)	AECOPD PPM vs AECOPD no PPM Sens 88%, Spec 60%	Cyranose 320	LDA; SLR
With vs without bacterial infection Sens 73%, Spec 76% AUC 0.72		Aeonose	ANN
		SpiroNose	PCA; Unsupervised Hierarchical clustering
n6MWD change predicted by eNose+GOLD Sens 81%, Spec 78% CVA 79%		BIONOTE	PLS-DA
Maintenance vs restart ICS Spirometry + breathprint prediction model AUC 0.857		Cyranose 320, Tor Vergata	Multilevel PLS; KNN
		BIONOTE	PLS-DA Unsupervised K-means clustering
During vs after Sens 79%, Spec 71% CVA 75%	Before vs after Sens 57%, Spec 64% CVA 61%	Cyranose 320, Tor Vergata, Comon Invent	PCA
Smoking vs air pollution associated Not significant	GOLD I-II vs GOLD III-IV Not significant	Cyranose 320	PCA; CDA; SVM
CF vs PCD Sens 84%, Spec 60% AUC 0.77	Exacerbation CF Sens 89%, Spec 56% AUC 0.76	Cyranose 320	PCA
CF vs PCD Not significant	Pseudomonas vs. non-infected CF Sens 71.4%, Spec 63.3% AUC 0.69 (0.52 – 0.86)	Cyranose 320	PCA
		Cyranose 320	PCA; CDA
		Aeonose	ANN

	Study participants		Outcome measures	Results
Interstitial lung disease (ILD)				
Dragonieri 2013 [59]	n=31 sarcoidosis • n=11 untreated • n=20 treated n=25 HC		Diagnostic accuracy	Untreated vs HC AUC 0.825, CVA 83.3%
Yang 2018 [60]	*Training:* 80% of n=34 pneumo-coniosis n=64 HC	*Validation*: 20% of n=34 pneumo-coniosis n=64 HC	Diagnostic accuracy	*Training:* Sens 64.3-67.9%, Spec 88.0-92.0% AUC 0.89-0.91 Acc 80.8-82.1%
Krauss 2019 [61]	n=174 ILD • n=51 IPF • n=25 CTD-ILD n=33 HC n=23 COPD		Diagnostic accuracy	IPF vs HC Sens 88%, Spec 85% AUC 0.95
Dragonieri 2020 [62]	n=32 IPF n=36 HC n=33 COPD		Diagnostic accuracy	IPF vs HC AUC 1.00 (1.00-1.00) CVA 98.5%
Moor 2020 [58]	*Training:* n=215 ILD • n=57 IPF • n=158 non-IPF n=32 HC	*Validation*: n=107 ILD • n=28 IPF • n=79 non-IPF n=15 HC	Diagnostic accuracy	*Training + validation*: ILD vs HC Sens 100%, Spec 100% AUC 1.00 Acc 100%
Lung cancer (LC)				
Machado 2005 [69]	*Training:* n=14 LC n=20 HC n=27 other lung disease	*Validation*: n=14 LC n=30 HC n=32 other lung disease	Diagnostic accuracy	*Training:* LC vs HC+other CVA 71.6% (CDA)
Hubers 2014 [71]	*Training:* n=20 LC n=31 HC	*Validation*: n=18 LC n=8 HC	Diagnostic accuracy	*Training:* Sens 80%, Spec 48%
Schmekel, 2014 [89]	n=22 LC • n=10 survival >1 year • n=12 survival <1 year n=10 HC		Disease course	<1 y vs HC R = 0.95-0.98
McWilliams 2015 [72]	n=25 LC n= 166 smoking HC		Diagnostic accuracy	Sens 84-96%, Spec 63.3-81.3% AUC 0.84
Gasparri 2016 [73]	*Training:* n=51 LC n=54 HC	*Validation*: n=21 LC n=20 HC	Diagnostic accuracy	*Training + validation*: Sens 81%, Spec 91% AUC 0.874
Rocco 2016 [16]	n=100 (former) smokers • n=23 LC		Diagnostic accuracy	Detection LC Sens 86%, Spec 95% AUC 0.87
Van Hooren 2016 [82]	n=32 LC n=52 head-neck SCC		Diagnostic accuracy	Sens 84-96%, Spec 85-88% AUC 0.88-0.98 Acc 85-93%

			eNose	Statistical breathprint analysis
	Untreated vs treated CVA 74.2%	Treated vs HC Not significant	Cyranose 320	PCA; CDA
	Validation: Sens 33.3-66.7%, Spec 71.4-78.6% AUC 0.61-0.86 Acc 65.0-70.0%		Cyranose 320	LDA; SVM
	CTD-ILD vs HC Sens 84%, Spec 85% AUC 0.90	IPF vs CTD-ILD Sens 86%, Spec 64% AUC 0.84	Aeonose	ANN
	IPF vs COPD AUC 0.85 (0.75-0.95) CVA 80.0%	IPF vs COPD+HC AUC 0.84 CVA 96.1%	Cyranose 320	PCA; CDA; LDA
	Training: IPF vs non-IPF ILD Sens 92%, Spec 88% AUC 0.91 (0.85-0.96) Acc 91%	*Validation*: IPF vs non-IPF ILD Sens 95%, Spec 79% AUC 0.87 (0.77-0.96) Acc 91%	SpiroNose	PLS-DA
	Validation: LC vs HC+other Sens 71.4%, Spec 91.9% Acc 85% (SVM)		Cyranose 320	SVM PCA CDA
	Validation: Sens 94%, Spec 13%		Cyranose 320	PCA
	<1 y vs >1 y R = 0.86-0.97	Prediction model survival days R = 0.99	Applied Sensor AB model 2010	PCA; PLS; ANN
			Cyranose 320	Classification and regression tree; DFA
	Training: Sens 90%, Spec 100%	*Validation*: Sens 81%, Spec 100%	Prototype (8 QMB sensors)	PLS-DA
			BIONOTE	PLS-Toolbox; PLS-DA
			Aeonose	ANN

	Study participants		Outcome measures	Results
Shlomi 2017 [68]	n=30 benign nodule n=89 LC • n=16 early stage LC • n=53 EGFR tested (n=19 mutation)		Diagnostic accuracy	Early stage LC vs benign Sens 75%, Spec 93.3% Acc 87.0
Tirzite 2017 [84]	n=165 LC n=79 HC n=91 other lung disease		Diagnostic accuracy	LC vs HC+other Sens 87.3-88.9%, Spec 66.7-71.2% CVV 72.8%
Huang 2018 [74]	*Training:* 80% of n=56 LC n=188 HC	*Validation*: 20% of n=56 LC n=188 HC *External:* n=12 LC n=29 HC	Diagnostic accuracy	*Validation*: LC vs HC Sens 100, 92.3%, Spec 88.6, 92.9% AUC 0.96, 0.95 Acc 90.2, 92.7%
Van de Goor 2018 [76]	*Training*: n=52 LC n=93 HC	*Validation*: n=8 LC n=14 HC	Diagnostic accuracy	*Training*: Sens 83%, Spec 84% AUC 0.84 Acc 83%
Tirzite 2019 [78]	n=119 LC smoker n=133 LC non-smoker n=223 HC+other lung disease • n=91 smoking		Diagnostic accuracy	LC non-smoker vs HC+other Sens 96.2%, Spec 90.6%
Kononov 2020 [79]	n=65 LC n=53 HC		Diagnostic accuracy	Sens 85.0-95.0%, Spec 81.2-100% CVA 88.9-97.2%, AUC 0.95-0.98
Krauss 2020 [81]	n=91 LC active disease • n=51 incident LC n=29 LC complete response n=33 HC n=23 COPD		Diagnostic accuracy	LC active vs HC Sens 84%, Spec 97% AUC 0.92
Lung cancer - (non-)small cell lung cancer ((N)SCLC)				
Dragonieri 2009 [70]	n=10 NSCLC n=10 COPD n=10 HC		Diagnostic accuracy	NSCLC vs HC CVV 90%
Kort 2018 [75]	n=144 NSCLC n=18 SCLC n=85 HC n=61 suspected, LC excluded		Diagnostic accuracy	NSCLC vs HC Sens 92.2%, Spec 51.2% AUC 0.85
De Vries 2019 [88]	*Training*: n=92 NSCLC • n=42 response • n=50 no response	*Validation*: n=51 NSCLC • n=23 response • n=28 no response	Therapeutic effect (anti-PD-1 therapy)	*Training*: CVV 82%, AUC 0.89 (0.82-0.96)
Mohamed 2019 [77]	n=50 NSCLC n=50 HC		Diagnostic accuracy	Sens 92.9%, Spec 90% Acc 97.7%
Kort 2020 [80]	n=138 NSCLC n=143 controls • n=59 suspected, LC excluded • n=84 HC		Diagnostic accuracy	NSCLC vs controls (eNose data only) Sens 94.2%, Spec 44.1% AUC 0.75

		eNose	Statistical breathprint analysis
EGFR mutation vs wild type Sens 79.0%, Spec 85.3% Acc 83.0%		Prototype (40 nanomaterial-sensors)	DFA
LC vs HC Sens 97.8-98.8%, Spec 68.8-81.0% CVV 69.7%	LC stages Not significant	Cyranose 320	SVM
External validation: LC vs HC Sens 75, 83.3%, Spec 96.6, 86.2% AUC 0.91, 0.90 Acc 85.4, 85.4%		Cyranose 320	LDA; SVM
Validation: Sens 88%, Spec 86% Acc 86%		Aeonose	ANN
LC smoker vs HC+other Sens 95.8%, Spec 92.3%		Cyranose 320	LRA
		Prototype (6 MOS)	PCA; Logistic regression; KNN; Random forest; LDA; SVM
Incident LC vs HC Sens 88%, Spec 79% AUC 89%		Aeonose	ANN
NSCLC vs COPD CVV 85%		Cyranose 320	PCA; CDA
NSCLC vs HC+LC excluded Sens 94.4%, Spec 32.9% AUC 0.76	SCLC vs HC Sens 90.5%, Spec 51.2% AUC 0.86	Aeonose	ANN
Validation: AUC 0.85 (0.7–0.96) Sens 43%, Spec 100%		SpiroNose	LDA
		PEN3	PCA; ANN
NSCLC vs controls (multivariate) Sens 94.2-95.7%, Spec 49.0-59.7% AUC 0.84-0.86		Aeonose	ANN; Multivariate logistic regression

	Study participants		Outcome measures	Results
Fielding 2020 [83]	n=20 bronchial SCC • n=10 in situ • n=10 advanced stage n=22 laryngeal SCC • n=12 in situ • n=10 advanced stage n=13 HC		Diagnostic accuracy	BSCC in situ vs HC Sens 77%, Spec 80% Misclassification rate 28%
Lung cancer – Malignant Pleural Mesothelioma (MPM)				
Chapman 2012 [85]	*Training*: n=10 MPM n=10 HC	*Validation*: n=10 MPM n=32 HC n=18 benign ARD	Diagnostic accuracy	MPM vs HC *Training:* CVA 95% *Validation*: Sens 90% Spec 91%
Dragonieri 2012 [86]	n=13 MPM • internal validation with *training set* n=8, *validation set* n=5 n=13 HC n=13 AEx		Diagnostic accuracy	MPM vs HC Sens 92.3%, Spec 69.2% AUC 0.893,CVA 84.6% *Validation*: AUC 0.83, CVA 85.0%
Lamote 2017 [87]	n=11 MPM n=12 HC n=15 AEx n=12 benign ARD		Diagnostic accuracy	MPM vs HC Sens 66.7% (37.7-88.4) Spec 63.6% (33.7-87.2) AUC 0.667 (0.434-0.900) Acc 65.2% (44.5-82.3)
Pulmonary infections				
De Heer 2016 [101]	n=168 bottles with strain • n=135 bacteria + yeast • n=30 medium only • n=62 mould (A. fumigatus and R. oryzae)		Diagnostic accuracy (in vitro)	Mould vs other Sens 91.9%, Spec 95.2% AUC 0.970 (0.949-0.991) Acc 92.9%
Suarez-Cuartin 2018 [102]	n=73 bronchiectasis • n=41 colonised (n=27 pseudomonas) • n=32 non-colonised		Diagnostic accuracy	Colonised vs non-colonised AUC 0.75, CVA 72.1%
Pulmonary infections – Ventilator-associated pneumonia (VAP)				
Hanson 2005 [105]	n=19 VAP (clinical pneumonia score, CPIS ≥6) n=19 controls (CPIS <6)		Diagnostic accuracy	Correlation CPIS -breathprint $R^2 = 0.81$
Hockstein 2005 [106]	n=15 VAP (pneumonia score ≥7) n=29 HC (ventilated)		Diagnostic accuracy	Acc 66-70%
Humphreys 2011 [100]	n=44 VAP suspected • 98 BAL samples • Groups: gram-positive, gram-negative, fungi, no growth n=6 HC (ventilated)		Diagnostic accuracy (in vitro)	Differentiation groups (LDA) Sens 74-95%, Spec 77-100% Acc 83%
Schnabel 2015 [107]	n=72 VAP suspected • n=33 BAL+ • n=39 BAL- n=53 HC (ventilated)		Diagnostic accuracy	BAL+ VAP vs HC Sens 88%, Spec 66% AUC 0.82 (0.73-0.91)
Chen 2020 [15]	*Training:* 80% of n=33 VAP n=26 HC (ventilated)	*Validation*: 20% of n=33 VAP n=26 HC (ventilated)	Diagnostic accuracy	*Training:* AUC 0.823 (0.70-0.94)

			eNose	Statistical breathprint analysis
	BSCC vs LSCC adv. Sens 100%, Spec 80% Misclassification rate 10%		Cyranose 320	Bootstrap forest
	MPM vs ARD *Validation*: Sens 90% Spec 83.3%	MPM vs ARD vs HC *Validation*: Sens 90%, Spec 88%	Cyranose 320	PCA
	MPM vs AEx Sens 92.3%, Spec 85.7% AUC 0.917, CVA 80.8% *Validation*: AUC 0.88, CVA 85.9%	MPM vs AEx vs HC AUC 0.885, CVA 79.5%	Cyranose 320	PCA; CDA
	MPM vs benign ARD Sens 75.0% (45.9-93.2) Spec 64% (33.7-87.2) AUC 0.758 (0.548-0.967) Acc 48.9-85.6% (48.9-85.6)	MPM vs benign ARD+AEx Sens 81.5% (63.7-92.9) Spec 54.5% (26.0-81.0) AUC 0.747 (0.582-0.913) Acc 73.7% (58.1-85.8)	Cyranose 320	PCA
			Cyranose 320	PCA; CDA
	Pseudomonas vs other PPM AUC 0.96, CVA 89.2%	Pseudomonas vs non-colonised AUC 0.82, CVA 72.7%	Cyranose 320	PCA
			Cyranose 320	PLS
			Cyranose 320	KNN
	Differentiation groups (cross-validation) Sens 56-84%, Spec 81-97% Acc 70%		Prototype (24 MOS)	PCA; LDA
	BAL+ vs BAL- VAP Sens 76%, Spec 56% AUC 0.69 (0.57-0.81)		DiagNose	Random Forest; PCA
	Validation: Sens 79% (±8), Spec 83% (±0) AUC 0.833 (0.70-0.94) Acc 0.81 (±0.04)		Cyranose 320	KNN; Naive Bayes; decision tree; neural network; SVM; random forest

	Study participants		Outcome measures	Results
Pulmonary infections – Tuberculosis (TB)				
Fend 2006 [108]	n=188 TB n=142 TB excluded		Diagnostic accuracy (in vitro)	Sens 89% (80-97) Spec 88% (85-97)
Bruins 2013 [109]	*Training:* n=15 TB n=15 HC	*Validation*: n=34 TB n=114 TB excluded n=46 HC	Diagnostic accuracy	*Training:* Sens 95.9% (92.9-97.7) Spec 98.5% (96.2-99.4)
Coronel Teixeira 2017 [110]	*Training:* n=23 TB n=46 HC	*Validation*: n=47 TB n=63 HC+asthma+ COPD	Diagnostic accuracy	*Training:* Sens 91% Spec 93%
Mohamed 2017 [111]	n=67 TB n=56 HC		Diagnostic accuracy	Sens 98.5% (92.1-100) Spec 100% (93.5-100) Accuracy 99.2%
Saktiawati 2019 [113]	*Training:* n=85 TB n=97 HC+TB excluded	*Validation*: n=128 TB n=159 TB excluded	Diagnostic accuracy	*Training:* Sens 85% (75-92) Spec 55% (44-65) AUC 0.82 (0.72-0.88)
Zetola 2017 [112]	n=51 TB n=20 HC		Diagnostic accuracy	Sens 94.1% (83.8-98.8) Spec 90.0% (68.3-98.8)
Pulmonary infections – Aspergillosis				
De Heer 2013 [103]	n=11 neutropenia • n=5 probable/proven aspergillosis • n=6 no aspergillus		Diagnostic accuracy	Sens 100% (48-100) Spec 83.3% (36-100) AUC 0.933, CVA 90.9% (59-100)
De Heer 2016 [55]	n=9 CF colonised A. fumigatus n=18 CF not colonised		Diagnostic accuracy	Sens 78%, Spec 94% AUC 0.80-0.89, CVA 88.9%
Pulmonary infections – Corona Virus Disease (COVID-19)				
Wintjens 2020 [115]	n=219 screened • n=57 COVID-19 positive		Diagnostic accuracy	Sens 86% (74-93), Spec 54% (46-62) AUC 0.74, CVA 62%
Obstructive sleep apnoea (OSA)				
Greulich 2013 [90]	n=40 OSA n=20 HC		Diagnostic accuracy	OSA vs HC Sens 93%, Spec 70% AUC 0.85
	n=40 OSA • 3 months CPAP ventilation		Therapeutic effect	Before vs after CPAP Sens 80%, Spec 65% AUC 0.82
Incalzi 2014 [96]	n=50 OSA • 1 night CPAP ventilation		Therapeutic effect	Change in breathprint (visually different, no statistical analysis)
Dragonieri 2015 [91]	n=19 OSA n=14 obese n=20 HC		Diagnostic accuracy	Obese OSA vs HC CVA% 97.4, AUC 1.00

			eNose	Statistical breathprint analysis
			Bloodhound BH-114	PSA; DFA; ANN
	Validation: TB vs HC Sens 93.5% (91.1-95.4) Spec 85.3% (82.7-87.5)	*Validation*: TB vs TB excl. Sens 76.5% (57.98-88.5) Spec 74.8% (64.5-82.9)	DiagNose	ANN
	Validation: Sens 88% Spec 92%		Aeonose	Tucker 3–like algorithm; ANN
			PEN3	PCA; ANN
	Validation: Sens 78% (70-85) Spec 42% (34-50) AUC 0.72 (0.66-0.78)		Aeonose	ANN
			Prototype (QMB sensors)	PCA; KNN
			Cyranose 320	PCA; CDA
			Cyranose 320	PCA; CDA
			Aeonose	ANN
			Cyranose 320	PCA
			BIONOTE	PCA; PLS-DA
	Obese OSA vs obese CVA% 67.6 AUC 0.77	Obese vs HC CVA% 94.1 AUC 0.94	Cyranose 320	PCA; CDA; KNN

	Study participants		Outcome measures	Results
Kunos 2015 [97]	n=17 OSA n=9 non-OSA sleep disorder n=10 HC • 7AM and 7PM sample n=26 HC • 7AM sample		Diagnostic accuracy	OSA 7AM vs 7PM Significantly different
Dragonieri 2016 [93]	*Training:* n=13 OSA n=15 COPD n=13 overlap	*Validation*: n=6 OSA n=6 COPD n=6 overlap	Diagnostic accuracy	*Training:* OSA vs overlap CVA 96.2%, AUC 0.98
Scarlata 2017 [92]	n=40 OSA • n=20 hypoxic n=20 obese n=20 COPD n=56 HC		Diagnostic accuracy	OSA vs HC Acc 98-100%
Other – Acute respiratory distress syndrome (ARDS)				
Bos 2014 [116]	*Training:* n=40 ARDS n=66 HC	*Validation*: n=18 ARDS n=26 HC	Diagnostic accuracy	*Training:* Sens 95%, Spec 42% AUC 0.72
Other - Lung transplantation (LTx)				
Kovacs 2013 [118]	n=16 LTx recipients n=33 HC		Diagnostic accuracy Therapeutic effect	LTx recipients vs HC Sens 63%, Spec 75% AUC 0.825 Correlation breathprint - tacrolimus levels R = -0.63
Other - Pulmonary embolism (PE)				
Fens 2010 [117]	n=20 PE • n=7 comorbidity n=20 PE excluded • n=13 comorbidity		Diagnostic accuracy	Comorbidity: PE vs excluded Acc 65%, AUC 0.55

An overview of eNose technology studies in lung diseases. Studies are divided per diagnosis and displayed in chronological order. Study results shown in sensitivity/specificity, AUC and CVA (if available). In case of a training and validation set, participant numbers and results of both set are shown. All presented results are statistical significant ($p<0.05$) unless stated otherwise. AATd = alfa-1-antitrypsin deficiency, acc = accuracy, AUC = area under the curve, AAR = extrinsic asthma with allergic rhinitis, AEx = asbestos exposure, ANN = artificial neural network, ARD = benign asbestos related disease, BMI = body mass index, CDA = canonical discriminant analysis, CVA/CVV = cross-validated accuracy/value, d = days, DFA = Discriminate function analysis, EBC = exhaled breath condensate, AECOPD = acute COPD exacerbation, EGFR = epidermal growth factor receptor, eos = eosinophils, FeNO = exhaled nitric oxide test, FVC = forced vital capacity, GOLD = global initiative for chronic obstructive lung disease, HC = healthy control (not suspected for studied disease, not diagnosed with other pulmonary disease), ICS = inhaled corticosteroids, IPF = idiopathic pulmonary fibrosis, KNN = k-nearest neighbours, LDA = linear discriminant analysis, MOS = metal oxide sensor, n6MWD = normalised six minute walking distance, OCS = oral corticosteroids, PAM = partitioning around medoids, PCA = principal component analysis, PEB = pure exhaled breath, PLS-DA = partial least squares discriminant analysis, PPM = potentially pathogenic microorganism, QMB = quartz microbalance, QoL = quality of life, ROC = receiver operator characteristics, SCC = squamous cell carcinoma (B = bronchial, L = laryngeal), sens = sensitivity, SLR = Sensor Logic Relations, spec = specificity, SVM = support vector machines, TLC = total lung capacity.

			eNose	Statistical breathprint analysis
	Non-OSA or HC 7AM vs 7PM Not significantly different	(Non-)OSA 7AM vs HC 7AM Significantly different Acc 77-81%	Cyranose 320	PCA
	Validation: OSA vs overlap CVA 91.7%, AUC 1.00	*Validation*: OSA vs COPD CVA 75%, AUC 0.83	Cyranose 320	PCA; CDA
	Non-hypoxic vs hypoxic OSA Acc 60-80%	HC vs COPD Acc 100%	BIONOTE	PLS-DA
	Validation: Sens 89%, Spec 50% AUC 0.71		Cyranose 320	Sparse-partial least square logistic regression
			Cyranose 320	PCA; Linear regression
			Cyranose 320	PCA; Linear regression
	No comorbidity: PE vs excluded Acc 85%, AUC 0.81	No comorbidity: PE vs excluded (breathprint + Wells) AUC 0.90	Cyranose 320	PCA

Asthma

Asthma is a chronic lung disease characterised by reversible airflow obstruction with airway inflammation and hyperresponsiveness. Common symptoms, such as cough, chest tightness, shortness of breath and wheezing, are variable in severity and often non-specific [17]. Various studies, both in children and adults, showed that eNose technology can differentiate asthma patients from healthy controls with a good accuracy [18-25]. Two studies also demonstrated that breathprints of asthma patients were significantly different than breathprints of chronic obstructive pulmonary disease (COPD) patients [19, 26]. Interestingly, two studies reported better performance of eNose technology than conventional investigations (spirometry or an exhaled nitric oxide (FeNO) test) for detecting asthma. These studies were performed in patients with established asthma diagnosis [21, 22]. Diagnostic performance further increased when eNose technology was combined with a FeNO test (accuracy 95.7%) [21]. Moreover, even after loss of control and reaching stable disease with oral corticosteroids (OCS) treatment eNose technology could differentiate asthma from healthy controls, while the diagnostic value of FeNO decreased. In the same study, breathprint significantly predicted response to subsequent OCS treatment, while sputum eosinophils, FeNO values and, hyperresponsiveness did not [22].

The existence of multiple asthma pheno- and endotypes with different underlying pathophysiological mechanisms is increasingly acknowledged [27]. In recent years, many eNose studies have attempted to identify different clusters of asthma patients, using both supervised and unsupervised methods [28-31]. For example, supervised clustering for eosinophilic, neutrophilic and paucigranulocytic phenotypes revealed significant differences in breathprints between groups [28]. One study identified three clusters using unsupervised breathprint analysis in a group of severe asthmatic patients, corresponding with different inflammatory profiles. During follow-up, 30 of 51 patients migrated to another cluster; migration was associated with changes in sputum eosinophil count [29]. Two other longitudinal studies showed changes in breathprint when asthma control was lost after withdrawal of corticosteroids in previously stable asthma patients, and also after recovery [22, 32]. A pilot study, in which bronchoconstriction was induced in stable asthma patients, found that changes in airway calibre did not alter breathprints. Moreover, breathprints remained stable during the day in individual patients [20]. This implies that inflammatory processes and not (acute) airway obstruction influence breathprints. Overall, these findings suggest that eNose technology is a promising tool for phenotyping and monitoring asthmatics. Longer follow-up studies are required to examine whether cluster-migration or change in breathprint are also related to actual clinical course.

A currently ongoing study is evaluating whether eNose technology can be used to predict response to monoclonal antibody therapy (NCT03988790).

Paediatric asthma

In general, the diagnosis of asthma in children is challenging. Lung function tests are often difficult to perform and do not always provide a diagnosis. Interestingly, a study in 45 children demonstrated that eNose measurements were fairly well repeatable, both in healthy and asthmatic participants [33]. Moreover, two studies showed that eNose technology distinguishes children with asthma from healthy controls [23, 34]. An eNose seemed to be more accurate for diagnosing asthma than spirometry with bronchodilation only [34]. Also, uncontrolled asthma could be differentiated from controlled asthma and healthy controls [35]. Furthermore, eNose technology accurately distinguished children with persistent asthma from healthy controls, but not the ones with intermittent asthma [34]. This was possibly due to more airway inflammation reflected in the breathprints of persistent asthmatics. Hence, eNose technology could potentially facilitate easier and earlier diagnosis of asthma in children, and guide therapy in clinical practice. However, large validation studies focusing on diagnosing asthma in children are currently lacking.

COPD

Although COPD is one of the major causes of death worldwide, epidemiological studies indicate that it remains largely underdiagnosed [36]. COPD is a complex, heterogeneous disease with several phenotypes, which can overlap with asthma and pulmonary infections, among others. Furthermore, the diagnosis is delayed in patients whose symptoms are attributed to (undiagnosed) heart failure [37]. Hence, there is an unmet clinical need for accurate timely diagnosis. Also better disease course prediction and therapy guidance is warranted.

Several studies have evaluated the ability of eNose technology to diagnose COPD. Exhaled breath analysis discriminated between COPD and (smoking) healthy controls with an accuracy of 66-100% [19, 38-42]. Even though these are promising results, most studies were relatively small and lacked a validation cohort. Several studies aimed to distinguish subgroups within COPD by performing unsupervised analyses on breathprint data [43-45]. De Vries et al. performed unsupervised cluster analysis in a combined group of asthma and COPD patients [44]. Interestingly, they identified and validated five clusters which mainly differed based on clinical and inflammatory characteristics (eosinophil and neutrophil count) rather than diagnosis. Two other

studies identified 3-4 unsupervised clusters based on breathprint data. The clusters differed regarding several clinical and demographic features [43, 45]. However, in both studies, clusters were determined by different clinical parameters, showing the need for further (validation) studies. A recent study indicated that breathprints of patients with COPD associated with air pollution did not differ from smoking-associated COPD [42]. Also, no differences in breathprint between Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage I-II versus GOLD stage III-IV were detected in another study [42]. The breathprint of patients with smoking-related COPD and patients with alpha-1-antitripsin, however, could be distinguished with an accuracy of 82% in a small single-centre study [38].

eNose technology can theoretically be useful in early detection of inflammation and acute exacerbation of COPD (AECOPD), as inflammatory processes influence breathprints. This hypothesis was confirmed in a cross-sectional study evaluating the association of breathprints with different inflammation markers in sputum; eNose breathprints highly correlated with inflammatory activity [46]. In patients with an AECOPD, presence of viral and bacterial infection was accurately detected by an eNose [47]. In another group of AECOPD patients, patients with colonisation of potentially pathogenic microorganisms had a significantly different breathprint than AECOPD patients that were not colonised. Besides, AECOPD patients' breathprints differed from stable COPD patients without microorganism colonisation [41]. Stable COPD patients with bacterial colonisation were also significantly different from those without (area under the curve (AUC) 0.922) [39]. Two prospective longitudinal studies indicated that the breathprint before, during and after recovery of an AECOPD differed [41, 48]. Confirming these results in larger cohort studies might lead the way to use breathprints for earlier detection and (targeted) treatment of infections and AECOPDs. This is interesting as treatment may improve outcomes and prevent hospitalizations [49].

Regarding prognostic value of eNose technology, one study demonstrated that eNose data correlated better to change in 6-minute walking distance over one year, than the current GOLD classification [50]. A few studies evaluated the effect of initiation and withdrawal of inhalation medication on breathprints. Two studies found significant changes in breathprint after start of inhalation therapy [51, 45]. A designed multidimensional model, combining eNose technology with spirometry, gave a better indication of treatment response (AUC 0.857) than spirometry only (AUC 0.561) [51]. This small pilot study shows the potential of integrating eNose technology in standard practice. However, it remains to be elucidated whether eNose technology can serve as a marker for therapy compliance of inhaled medication.

Cystic fibrosis

Cystic fibrosis (CF) is associated with bronchiectasis, recurrent infectious exacerbations, and progressive deterioration of lung function due to exacerbations [52].

A few studies using different eNoses showed that patients with CF could accurately be distinguished from healthy controls and asthma patients based on their breathprint [53, 54, 23]. Two studies showed conflicted results regarding differentiation of CF from primary ciliary dyskinesia (PCD) patients, a bronchiectatic lung disease that mimics symptoms of CF [54]. While Paff et al. showed that CF and PCD could be adequately discriminated, Joensen et al. found no significant differences [53, 54]. This was possibly be due to methodological differences, such as different breath collection methods and a more heterogeneous patient population in the latter study. Furthermore, eNose technology adequately discriminated between patients with and without exacerbations, with and without chronic Pseudomonas aeruginosa colonisation, and patients with and without Aspergillus fumigatus colonisation [53-55]. It would be of great interest to investigate whether early stage respiratory infections and exacerbations can also be detected and eventually be predicted by eNose technology. This will possibly increase the chance of successful eradication and slowing down pulmonary function decline.

Interstitial lung disease

Interstitial lung disease (ILD) is a heterogeneous group of relatively uncommon diseases causing fibrotic and/or inflammatory changes in interstitial lung tissue. Disease course and treatment strategies widely vary for different ILDs, and even within individual ILDs disease course often varies. Diagnosis is based on integration of clinical data with imaging and if needed pathology data. Diagnosis is often complex and diagnostic delays are common [56, 57]. eNose technology has the potential to replace invasive procedures, and aid the diagnostic process to facilitate timely and accurate diagnosis.

A large single centre cohort, including various ILDs, found that breathprints of ILD patients could be distinguished from healthy controls with 100% accuracy. Results were confirmed in a validation cohort [58]. A few other studies compared individual ILDs with healthy controls and COPD patients [59-62]. Breathprints of patients with idiopathic pulmonary fibrosis (IPF), ILD associated with connective tissue disease and pneumoconiosis were significantly different from healthy controls [60-62]. In sarcoidosis patients, the breathprint of patients with untreated sarcoidosis differed from healthy controls, implying that eNose technology may be used for initial diagnosis. This study found that breathprints of treated sarcoidosis patients were not significantly different from healthy controls, but the number of participants was small [59]. Comparing

different ILDs, eNose technology distinguished IPF from non-IPF ILD patients with an accuracy of 91% in both training and validation cohort. Exploratory analyses indicated that individual ILDs can also be discriminated adequately [58]. However, groups were relatively small and, thus, results should be validated and confirmed in larger cohorts. A currently ongoing large multicentre study is investigating the potential of eNose technology to identify individual diseases, predict disease course, and response to treatment in fibrotic ILDs (NCT04680832).

Lung cancer

Worldwide, lung cancer is the leading cause of cancer deaths and has the highest incidence of all cancer types. More than 80% of patients suffering from lung cancer are former or current tobacco smokers [63]. Early diagnosis is clearly associated with better outcomes, and lung cancer screening has shown to reduce mortality [64, 65]. Nevertheless, early diagnosis remains challenging, since initial clinical presentation often overlaps with COPD or other smoking-related diseases, and symptoms often only appear in late stages [66]. Low-dose CT scan is currently the best available tool for screening. However, this type of screening is only cost-effective in a selected group of former and current smokers [67]. Also, differentiation of benign from malignant nodules is not possible with CT scan results; therefore, detected nodules warrant further invasive investigations. eNose could possibly serve as non-invasive and less costly screening tool to identify malign pulmonary neoplasms. Two studies used eNose technology in high-risk patients enrolled for lung cancer screening. Both studies found a higher specificity for detecting lung cancer with eNose compared to low-dose CT scan; thus, the use of eNose technology as screening tool can potentially reduce the false-positive rate and prevent unnecessary (invasive) testing [16, 68]. It is important to note that not all lesions classified as benign were histologically proven in these studies.

Whether an eNose can differentiate lung cancer patients from healthy controls, patients with benign lung nodules or (former) smokers, has been investigated in different cohorts. All studies in (non-) small cell lung cancer ((N)SCLC) showed significant results, albeit with a wide range in reported sensitivity (71-99%) and specificity (13-100%) [69-81]. Smoking status of participants did not seem to influence accuracy of an eNose for detecting cancer [78]. One small study showed that patients with and without an EGFR (epidermal growth factor receptor) mutation had distinct breathprints [68]. It has not been evaluated whether eNoses can recognize specific types of lung cancer in a cohort with different subtypes. Recognition of subtypes seems plausible, as differentiation of lung cancer from head-neck cancer was possible with eNose technology [82, 83]. eNose technology did not discriminate between different stages

of lung cancer [84]. One recent study in NSCLC combined eNose data with relevant clinical parameters (such as age, number of pack years, and presence of COPD), and showed a higher accuracy for lung cancer detection than using eNose data only. These results highlight the potential of eNose technology as additional diagnostic procedure [80]. Some small studies indicated that eNose technology was also able to differentiate patients suffering from malignant pleural mesothelioma (MPM) and healthy controls. Differentiation of MPM from benign asbestosis disease and asymptomatic asbestos exposure had a high sensitivity too [85-87].

Prediction of response to therapy is investigated for anti-programmed death (PD)-1 receptor therapy in NSCLC patients. Breathprints were collected before start of pembrolizumab or nivolumab therapy. Exhaled breath data could predict which patients would respond to therapy with an AUC of 0.89, confirmed in a validation cohort. By setting a cut-off value to obtain 100% specificity, the investigators were able to detect 24% of non-responders to anti-PD-1 therapy. In this regard, eNose seems to be more accurate than the currently used biomarker PD-L1 [88]. Another study is currently registered for recruiting until July 2021 and will evaluate the effect of immunotherapy on breathprints of exhaled breath and sweat in lung cancer patients (NCT03988192).

Schmekel et al. investigated the ability of eNose to predict prognosis in patients with end stage lung cancer. They collected breathprints before start and several times after start of palliative chemotherapy and applied different prediction models. Patients with less than one year survival and more than one year survival could be separated based on breathprint [89]. The authors suggest to use this eNose-based prediction for choosing a certain treatment strategy, but this needs confirmation in studies first.

Obstructive sleep apnoea

At the moment, the gold standard for diagnosing obstructive sleep apnoea (OSA) is poly(somno)graphy which is a costly and time-consuming test. eNose technology has been investigated as an alternative modality to diagnose this condition and assess treatment effect.

It was shown that breathprints from OSA patients and healthy controls can be distinguished reliably [90-92]. However, it remains questionable whether breathprints distinguishes true OSA, or if the breathprint is just a reflection of a metabolic syndrome or underlying inflammation caused by obesity. In one of the studies this question was more apparent as groups were not matched for body mass index [90]. Dragonieri et al.

found that eNose technology did discriminate obese patients with and without OSA, with moderate accuracy [91].Nevertheless, another study could not confirm those results [92].

Other researchers investigated OSA, OSA-COPD overlap syndrome and COPD. OSA could be distinguished from the overlap syndrome, but eNose technology could not discriminate well between the overlap syndrome and COPD. Also here it is not clear whether true OSA can be detected or other factors, such as COPD, are picked up [93, 92]. Whether included patients also suffer from heart failure is not clearly displayed in these studies, although it is known that many heart failure patients suffer from OSA and that heart failure might influence breathprint [94, 95].

The effects of CPAP treatment in patients with OSA has also been studied. The breathprint of OSA patients changed significantly already after one night of CPAP treatment [96]. Significant difference in breathprint was also found before and after three months of CPAP treatment [90]. It remains to be elucidated what this change in breathprint indicates. Possibly, the alteration in breathprint could serve as a marker for metabolic success, therapeutic benefit or treatment adherence. Furthermore, it must be noted that the breathprints of patients with OSA differed between morning and evening [97]. Hence, diurnal variance must be taken into account when using an eNose for patients with OSA.

Pulmonary infections

Pathogenic micro-organisms, such as viruses, bacteria or fungi, can cause severe pulmonary infections. Identification of specific micro-organisms with sputum cultures can take up to several days, and is only possible if a specimen with sufficient quality is obtained. Specificity and sensitivity also depend on the causative micro-organism, experience of laboratory observer, and prior treatment [98]. Therefore, reported sensitivity of detecting bacteria in sputum culture ranges between 57 and 95%, and specificity between 48 and 87% [99]. Detection of specific micro-organisms using eNose technology can potentially reduce misuse of antibiotics and facilitate timely start of guided therapy.

Until now, two *in vitro* studies aimed to differentiate micro-organisms by analysing breathprints of their headspace air [100, 101]. Mould species were discriminated from other samples (bacteria, yeasts, and control medium) with a high accuracy (92.9%). Furthermore, different mould species seemed to have different breathprints [101]. Another study performed eNose analyses on bronchoalveolar lavage samples,

and demonstrated accurate discrimination between Gram-positive bacteria, Gram-negative bacteria, fungi, and samples without growth of micro-organisms [100]. *In vivo,* breathprints of bronchiectasis patients significantly differed between those colonised with Pseudomonas Aeruginosa and those colonised with other pathogenic micro-organisms or non-colonised [102]. For detection of aspergillus colonisation or invasive aspergillosis in specific patient groups (CF and neutropenic patients), studies revealed a high accuracy of eNose breathprint analysis [103, 55]. These studies did not include a validation cohort or healthy control group.

Ventilator-associated pneumonia (VAP) is a common nosocomial infection in ventilated patients and has an incidence and mortality around 9% [99, 104]. In most eNose studies, bacterial growth in sputum or a clinical pneumonia score was used to define VAP [105-107, 15]. Two studies showed that obtained breathprints highly correlated with a clinical pneumonia score, implying that eNose technology might be used to predict the probability of a VAP [105, 106]. Two case-control studies in patients with VAP and ventilated patients without pneumonia showed conflicting results; Schnabel and colleagues concluded that eNose technology lacked sensitivity and specificity, whereas a recently published study of Chen and colleagues found a good accuracy for detecting VAP [107, 15]. This shows the need for more research on this topic before eNose can be used to determine the need for more (invasive) diagnostics in ill patients, such as performing bronchoscopy.

In pulmonary tuberculosis (TB) patients, detection and screening with eNose technology has been studied in different countries and compared to different control groups [108-113]. As TB is the leading cause of death from an infection caused by a single micro-organism, and as it has a high prevalence in developing countries, establishing a fast non-invasive cheap screening tool is much needed [114]. In one study, eNose technology differentiated TB from non-TB quite accurately, suggesting that it can potentially serve as a screening tool. Detection of TB had a sensitivity of 89% and a specificity of 91% compared to positive cultures. This sensitivity and specificity exceeded Ziehl-Neelsen staining [108]. However, all studies with proven TB and healthy participants in the training cohort, had a lower accuracy when validating the results in a cohort also including suspected TB patients [109, 110, 113]. Thus, more research is necessary before eNose technology can be used as a population-wide screening tool.

Due to the Corona Virus Disease (COVID-19) pandemic, much research effort is being put in the evaluation of eNose technology as a fast and non-invasive tool for the detection of COVID-19 (NCT04475562, NCT04475575, NCT04558372, NCT04379154, NCT04614883, NL8694). To date, one study tested the accuracy of eNose technology

for COVID-19 screening prior to surgery in non-symptomatic patients and found a negative predictive value up to 0.96. Reverse transcription-polymerase chain reaction on a pharyngeal swab and antibody testing were used to confirm presence or absence of COVID-19 [115].

Other

A number of eNose studies have been performed in other lung diseases. In acute respiratory distress syndrome (ARDS), eNose technology could discriminate between mechanically ventilated patients with and without ARDS, with moderate accuracy in a training and validation cohort [116].

One small proof-of-principle study has been performed in patients with suspected pulmonary embolism, defined as a high clinical probability according to the Well's score or elevated D-dimer. Breathprints of non-comorbid patients with and without pulmonary embolism could be distinguished with an accuracy of 85%. However, in patients with comorbidities known to influence VOCs (e.g. cancer, diabetes) the accuracy dropped [117].

Finally, eNose technology could be useful for follow-up and monitoring lung transplant recipients. One study found a significant association between breathprint and plasma tacrolimus levels, suggesting that eNoses might be used for non-invasive therapeutic drug monitoring [118].

A clinical trial in lung transplant recipients is currently conducted (NL9251) looking at discrimination of stable lung transplant recipients, acute cellular rejection, and chronic lung allograft rejection.

Discussion

In the past decades, multiple eNoses have been developed and tested in numerous clinical studies for a wide spectrum of lung diseases. So far, the vast majority of studies evaluated the ability of eNose technology to distinguish lung diseases from healthy controls, and to discriminate between different diagnoses. A small number of studies have been performed for prognostic or therapeutic purposes, and only a handful of studies have focused on clustering patients by breathprint and identifying phenotypes. Results in lung diseases are overall very promising, but several issues should be addressed before eNoses can be implemented in daily clinical practice.

One of the issues is the use of various eNose devices with different qualifications, types of sensors and breath sample collection methods as summarised in **Table 1**. It is not possible to point out the best eNose device or select one optimal sensor type, as each setting, disease and research aim can require different features. For example, a portable device might be optimal for an acute care setting, direct sampling without collection bags might be useful in low resource areas and as point-of-care technique, and a device that corrects for ambient air will probably generate more comparable results in multicentre use and settings with unstable or varying environmental conditions.

Given important differences between the various devices, it is difficult to compare data of the different eNose devices. Hence, each eNose needs to be validated for every clinical application. This implies that knowledge about characteristics of eNose devices is essential before initiating eNose research, as the type of device cannot easily be changed during the trajectory of developing a clinical tool. Additionally, the influence of endogenous (e.g. comorbidities, ethnicity, age) and exogenous factors (e.g. smoking, nutrition, drug use, measurement environment) on breathprints needs to be further elucidated.

Furthermore, studies differ significantly with regards to study design (e.g. patient selection, number of participants, and presence of a validation cohort). As illustrated in **Figure 2**, the majority of studies so far can be considered as pilot or exploratory studies, and have small numbers of participants. The most important goal of these studies is to test new hypotheses, which can be further assessed and confirmed in larger studies with external validation. However, these validation studies are not often conducted. This lack of validation is a major issue in development of a clinical useful breath biomarker, as breath analysis results are not always interchangeable between research settings due to a combination of the above mentioned factors. To ensure optimal outcomes, comparison and generalisability of eNose studies, the design and analysis methods should ideally be based on specific predefined research aims.

Moreover, most studies do not explain the rationale for choosing a certain machine learning model for analysing eNose data. This prevents insights in and discussion regarding the optimal analysis techniques and algorithms. Machine learning models are complex to execute and interpret, and if not used in the right way are prone for overfitting. To avoid inadequate modelling, data scientists should always be involved in these complex analyses and models should be validated independently to exclude overfitting. To allow for comparison of different modelling techniques, we recommend an extensive world-wide shared database per eNose with FAIR (findable, accessible, interoperable, and reusable) and open source data, including patient characteristics

and other pre-test probabilities. This database would ensure optimal training, validation, and application of models.

Finally, a factor that hampers eNose implementation is the need for a strong gold standard to establish a diagnosis or to evaluate therapeutic effect. High quality data input is required for optimal validity when developing a new technique. Some of the diseases mentioned in this review lack a gold standard, and even if a gold standard does exist, there is always a range of uncertainty. There is a potential for unsupervised machine learning models in this regard, as such analyses could help to identify previously unrecognised phenotype clusters. Discovering such new clusters can help to generate hypotheses about the existence of unravelled disease subtypes or overlap between diagnoses, and might eventually guide new diagnostic standards.

In conclusion, eNose technology in the field of lung diseases is promising and at the doorstep of the pulmonologist's office. To facilitate clinical implementation, we recommend conducting prospective multicentre trials including validation in external cohorts with a study design and analysis method relevant for the research aim, and sharing databases on open source platforms. If supported by sufficient evidence, research can subsequently be extended to clinical implementation studies, and finally, use in daily practice.

We believe that eNose technology has the potential to facilitate personalised medicine in lung diseases through establishing early, accurate diagnosis and monitoring disease course and therapeutic effects.

Acknowledgements

The authors wish to thank dr. Sabrina T.G. Meertens-Gunput from the Erasmus MC Medical Library for developing and updating the search strategies.

References

1. van de Kant KDG, van der Sande LJTM, Jöbsis Q, van Schayck OCP, Dompeling E. Clinical use of exhaled volatile organic compounds in pulmonary diseases: a systematic review. Respiratory Research. 2012 2012/12/21;13(1):117.
2. van der Schee MP, Paff T, Brinkman P, van Aalderen WMC, Haarman EG, Sterk PJ. Breathomics in Lung Disease. Chest. 2015 2015/01/01/;147(1):224-31.
3. Potter P. Hippocrates Vol VI. Diseases, internal affections. Cambridge, MA/London: Harvard University Press; 1988.
4. Španěl P, Smith D. Volatile compounds in health and disease. Curr Opin Clin Nutr Metab Care. 2011 Sep;14(5):455-60.
5. Lourenço C, Turner C. Breath analysis in disease diagnosis: methodological considerations and applications. Metabolites. 2014;4(2):465-98.
6. Gardner JW, Bartlett PN. A brief history of electronic noses. Sensors and Actuators B: Chemical. 1994 1994/03/01/;18(1):210-11.
7. Liu X, Cheng S, Liu H, Hu S, Zhang D, Ning H. A survey on gas sensing technology. Sensors (Basel, Switzerland). 2012;12(7):9635-65.
8. Nag A, Mitra A, Mukhopadhyay SC. Graphene and its sensor-based applications: A review. Sensors and Actuators A: Physical. 2018 2018/02/01/;270:177-94.
9. Kovalska E, Lesongeur P, Hogan BT, Baldycheva A. Multi-layer graphene as a selective detector for future lung cancer biosensing platforms. Nanoscale. 2019;11(5):2476-83.
10. Chen Q, Chen Z, Liu D, He Z, Wu J. Constructing E-Nose Using Metal-Ion Induced Assembly of Graphene Oxide for Diagnosis of Lung Cancer via Exhaled Breath. ACS Appl Mater Interfaces. 2020.
11. Liu X, Cheng S, Liu H, Hu S, Zhang D, Ning H. A Survey on Gas Sensing Technology. Sensors. 2012;12(7):9635-65.
12. Santos JP LJ, Aleixandre M. Electronic noses applications in beer technology. Brewing Technology. 2017.
13. Shobha G, Rangaswamy S. Chapter 8 - Machine Learning. In: Gudivada VN, Rao CR, editors. Handbook of Statistics. Elsevier; 2018. p. 197-228.
14. Liao YH, Wang ZC, Zhang FG, Abbod MF, Shih CH, Shieh JS. Machine Learning Methods Applied to Predict Ventilator-Associated Pneumonia with Pseudomonas aeruginosa Infection via Sensor Array of Electronic Nose in Intensive Care Unit. Sensors (Basel). 2019;19(8).
15. Chen CY, Lin WC, Yang HY. Diagnosis of ventilator-associated pneumonia using electronic nose sensor array signals: Solutions to improve the application of machine learning in respiratory research. Respir Res. 2020;21(1).
16. Rocco R, Incalzi RA, Pennazza G, Santonico M, Pedone C, Bartoli IR, et al. BIONOTE e-nose technology may reduce false positives in lung cancer screening programmes. Eur J Cardio-thorac Surg. 2016;49(4):1112-17.
17. National Asthma E, Prevention P. Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007. J Allergy Clin Immunol. 2007 Nov;120(5 Suppl):S94-138.

18. Dragonieri S, Schot R, Mertens BJA, Le Cessie S, Gauw SA, Spanevello A, et al. An electronic nose in the discrimination of patients with asthma and controls. J Allergy Clin Immunol. 2007;120(4):856-62.

19. Fens N, Zwinderman AH, Van Der Schee MP, De Nijs SB, Dijkers E, Roldaan AC, et al. Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma. Am J Respir Crit Care Med. 2009;180(11):1076-82.

20. Lazar Z, Fens N, van der Maten J, van der Schee MP, Wagener AH, de Nijs SB, et al. Electronic nose breathprints are independent of acute changes in airway caliber in asthma. Sensors (Basel). 2010;10(10):9127-38.

21. Montuschi P, Santonico M, Mondino C, Pennazza G, Maritini G, Martinelli E, et al. Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma. Chest. 2010;137(4):790-96.

22. Van der Schee MP, Palmay R, Cowan JO, Taylor DR. Predicting steroid responsiveness in patients with asthma using exhaled breath profiling. Clin Exp Allergy. 2013;43(11):1217-25.

23. Bannier MAGE, Van De Kant KDG, Jöbsis Q, Dompeling E. Feasibility and diagnostic accuracy of an electronic nose in children with asthma and cystic fibrosis. J Breath Res. 2019;13(3).

24. Dragonieri S, Quaranta VN, Carratu P, Ranieri T, Resta O. Exhaled breath profiling by electronic nose enabled discrimination of allergic rhinitis and extrinsic asthma. Biomarkers. 2019;24(1):70-75.

25. Tenero L, Sandri M, Piazza M, Paiola G, Zaffanello M, Piacentini G. Electronic nose in discrimination of children with uncontrolled asthma. Journal of breath research. 2020;14(4):046003.

26. Fens N, Roldaan AC, van der Schee MP, Boksem RJ, Zwinderman AH, Bel EH, et al. External validation of exhaled breath profiling using an electronic nose in the discrimination of asthma with fixed airways obstruction and chronic obstructive pulmonary disease. Clin Exp Allergy. 2011;41(10):1371-78.

27. Pembrey L, Barreto ML, Douwes J, Cooper P, Henderson J, Mpairwe H, et al. Understanding asthma phenotypes: the World Asthma Phenotypes (WASP) international collaboration. ERJ Open Research. 2018;4(3):00013-2018.

28. Plaza V, Crespo A, Giner J, Merino JL, Ramos-Barbón D, Mateus EF, et al. Inflammatory asthma phenotype discrimination using an electronic nose breath analyzer. J Invest Allergol Clin Immunol. 2015;25(6):431-37.

29. Brinkman P, Wagener AH, Hekking PP, Bansal AT, Maitland-van der Zee AH, Wang Y, et al. Identification and prospective stability of electronic nose (eNose)–derived inflammatory phenotypes in patients with severe asthma. J Allergy Clin Immunol. 2019;143(5):1811-20.e7.

30. Abdel-Aziz MI, Brinkman P, Vijverberg SJH, Neerincx AH, de Vries R, Dagelet YWF, et al. eNose breath prints as a surrogate biomarker for classifying patients with asthma by atopy. J Allergy Clin Immunol. 2020.

31. Farraia M, Cavaleiro Rufo J, Paciência I, Castro Mendes F, Rodolfo A, Rama T, et al. Human volatilome analysis using eNose to assess uncontrolled asthma in a clinical setting. Allergy Eur J Allergy Clin Immunol. 2020.

32. Brinkman P, van de Pol MA, Gerritsen MG, Bos LD, Dekker T, Smids BS, et al. Exhaled breath profiles in the monitoring of loss of control and clinical recovery in asthma. Clin Exp Allergy. 2017;47(9):1159-69.

33. Fasola S, Ferrante G, Sabatini A, Santonico M, Zompanti A, Grasso S, et al. Repeatability of exhaled breath fingerprint collected by a modern sampling system in asthmatic and healthy children. J Breath Res. 2019;13(3).

34. Cavaleiro Rufo J, Paciência I, Mendes FC, Farraia M, Rodolfo A, Silva D, et al. Exhaled breath condensate volatilome allows sensitive diagnosis of persistent asthma. Allergy Eur J Allergy Clin Immunol. 2019;74(3):527-34.

35. Tenero L, Sandri M, Piazza M, Paiola G, Zaffanello M, Piacentini G. Electronic nose in discrimination of children with uncontrolled asthma. J Breath Res. 2020;14(4).

36. Lamprecht B, Soriano JB, Studnicka M, Kaiser B, Vanfleteren LE, Gnatiuc L, et al. Determinants of underdiagnosis of COPD in national and international surveys. Chest. 2015 Oct;148(4):971-85.

37. Hanlon P, Daines L, Campbell C, McKinstry B, Weller D, Pinnock H. Telehealth Interventions to Support Self-Management of Long-Term Conditions: A Systematic Metareview of Diabetes, Heart Failure, Asthma, Chronic Obstructive Pulmonary Disease, and Cancer. J Med Internet Res. 2017 2017/05/17;19(5):e172.

38. Hattesohl ADM, Jörres RA, Dressel H, Schmid S, Vogelmeier C, Greulich T, et al. Discrimination between COPD patients with and without alpha 1-antitrypsin deficiency using an electronic nose. Respirology. 2011;16(8):1258-64.

39. Sibila O, Garcia-Bellmunt L, Giner J, Merino JL, Suarez-Cuartin G, Torrego A, et al. Identification of airway bacterial colonization by an electronic nose in Chronic Obstructive Pulmonary Disease. Respir Med. 2014;108(11):1608-14.

40. Cazzola M, Segreti A, Capuano R, Bergamini A, Martinelli E, Calzetta L, et al. Analysis of exhaled breath fingerprints and volatile organic compounds in COPD. COPD Res Pract. 2015;1(1).

41. Shafiek H, Fiorentino F, Merino JL, López C, Oliver A, Segura J, et al. Using the electronic nose to identify airway infection during COPD exacerbations. PLoS ONE. 2015;10(9).

42. Rodríguez-Aguilar M, Díaz de León-Martínez L, Gorocica-Rosete P, Padilla RP, Thirión-Romero I, Ornelas-Rebolledo O, et al. Identification of breath-prints for the COPD detection associated with smoking and household air pollution by electronic nose. Respir Med. 2020;163.

43. Fens N, Van Rossum AGJ, Zanen P, Van Ginneken B, Van Klaveren RJ, Zwinderman AH, et al. Subphenotypes of mild-to-moderate COPD by factor and cluster analysis of pulmonary function, CT imaging and breathomics in a population-based survey. COPD J Chronic Obstructive Pulm Dis. 2013;10(3):277-85.

44. De Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. Eur Respir J. 2018;51(1).

45. Scarlata S, Finamore P, Santangelo S, Giannunzio G, Pennazza G, Grasso S, et al. Cluster analysis on breath print of newly diagnosed COPD patients: Effects of therapy. J Breath Res. 2018;12(3).

46. Fens N, De Nijs SB, Peters S, Dekker T, Knobel HH, Vink TJ, et al. Exhaled air molecular profiling in relation to inflammatory subtype and activity in COPD. Eur Respir J. 2011;38(6):1301-09.

47. Van Geffen WH, Bruins M, Kerstjens HAM. Diagnosing viral and bacterial respiratory infections in acute COPD exacerbations by an electronic nose: A pilot study. J Breath Res. 2016;10(3).

48. van Velzen P, Brinkman P, Knobel HH, van den Berg JWK, Jonkers RE, Loijmans RJ, et al. Exhaled Breath Profiles Before, During and After Exacerbation of COPD: A Prospective Follow-Up Study. COPD J Chronic Obstructive Pulm Dis. 2019.

49. Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. 2019.

50. Finamore P, Pedone C, Scarlata S, Di Paolo A, Grasso S, Santonico M, et al. Validation of exhaled volatile organic compounds analysis using electronic nose as index of COPD severity. Int J COPD. 2018;13:1441-48.

51. Montuschi P, Santini G, Mores N, Vignoli A, Macagno F, Shoreh R, et al. Breathomics for assessing the effects of treatment and withdrawal with inhaled beclomethasone/formoterol in patients with COPD. Front Pharmacol. 2018;9(APR).

52. Castellani C, Duff AJA, Bell SC, Heijerman HGM, Munck A, Ratjen F, et al. ECFS best practice guidelines: the 2018 revision. Journal of Cystic Fibrosis. 2018 2018/03/01/;17(2):153-78.

53. Paff T, van der Schee MP, Daniels JMA, Pals G, Postmus PE, Sterk PJ, et al. Exhaled molecular profiles in the assessment of cystic fibrosis and primary ciliary dyskinesia. J Cyst Fibrosis. 2013;12(5):454-60.

54. Joensen O, Paff T, Haarman EG, Skovgaard IM, Jensen PØ, Bjarnsholt T, et al. Exhaled breath analysis using electronic nose in cystic fibrosis and primary ciliary dyskinesia patients with chronic pulmonary infections. PLoS ONE. 2014;9(12).

55. De Heer K, Kok MGM, Fens N, Weersink EJM, Zwinderman AH, Van Der Schee MPC, et al. Detection of Airway Colonization by Aspergillus fumigatus by Use of Electronic Nose Technology in Patients with Cystic Fibrosis. J Clin Microbiol. 2016;54(3):569-75.

56. Antoniou KM, Margaritopoulos GA, Tomassetti S, Bonella F, Costabel U, Poletti V. Interstitial lung disease. European Respiratory Review. 2014;23(131):40-54.

57. Pritchard D, Adegunsoye A, Lafond E, Pugashetti JV, DiGeronimo R, Boctor N, et al. Diagnostic test interpretation and referral delay in patients with interstitial lung disease. Respiratory Research. 2019 2019/11/12;20(1):253.

58. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2020.

59. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratú P, Resta O, et al. An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. Respir Med. 2013;107(7):1073-78.

60. Yang HY, Peng HY, Chang CJ, Chen PC. Diagnostic accuracy of breath tests for pneumoconiosis using an electronic nose. J Breath Res. 2018;12(1).

61. Krauss E, Haberer J, Maurer O, Barreto G, Drakopanagiotakis F, Degen M, et al. Exploring the Ability of Electronic Nose Technology to Recognize Interstitial Lung Diseases (ILD) by Non-Invasive Breath Screening of Exhaled Volatile Compounds (VOC): A Pilot Study from the European IPF Registry (eurIPFreg) and Biobank. J Clin Med. 2019 Oct 16;8(10).

62. Dragonieri S, Scioscia G, Quaranta VN, Carratu P, Venuti MP, Falcone M, et al. Exhaled volatile organic compounds analysis by e-nose can detect idiopathic pulmonary fibrosis. J Breath Res. 2020 Jul 21;14(4):047101.

63. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018;68(6):394-424.

64. Wender R, Fontham ETH, Barrera Jr E, Colditz GA, Church TR, Ettinger DS, et al. American Cancer Society lung cancer screening guidelines. CA: A Cancer Journal for Clinicians. 2013;63(2):106-17.

65. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. New England Journal of Medicine. 2020;382(6):503-13.

66. Planchard D, Popat S, Kerr K, Novello S, Smit EF, Faivre-Finn C, et al. Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018 Oct 1;29(Suppl 4):iv192-iv237.

67. Cressman S, Peacock SJ, Tammemägi MC, Evans WK, Leighl NB, Goffin JR, et al. The Cost-Effectiveness of High-Risk Lung Cancer Screening and Drivers of Program Efficiency. Journal of Thoracic Oncology. 2017 2017/08/01/;12(8):1210-22.

68. Shlomi D, Abud M, Liran O, Bar J, Gai-Mor N, Ilouze M, et al. Detection of Lung Cancer and EGFR Mutation by Electronic Nose System. J Thorac Oncol. 2017;12(10):1544-51.

69. Machado RF, Laskowski D, Deffenderfer O, Burch T, Zheng S, Mazzone PJ, et al. Detection of lung cancer by sensor array analyses of exhaled breath. Am J Respir Crit Care Med. 2005;171(11):1286-91.

70. Dragonieri S, Annema JT, Schot R, van der Schee MPC, Spanevello A, Carratú P, et al. An electronic nose in the discrimination of patients with non-small cell lung cancer and COPD. Lung Cancer. 2009;64(2):166-70.

71. Hubers AJ, Brinkman P, Boksem RJ, Rhodius RJ, Witte BI, Zwinderman AH, et al. Combined sputum hypermethylation and eNose analysis for lung cancer diagnosis. J Clin Pathol. 2014;67(8):707-11.

72. McWilliams A, Beigi P, Srinidhi A, Lam S, MacAulay CE. Sex and smoking status effects on the early detection of early lung cancer in high-risk smokers using an electronic nose. IEEE Trans Biomed Eng. 2015;62(8):2044-54.

73. Gasparri R, Santonico M, Valentini C, Sedda G, Borri A, Petrella F, et al. Volatile signature for the early diagnosis of lung cancer. J Breath Res. 2016;10(1).

74. Huang CH, Zeng C, Wang YC, Peng HY, Lin CS, Chang CJ, 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).

75. Kort S, Tiggeloven MM, Brusse-Keizer M, Gerritsen JW, Schouwink JH, Citgez E, et al. Multi-centre prospective study on diagnosing subtypes of lung cancer by exhaled-breath analysis. Lung Cancer. 2018;125:223-29.

76. van de Goor R, van Hooren M, Dingemans AM, Kremer B, Kross K. Training and Validating a Portable Electronic Nose for Lung Cancer Screening. J Thorac Oncol. 2018;13(5):676-81.

77. Mohamed EI, Mohamed MA, Abdel-Mageed SM, Abdel-Mohdy TS, Badawi MI, Darwish SH. Volatile organic compounds of biofluids for detecting lung cancer by an electronic nose based on artificial neural network. J App Biomed. 2019;17(1):61-67.

78. Tirzïte M, Bukovskis M, Strazda G, Jurka N, Taivans I. Detection of lung cancer with electronic nose and logistic regression analysis. J Breath Res. 2019;13(1).

79. Kononov A, Korotetsky B, Jahatspanian I, Gubal A, Vasiliev A, Arsenjev A, et al. Online breath analysis using metal oxide semiconductor sensors (electronic nose) for diagnosis of lung cancer. J Breath Res. 2020;14(1).

80. Kort S, Brusse-Keizer M, Gerritsen JW, Schouwink H, Citgez E, de Jongh F, et al. Improving lung cancer diagnosis by combining exhaled-breath data and clinical parameters. 2020 Jan.

81. Krauss E, Haberer J, Barreto G, Degen M, Seeger W, Guenther A. Recognition of breathprints of lung cancer and chronic obstructive pulmonary disease using the Aeonose®electronic nose. J Breath Res. 2020;14(4).

82. van Hooren MRA, Leunis N, Brandsma DS, Dingemans AMC, Kremer B, Kross KW. Differentiating head and neck carcinoma from lung carcinoma with an electronic nose: a proof of concept study. Eur Arch Oto-Rhino-Laryngol. 2016;273(11):3897-903.

83. Fielding D, Hartel G, Pass D, Davis M, Brown M, Dent A, et al. Volatile organic compound breath testing detects in-situ squamous cell carcinoma of bronchial and laryngeal regions and shows distinct profiles of each tumour. J Breath Res. 2020 Oct 06;14(4):046013.

84. Tirzite M, Bukovskis M, Strazda G, Jurka N, Taivans I. Detection of lung cancer in exhaled breath with an electronic nose using support vector machine analysis. J Breath Res. 2017;11(3).

85. Chapman EA, Thomas PS, Stone E, Lewis C, Yates DH. A breath test for malignant mesothelioma using an electronic nose. Eur Respir J. 2012;40(2):448-54.

86. Dragonieri S, Van Der Schee MP, Massaro T, Schiavulli N, Brinkman P, Pinca A, et al. An electronic nose distinguishes exhaled breath of patients with Malignant Pleural Mesothelioma from controls. Lung Cancer. 2012;75(3):326-31.

87. Lamote K, Brinkman P, Vandermeersch L, Vynck M, Sterk PJ, Van Langenhove H, et al. Breath analysis by gas chromatography-mass spectrometry and electronic nose to screen for pleural mesothelioma: A crosssectional case-control study. Oncotarget. 2017;8(53):91593-602.

88. De Vries R, Muller M, Van Der Noort V, Theelen WSME, Schouten RD, Hummelink K, et al. Prediction of response to anti-PD-1 therapy in patients with non-small-cell lung cancer by electronic nose analysis of exhaled breath. Ann Oncol. 2019;30(10):1660-66.

89. Schmekel B, Winquist F, Vikström A. Analysis of breath samples for lung cancer survival. Anal Chim Acta. 2014;840:82-86.

90. Greulich T, Hattesohl A, Grabisch A, Koepke J, Schmid S, Noeske S, et al. Detection of obstructive sleep apnoea by an electronic nose. Eur Respir J. 2013;42(1):145-55.

91. Dragonieri S, Porcelli F, Longobardi F, Carratu P, Aliani M, Ventura VA, et al. An electronic nose in the discrimination of obese patients with and without obstructive sleep apnoea. J Breath Res. 2015;9(2).

92. Scarlata S, Pennazza G, Santonico M, Santangelo S, Rossi Bartoli I, Rivera C, et al. Screening of Obstructive Sleep Apnea Syndrome by Electronic-Nose Analysis of Volatile Organic Compounds. Sci Rep. 2017;7(1):11938.

93. Dragonieri S, Quaranta VN, Carratu P, Ranieri T, Resta O. Exhaled breath profiling in patients with COPD and OSA overlap syndrome: A pilot study. J Breath Res. 2016;10(4).

94. Finamore P, Pedone C, Lelli D, Costanzo L, Bartoli IR, De Vincentis A, et al. Analysis of volatile organic compounds: an innovative approach to heart failure characterization in older patients. Journal of Breath Research. 2018 2018/02/06;12(2):026007.

95. Costanzo MR. Central Sleep Apnea in Patients with Heart Failure-How to Screen, How to Treat. Curr Heart Fail Rep. 2020 Oct;17(5):277-87.

96. Incalzi RA, Pennazza G, Scarlata S, Santonico M, Vernile C, Cortese L, et al. Comorbidity modulates non invasive ventilation-induced changes in breath print of obstructive sleep apnea syndrome patients. Sleep Breathing. 2014.

97. Kunos L, Bikov A, Lazar Z, Korosi BZ, Benedek P, Losonczy G, et al. Evening and morning exhaled volatile compound patterns are different in obstructive sleep apnoea assessed with electronic nose. Sleep Breathing. 2015;19(1):247-53.

98. Skerrett SJ. DIAGNOSTIC TESTING FOR COMMUNITY-ACQUIRED PNEUMONIA. Clinics in Chest Medicine. 1999 1999/09/01/;20(3):531-48.

99. Rello J, Ollendorf DA, Oster G, Vera-Llonch M, Bellm L, Redman R, et al. Epidemiology and outcomes of ventilator-associated pneumonia in a large US database. Chest. 2002 Dec;122(6):2115-21.

100. Humphreys L, Orme RML, Moore P, Charaklias N, Sahgal N, Planas Pont N, et al. Electronic nose analysis of bronchoalveolar lavage fluid. Eur J Clin Invest. 2011;41(1):52-58.

101. De Heer K, Vonk SI, Kok M, Kolader M, Zwinderman AH, Van Oers MHJ, et al. ENose technology can detect and classify human pathogenic molds in vitro: A proof-of-concept study of Aspergillus fumigatus and Rhizopus oryzae. J Breath Res. 2016;10(3).

102. Suarez-Cuartin G, Giner J, Merino JL, Rodrigo-Troyano A, Feliu A, Perea L, et al. Identification of Pseudomonas aeruginosa and airway bacterial colonization by an electronic nose in bronchiectasis. Respir Med. 2018;136:111-17.

103. De Heer K, Van Der Schee MP, Zwinderman K, Van Den Berk IAH, Visser CE, Van Oers R, et al. Electronic nose technology for detection of invasive pulmonary aspergillosis in prolonged chemotherapy-induced neutropenia: A proof-of-principle study. J Clin Microbiol. 2013;51(5):1490-95.

104. Melsen WG, Rovers MM, Koeman M, Bonten MJ. Estimating the attributable mortality of ventilator-associated pneumonia from randomized prevention studies. Crit Care Med. 2011 Dec;39(12):2736-42.

105. Hanson ICW, Thaler ER. Electronic nose prediction of a clinical pneumonia score: Biosensors and microbes. Anesthesiology. 2005;102(1):63-68.

106. Hockstein NG, Thaler ER, Lin Y, Lee DD, Hanson CW. Correlation of pneumonia score with electronic nose signature: A prospective study. Ann Otol Rhinol Laryngol. 2005;114(7):504-08.

107. Schnabel RM, Boumans MLL, Smolinska A, Stobberingh EE, Kaufmann R, Roekaerts PMHJ, et al. Electronic nose analysis of exhaled breath to diagnose ventilator-associated pneumonia. Respir Med. 2015;109(11):1454-59.

108. Fend R, Kolk AHJ, Bessant C, Buijtels P, Klatser PR, Woodman AC. Prospects for clinical application of electronic-nose technology to early detection of Mycobacterium tuberculosis in culture and sputum. J Clin Microbiol. 2006;44(6):2039-45.

109. Bruins M, Rahim Z, Bos A, van de Sande WW, Endtz HP, van Belkum A. Diagnosis of active tuberculosis by e-nose analysis of exhaled air. Tuberculosis (Edinb). 2013 Mar;93(2):232-38.

110. Coronel Teixeira R, Rodríguez M, Jiménez de Romero N, Bruins M, Gómez R, Yntema JB, et al. The potential of a portable, point-of-care electronic nose to diagnose tuberculosis. J Infect. 2017;75(5):441-47.

111. Mohamed EI, Mohamed MA, Moustafa MH, Abdel-Mageed SM, Moro AM, Baess AI, et al. Qualitative analysis of biological tuberculosis samples by an electronic nose-based artificial neural network. Int J Tuberc Lung Dis. 2017;21(7):810-17.

112. Zetola NM, Modongo C, Matsiri O, Tamuhla T, Mbongwe B, Matlhagela K, et al. Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. J Infect. 2017;74(4):367-76.

113. Saktiawati AMI, Stienstra Y, Subronto YW, Rintiswati N, Sumardi, Gerritsen JW, et al. Sensitivity and specificity of an electronic nose in diagnosing pulmonary tuberculosis among patients with suspected tuberculosis. PLoS ONE. 2019;14(6).

114. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/tuberculosis.

115. Wintjens A, Hintzen KFH, Engelen SME, Lubbers T, Savelkoul PHM, Wesseling G, et al. Applying the electronic nose for pre-operative SARS-CoV-2 screening. Surg Endosc. 2020 Dec 2:1-8.

116. Bos LDJ, Schultz MJ, Sterk PJ. Exhaled breath profiling for diagnosing acute respiratory distress syndrome. BMC Pulm Med. 2014;14(1).

117. Fens N, Douma RA, Sterk PJ, Kamphuisen PW. Breathomics as a diagnostic tool for pulmonary embolism. J Thromb Haemost. 2010 Dec;8(12):2831-33.

118. Kovacs D, Bikov A, Losonczy G, Murakozy G, Horvath I. Follow up of lung transplant recipients using an electronic nose. J Breath Res. 2013;7(1).

119. Abdel-Aziz M, Brinkman P, Vijverberg S, Neerincx A, Vries R, Dagelet Y, et al. eNose breathprints as a surrogate biomarker for classifying asthma patients by atopy. Journal of Allergy and Clinical Immunology. 2020.

ADDITIONAL FILE 1

In-depth explanation on sensor technology used in eNoses

In this supplementary text, we provide for each type of sensor a more in-depth explanation how the technology works, the sensor characteristics, advantages and disadvantages, and current use in breath analysis.

Electrical sensors

A form of electrical sensors often used for gas sensing are conductometric sensors. In conductometric sensors, at the moment, metal oxide semiconductors (MOS) are the most common sensing materials, mainly suitable for gases. Redox reactions between the oxide surface of the sensor and the target gas, induces a reaction on the sensor, resulting in an electronic variation of the oxide surface, which is transduced into a variation of electrical resistance within the sensors. This can be detected by measuring, for example, a change in capacitance, mass or reaction energy [1]. Sensors based on MOS are low cost and have a high sensitivity. However, MOS sensors need to be operated at high temperatures, thus requiring a heating component and leading to high power consumption.

Feasibility studies are conducted for the use of graphene in constructing a MOS sensor [2, 3]. Graphene does not degrade over time and is stable under environmental conditions. Furthermore, graphene is highly sensitive at room temperature, making the use of a heating component unnecessary [4]. However, the fabrication of graphene is an expensive and complex process [4]. Chen et al. constructed an eNose using a metal-ion induced assembly of graphene oxide. They managed to obtain a homogeneous coating of the graphene oxide, creating more excellent gas sensing performances at room temperature [3].

Another used electrical sensor for electronic sensors are (conducting) polymers. Conducting polymer sensors operate based on a change in electrical resistance, caused by the adsorption of an analyte on the sensor surface. Conducting polymers can operate at ambient temperature, thus leading to a lower power consumption than MOS sensors, and are sensitive for an abundance of VOCs. These sensors are, however, easily influenced by humidity and temperature and possess a limited sensor life.

Gravimetric sensors

Gravimetric sensors operate based on a change in mass, leading to a frequency shift as a sensor response. Gravimetric sensors could be based on either quartz crystal microbalance (QCM), surface acoustic wave (SAW) propagation, or microcantilever. Both the QCM and SAW propagation based sensors use an acoustic wave to detect

analytes. The difference is that SAW propagation sensors operate using acoustic waves that travel across the surface of the sensing membrane, while QCM sensors operate based on the piezoelectric effect. For microcantilever based sensors, the presence of a specific analyte causes the cantilever to bend, leading to a frequency shift.

Gravimetric sensors possess a high sensitivity, but also contains a complex circuitry and are sensitive to humidity and temperature.

Optical sensors

Optical sensors operate based on optical phenomena, for example such as fluorescence and absorbance, caused by the response upon analyte binding. Optical sensors possess a very high sensitivity and specificity, but are in need of complex sensor-array systems, are hardly portable due to breakable optics and components, and are more expensive to use. Therefore, the use of optical sensors in (medical) eNose technology is limited.

References

1. Liu X, Cheng S, Liu H, Hu S, Zhang D, Ning H. A survey on gas sensing technology. Sensors (Basel, Switzerland). 2012;12(7):9635-65.
2. Kovalska E, Lesongeur P, Hogan BT, Baldycheva A. Multi-layer graphene as a selective detector for future lung cancer biosensing platforms. Nanoscale. 2019;11(5):2476-83.
3. Chen Q, Chen Z, Liu D, He Z, Wu J. Constructing E-Nose Using Metal-Ion Induced Assembly of Graphene Oxide for Diagnosis of Lung Cancer via Exhaled Breath. ACS Appl Mater Interfaces. 2020.
4. Nag A, Mitra A, Mukhopadhyay SC. Graphene and its sensor-based applications: A review. Sensors and Actuators A: Physical. 2018 2018/02/01/;270:177-94.

ADDITIONAL FILE 2

Search strategy

A systematic literature search was performed in three online databases on the 21st of October 2020, as stated below.

Only original articles were included. Articles with no full text available, reviews, abstracts, editorials, congress articles and animal studies were excluded. Moreover, articles were restricted to those investigating eNose technology for clinical purpose; articles only describing techniques as GC or MS, as well as early prototypes of eNose sensor technology were excluded.

Embase.com

('electronic nose'/de OR ('mass fragmentography'/de AND ('nose'/de OR 'breath analysis'/exp)) OR 'volatile organic compound'/de OR (eNOSE* OR e-NOSE* OR cyranose* OR spironose* OR ((electronic* OR artificial* OR spectromet* OR GC-MS) NEAR/3 (nose*)) OR volatile-organic-compound* OR VOC):ti,ab,kw) AND ('respiratory tract disease'/exp OR (lung* OR pulmonar* OR respirator*-tract* OR Pneumonolog* OR asthma* OR COPD OR Sarcoidos*):ab,ti,kw) NOT ([Conference Abstract]/lim AND [1800-2017]/py)

Medline (Ovid)

(Electronic Nose/ OR (Gas Chromatography-Mass Spectrometry/ AND (Nose/ OR Breath Tests/)) OR volatile organic compound/ OR (eNOSe* OR cyranose* OR spironose* OR ((electronic* OR artificial* OR spectromet* OR GC-MS) ADJ3 (nose*)) OR volatile-organic-compound* OR VOC).ab,ti,kf.) AND (exp Respiratory Tract Diseases/ OR (lung* OR pulmonar* OR respirator*-tract* OR Pneumonolog* OR asthma* OR COPD OR Sarcoidos*).ab,ti,kf.) NOT (news OR congres* OR abstract* OR book* OR chapter* OR dissertation abstract*).pt.

Cochrane Central

((eNOSE* OR e-NOSE* OR cyranose* OR spironose* OR ((electronic* OR artificial* OR spectromet* OR GC-MS) NEAR/3 (nose*)) OR volatile-organic-compound* OR VOC):ti,ab,kw) AND ((lung* OR pulmonar* OR respirator* NEXT/1 tract* OR Pneumonolog* OR asthma* OR COPD OR Sarcoidos*):ab,ti,kw)

CHAPTER 2

Exhaled breath analysis in interstitial lung disease

Curr Opin Pulm Med. 2023 Sep 1;29(5):443-450.

Iris G. van der Sar, Marlies S. Wijsenbeek, Catharina C. Moor

Erasmus Medical Center, Rotterdam (the Netherlands)

Abstract

Purpose of review

There is a need for better non-invasive tools to diagnose interstitial lung disease (ILD) and predict disease course. Volatile organic compounds present in exhaled breath contain valuable information on a person's health and may be a novel biomarker in ILD. In this review, we will give an overview of the basic principles of breath analysis, summarize the available evidence in ILD, and discuss future perspectives.

Recent findings

An increasing number of studies on exhaled breath analysis were performed over the last decade in patients with ILD, using two methods for exhaled breath analysis: gas chromatography-mass spectrometry and electronic nose technology. Most studies showed high accuracy for diagnosis of ILD, but study design and methods widely varied. Studies investigating the potential of electronic nose technology to predict treatment response and disease behavior are ongoing.

Summary

The majority of studies using exhaled breath analysis in ILD show promising results for diagnostic purposes, but validation studies are lacking. Larger prospective longitudinal studies using standardized methods are needed to collect the evidence required for developing an approved diagnostic medical test.

Keywords

Breath test, interstitial lung disease, gas chromatography-mass spectrometry, electronic nose technology, biomarker

Key points

- Volatile organic compounds present in exhaled breath might serve as future biomarkers for diagnosing ILD.
- Breath analysis by GC-MS is useful for individual compound analysis and might lead to new insights in pathophysiology.
- Breath analysis by eNose technology is promising as point-of-care medical tool because of real-time breath profiling.
- Available evidence on exhaled breath analysis shows generally high accuracies for detection of ILD, but externally validated results are still lacking.

Introduction

Around 400 BC Hippocrates already mentioned the importance of the human nose as diagnostic tool. He related the typical smell of various body secretions, like breath, sputum, urine and stool, to a certain diagnosis [1]. In the past few years, analysis of exhaled breath has increasingly been studied as potential diagnostic marker in a wide range of (respiratory) disorders, including interstitial lung disease (ILD) [2-4].

ILDs form a heterogeneous group of >200 different lung diseases in which the interstitium of the lung is affected by fibrosis, inflammation, or a combination of both [5]. Symptoms as dyspnea, cough, and fatigue are non-specific, and there is no single non-invasive diagnostic test for ILD. Hence, delay during the diagnostic process and referral to specialized hospitals is common [6]. Therefore, better screening and diagnostic tools are needed. Disease course of different ILDs is highly variable and even within specific diagnoses, disease behavior and response to therapy varies between patients. This highlights the importance of new prognostic and predictive biomarkers. However to date, no reliable blood biomarkers have been found in ILD [7]. As exhaled breath provides additional information about a person's health status, this is an interesting new biomarker source for ILD.

Compared to ILD, exhaled breath analysis has more extensively been studied in other lung diseases, with lung cancer being the main area of research in the last years. Kort et al. recently reported results from a multicenter validation study of breath analysis in lung cancer [8]. The robust results on differentiating patients with and without lung cancer show the potential value of using eNose technology as a diagnostic tool in medical practice. Strikingly, eNose technology can accurately predict response to treatment in patients with stage 4 non-small cell lung carcinoma [9, 10]. Validation studies are currently ongoing. More detailed information on eNose technology in other lung diseases can be found elsewhere [4].

In this review, we will focus on the potential of exhaled breath analysis in ILD, describe basic principles of different analysis methods, summarize available evidence in patients with ILD, and discuss future perspectives of exhaled breath analysis in ILD.

Exhaled breath analysis

Exhaled breath contains different types of compounds from exogenous and endogenous origin. Compounds range from large (e.g. microorganisms) to smaller compounds. The smaller compounds can be categorized as volatile (i.e. evaporates easily) or non-volatile, and as organic (i.e. contains carbon) or non-organic. For each category, different breath sample and analysis methods are required to capture the compounds. An overview can be found in **Table 1**.

Table 1: Examples of collection and analysis methods of exhaled breath compounds.

Target compound	Example	Breath sampling	Breath analysis
Non-volatile organic compounds and water soluble volatile molecules	Lipids, amino acids	Exhaled breath condensate	Spectrometry or enzyme immunoassay
Volatile organic compounds	Acetone, ethanol	Exhaled air*	Spectrometry or (cross-reactive) sensors
Volatile non-organic compounds	Nitric oxide	Exhaled air*	Specific sensor

*This table includes the most common ways of sampling and analyzing breath, and is not intended being a complete overview as no standard approach exists. *Exhaled air can be collected and processed in a sampling bag prior to compound analysis or can be captured and stored directly by exhaling through a device.*

Especially the analysis of volatile organic compounds (VOCs) is of interest in biomarker research. The concentration and type of VOCs (i.e. the volatilome) are affected by various (patho)physiological processes in the body and are unique for all individuals. The majority of endogenous VOCs originate from metabolic activity of organs or human microbiota, and from pathologic processes [11]. Subsequently, VOCs are excreted to the blood stream, diffused to and exhaled via the alveoli, or excreted by other organs such as the gut, kidneys or skin. As breath is the main source of VOCs and the lung tissue itself also excretes VOCs, breath analysis is mostly studied in respiratory diseases [12, 11].

Researchers can either choose a targeted or non-targeted approach when analyzing VOCs in breath. A targeted approach is hypothesis-based and aims to identify one or more predefined VOCs. Non-targeted analysis looks for differentiating VOCs or patterns in the full volatilome without prior knowledge or assumptions. This non-targeted approach is often called ‘breathomics’, as it shares similarities with the field of genomics, proteomics and metabolomics. In general, two different methods can be used to analyze VOCs: gas chromatography-mass spectrometry (GC-MS) or a sensor-based technique (so-called electronic nose, eNose). GC-MS analysis can be either targeted or non-targeted, but eNose research follows a non-targeted approach. **Figure 1** shows a schematic overview of similarities and differences of these methods.

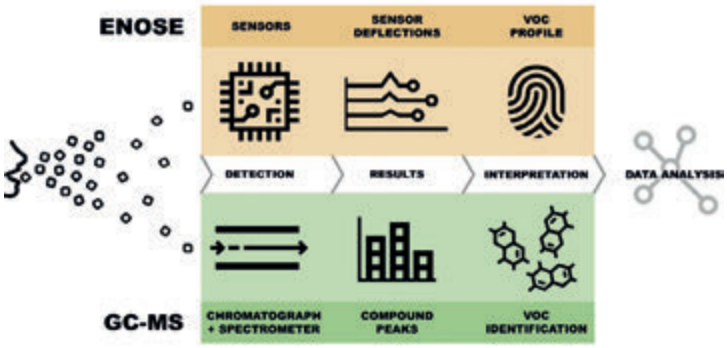


Figure 1: Overview and comparison of GC-MS and eNose breath analysis.
eNose = electronic nose; GC-MS = gas chromatography combined with mass spectrometry; VOC = volatile organic compounds.

Gas chromatography-mass spectrometry

The use of GC-MS to analyze VOCs in exhaled breath originates from the 1970s [13]. This analytical method combines two steps to identify compounds in gas mixtures. In short, during gas chromatography gaseous compounds are separated into molecules by sending the breath sample through a capillary column. All molecules leave the column at different times, resulting in a specific retention time. Subsequently, a mass spectrometer is used to ionize the molecules and calculate a mass-to-charge ratio of ionized molecules. The ratio can be used to identify specific VOCs by comparison with mass spectral libraries. Results are usually presented in a chromatogram, showing intensity peaks to indicate the concentration of all detected compounds. Technical and analytical variations exist for each step of GC-MS [14].

In the medical field, GC-MS could especially be useful for two purposes. First, this analysis method allows to identify individual compounds of exhaled breath, which might unravel pathophysiological processes. Second, many GC-MS studies evaluate the potential of specific VOCs as a new biomarker to diagnose or monitor specific conditions.

GC-MS in ILD

To date, only a small number of studies evaluated whether GC-MS analysis can detect ILD (**Table 2**). The first small pilot study in sarcoidosis compared VOC profiles of patients with those suspected of sarcoidosis. Suspected sarcoidosis was defined as the presence of enlarged mediastinal lymph nodes, without a confirmed diagnosis of sarcoidosis. There seemed to be differences between breath profiles of the two groups

based on 13 discriminative chromatogram peaks. However, the authors only provided visual plots, and did not perform statistical tests to evaluate whether breath profiles of the two groups were actually significantly different [15]. In 2017, a larger study found differences in sarcoidosis VOC profiles compared to healthy controls [16]. In both studies, not all patients had lung parenchymal involvement.

Two studies were conducted in patients with occupational lung diseases. Yang et al. studied stone workers with and without a pneumoconiosis diagnosis [17]. Jalali et al. included subjects exposed to silica, either with or without a diagnosis of silicosis [18]. Both studies identified several VOCs that differentiated patients with ILD from exposed patients without ILD, but it is unclear whether these VOCs were overlapping.

A recent study showed differences in breath profiles of patients with idiopathic pulmonary fibrosis (IPF) and connective tissue disease (CTD)-associated ILD using GC-MS analysis. Breath profiles of the patient groups differed significantly, with 16 discriminative VOCs being identified [19]. This was the first breathomics study using GC-MS indicating that VOC profiles in pulmonary fibrosis depend on the underlying condition. However, no test or validation cohort was applied, so further research should elucidate whether results can be replicated and validated. Additionally, this paper described 34 discriminatory VOCS between patients with IPF and healthy controls, of which five VOCs were most contributing. These five VOCs were different from the four identified significant VOCs detected by Yamada et al. in a similar analysis between IPF and healthy subjects conducted in 2017 [20]. Several factors could have contributed to this discrepancy, including differences in methodology (e.g. breath collection, breath and data analysis methods) and included patient cohorts (e.g. sample size, patient characteristics, matching of controls). Alternatively, these results may be exemplary for the limited performance of individual VOCs as disease specific biomarkers in ILD.

Preliminary data from conference abstracts during the last three years reported on new applications of GC-MS, such as prediction and screening. In a longitudinal cohort of patients with IPF, one specific VOC predicted disease progression after six months [21]. A study in patients with systemic sclerosis evaluated whether GC-MS analysis could be used for early detection of ILD in patients with systemic sclerosis. However, in this small cohort there were no differences in VOCs between systemic sclerosis patients with or without ILD [22].

Table 2: Main results of VOC breath analysis in patients with ILD using chromatography and spectroscopy.

Author	Year	Patient groups (n=)	Technique	Comparison	Discriminative VOCs	Performance (AUC / accuracy)
Plantier et al. [19]	2022	IPF (53) CTD-ILD (51) HC (51)	GC-tof-MS	IPF vs HC CTD-ILD vs HC IPF vs CTD-ILD	34 11 16	0.91 / 84.6% 0.84 / 77.5% 0.84 / 76.9%
Yamada et al. [20]	2017	IPF (40) HC (55)	MCC-IMS	IPF vs HC	5	- / 76.8-83.2%~
Yang et al. [17]	2017	Pneumoconiosis (25) Stone workers (154)	GC-MS	Pneumoconiosis vs exposed	9	0.90 / - *
Fijten et al. [16]	2017	Sarcoidosis (87)[#] HC (26)	GC-tof-MS	Sarcoidosis vs HC	9	0.76 / 74.1%*
Jalali et al. [18]	2016	Silicosis (4) HC (45)^ Silica exposed (20)	GC-MS	Silicosis vs exposed Silicosis vs HC	Multiple results	Not reported
Westhoff et al. [15]	2007	Sarcoidosis (5)[$] Sarcoidosis suspected (4)[+]	MCC-IMS	Sarcoidosis vs suspected	13	No statistical test results reported
Hayton et al. [21]	*2020*	*IPF (46)*	*GC-MS*	*Stable vs disease progression at 6 months*	*1*	*N/A*
Guiot et al. [22]	*2020*	*SSc (27, of which 17 with SSc-ILD)*	*Unknown*	*SSc-ILD vs SSc without ILD*	*0*	*N/A*

*Main results of the cited papers are displayed. Conference abstracts are shown in italic. ~A separate accuracy for each of the discriminative VOCs was calculated. ^Group was split in 20 non-smoking and 25 smoking subjects. [#]n=18 had Scadding stage 0. *Results of test/validation cohorts or cross-validation analyses are displayed here. [$]Not all patients had ILD. [+]Sarcoidosis excluded after biopsy of mediastinal lymphadenopathy. AUC = area-under-the-curve; CTD = connective tissue disease; GC = gas chromatography; MCC = multi-capillary column (i.e. variation of a capillary gas chromatograph); MS = mass spectrometry; HC = healthy control; ILD = interstitial lung disease; IMS = ion mobility spectrometry; IPF = idiopathic pulmonary fibrosis; SSc = systemic sclerosis; tof = time of flight (i.e. type of mass analyzer).*

Electronic Nose technology

eNose technology is a sensor-based technique for gas analysis based on the mammalian olfactory system. Exhaled VOCs are captured by an eNose device that contains multiple sensors (similar to the olfactory receptors in a human nose). These sensors have different sensitivities for ranges of VOCs, leading to specific sensor deflections that are subsequently pooled and processed to create a breath profile (**Figure 1**). By analyzing breath data with pattern recognition algorithms specific diseases can be distinguished, as previously shown by eNose studies in a wide range of respiratory and non-respiratory diseases [4, 23, 24]. The most important difference with GC-MS is that eNoses do not identify individual VOCs. Consequently, the purpose of eNose breath analysis is not to elucidate disease pathophysiology, but rather to use as a point-of-care diagnostic tool in clinical practice.

The first eNose was developed in 1964, but it was not until the 1980s that the first studies on the use of eNose in the medical field were published, and that the term electronic nose was used for the first time [25]. Since then, eNose technology has received increasing attention, and a variety of eNose devices has been developed and is currently available on the market for research purposes. These devices differ in type and number of sensors (electrical, gravimetric, and optical sensors), portability, method of breath collection (e.g., direct online analysis, or collection and storage on-site), correction for ambient air or other possible confounders, and technology readiness level [4]. To our knowledge, there are no studies available that directly compare the performance of different eNose devices, and hence, the choice for a device may depend on research setting, costs, and availability.

eNose technology in ILD

Several single-center studies on the potential of eNose technology for identification of ILDs have been published over the last ten years (**Table 3**). In these studies, different patient populations, eNose devices, and analysis techniques have been used. The first small pilot study in 2013 found that breath profiles of patients with untreated pulmonary sarcoidosis differed from healthy controls, with a cross-validated accuracy of 83.3% [26]. However, breath profiles of patients receiving immunosuppressive medication for sarcoidosis could not be distinguished from healthy controls. This implies that inflammation influences the breath profile in patients with sarcoidosis, since adequately treated patients were less likely to have ongoing inflammation. The potential of eNose technology to separate patients with sarcoidosis from healthy controls was confirmed by a larger single-center study using a different type of eNose [27]. In this cohort there was 100% discrimination between patients with sarcoidosis and healthy

controls, in both a training and test set, irrespective of the use of immunosuppressive medication and organ involvement. Patients with pulmonary sarcoidosis were adequately distinguished from patients with other ILDs, and in particular from patients with hypersensitivity pneumonitis, which is also characterized by granulomatous inflammation. External validation studies should further assess the ability of eNose to differentiate sarcoidosis from other granulomatous diseases. Within the group of patients with sarcoidosis there were no distinctive differences in breathprint, except between patients with a normal and elevated serum soluble IL2 receptor level. As the soluble IL2 receptor is a marker for inflammatory activity in sarcoidosis, this result also suggests an influence of systemic inflammation on breath profiles.

The potential of eNose technology in pneumoconiosis has been assessed in two studies [28, 29]. Yang and colleagues, who also studied GC-MS in this population, found a relatively high area under the curve (AUC) for differentiating patients with pneumoconiosis from a control group of stone workers [28]. A larger study published in 2022 evaluated breath profiles in a cohort of miners, with and without silicosis [29]. Their customized eNose system showed a good accuracy in a training and an external validation set, also for patients with early-stage disease. A strength of these studies is that they compared breath data of patients with a cohort at-risk for developing pneumoconiosis, suggesting that eNose technology has potential as a screening tool in this population.

Three research groups, each using a different eNose, showed that the breath profile of patients with IPF could be very well discriminated from healthy controls [30-32]. The first study from 2019 also showed a high accuracy when comparing CTD-ILD with healthy controls. Nevertheless, the accuracy to detect differences within the group of ILDs was slightly lower, and data were not validated [30]. A large single-center study found that patients with IPF had significantly different breath profiles than patients with other forms of pulmonary fibrosis (accuracy 91%, confirmed in a test set) [32]. There were also distinctive differences between individual ILDs, but group sizes were small and results need external validation. Dragonieri et al. found an accurate distinction between IPF and COPD in a training and external validation cohort, and a significant correlation between total cell count in bronchoalveolar lavage and eNose sensor data [31]. The current available data imply that eNose technology can be used as a non-invasive tool for screening and diagnostic purposes: i) to distinguish ILD from other chronic respiratory diseases and ii) to classify and phenotype individual ILDs.

Table 3: Main results of VOC breath analysis in patients with ILD using electronic nose technology.

Author	Year	Patient groups (n=)	eNose device	Comparison	Performance (AUC / accuracy)
Van der Sar et al. [27]	2022	Sarcoidosis (252, of which 224 pulmonary) ILD (317, of which 50 HP) HC (48)	SpiroNose	Sarcoidosis vs HC	1.00 / 100%*
				Pulmonary sarcoidosis vs ILD	0.87 / 83.2%*
				Pulmonary sarcoidosis vs HP	0.88 / 87.8%*
Xuan et al. [29]	2022	Silicosis (221, of which 85 stage I disease) Miners (398)	Customized system^	Silicosis vs miners	0.77-0.89 / 78.5-84.3%
				Stage I silicosis vs miners	0.78-0.94 / 70.8-91.7%
Moor et al. [30]	2021	ILD (215, of which 85 IPF) HC (48)	SpiroNose	ILD vs HC	1.00 / 100%*
				IPF vs non-IPF	0.87 / 91%*
Dragonieri et al. [31]	2020	IPF (42) COPD (43) HC (46)	Cyranose 320	IPF vs HC	1.00 / 98.5%
				IPF vs COPD	0.85 / 80.0%
Krauss et al. [32]	2019	ILD (174, of which 51 IPF and 25 CTD-ILD) COPD (23) HC (33)	Aeonose	IPF vs HC	0.95 / -
				CTD-ILD vs HC	0.90 / -
				IPF vs CTD-ILD	0.84 / -
				CTD-ILD vs COPD	0.85 / -
Yang et al. [28]	2017	Pneumoconiosis (34) Stone workers (64)	Cyranose 320	Pneumoconiosis vs workers	0.86-0.89 / 65.0-70.0%*
Dragonieri et al. [26]	2013	Pulmonary sarcoidosis (31, of which 11 untreated) HC (25)	Cyranose 320	Sarcoidosis untreated vs HC	0.83 / 83.3%
				Sarcoidosis untreated vs treated	- / 74.2%
Van der Sar et al. [34]	*2022*	*ILD (42, of which 22 starting immunosuppressive and 20 antifibrotic treatment)*	*SpiroNose*	*Yes vs no response to immunosuppressants*	*0.84 / -*
				Yes vs no response to antifibrotics	*0.75 / -*
Van der Sar et al. [33]	*2021*	*Pulmonary fibrosis (304)*	*SpiroNose*	*N/A (unsupervised analysis)*	*3 distinct clusters identified*

*Not all group comparisons and results described in the cited papers are displayed. Conference abstracts are shown in italic. Displayed subject numbers are the sum of training and test/validation cohorts, if applicable. *Results of independent test/validation cohorts. ^Based on Pilot (Vaporsense) sensor array. AUC = area-under-the-curve; COPD = chronic obstructive pulmonary disease; CTD = connective tissue disease; eNose = electronic nose; HC = healthy control; HP = hypersensitivity pneumonitis; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis.*

An exploratory study, of which results were presented as conference abstract in 2021, analyzed the potential of unsupervised analysis in a pulmonary fibrosis cohort [33]. In a group of 304 patients, three different clusters could be identified based on breath profiles. Clusters significantly differed with regard to diagnosis, gender, and immunosuppressant use, again indicating that breath profiles are influenced by inflammation. Longitudinal follow-up is needed to evaluate whether these clusters are associated with disease behavior and progression. Another application of eNose data is the prediction of disease behavior. A study in a small cohort of ILD patients suggested that eNose technology has the potential to predict treatment response in patients before starting on antifibrotic treatment (AUC 0.75) and immunosuppressive treatment (AUC 0.84) [34].

Future challenges and perspectives

The summarized evidence in this review shows that VOCs in exhaled breath hold valuable information for diagnosing ILD and potentially for prediction of disease course in individual patients. eNose breath tests hold great promise as a non-invasive, quick, and relatively low-cost medical application for ILD. Further validation in different cohorts and other important challenges need to be addressed before current research findings can be translated into an approved and validated medical test.

To date, there are no breath analysis studies in ILD published that replicate and validate previous findings in new patient cohorts. Moreover, available results are difficult to compare, which is partly due to differences in study design or healthcare setting. Many different methods and devices exist for breath collection and processing, VOC identification or VOC profile creation, and data analysis. Validation studies with new patient cohorts following similar standardized procedures are highly needed to test and validate various GC-MS and eNose applications.

GC-MS has already been studied for decades, but this technique has not made it to clinical practice in any medical field yet. There might be several reasons for this. Breath analysis with a chromatograph and spectrometer is a complex technique. The procedure is precise, elaborative and requires experienced investigators. Many labs have their own methods for breath collection and analysis, and approach to correct for possible confounders such as ambient air, environment or patient-related factors. Another reason why GC-MS studies have failed in finding a reliable biomarker for ILD might be that studies mainly focus on a combination of one or more significant individual VOCs. Single VOCs can provide valuable information on pathophysiological processes, but can be influenced by various endogenous or exogenous factors that are

difficult to identify or eliminate. Therefore, an approach that identifies a breath profile rather than individual VOCs may be more suitable when aiming to find a biomarker or medical test [3]. eNose technology has the advantage of creating this breath profile instantly by combining multiple sensor deviations. Besides, compared to GC-MS, measurements are less time-consuming and easier to perform. Moreover, there is immediate feedback on the measurement quality when using a device connected to an online platform. An online device facilitates analysis of breath data in real-time, which makes eNose technology suitable as point-of-care medical test. Especially when an eNose device corrects for known confounders, such as ambient air, it can be expected that findings can be replicated in various locations and health care settings.

Until now, the majority of exhaled breath studies in ILD focused on differentiating patient groups, to develop a diagnostic tool for ILD. Data on other applications as disease phenotyping, prediction of disease course, or response to treatment are preliminary. **Figure 2** shows an overview of the current status of developing clinical breath tests for ILD, with evidence from eNose studies in ILD categorized by phase of the diagnostic trajectory and clinical application. This figure highlights that none of the outcomes in ILD are externally validated and no implementation studies have been performed yet. To collect robust evidence for a clinically applicable breath test, all research steps need to be completed for each specific application and individual ILD diagnosis. To make this process more efficient and less costly, we need multi-national collaboration in large research projects. An ongoing multicenter longitudinal trial in four European countries will evaluate diagnostic accuracy for individual ILDs and assess the value of eNose technology as biomarker for disease progression and response to treatment (NCT04680832).

The ultimate future diagnostic breath test would profile the full human volatilome in real-time following a standardized procedure, correct for confounders, and be connected to an online database. The output of this test could be a probability score of individual ILD diagnoses for a particular patient (e.g. 85% probability that this patient has IPF) to support decision making by physicians and multi-disciplinary team discussions. Such a test might prevent invasive procedures in the diagnostic work-up of patients. A breath test using eNose technology is likely to be more suitable for this purpose than GC-MS. Nevertheless, comparison with GC-MS data might be of additional value to gain more insights in pathophysiological processes, and for the calibration or optimization of the medical test.

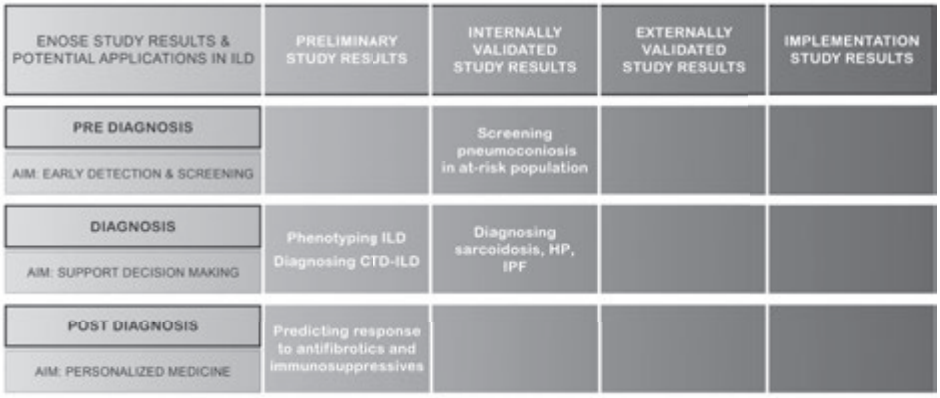


Figure 2: Overview of available evidence on eNose technology in ILD for each research step towards clinically applicable breath tests

Evidence is categorized per different phases and corresponding applications within the patient journey (before, during or after the diagnostic phase). No studies published externally validated data or implementation study data. CTD = connective tissue disease; HP = hypersensitivity pneumonitis; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis.

Conclusion

Since Hippocrates alluded to the nose as important diagnostic tool more than 2000 years ago, different techniques have been developed for exhaled breath analysis. Studies on eNose technology in ILD showed promising results for various clinical applications in ILD, but its value as a diagnostic and prognostic biomarker should be further explored and validated in the upcoming years.

References

1. Hippocrates, Potter P, Hippocrates. Hippocrates. Volume 1 [Internet]. 2022 Cambridge, Massachusetts ;: Harvard University Press Available from: https://www.loebclassics.com/view/LCL147/2022/volume.xml
2. Shirasu M, Touhara K. The scent of disease: volatile organic compounds of the human body related to disease and disorder. The Journal of Biochemistry. 2011;150(3):257-66.
3. Hayton C, Terrington D, Wilson AM, Chaudhuri N, Leonard C, Fowler SJ. Breath biomarkers in idiopathic pulmonary fibrosis: a systematic review. Respir Res. 2019 Jan 11;20(1):7.
4. van der Sar IG, Wijbenga N, Nakshbandi G, Aerts JGJV, Manintveld OC, Wijsenbeek MS, et al. The smell of lung disease: a review of the current status of electronic nose technology. Respiratory Research. 2021 2021/09/17;22(1):246.
5. Wijsenbeek M, Suzuki A, Maher TM. Interstitial lung diseases. Lancet. 2022 Sep 3;400(10354):769-86.
6. Spagnolo P, Ryerson CJ, Putman R, Oldham J, Salisbury M, Sverzellati N, et al. Early diagnosis of fibrotic interstitial lung disease: challenges and opportunities. Lancet Respir Med. 2021 Sep;9(9):1065-76.
7. Guiot J, Moermans C, Henket M, Corhay JL, Louis R. Blood Biomarkers in Idiopathic Pulmonary Fibrosis. Lung. 2017 Jun;195(3):273-80.
8. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh FH, van Putten JWG, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2023 Mar;163(3):697-706.
9. de Vries R, Muller M, van der Noort V, Theelen W, Schouten RD, Hummelink K, et al. Prediction of response to anti-PD-1 therapy in patients with non-small-cell lung cancer by electronic nose analysis of exhaled breath. Ann Oncol. 2019 Oct 1;30(10):1660-66.
10. Buma AIG, Muller M, de Vries R, Sterk PJ, van der Noort V, Wolf-Lansdorf M, et al. eNose analysis for early immunotherapy response monitoring in non-small cell lung cancer. Lung Cancer. 2021 Oct;160:36-43.
11. Ruzsanyi V, Lochmann F, Jürschik S, Mochalski P, Unterkofler K, Mayhew CA. Chapter 1 Origin and Emission of Volatile Biomarkers in Breath: End-tidal Perspective. Volatile Biomarkers for Human Health: From Nature to Artificial Senses. The Royal Society of Chemistry; 2023. p. 1-21.
12. de Lacy Costello B, Amann A, Al-Kateb H, Flynn C, Filipiak W, Khalid T, et al. A review of the volatiles from the healthy human body. J Breath Res. 2014 Mar;8(1):014001.
13. Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. Proc Natl Acad Sci U S A. 1971 Oct;68(10):2374-6.
14. Beale DJ, Pinu FR, Kouremenos KA, Poojary MM, Narayana VK, Boughton BA, et al. Review of recent developments in GC-MS approaches to metabolomics-based research. Metabolomics. 2018 Nov 17;14(11):152.
15. Westhoff M, Litterst P, Freitag L, Baumbach JI. Ion mobility spectrometry in the diagnosis of sarcoidosis: results of a feasibility study. J Physiol Pharmacol. 2007 Nov;58 Suppl 5(Pt 2):739-51.
16. Fijten RRR, Smolinska A, Drent M, Dallinga JW, Mostard R, Pachen DM, et al. The necessity of external validation in exhaled breath research: a case study of sarcoidosis. J Breath Res. 2017 Nov 29;12(1):016004.

17. Yang HY, Shie RH, Chang CJ, Chen PC. Development of breath test for pneumoconiosis: a case-control study. Respir Res. 2017 Oct 17;18(1):178.

18. Jalali M, Zare Sakhvidi MJ, Bahrami A, Berijani N, Mahjub H. Oxidative Stress Biomarkers in Exhaled Breath of Workers Exposed to Crystalline Silica Dust by SPME-GC-MS. J Res Health Sci. 2016 Summer;16(3):153-61.

19. Plantier L, Smolinska A, Fijten R, Flamant M, Dallinga J, Mercadier JJ, et al. The use of exhaled air analysis in discriminating interstitial lung diseases: a pilot study. Respir Res. 2022 Jan 20;23(1):12.

20. Yamada YI, Yamada G, Otsuka M, Nishikiori H, Ikeda K, Umeda Y, et al. Volatile Organic Compounds in Exhaled Breath of Idiopathic Pulmonary Fibrosis for Discrimination from Healthy Subjects. Lung. 2017 Apr;195(2):247-54.

21. Hayton C, Terrington D, Ahmed W, White I, Wilkinson M, Vekaria K, et al. Exhaled volatile organic compounds in idiopathic pulmonary fibrosis and disease progression. European Respiratory Journal. 2020;56((Hayton C., conalhayton@doctors.org.uk; Ahmed W.; White I.; Wilkinson M.; Vekaria K.; Fowler S.) University of Manchester, Manchester, United Kingdom).

22. Guiot J, Zanella D, Stefanuto P-H, Giltay L, Henket M, Guissard F, et al. Exhaled Volatile Organic Compounds are Able to Diagnose Systemic Sclerosis. European Respiratory Journal. 2020;56(suppl 64):3545.

23. Yang HY, Chen WC, Tsai RC. Accuracy of the Electronic Nose Breath Tests in Clinical Application: A Systematic Review and Meta-Analysis. Biosensors (Basel). 2021 Nov 22;11(11).

24. Scheepers M, Al-Difaie Z, Brandts L, Peeters A, van Grinsven B, Bouvy ND. Diagnostic Performance of Electronic Noses in Cancer Diagnoses Using Exhaled Breath: A Systematic Review and Meta-analysis. JAMA Netw Open. 2022 Jun 1;5(6):e2219372.

25. Gardner JW, Bartlett PN. A brief history of electronic noses. Sensors and Actuators B: Chemical. 1994 1994/03/01/;18(1):210-11.

26. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratú P, Resta O, et al. An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. Respiratory Medicine. 2013 2013/07/01/;107(7):1073-78.

27. van der Sar IG, Moor CC, Oppenheimer JC, Luijendijk ML, Daele PLA, Maitland-van der Zee AH, et al. Diagnostic Performance of Electronic Nose Technology in Sarcoidosis. Chest. 2021 2021;161.

28. Yang HY, Peng HY, Chang CJ, Chen PC. Diagnostic accuracy of breath tests for pneumoconiosis using an electronic nose. J Breath Res. 2017 Nov 29;12(1):016001.

29. Xuan W, Zheng L, Bunes BR, Crane N, Zhou F, Zang L. Engineering solutions to breath tests based on an e-nose system for silicosis screening and early detection in miners. Journal of Breath Research. 2022 2022/4;16(3):036001.

30. Krauss E, Haberer J, Maurer O, Barreto G, Drakopanagiotakis F, Degen M, et al. Exploring the Ability of Electronic Nose Technology to Recognize Interstitial Lung Diseases (ILD) by Non-Invasive Breath Screening of Exhaled Volatile Compounds (VOC): A Pilot Study from the European IPF Registry (eurIPFreg) and Biobank. J Clin Med. 2019 Oct 16;8(10).

31. Dragonieri S, Scioscia G, Quaranta VN, Carratu P, Venuti MP, Falcone M, et al. Exhaled volatile organic compounds analysis by e-nose can detect idiopathic pulmonary fibrosis. Journal of Breath Research. 2020 2020/7;14(4):047101.

32. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).

33. van der Sar IG, Moor CC, Luijendijk ML, Brinkman P, Maitland-Van Der Zee AH, Aerts JGJV, et al. Unsupervised analysis of electronic nose data identifies patient clusters in fibrosing interstitial lung disease. European Respiratory Journal. 2021;58(suppl 65):PA475.

34. van der Sar IG, Moor CC, Vellekoop BP, Wijsenbeek MS. Predicting treatment response in patients with interstitial lung disease using electronic nose technology. European Respiratory Journal. 2022;60(suppl 66):345.

CHAPTER 3

Patient reported experiences and delays during the diagnostic pathway for pulmonary fibrosis: a multinational european survey

Front Med (Lausanne). 2021 Aug 4;8:711194.

Iris G. van der Sar[1#], Steven Jones[2#], Deborah L. Clarke[3], Francesco Bonella[4], Jean Michel Fourrier[5], Katarzyna Lewandowska[6], Guadalupe Bermudo[7], Alexander Simidchiev[8], Irina R. Strambu[9], Marlies S. Wijsenbeek[1], Helen Parfrey[10]

[1]Erasmus Medical Center, Rotterdam (the Netherlands); [2]Action for Pulmonary Fibrosis, Lichfield (United Kingdom); [3]Galapagos NV, Mechelen (Belgium); [4]Ruhrlandklinik, University of Duisburg-Essen, Essen (Germany); [5]Association Pierre Enjalran Fibrose Pulmonaire Idiopathique, Meyzieu (France); [6]National Research Institute of Tuberculosis and Lung Diseases, Warsaw (Poland); [7]Hospital Universitari de Bellvitge, Barcelona (Spain); [8]Medical Institute MVR, Sofia (Bulgaria); [9]Carol Davila University of Medicine and Pharmacy, Bucharest (Romania); [10]Royal Papworth Hospital, Cambridge (United Kingdom); [#]Authors share first authorship

Abstract

Introduction

Pulmonary fibrosis includes a spectrum of diseases and is incurable. There is a variation in disease course, but it is often progressive leading to increased breathlessness, impaired quality of life and decreased life expectancy. Detection of pulmonary fibrosis is challenging, which contributes to considerable delays in diagnosis and treatment. More knowledge about the diagnostic journey from patients' perspective is needed to improve the diagnostic pathway. The aims of this study were to evaluate the time to diagnosis of pulmonary fibrosis, identify potential reasons for delays, and document patients emotions.

Methods

Members of European patient organisations, with a self-reported diagnosis of pulmonary fibrosis, were invited to participate in an online survey. The survey assessed the diagnostic pathway retrospectively, focusing on four stages: (1) time from initial symptoms to first appointment in primary care; (2) time to hospital referral; (3) time to first hospital appointment; (4) time to final diagnosis. It comprised open-ended and closed questions focusing on time to diagnosis, factors contributing to delays, diagnostic tests, patient emotions, and information provision.

Results

273 participants (214 idiopathic pulmonary fibrosis, 28 sarcoidosis, 31 other) from 13 countries responded. Forty percent of individuals took ≥ 1 year to receive a final diagnosis. Greatest delays were reported in stage 1, with only 50.2% making an appointment within three months. For stage 2, 73.3% reported a hospital referral within three primary care visits. However, 9.9% reported six or more visits. After referral, 76.9% of patients were assessed by a specialist within three months (stage 3) and 62.6% received a final diagnosis within three months of their first hospital visit (stage 4). Emotions during the journey were overall negative. A major need for more information and support during and after the diagnostic process was identified.

Conclusion

The time to diagnose pulmonary fibrosis varies widely across Europe. Delays occur at each stage of the diagnostic pathway. Raising awareness about pulmonary fibrosis amongst the general population and healthcare workers is essential to shorten the time to diagnosis. Furthermore, there remains a need to provide patients with sufficient information and support at all stages of their diagnostic journey.

Introduction

Interstitial lung disease (ILD) describes a relatively uncommon group of diseases characterised by inflammation and fibrosis of the lung interstitium. Pulmonary fibrosis is a chronic, and often progressive condition. There is, however, considerable variation amongst patients in terms of aetiology, treatment strategies and disease course [1]. Amongst all types of pulmonary fibrosis, idiopathic pulmonary fibrosis (IPF) is the most prevalent and accounts for about two-thirds of cases. It has the worst prognosis due to rapid disease progression with a mean survival of 4 years from diagnosis without anti-fibrotic therapy [2]. Other types of progressive pulmonary fibrosis include chronic hypersensitivity pneumonitis, auto-immune disease related ILD and occupational diseases such as asbestosis [1]. Epidemiological data for all types of pulmonary fibrosis are limited as most registries and studies have focused on IPF or progressive phenotypes only [3]. The reported prevalence (per 100,000 persons) of the ILDs that most often result in pulmonary fibrosis is 30.2 for sarcoidosis, 12.1 for ILD related to a connective tissue disease and 8.2 for IPF. Overall, the proportion of ILD patients who develop pulmonary fibrosis varies from 13 to 100% per individual disease [1].

The diagnostic journey usually starts with patients presenting to their primary care physicians with initial symptoms of cough or mild dyspnoea. These non-specific symptoms, combined with the heterogeneity and rarity of pulmonary fibrosis, as well as requirement for multiple diagnostic investigations, results in a prolonged time to diagnosis with potential delays related to patient factors and healthcare systems [4]. Reported time to diagnosis from the onset of initial symptoms varies in different studies but may be up to a median of 2.1 years (IQR 0.9-5.0) [5]. Longer time to diagnosis is associated with worse outcomes in IPF [6, 7], causes delayed treatment, leads to more extensive fibrosis [8] and affects patients' wellbeing. Therefore, it is important to get better insights into patients' experiences during the diagnostic journey to identify reasons for potential delays. Understanding patients' experiences will also help healthcare workers guide and support patients during their diagnosis journey. However, to date, only a few studies have explored the reasons for diagnostic delays using data reported by pulmonary fibrosis patients [9-13]. Most analyses are based on retrospective data obtained from healthcare records [6, 14, 5, 15, 8, 16-18].

In this paper, we present data obtained from a multinational patient survey regarding time to diagnosis and potential causes for diagnostic delays, together with patient experiences on the pathway to diagnosis. Based upon these findings, we provide general recommendations to improve the diagnostic process.

Methods

Survey design and distribution

A survey was designed to collect quantitative and qualitative data from patients diagnosed with pulmonary fibrosis across Europe. This survey was developed based upon a market research survey on the IPF patient journey (unpublished data) carried out using a mixture of in-depth telephone interviews with 28 patients and 30 pulmonologists, and online interviews with 315 pulmonologists spanning USA, France, Germany, Italy, Spain, United Kingdom, Australia, Brazil, Canada and Japan. The patient survey was developed jointly between Galapagos and two patient organisations: Action for Pulmonary Fibrosis (APF, based in the United Kingdom) and the European Idiopathic Pulmonary Fibrosis and Related Disorders Federation (EU-IPFF). Insights from this patient journey research resulted in a questionnaire incorporating both closed and open-ended questions, which focused on the following four stages of the patient journey to identify key points in the delay to diagnosis. The first stage was the time from first onset of symptoms at home, before seeking medical attention in a primary care setting; the second the amount of visits in primary care before being referred to a hospital specialist; the third the time taken to be seen in a hospital by a specialist; and the last the time taken to receive a diagnosis (**Figure 1A**). The survey also gathered data on the overall time from first onset of symptoms to diagnosis and information provided by healthcare workers. Patients were also asked about their feelings throughout the diagnostic journey and to provide advice for patients navigating this journey in the future. No personalised data were collected and all data were anonymised. The questionnaire was designed in English and translated into seven languages (Bulgarian, Dutch, French, German, Hungarian, Italian and Spanish) by a certified translation agency. It was created using the Typeform® platform. Patients were invited to complete the questionnaire by an e-mail containing a link to the platform. The complete survey in English can be found in the **Supplementary Material 1**.

The survey was disseminated by the EU-IPFF through its member patient organisations in Europe; these organisations distributed the survey to members and other patients through email and social media. Patients with a self-reported diagnosis of pulmonary fibrosis, and who had an email address and internet access were eligible to participate. The survey was sent out on 7th June 2020 with a reminder after two weeks. It closed on 1st July 2020. Ethical review was not required for this online questionnaire. Patients agreed with the use of their responses for further analysis without collection of personal data and were informed that all data was anonymised.

Data analysis

Responses in languages other than English were translated into English by a certified translation agency. Open-ended questions were assessed qualitatively and coded or categorised for interpretation. Data were uploaded and calculations were performed in Excel (Microsoft, Redmond, WA, USA). R version 4.0.3 for Mac OS X GUI (PBC, Boston, MA, USA) was used for creating a word cloud. All responses were included in the analysis, except for blank responses.

Literature search

In addition to the survey, a literature search on diagnostic delays in ILD, with a focus on pulmonary fibrosis, was conducted in order to provide a complete overview of the available evidence from patient surveys, physician surveys and medical file analysis.

The systematic literature search was performed in Embase, Medline, Web of science, Cochrane and Google scholar databases. The following search terms were used: diagnostic delay, time to diagnosis, interstitial lung disease (including sarcoidosis, vasculitis, interstitial pneumonia). Full search and outcome can be found in the **Supplementary Material 2**. Animal studies, paediatric subjects and articles in languages other than English were excluded. The reference list was screened for relevance by title and abstract. Letters to the editor, abstracts, posters and articles without available full text were excluded.

Results

Respondent characteristics

273 patients from thirteen different countries responded. The largest group of respondents were IPF patients (n=214, 78.4%), followed by sarcoidosis (n=28, 10.3%). Other types of pulmonary fibrosis diagnoses accounted for 31 respondents (11.4%) and included patients with autoimmune related disorders, chronic hypersensitivity pneumonitis and other conditions. The majority of respondents received a diagnosis of pulmonary fibrosis in Spain (21.6%), Belgium (20.1%), United Kingdom (18.3%), Italy (17.2%) or Germany (10.6%). A smaller number of respondent were diagnosed in the Netherlands (3.3%), Bulgaria (2.6%), France (1.8%), Poland (1.8%), Austria (1.5%), Ireland (0.4%), Norway 0.4%) and Romania (0.4%). Shortness of breath, dry cough, and tiredness were the most common initial symptoms in all diagnosis groups (**Figure 2A**).

The total time from initial symptom onset to a final diagnosis of pulmonary fibrosis, varied greatly amongst patients (**Figure 1C**). Overall, nearly 30% received a diagnosis within 3 months, with 31.3% patients with IPF receiving a diagnosis within three months, compared to 14.3% for sarcoidosis and 19.4% for other types of pulmonary fibrosis. Moreover, 40.2% of all patients had to wait a year or more to be diagnosed, with the largest difference between the proportion of patient with IPF (36.4%) and other types of pulmonary fibrosis (58.1%).

Stages of the diagnostic process

Stage 1: From initial symptom onset to first primary care assessment

More than half of respondents made a first appointment with a primary care physician within 3 months of symptom onset (52.0%), but nearly 30% waited more than 6 months (**Figure 1B, stage 1**). A number of patients responded that they did not delay visiting their doctor (26.7%). Of all patients with a delay in stage 1 of six months or less (n=177), 65.0% reported a total time to diagnosis of 1 year or less. Where patients with a delay of more than six months (n=72) in this stage, only 34.7% reported being diagnosed within a year.

There were a variety of reasons for delays (**Figure 2B**). In a large number of cases, patients delayed seeking medical advice because they were not concerned about their symptoms. Patients believed symptoms were related to other causes (e.g., cold, smoking, stress; 35.2%), related to age (25.6%), or due to another established disease (5.1%). The main reasons that triggered patients to make an appointment with their primary care physician were worries about their symptoms, including shortness of breath (45.1%), cough (31.9%) and fatigue (20.9%) (**Figure 2C**). For 18.7% of patients, it was the impact of symptoms on their daily activities, especially on physical activity (e.g., sports, climbing stairs, walking, household, gardening) and work-related activities that led them to consult their primary care physician. In addition, some patients were prompted to make an appointment following the suggestion from family members or friends (22.7%), or another physician (7%).

Stage 2: From start of primary care assessment to referral to pulmonologist

At the first primary care appointment, a variety of actions were taken by the treating physicians. Almost half of all patients were referred to a pulmonologist (**Figure 3**). Other reported physician's actions included additional tests (19.0%), treatment for another disease (16.5%) and referral to other specialists rather than a pulmonologist (10.3%). Overall, the majority (73.3%) of patients were referred to a pulmonologist within three primary care visits, but for 9.9% of patients it took six or more appointments (**Figure 1B, stage 2**).

Comparing the different diagnosis groups, 43.2% of IPF patients were referred to a pulmonologist after one primary care visit. This was lower for those with sarcoidosis (28.6%) and other types of pulmonary fibrosis (25.8%). Furthermore, 39.3% of sarcoidosis patients were referred after six or more primary care visits, compared to 6.6% of IPF and 6.7% of other fibrosis types in this cohort.

Figure 1: Diagnostic pathway and time to diagnosis.
(A) Schematic overview of the diagnostic pathway for pulmonary fibrosis, including stages and topics assessed in the survey. (B) Patient reported time per stage. (C) Patient reported overall time to diagnosis. PF, pulmonary fibrosis.

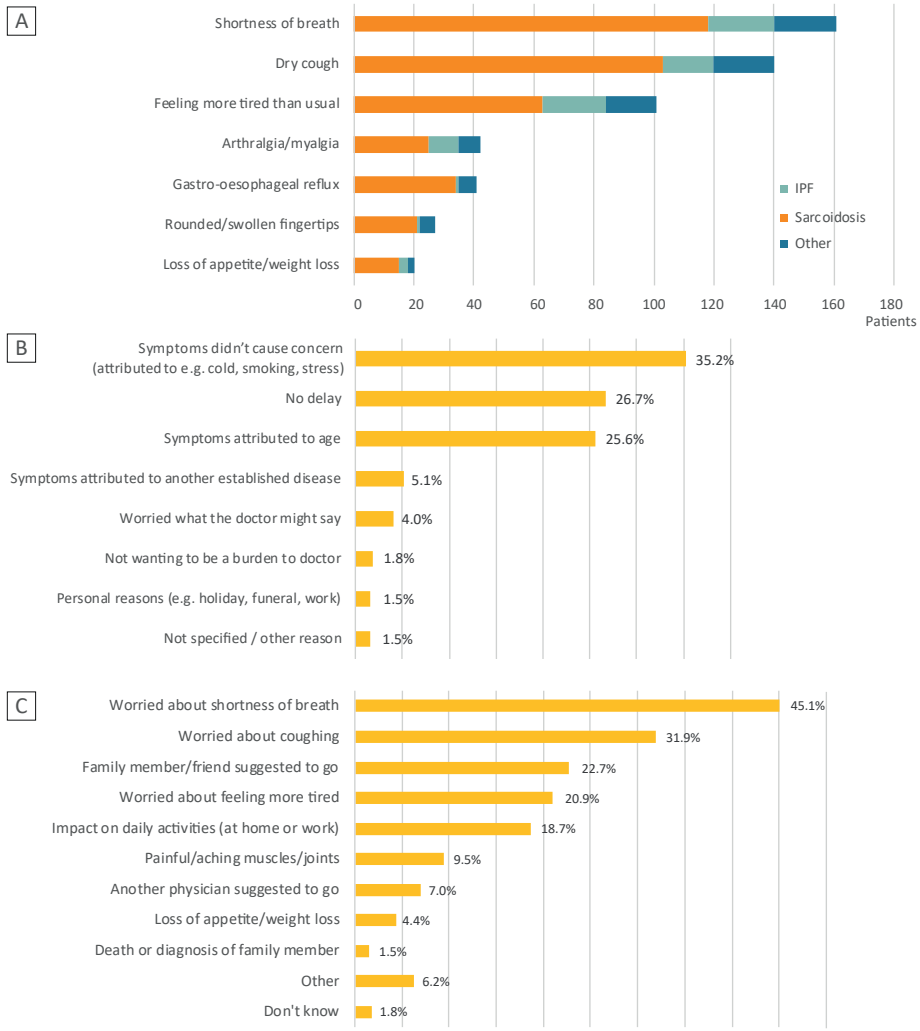


Figure 2: Patient symptoms and motives in stage 1.
(A) Number of patients (n =) reporting a specific symptom at onset. Bars are divided into diagnosis groups (total responses n = 532). (B) Reason to delay the initial primary care appointment (n = 277). (C) Reason to schedule the initial primary care appointment (n = 463). Percentages do not add up to 100% as more than one response was allowed. IPF, idiopathic pulmonary fibrosis.

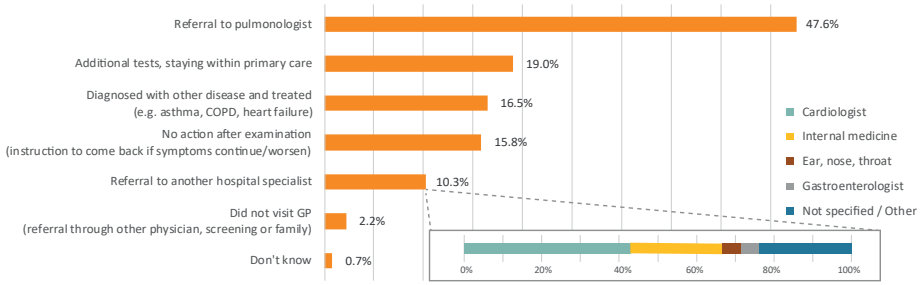


Figure 3: Action of physician at first visit primary care.
Percentages do not add up to 100% as more than one response was allowed. Total responses n=306. COPD = chronic obstructive pulmonary disease; GP = general practitioner.

Stage 3: From referral to first hospital appointment

Once patients were referred to a pulmonary specialist, 76.9% of all patients had their first visit within three months (**Figure 1B, stage 3**). This was lower for the subgroup of sarcoidosis patients (50.0%) compared to IPF (79.9%), and other types of pulmonary fibrosis (80.6%). Few IPF patients (2.3%) had a delay of more than a year from referral to first hospital appointment, in contrast to almost a third of the sarcoidosis patients (32.1%). All patients with other types of pulmonary fibrosis were assessed within a year of the referral.

Stage 4: From first hospital appointment to diagnosis pulmonary fibrosis

The 273 respondents underwent a total of 1,232 diagnostic tests in the hospital (**Table 1**). The majority of patients reported having performed spirometry (n=246), blood tests (n=222) and chest imaging (X-ray n=209; CT scan n=201) without large differences in proportions between the diagnosis subgroups. Other tests reported included assessment of 6-minute walk test (n=149), lung biopsy (n=125) and bronchoaveolar lavage (n=74). Lung biopsy was more frequently reported by sarcoidosis patients compared to the other subgroups.

Although the final diagnosis was made within three months of the first hospital appointment for 62.6% of the 273 patients (**Figure 1B, stage 4**), 21.6% took between 3 months and 1 year, and 13.2% took over one year; 2.6% did not know how long this took. Small differences were found between the proportion of patients in each diagnosis group who were diagnosed within 3 months (IPF 64.5%, sarcoidosis 50.0%, and other pulmonary fibrosis types 61.3%) and more than 1 year after the first hospital appointment (IPF 11.2%, sarcoidosis 21.4%, and other pulmonary fibrosis types 19.4%).

Table 1: Performed tests in hospital before diagnosis.

Tests	IPF (n=214)		Sarcoidosis (n=28)		Other type (n=31)	
	n=	% of patients in subgroup	n=	% of patients in subgroup	n=	% of patients in subgroup
Spirometry	194	90.7%	24	85.7%	28	90.3%
Blood tests	168	78.5%	26	92.9%	28	90.3%
Chest X-ray	161	75.2%	22	78.6%	26	83.9%
CT scan	156	72.9%	19	67.9%	26	83.9%
6-minute walk test	120	56.1%	10	35.7%	19	61.3%
Lung biopsy	93	43.5%	19	67.9%	13	41.9%
Bronchoaveolar lavage	49	22.9%	11	39.3%	14	45.2%
Other / Don't know	5	2.3%	1	3.6%	-	-
Tests per patient (mean)	*4.4*		*4.7*		*5.0*	

Number of patients (n=) reporting a specific diagnostic test. Percentages do not add up to 100% as more than one response was allowed. CT = Computed tomography; IPF = idiopathic pulmonary fibrosis.

Experiences and recommendations

Information provision

We assessed the patient perceptions on the information provided at the different stages in the diagnostic pathway. During assessment at the hospital (stage 4), 13.6% of patients reported not knowing why certain diagnostic tests were being performed. Almost a quarter (23.6%) of all patients felt they received insufficient information. At diagnosis, most patients (75.6%) received an explanation about their diagnosis from a physician and/or specialist nurse during a consultation. However, only 6.0% percent of patients received educational materials and 6.0% received information related to support groups. A small number (3.0%) reported not having received any information at the time of diagnosis. In response to an open-ended question, patients reported that the discussion with their doctor or nurse was particularly valuable, as well as ongoing follow up appointments at the hospital and contact details to enable them to ask questions or reach out if they were feeling unwell.

The patients stated that they would have benefitted from more information during the diagnostic process, not only after the diagnosis was established. They would have welcomed more information before, at and after diagnosis on the following topics: differential diagnosis, diagnostic tests, available pharmacological and non-pharmacological therapies, disease course and prognosis. Respondents would have also liked more information on living with pulmonary fibrosis day-to-day, future perspectives, access to a psychologist, and information on peer support groups for patients and carers.

Emotional experiences

Patients' perceptions and experiences were retrospectively assessed at different time points during their diagnostic journey. When describing their feelings after the onset of symptoms before their first doctor's visit (n=179 responses), 65.4% of the respondents experienced negative emotions, 5.6% positive emotions, and the remainder (29.1%) were neutral. When asked to describe feelings after referral to the hospital (n=240 responses), 74.6% of the responding patients experienced negative emotions at that time (16.7% neutral, 8.8% positive) (**Figure 4**).

Figure 4: Reported feelings during stage 3.
Words grouped after coding, ones with minimum frequency of 2 are included in figure (n=28). Full list (n=62) can be found in Supplementary Material 3.

Recommendations to patients

Overall, the advice and tips offered by patients to those undiagnosed or living with pulmonary fibrosis were: seeking help early when you experience symptoms, pushing for a speedier diagnosis, seeking as much information as possible from healthcare professionals at all stages, taking regular exercise, joining pulmonary rehabilitation classes to assist with breathlessness, joining patient support groups, remaining positive, pacing themselves, and making the most of their time. General tips for fellow patients regarding mental wellbeing contained phrases such as: stay calm, stay positive, no stress, don't despair, don't give up, focus on the present, and don't get agitated, frustrated or anxious.

Recommendations to healthcare

Advices to healthcare workers included performing tests earlier, providing more information and lifestyle advice, gaining more knowledge about pulmonary fibrosis,

improving communication between healthcare workers, structuring the diagnostic process better, and earlier start of pharmacological and palliative treatment. More recommendations are listed as quotes in **Supplementary Material 4**.

Discussion

The purpose of this survey was to document the time taken to diagnosis and to identify potential causes of delays at different stages of the diagnostic pathway for pulmonary fibrosis patients in Europe. The second aim was to describe patients' experiences during this journey.

We found that the time to diagnosis varies widely. Only 30% of patients were diagnosed with pulmonary fibrosis within 3 months of symptom onset, while for over 40% of patients it took more than one year to be diagnosed. Other studies observed a median time from onset of first symptoms to diagnosis of 7 months (range 0-252) based on a patient survey [12] and 2.1 years (IQR 0.9–5.0) from a retrospective cohort study [5]. In 2020, a group of ILD specialists reported a mean time from symptom onset to pulmonary fibrosis diagnosis of 2.3 years (Q1-Q3: 2-3) [19]. The proportion of patients in our cohort who took more than a year to be diagnosed is smaller than that reported by other studies of pulmonary fibrosis patients [9, 12]. Moreover, in a study of IPF patients, the median time to diagnosis was 13.6 months (range 5.9 to 39.5; max. 274.3) but 49% of the cohort received a diagnosis after more than one year [18]. In another study, the median time for establishing a diagnosis was 1.5 years (range <1 week to 12 years) but this was calculated from the time of the first doctors' appointment rather than onset of symptoms [10]. Compared to these historical studies, our results suggest fewer patients had such long delays from symptom onset to diagnosis.

Delays in diagnosis can occur at each stage of the patient journey and may be due to both patient- and healthcare-related causes. The longest delay we observed occurred in stage 1 (**Figure 1B**). More delay in this stage translated into a prolonged time to the final diagnosis. Our results show that only a quarter (26.7%) of all patients did not delay their initial appointment with their primary care physician. These findings are similar to results from a patient survey conducted in 2015 [12]. A more recent survey amongst IPF patients reported a median delay of 0.1 years for this stage [5]. From our survey, those who delayed their appointment reported they had not been concerned about their symptoms. This highlights the need to raise awareness of pulmonary fibrosis amongst the general public, so that individuals seek medical assistance earlier.

The time taken by people being treated in primary care (Stage 2) varies. In our survey, almost 40% of patients were referred to a hospital specialist after their first primary care appointment, which is greater than that observed in a study conducted in the USA in 2015 (27.8%) [12]. However, Hoyer et al. found that 80% of patients in Denmark (between 2016-2019) were referred after 1 or 2 visits to their general practitioner [5]. These observations may reflect differences in healthcare systems or in awareness of pulmonary fibrosis between countries.

Of all respondents, 15.3% were referred after 4 or more appointments. Several factors may contribute to delays in primary care. Firstly, initial symptoms in the early stage of the disease can be non-specific and not yet known to be life threatening. In support of this, 42% of IPF patients had a normal lung function when initially assessed in primary care [14]. Secondly, primary care physicians may suspect the symptoms to be due to more common respiratory diseases (such as asthma, pneumonia, bronchitis, allergies and COPD [12]) and decide on a period of observation [20]. Such misdiagnosis occurs in up to 41% of patients [5] and can prolong time to establish an ILD diagnosis [12, 13]. Thirdly, primary care physicians may lack knowledge about pulmonary fibrosis. A study in Finland found almost half of referral letters lack key information related to possible ILD diagnosis [14]. An e-learning for General Practitioners has recently been launched by the Royal College of General Practitioners in the United Kingdom and patient organisation APF to increase knowledge about symptoms and treatment of pulmonary fibrosis [21]. In other countries, similar initiatives are evolving.

Stage 3 is the time between being referred and the patient's actual hospital appointment. Based on our data, 76.9% were assessed by a pulmonologist within three months, compared to 91% reported from a Finnish cohort [14]. In this Finnish study only referral letters to tertiary care centres were evaluated, which may explain the higher percentage. However, in the United Kingdom and Ireland the time to secondary care respiratory clinic visit (47 days (25–84)) was significantly less than the time to an ILD specialist clinic visit (290 days (133–773)) [17]. Given differences in the structure and complexities of healthcare systems, it is difficult to compare data from different countries. To our knowledge, there are no published data as to why delays in stage 3 occur. It may reflect waiting times or patients postponing a hospital clinic appointment.

Delays occurring from the first hospital appointment to final diagnosis (stage 4) can, be partly explained by the number of diagnostic tests, access to them [22] and challenges in confirming a specific diagnosis accurately. Patients in our survey underwent on average 4.5 tests per person. The most common were spirometry, blood tests and radiological chest imaging, similar to those reported by others [12, 15]. The proportion

of reported lung biopsies was surprisingly high in our cohort (41.9-67.9%), which may reflect variation in healthcare practices, as biopsy rates differ between countries (16.1 to 1.2% (2013 to 2019) in England [23], 34.1% in Germany (2012-2014) [24], 20.1% in Italy (2015-2017) [25]).

Several parameters may predict potential delays, as they are associated with an increased time to diagnosis. In our cohort patients with a final diagnosis of IPF experience shorter delays and undergo less invasive diagnostic testing than patients with other diagnoses. These differences may be due to IPF patients presenting with more severe symptoms initially, availability of the IPF international diagnostic guidelines, or availability of tests [22, 26]. We can only speculate on this as we did not collect data on disease severity nor have powered for separate subgroup analyses. Another parameter that may influence time to diagnosis are the specific presenting symptoms. When patients present with dyspnoea, the median time to confirm an ILD diagnosis was 307 days, which increased for symptoms as cough and fatigue, to 563 and 639 days respectively [16]. Similarly, Pritchard et al. found an association between dyspnoea and a shorter time to hospital referral, which was not observed for lung crackles or chronic cough [8]. Other factors that may contribute to a delayed diagnosis include presence of specific comorbidities, male sex, increased body mass index, older age, previous inhalation therapy use, preserved diffusing capacity and better St. George's Respiratory Questionnaire scores [6, 5, 17, 18]. Lastly, abnormal chest imaging is one of the main reasons to initiate a hospital referral from primary care [14, 8] and naming ILD on the thoracic CT radiologic report doubled the likelihood of a referral to a pulmonologist within 6 months [8]. Interestingly, performing lung function tests in primary care, which indicated the possibility of ILD did not significantly influence time to CT scan or hospital referral [8].

Patients' experiences

The pulmonary fibrosis journey to diagnosis generally involves extensive, repetitive and sometimes invasive testing. Most patients in the survey reported that this causes a considerable burden, which can impact on emotional health, finances, and personal and professional life [12]. Shortening the diagnostic journey and assessment at an ILD expert centre results in higher patient satisfaction [10]. In addition, our survey highlighted the need to better inform patients during their diagnostic journey, to provide information on how to live with pulmonary fibrosis and advice on lifestyle changes at diagnosis. After diagnosis, providing information on perspectives and options and discussions concerning symptom management should also be a priority as identified by our respondents. These observations are similar to those reported from surveys and in-

depth patient interviews [27, 28]. In one paper, authors highlighted that patients need time to come to terms with their diagnosis and that repeated provision of information was essential to fully understand the consequences and implications of their disease [9]. However, a survey of ILD professionals in Europe showed that although two-thirds of specialist centres offered patient education only a few patients attended these existing programmes [13]. Furthermore, only 6% of patients from our survey were informed about support groups, despite the value of peer support to patients and carers reported not only by our respondents but also from a previous patient survey [10]. However, scientific evidence for the benefits of peer support is scarce [29]. Regarding caregivers' needs, several patients in our survey highlighted the need to provide them with more information on the patient's experience and practical help on how best to support them [30]. Finally, providing details of websites which offer reliable and accurate information is important as many websites contain incorrect or outdated information [31].

Limitations

In this study, we used a variety of survey methods, which resulted in a good understanding of patients' perceptions and experiences. Nevertheless, using patient reported data is also a weakness of this study. A general limitation of open-ended questions is the variety of responses, which could not be included in the quantitative analysis. Limitations also include patient recall, non-response and misinformation bias. These factors could have influenced the lung biopsies reported in our cohort, as patients may not differentiate between procedures such as endobronchial biopsies, surgical biopsies or only bronchoscopy. As the responses were anonymous, we could not confirm information from medical records.

Several factors prevent generalization of these results to the overall population of patients with pulmonary fibrosis. We used a non-random sample of self-selecting pulmonary fibrosis patients invited via patient associations without a pre-defined number of invited patients, target or countries. Most organisations have, until recently, focused on supporting and representing IPF patients, which likely accounts for the high number of IPF participants in this survey. Furthermore, patient characteristics, such as gender, age, comorbidities and stage and/or severity of disease were not collected.

Although there are European guidelines for the diagnostic pathway of IPF and other ILDs, differences exist between countries [13]. This may be related to the organisation of healthcare and options for primary care physicians to refer for CT scans or to ILD expertise centres. In our survey, we did not take these differences into account nor collect information on whether a CT chest scan was performed in primary care.

Recommendations clinical practice

There is an urgent need to improve the diagnostic journey and recommendations on how to achieve this have been raised in several papers [10, 11, 13]. Our findings on patient satisfaction and diagnostic delay endorse this and encourage further improvement. Rapid diagnosis is becoming increasingly important because several treatments are currently available to slow disease progression, improve quality of life and may extend life expectancy [32-34]. Although there are guidelines and other guidance documents on features, diagnosis and management of ILD [35, 26, 36, 37] many patients have a diagnosis that is not confirmed by a multidisciplinary discussion and do not receive treatment [38]. Additionally, geographical differences that may influence time to diagnosis and access to treatment still exists between countries [13].

Table 2: Strategies for improving the diagnostic pathway of pulmonary fibrosis patients.

	Stage 1	**Stage 2**	**Stage 3**	**Stage 4**	**After diagnosis**
Education and information	Increase awareness of PF amongst the general public	Increase awareness of PF symptoms amongst primary care physicians and nurses	Inform patients and policy makers on the need for urgency in hospital referral	Inform patients about the reasons for diagnostic investigation and the differential diagnosis	Inform patients about drug treatment, non-pharmaceutical treatment (rehabilitation, oxygen therapy, palliative care, lung transplant), prognosis and lifestyle
Improving standard care		Develop criteria for referral for chest CT scan or to a specialist when abnormalities on examination suggest PF	Regular (virtual) MDDs between general hospital specialist and ILD experts	Day case assessment with diagnostic investigations and clinical assessment	Introduce psychological support, helplines and peer groups for patients as part of standard care
		Better communication between primary care physician and ILD specialist	Increase the number of ILD specialists in general hospitals	Availability of DLCO measurement in all hospitals	Discuss duration and frequency of follow-up visit
Research areas	Identify the optimum way to provide information about PF to the general population	Cost-effectiveness of performing chest CT scan in primary care or at community facilities	Comparing waiting times and diagnostic pathway of PF to other uncommon diseases or disorders with poor prognosis (e.g. cancer [40])		Assess caregivers' needs on counselling and support

Content is based on survey outcomes, available literature and authors' opinions. CT = computed tomography; DLCO = diffusing capacity for carbon monoxide; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis; MDD = multidisciplinary discussion; PF = pulmonary fibrosis.

In **Table 2**, we provide concrete strategies for each stage of the diagnostic journey to improve the standard clinical practice and patient satisfaction in order to promote a more rapid pathway for patients with pulmonary fibrosis throughout Europe. These strategies are based upon our survey outcomes, available literature, and expert authors' opinions. Awareness and education in general public, patients and healthcare workers is a major topic in this field, as well as for other rare lung diseases [39].

Conclusion

From the onset of symptoms to diagnosis of pulmonary fibrosis, the patient journey involves delays at each stage of the diagnostic pathway. Most of these delays are avoidable. Based upon our findings, there is a particular need to raise awareness of pulmonary fibrosis in the general population. Additionally, patients' experiences highlight the need for understandable information concerning the diagnostic tests performed, differential diagnosis, final diagnosis and treatments as well as peer support groups. Improving several aspects of the diagnostic pathway for pulmonary fibrosis is therefore warranted to minimise delays and improve patient satisfaction throughout Europe.

Acknowledgments

The authors wish to thank W.M. Bramer from the Erasmus Medical Center Medical Library for developing the search strategies. The authors also wish to thank all patients and their carers for participating in the survey.

References

1. Wijsenbeek M, Cottin V. Spectrum of Fibrotic Lung Diseases. N Engl J Med. 2020 Sep 3;383(10):958-68.
2. Khor YH, Ng Y, Barnes H, Goh NSL, McDonald CF, Holland AE. Prognosis of idiopathic pulmonary fibrosis without anti-fibrotic therapy: a systematic review. European Respiratory Review. 2020;29(157):190158.
3. Olson AL, Gifford AH, Inase N, Fernandez Perez ER, Suda T. The epidemiology of idiopathic pulmonary fibrosis and interstitial lung diseases at risk of a progressive-fibrosing phenotype. Eur Respir Rev. 2018 Dec 31;27(150).
4. Gulati M. Diagnostic assessment of patients with interstitial lung disease. Primary Care Respiratory Journal. 2011 2011/06/01;20(2):120-27.
5. Hoyer N, Prior TS, Bendstrup E, Wilcke T, Shaker SB. Risk factors for diagnostic delay in idiopathic pulmonary fibrosis. Respiratory Research. 2019 2019/05/24;20(1):103.
6. Lamas DJ, Kawut SM, Bagiella E, Philip N, Arcasoy SM, Lederer DJ. Delayed access and survival in idiopathic pulmonary fibrosis: a cohort study. Am J Respir Crit Care Med. 2011 Oct 01;184(7):842-47.
7. Vasakova M, Mogulkoc N, Sterclova M, Zolnowska B, Bartos V, Plackova M, et al. Does timeliness of diagnosis influence survival and treatment response in idiopathic pulmonary fibrosis? Real-world results from the EMPIRE registry. Eur Respir J. 2017;50.
8. Pritchard D, Adegunsoye A, Lafond E, Pugashetti JV, Digeronimo R, Boctor N, et al. Diagnostic test interpretation and referral delay in patients with interstitial lung disease. Respir Res. 2019;20(1).
9. Collard HR, Tino G, Noble PW, Shreve MA, Michaels M, Carlson B, et al. Patient experiences with pulmonary fibrosis. Respir Med. 2007 Jun;101(6):1350-54.
10. Schoenheit G, Becattelli I, Cohen AH. Living with idiopathic pulmonary fibrosis:an in-depth qualitative survey of European patients. Chronic Respiratory Disease. 2011;8(4):225-31.
11. Bonella F, Wijsenbeek M, Molina-Molina M, Duck A, Mele R, Geissler K, et al. European IPF Patient Charter: unmet needs and a call to action for healthcare policymakers. European Respiratory Journal. 2016;47(2):597-606.
12. Cosgrove GP, Bianchi P, Danese S, Lederer DJ. Barriers to timely diagnosis of interstitial lung disease in the real world: The INTENSITY survey. BMC Pulm Med. 2018;18(1).
13. Moor CC, Wijsenbeek MS, Balestro E, Biondini D, Bondue B, Cottin V, et al. Gaps in care of patients living with pulmonary fibrosis: A joint patient and expert statement on the results of a europe-wide survey. ERJ Open Res. 2019;5(4).
14. Purokivi M, Hodgson U, Myllärniemi M, Salomaa ER, Kaarteenaho R. Are physicians in primary health care able to recognize pulmonary fibrosis? Eur Clin Respir J. 2017;4(1).
15. Mooney J, Chang E, Lalla D, Papoyan E, Raimundo K, Reddy SR, et al. Potential delays in diagnosis of idiopathic pulmonary fibrosis in Medicare beneficiaries. Ann Am Thorac Soc. 2019;16(3):393-96.
16. Sköld CM, Arnheim-Dahlström L, Bartley K, Janson C, Kirchgaessler KU, Levine A, et al. Patient journey and treatment patterns in adults with IPF based on health care data in Sweden from 2001 to 2015. Respir Med. 2019;155:72-78.
17. Brereton CJ, Wallis T, Casey M, Fox L. Time taken from primary care referral to a specialist centre diagnosis of idiopathic pulmonary fibrosis: an opportunity to improve patient outcomes? ERJ Open …. 2020.

18. Snyder LD, Mosher C, Holtze CH, Lancaster LH, Flaherty KR, Noth I, et al. Time to diagnosis of idiopathic pulmonary fibrosis in the IPF-pro registry. Am J Respir Crit Care Med. 2020;201(1).

19. Wuyts WA, Papiris S, Manali E, Kilpeläinen M, Davidsen JR, Miedema J, et al. The Burden of Progressive Fibrosing Interstitial Lung Disease: A DELPHI Approach. Adv Ther. 2020 Jul;37(7):3246-64.

20. Heins MJ, Schermer TRJ, de Saegher MEA, van Boven K, van Weel C, Grutters JC. Diagnostic pathways for interstitial lung diseases in primary care. Prim Care Respir J. 2012;21(3):253-54.

21. Royal College of General Practitioners. Available online at: https://elearning.rcgp.org.uk/course/view.php?id=409

22. Cottin V. Current approaches to the diagnosis and treatment of idiopathic pulmonary fibrosis in Europe: the AIR survey. European Respiratory Review. 2014;23(132):225-30.

23. Spencer LG, Loughenbury M, Chaudhuri C, Spiteri M, Parfrey H. Idiopathic Pulmonary Fibrosis in the United Kingdom: Analysis of the British Thoracic Society Electronic Registry between 2013 and 2019. ERJ Open Research. 2020:00187-2020.

24. Behr J, Kreuter M, Hoeper MM, Wirtz H, Klotsche J, Koschel D, et al. Management of patients with idiopathic pulmonary fibrosis in clinical practice: the INSIGHTS-IPF registry. The European respiratory journal. 2015;46(1):186-96.

25. Poletti V, Vancheri C, Albera C, Harari S, Pesci A, Metella RR, et al. Clinical course of IPF in Italian patients during 12 months of observation: results from the FIBRONET observational study. Respiratory Research. 2021 2021/02/24;22(1):66.

26. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2018;198(5):e44-e68.

27. Overgaard D, Kaldan G, Marsaa K, Nielsen TL, Shaker SB, Egerod I. The lived experience with idiopathic pulmonary fibrosis: a qualitative study. European Respiratory Journal. 2016;47(5):1472-80.

28. Senanayake S, Harrison K, Lewis M, McNarry M, Hudson J. Patients' experiences of coping with Idiopathic Pulmonary Fibrosis and their recommendations for its clinical management. PLOS ONE. 2018;13(5):e0197660.

29. Magnani D, Lenoci G, Balduzzi S, Artioli G, Ferri P. Effectiveness of support groups to improve the quality of life of people with idiopathic pulmonary fibrosis a pre-post test pilot study. Acta Biomed. 2017 Nov 30;88(5S):5-12.

30. Ramadurai D, Corder S, Churney T, Graney B, Harshman A, Meadows S, et al. Understanding the informational needs of patients with IPF and their caregivers: 'You get diagnosed, and you ask this question right away, what does this mean?'. BMJ Open Quality. 2018;7(1):e000207.

31. Fisher JH, O'Connor D, Flexman AM, Shapera S, Ryerson CJ. Accuracy and Reliability of Internet Resources for Information on Idiopathic Pulmonary Fibrosis. American Journal of Respiratory and Critical Care Medicine. 2016;194(2):218-25.

32. Bolton CE, Bevan-Smith EF, Blakey JD, Crowe P, Elkin SL, Garrod R, et al. British Thoracic Society guideline on pulmonary rehabilitation in adults: accredited by NICE. Thorax. 2013;68(Suppl 2):ii1-ii30.

33. Flaherty KR, Wells AU, Cottin V, Devaraj A, Walsh SLF, Inoue Y, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. New England Journal of Medicine. 2019;381(18):1718-27.

34. Susan SJ, Jerry AK, David JL, Marya G, Tanzib H, Ai-Yui MT, et al. Home Oxygen Therapy for Adults with Chronic Lung Disease. An Official American Thoracic Society Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2020;202(10):e121-e41.

35. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med. 2013 Sep 15;188(6):733-48.

36. Raghu G, Remy-Jardin M, Ryerson CJ, Myers JL, Kreuter M, Vasakova M, et al. Diagnosis of Hypersensitivity Pneumonitis in Adults. An Official ATS/JRS/ALAT Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2020;202(3):e36-e69.

37. Hoffmann-Vold AM, Maher TM, Philpot EE, Ashrafzadeh A, Distler O. Assessment of recent evidence for the management of patients with systemic sclerosis-associated interstitial lung disease: a systematic review. ERJ Open Res. 2021 Jan;7(1).

38. Maher TM, Molina-Molina M, Russell A-M, Bonella F, Jouneau S, Ripamonti E, et al. Unmet needs in the treatment of idiopathic pulmonary fibrosis—insights from patient chart review in five European countries. Bmc Pulmonary Medicine. 2017 2017/09/15;17(1):124.

39. Alfaro TM, Wijsenbeek MS, Powell P, Stolz D, Hurst JR, Kreuter M, et al. Educational aspects of rare and orphan lung diseases. Respiratory Research. 2021 2021/03/24;22(1):92.

40. van Harten WH, Goedbloed N, Boekhout AH, Heintzbergen S. Implementing large scale fast track diagnostics in a comprehensive cancer center, pre- and post-measurement data. BMC Health Services Research. 2018 2018/02/07;18(1):85.

SUPPLEMENTARY MATERIAL 1

Survey: Uncovering insights into the pulmonary fibrosis patient journey

Introductory questions

1. All of the information you have supplied in the survey is anonymous and your personal data will not be collected. Please confirm if you are happy to receive a report of the survey including key findings and top tips shared by other patients.
    a. I accept
    b. I don't accept

2. Where are you from/which healthcare system did you go through?
    a. UK
    b. Ireland
    c. France
    d. Spain
    e. Italy
    f. Austria
    g. Belgium
    h. Bulgaria
    i. Poland
    j. The Netherlands
    k. Hungary
    l. Norway
    m. Germany
    n. Other

3. What is your diagnosis?
    a. Idiopathic pulmonary fibrosis
    b. Chronic hypersensitivity pneumonitis
    c. Autoimmune-related pulmonary fibrosis (e.g. scleroderma or rheumatoid related)
    d. Sarcoidosis
    e. Non classified
    f. Other

At home

1. When thinking about your journey to a diagnosis of pulmonary fibrosis, which symptoms were you experiencing before making an initial doctor's appointment? Tick as many as apply.
    a. Dry cough
    b. Shortness of breath
    c. Acid reflux
    d. Feeling more tired than usual
    e. Loss of appetite/weight loss
    f. Rounded/swollen fingertips
    g. Aching muscles/joints
    h. Other

2. Did you visit your pharmacist prior to visiting your doctor?
    a. Yes
    b. No

3. How long did you wait from the onset of your symptoms to making an appointment with your doctor?
    a. Less than 1 week
    b. Less than 1 month
    c. 1 to 3 months
    d. 3 to 6 months
    e. 6 months to 1 year
    f. 1 year to 2 years
    g. Over 2 years
    h. Don't know/can't remember

4. If you delayed going to see your doctor, why? Tick as many as apply.
    a. You thought your symptoms were related to your age
    b. You didn't want to be a burden to your doctor as they are busy
    c. Your symptoms didn't cause you concern
    d. You were worried what the doctor might say
    e. You did not delay
    f. Other

5. What prompted you to finally go? Tick as many as apply.
    a. You were worried about your cough
    b. You were worried about your shortness of breath
    c. You were worried about feeling more tired than usual
    d. Your loss of appetite/weight loss
    e. Your painful/aching muscles/joints
    f. A family member/friend suggested you go
    g. Your symptoms were impacting on your daily activities (at home or work; please give details below after clicking 'OK')
    h. You don't know/can't remember
    i. Other

6. Your symptoms were impacting on your daily activities (at home or work; please give details here).
7. What were your feelings at this stage of your journey?
8. Is there anything that could have helped you at this stage?

General practitioner/primary care practice

1. What happened during your first appointment at your general practitioner/primary care practice, and what was the approach recommended by your doctor? Tick as many as apply.
    a. I was examined but no action was taken, and was told to come back if my symptoms continued or worsened
    b. The doctor performed some tests, but I stayed within their care
    c. I was diagnosed with and treated for another condition, e.g. asthma, chronic obstructive pulmonary disease, heart failure
    d. I was referred to a pulmonologist in a hospital
    e. I was referred to another specialist in a hospital (please provide further details after clicking 'OK')
    f. I don't know/can't remember
    g. Other

2. I was referred to another specialist in a hospital (please provide further details below).
3. How many times did you see your general practitioner before you were referred to a hospital?
    a. Once
    b. Two to three times
    c. Four to five times
    d. Over six times
    e. Don't know/can't remember

5. How long did it take until you saw a specialist at a hospital from when your doctor referred you?
    a. Less than 1 week
    b. Less than 1 month
    c. 1 to 3 months
    d. 3 to 6 months
    e. 6 months to 1 year
    f. 1 year to 2 years
    g. Over 2 years
    h. Don't know/can't remember

7. What were your feelings at this stage of your journey?
8. What worked well or what could have helped you?

Hospital setting

1. What investigations/tests did you undergo at the hospital? Tick as many as apply.
    a. Lung function test (spirometry)
    b. Blood tests
    c. Chest x-ray
    d. Computed tomography (CT) scan/lung imaging
    e. Lung biopsy
    f. Lung wash (lavage)
    g. 6-minute walk test
    h. Don't know/can't remember
    i. Other

2. Were you told why these different tests were needed?
    a. Yes
    b. No

3. Do you feel you were given sufficient information?
    a. Yes
    b. No

4. What information would have been helpful?
5. In addition to your respiratory doctor, which other healthcare professionals did you see at the hospital? Tick as many as apply.
    a. Respiratory nurse
    b. Cardiologist
    c. Radiologist
    d. Don't know/can't remember
    e. Other

6. Which of the following correctly describes the outcome of the tests at the hospital?
    a. Diagnosed with pulmonary fibrosis at this hospital
    b. Referred to another hospital that specializes in pulmonary fibrosis
    c. Don't know/can't remember

7. What information did you receive at this stage in your journey?
8. What additional information do you wish you had been given?

9. How long did it take for you to be given a confirmed diagnosis of pulmonary fibrosis from your first appointment at a hospital?
    a. Less than 1 week
    b. Less than 1 month
    c. One to 3 months
    d. 3 to 6 months
    e. 6 months to 1 year
    f. 1 year to 2 years
    g. Over 2 years
    h. Don't know/can't remember

10. What information were you provided with at diagnosis, and after?
    a. Explanation by the doctor and/or specialist nurse during consultation
    b. Printed educational materials
    c. Educational materials to help explain my diagnosis to friends/family
    d. Support group recommendation
    e. Website recommendation
    f. Don't know/can't remember
    g. Other

11. What did you find especially helpful?
12. What could have been helpful for you and your carers?
13. When thinking about your entire journey, from first symptoms to your diagnosis, how long did it take?
    a. Less than 1 week
    b. Less than 1 month
    c. One to 3 months
    d. Three to 6 months
    e. 6 months to 1 year
    f. 1 year to 2 years
    g. 2 years to 5 years
    h. Over 5 years
    i. Don't know/can't remember

To help other patients with pulmonary fibrosis

1. What piece of advice would you give to patients navigating the route to diagnosis in future?
2. What top tips could you provide on adjusting your lifestyle to live with pulmonary fibrosis?

Thank you for taking the time to complete this survey. Your input and insights are extremely valuable. Watch out for a future email in which we will circulate key findings and share the top tips provided by everyone who has taken part.

SUPPLEMENTARY MATERIAL 2

Literature search

Embase.com

('delayed diagnosis'/de OR (((delay* OR time-to) NEAR/3 diagnos*)):ab,ti) AND ('interstitial lung disease'/exp OR 'lung fibrosis'/exp OR 'lung sarcoidosis'/de OR ((interstitial* NEAR/3 (lung OR pulmonary*) NEAR/3 disease*) OR ((eosinophil* OR interstitial*) NEAR/3 pneumon*) OR (idiopathic* NEAR/3 (lung OR pulmonary*) NEAR/3 fibros*) OR (fibros* NEAR/3 alveolit*) OR (ANCA NEAR/3 vasculitide*) OR (Wegener* NEAR/3 granulomato*) OR ((lung OR pulmonary*) NEAR/3 sarcoidos*)):ab,ti) AND [english]/lim NOT ([animals]/lim NOT [humans]/lim)

Medline ALL Ovid

(Delayed Diagnosis / OR (((delay* OR time-to) ADJ3 diagnos*)).ab,ti.) AND (exp Lung Diseases, Interstitial/ OR Pulmonary Fibrosis/ OR Sarcoidosis, Pulmonary/ OR ((interstitial* ADJ3 (lung OR pulmonary*) ADJ3 disease*) OR ((eosinophil* OR interstitial*) ADJ3 pneumon*) OR (idiopathic* ADJ3 (lung OR pulmonary*) ADJ3 fibros*) OR (fibros* ADJ3 alveolit*) OR (ANCA ADJ3 vasculitide*) OR (Wegener* ADJ3 granulomato*) OR ((lung OR pulmonary*) ADJ3 sarcoidos*)).ab,ti.) AND english.la. NOT (exp animals/ NOT humans/)

Web of science (Science Citation Index Expanded & Social Sciences Citation Index)

TS=(((((delay* OR time-to) NEAR/2 diagnos*))) AND (((interstitial* NEAR/2 (lung OR pulmonary*) NEAR/2 disease*) OR ((eosinophil* OR interstitial*) NEAR/2 pneumon*) OR (idiopathic* NEAR/2 (lung OR pulmonary*) NEAR/2 fibros*) OR (fibros* NEAR/2 alveolit*) OR (ANCA NEAR/2 vasculitide*) OR (Wegener* NEAR/2 granulomato*) OR ((lung OR pulmonary*) NEAR/2 sarcoidos*)))) AND LA=(english)

Cochrane CENTRAL register of trials

((((delay* OR time next to) NEAR/3 diagnos*)):ab,ti) AND (((interstitial* NEAR/3 (lung OR pulmonary*) NEAR/3 disease*) OR ((eosinophil* OR interstitial*) NEAR/3 pneumon*) OR (idiopathic* NEAR/3 (lung OR pulmonary*) NEAR/3 fibros*) OR (fibros* NEAR/3 alveolit*) OR (ANCA NEAR/3 vasculitide*) OR (Wegener* NEAR/3 granulomato*) OR ((lung OR pulmonary*) NEAR/3 sarcoidos*)):ab,ti)

Google Scholar

"delayed diagnosis"|"time to diagnosis"|"diagnostic delay" "interstitial|idiopathic lung|pulmonary disease|fibrosis"|"eosinophilic|interstitial pneumonia"|"fibrosing alveolitis"|"lung|pulmonary sarcoidosis"

Result

	Number of papers	Number after deduplication
Embase.com	764	753
Medline ALL Ovid	717	524
*Web of science**	176	34
Cochrane CENTRAL register	4	0
Google Scholar	200	128
Total	**1861**	**1439**

**(Science Citation Index Expanded & Social Sciences Citation Index)*

SUPPLEMENTARY MATERIAL 3

Reported feelings

Feelings in period after onset of first symptoms (not displayed in Figure 4)

Coded described feeling	Count
Worried	20
Concerned	18
Fear	11
Uncertainty	7
Frustration	5
Discouraged	4
Distress	4
Surprised	3
Terrible	2
Disbelief	2
Panic	2
Anxiety	2
Unsure	2
Scared	2
Sadness	2
Helpless	2
Bad	1
Uneasy	1
Angry	1
Bewildered	1
Unwell	1
Frightened	1
Shocked	1
Annoyed	1
Discomfort	1
Guilty	1
Unsettled	1
Healthy	1
Dissatisfied	1
Heart problems	1
Despair	1
Anger	1
Disappointed	1
Incredulity	1
Troublesome	1
Innovative	1
Uncomfortable	1
Misunderstood	1
Unsatisfied	1
Not right	1
Apprehension	1
Confident	1
Weak	1
Perfect	1
Powerless	1
Confused	1
Grand Total	***118***

Feelings in period after referral to hospital (displayed in Figure 4)

Coded described feeling	Count
Worried	47
Concerned	21
Anxiety	10
Fear	9
Uncertainty	9
Nervous	4
Good	4
Confused	4
Fine	4
Disappointed	4
Helpless	4
Annoyed	3
Bad	3
Afraid	3
Distressed	3
Frustrated	3
Not right	2
Inquisitiveness	2
Frightened	2
Terrible	2
Misunderstood	2
Alone	2
Scared	2
Disbelief	2
Curious	2
Uneasy	2
Collapsed world	2
Hopefull	2
Vulnerable	1
Discomfort	1
Bored	1
Great	1
Frustration	1
Healthy	1
Resigned	1
Doubts	1
Stressful	1
Helplessness	1
Trust	1
Hope	1
Unsure	1
Hopeful	1
Doubt	1
Dread	1
Sadness	1
Desperate	1
Shocked	1
Exhausted	1
Surprised	1
Expectation	1
Tired	1
Mixed feelings	1
Devastated	1
Angry	1
Unsettled	1
No understanding	1
Upset	1
Devestated	1
Well	1
Overwhelmed	1
Relieved	1
Incredulous	1
Grand Total	***193***

SUPPLEMENTARY MATERIAL 4

Recommendations and experiences

Supplementary Table 1: Examples of recommendations from patients to healthcare providers (quotes from survey responses).

Recommendations for diagnostic pathway			**Positive experiences**
Before first GP visit	**After referral to hospital**	**At time of diagnosis**	**During diagnostic pathway**
Get an earlier appointment	Earlier referral to specialist center	More extensive information about what to expect	Taking rest
Earlier referral to hospital	Earlier appointment and testing	Offering psychological help	Receiving treatment with antibiotics, steroids, oxygen or antifibrotics
Getting the correct diagnosis	Earlier diagnosis	Making aware of patient associations and support groups	Support or concern from partner and family
Earlier start of treatment	Earlier start of (palliative) treatment	Advice regarding lifestyle	Adequate doctor: quick referral, information and discussion
More extensive examination	Getting more information from doctor	Advice on how to manage symptoms	Staying calm and positive
Getting more information and answers from doctor	More explanation by and discussion with doctor		Change specialist
Doctor with more knowledge about PF	Doctor with more knowledge about PF		Exercise
More knowledge about PF in general population	IPF training for general practitioners		Diagnostic testing during hospitalization
Better coordination between doctors	More knowledge about PF in general population		Information from doctor
Lifestyle advice from doctors (e.g. Lose weight, quit smoking)	Psychological help		Adequate referral to specialist center
	More structured approach to diagnosis		Fast testing
	Semi-annual lung function test		Participation in a clinical trial
	Concerned doctor		

Supplementary Table 2: Examples of recommendations from patients to future patients (quotes from survey responses).

Advice to future patients	
During diagnostic pathway	**Lifestyle after PF diagnosis**
Make an appointment with a GP sooner	Keep moving
Take enough rest	Have a healthy diet
Look for psychological support (professional psychologist or peer groups)	Look for (psychologic) support
Look for psychological support for family	Stay positive
Make early/immediate appointment when (persisting) symptoms of cough or shortness of breath	Continue life as normal as possible
Feel comfortable with your doctor, don't hesitate asking for a second opinion	Adjust your pace when listening to your body
Consult other doctor if you don't feel comfortable or taken seriously	Accept your limitations, but don't lock yourself at home
Ask questions and explanation about the tests, disease and therapy	Continue social activities
Start treatment (drugs, oxygen or transplant) as soon as possible	Find a balance with enough rest, but keep doing physical exercise as much as possible
Ask about ongoing clinical trials and new medication.	Do breathing exercises
Look up information yourself only in reliable sources	Do not overexert
Insist on further testing and/or referral	Find new hobbies
Go to specialised pulmonologist and hospital for adequate diagnosis	Quit smoking
Look for support groups and patient associations	Find support with fellow patients
Keep exercising	Look for psychological help
Join physiotherapy and/or rehabilitation courses	
Live healthy, eat good food	
Listen to the doctor's advice	
Look for professional psychological support	
Look for good people around you	
Distraction	

PART II

ENOSE FOR DIAGNOSING

CHAPTER 4

Differentiating interstitial lung diseases from other respiratory diseases using electronic nose

Respir Res. 2023 Nov 6;24(1):271.

Iris G. van der Sar[1], Marlies S. Wijsenbeek[1], Gert-Jan Braunstahl[1,2], Jason O. Loekabino[1], Anne-Marie C. Dingemans[1], Johannes C.C.M. in 't Veen[1,2], Catharina C. Moor[1]

[1]Erasmus Medical Center, Rotterdam (the Netherlands); [2]Franciscus Gasthuis & Vlietland, Rotterdam (the Netherlands)

Abstract

Introduction

Interstitial lung disease (ILD) may be difficult to distinguish from other respiratory diseases due to overlapping clinical presentation. Recognition of ILD is often late, causing delay which has been associated with worse clinical outcome. Electronic nose (eNose) sensor technology profiles volatile organic compounds in exhaled breath and has potential to detect ILD non-invasively. We assessed the accuracy of differentiating breath profiles of patients with ILD from patients with asthma, chronic obstructive pulmonary disease (COPD), and lung cancer using eNose technology.

Methods

Patients with ILD, asthma, COPD, and lung cancer, regardless of stage or treatment, were included in a cross-sectional study in two hospitals. Exhaled breath was analysed using an eNose (SpiroNose) and clinical data were collected. Datasets were split in training and test sets for independent validation of the model. Data were analyzed with partial least squares discriminant and receiver operating characteristic analyses.

Results

161 patients with ILD and 161 patients with asthma (n=65), COPD (n=50) or lung cancer (n=46) were included. Breath profiles of patients with ILD differed from all other diseases with an area under the curve (AUC) of 0.99 (95%CI 0.97-1.00) in the test set. Moreover, breath profiles of patients with ILD could be accurately distinguished from the individual diseases with an AUC of 1.00 (95%CI 1.00-1.00) for asthma, AUC of 0.96 (95%CI 0.90-1.00) for COPD, and AUC of 0.98 (95%CI 0.94-1.00) for lung cancer in test sets. Results were similar after excluding patients who never smoked.

Conclusions

Exhaled breath of patients with ILD can be distinghuished accurately from patients with other respiratory diseases using eNose technology. eNose has high potential as an easily accessible point-of-care medical test for identification of ILD amongst patients with respiratory symptoms, and could possibly facilitate earlier referral and diagnosis of patients suspected of ILD.

Keywords

Breath test, diagnostic test, biomarker, electronic nose, interstitial lung diseases, obstructive lung disease, lung cancer

Background

Worldwide, over 500 million people suffer from a respiratory disease and numbers are increasing, including numbers of patients with interstitial lung disease (ILD). However, ILDs still remain rare diagnoses. The overal global prevalence of ILD is approximately 0.09% [1]. Due to the lack of knowledge on ILD and the non-specific symptoms, recognizing patients suspected for ILD is poor amongst primary care physicians and community hospitals [2, 3]. Besides aspecific disease presentation, various patient and healthcare related factors play a role [4]. Moreover, lung function is often still preserved in early ILD. A median delay of up to 2,1 years from start of symptoms until diagnosis has been reported and has been associated with worse outcomes [5, 2, 6]. Therefore, a non-invasive, less costly, accessible and reliable test to improve the diagnostic process is highly needed [7].

An electronic nose (eNose) device is a sensor-based technique that detects and profiles volatile organic compounds of exhaled breath non-invasively, without identification of the individual compounds. Both physiological and pathophysiological processes in the human body influence the volatile organic compounds; thus, exhaled breath provides valuable information about a person's health.

Previous studies found that eNose technology can be used to accurately identify respiratory diseases, including ILD, lung cancer, asthma and chronic obstructive pulmonary disease (COPD) [8, 9]. In ILD, breath profiles of patients could be differentiated from healthy controls [10-14] and individual ILDs from COPD [11, 12]. Exploratory studies in pneumoconiosis show the potential of using an eNose for screening purposes in ILD [15, 16].

The aim of the current study is to investigate whether exhaled breath analysis using an eNose has potential as application for early detection of ILD amongst patients with respiratory symptoms. We assessed the accuracy of differentiating breath profiles of patients with ILD from patients with asthma, COPD, and lung cancer.

Methods

Study design

In this cross-sectional multicenter study patients were included at the outpatient clinic of the department of respiratory medicine of two hospitals in Rotterdam, the Netherlands: Erasmus University Medical Center (recognized expert center for ILD and lung cancer) and Franciscus Gasthuis & Vlietland (recognized expert center for asthma and COPD).

Patients with a diagnosis of ILD, asthma, COPD or lung cancer, regardless of stage or treatment were included in both hospitals between January 2019 and December 2022. ILD diagnosis was established by a multidisciplinary team according to the most recent guidelines [17-19]. At time of diagnosis, patients were diagnosed and classified for asthma following the applicable Global Initiative for Asthma guidelines [20], and for COPD following the Global initiative for chronic obstructive lung disease (GOLD) guidelines [21]. All patients with lung cancer had a pathology proven diagnosis. Patients with another lung disease, lung carcinoma in situ, current pulmonary infection or recent alcohol intake (< 8 hours) were excluded.

Data collection

The eNose used for exhaled breath analysis was the SpiroNose (Breathomix, Leiden, the Netherlands). This eNose contains seven different metaloxide semiconductor sensors in various arrays on both the inside and outside of the device [22, 23]. Each included patient performed one measurement that consisted of two breath maneuvers. One maneuver comprises five tidal breaths, an inhalation to total lung capacity, followed by a 5 s breath hold and a slow maximum expiration. Data were collected in an online platform that has a secured certified database (BreathBase). More details about the breath maneuver and breath data collection were described previously [23].

Participants completed a short questionnaire, including demographics, smoking history, and recent medication, food or drink intake. Other patient characteristics, medical history, medication use, and most recent available diagnostic test results (e.g., spirometry, chest imaging, pathologic assessment, blood samples) were collected from medical files.

Data analysis

Pre-processing

Sensor data was extracted from the BreathBase platform and pre-processed before analysis. Pre-processing includes selection of the best breath maneuver, data correction for ambient air, data scaling to the most stable sensor, and reduction of inter-array differences [22, 23]. For each sensor, the peak value and the ratio between peak value and breath hold are used for statistical analyses. The peak value of the most stable sensor is excluded, resulting in 13 values per measurement (i.e. the breath profile) labeled with the collected patient characteristics. Measurements of insufficient quality caused for example by wrong breathing technique or unstable ambient air are removed.

Dataset and analysis groups

To answer the main aim of the study, breath profiles of patients with ILD were compared to the whole group of patients with another respiratory diagnosis (asthma, COPD or lung cancer). The four diagnosis groups were also compared separately. Moreover, patients with lung diseases that often have similar patient characteristics and risk factors (idiopathic pulmonary fibrosis (IPF), COPD, lung cancer) were compared. A thorough power calculation was not possible, as data from previous similar studies were not available. We aimed to include enough patients in each diagnosis group to be able to split the groups in a training and test set in order to independently validate the results. Looking at eNose studies, a dataset size of ≥30 patients is generally sufficient to split [8]. To avoid imbalance between groups and reduction of statistical power of the model, larger groups were reduced by random patient selection using the function 'sample' in R [24].

To assess the influence of smoking on the accuracy of findings, comparison of breath profiles from patients with ILD versus all other diseases was repeated in patients who ever smoked. Moreover, the possible influence of medical center was assessed by comparing breath data of patients with asthma and COPD who were included in Erasmus Medical Center versus Franciscus Gasthuis&Vlietland. Lastly, breath data of all patients were compared based on their sex (males versus females) or smoking history (ever versus never, current versus former smokers) to test for the influence of these potential confounders. Descriptive statistics were used to analyze baseline data, including $\chi 2$, Student's t, and Mann Whitney tests to compare groups. We displayed normally distributed data as mean values (± standard deviation) and non-normally distributed data as median values (interquartile range). R version 4.2.1 for Windows with mixOmics version 6.20.0 package was used for analysis.

Data classification

The supervised classification technique partial least squares discriminant analysis (PLS-DA) was used to reduce dimensionality of breath profiles, and to classify and compare groups. Dimensionality reduction resulted in multiple principal components (PCs), which are weighted combinations of input variables (i.e. sensor values). If data was split in a training and test set, the first two PCs were used to assess the discriminative ability of eNose technology. If a dataset was not split, one PC was used to avoid model overfitting. Receiver operating characteristics analysis was applied to calculate the corresponding area under the curve (AUC) accuracy with 95% confidence interval (CI), sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV).

The presence of outliers in the values of PC1 and 2 were assessed. Outliers were defined as measurements outside upper and lower limits of a box-and-whisker plot. Limits were calculated as quartile 1 and 3 +/- 1.5 * interquartile range. The main analysis was repeated without outliers to assess influence of outliers on the main results.

Probability score prediction of individual patients

To show the eNose performance in clinical practice based on our trained PLS-DA model, one hundred patients with ILD were randomly selected from the part of our dataset that was left out from previous analyses and not used for training or testing the model (i.e. unseen data). For each patient an individual probability score was calculated (range 0-1) using the 'predict' function in R. This function predicts how well the new patient data fit the average ILD breath profile that resulted from the trained PLS-DA model (i.e. PC1 and PC2). The higher the individual probability score, the better the breath profile of the patient fits the ILD breath profile. A density plot (i.e. relative likelihood against probability score*100%) was created to display the distribution of probability scores for all one hundred unseen dataset of patients with ILD.

Results

Baseline characteristics

322 patients were included in this study; 161 patients with ILD were selected (from a total cohort of n=349) to compare to 161 patients with other respiratory diseases (65 with asthma, 50 with COPD, and 46 with lung cancer). For comparing ILD with individual diagnoses, a subset of 55 randomly selected patients with ILD was used. Baseline characteristics of the overall cohort and individual diagnosis groups are shown in **Table 1**. An overview of the selected patient cohorts for the main analyses is shown in a flowchart (**Figure 1**).

Main results

Breath profiles of patients with ILD differed from all other respiratory diseases with an AUC of 0.97 (95%CI 0.95-0.99) in the training and 0.99 (95%CI 0.97-1.00) in the test set (**Figure 2A**). Comparison of ILD with asthma (AUC 1.00, 95%CI 1.00-1.00), with COPD (AUC 0.96, 95%CI 0.90-1.00) and with lung cancer (AUC 0.98, 95%CI 0.94-1.00) showed similar results in the test sets. Additionally, breath profiles of patients with COPD and lung cancer (AUC 0.97, 95%CI 0.90-1.00) and COPD and asthma (AUC 0.90, 95%CI 0.79-1.00) could be distinguished with high accuracies. A scatter plot in **Figure 2B** visualizes how breath profiles of all individual disease groups relate to each other.

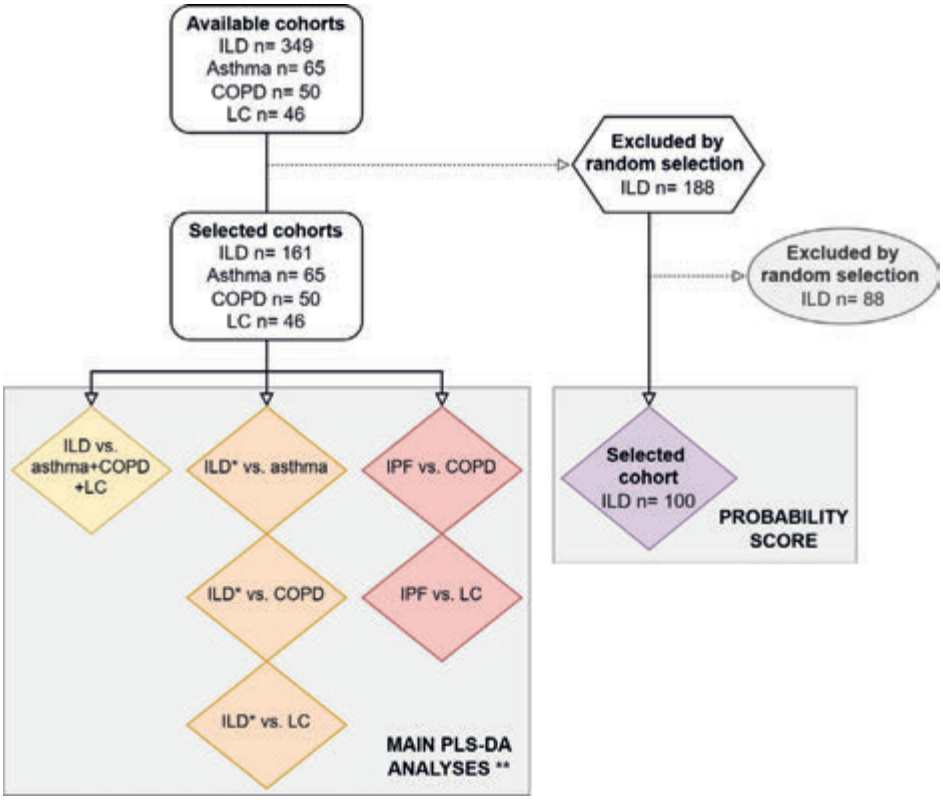


Figure 1: Flowchart of cohort selection for main analyses.
*Subgroup analyses are not included in this flowchart. IPF cohort existed of n=61 patients. *Subgroups reduced to size n=55 by random selection. **If group size ≥30, cohorts were split in a training and test set.*

Figure 2C shows the distribution of breath profiles of patients with IPF, COPD, and lung cancer. Comparing IPF with COPD resulted in an AUC of 0.93 (95% CI 0.86-1.00), and IPF with lung cancer in an AUC of 0.93 (95% CI 0.82-1.00) in the test sets. Corresponding specificity, sensitivity, accuracy, NPV and PPV of all group comparisons can be found in **Table 2**.

There were 25 outliers in the dataset. The outliers had no significant effect on the main results (see **Figure S1** and **Table S1** in **Additional file 1**).

Table 1: Baseline characteristics.

	Overall	ILD	Asthma	COPD	Lung Cancer	p-value
Subjects (n)	322	161	65	50	46	
Females (n)	154 (47.8)	60 (37.3)	45 (69.2)	23 (46.0)	26 (56.5)	<0.01
Age (years)	68.0 [58.0, 75.0]	71 [62, 76]	56 [42, 67]	66 [61, 74]	69 [63, 75]	<0.01
Smoking amount ~ (py)	32.6 (31.5)*	25.8 (23.3)	16.3 (15.9)	49.6 (38.7)	41.9 (38.4)	<0.01
Smoking status						<0.01
Never	96 (30.7)	48 (30.4)	37 (56.9)	0 (0.0)	13 (28.3)	
Former	193 (59.9)	110 (68.3)	23 (35.4)	31 (66.0)	27 (58.7)	
Current	30 (9.3)	2 (1.2)	5 (7.7)	17 (34.0)	6 (13.0)	
FVC (%pred)	84.7 (20.6)**	80.7 (20.8)	93.2 (16.7)	84.2 (20.4)	94.2 (22.1)	<0.01
FEV1 (%pred)	77.3 (22.4)**	81.7 (19.0)	81.6 (21.0)	54.5 (21.3)	85.3 (21.1)	<0.01
DLCOc (%pred)		51.0 (16.2)#				
Diagnosis or Stage (n)		IPF 61 (37.9) HP 27 (16.8) CTD-ILD 27 (16.8) iNSIP 11 (6.8) CPFE 7 (4.3) COP 6 (3.7) Other ILD 22 (13.7)		GOLD I 16 (32.0) GOLD II 20 (40.0) GOLD III 7 (14.0) GOLD IV 7 (14.0)	SCLC 4 (8.7) NSCLC 42 (91.3) ---------- Stage I 2 (4.3) Stage II 0 (0.0) Stage III 5 (10.9) Stage IV 39 (84.8)	
Eosinophil count (10^9/L)			0.2 [0.1, 0.4]##			
Use of immune-suppressants (n)		55 (34.2)^	7 (10.8)	4 (8.0)	9 (19.6)	
Use of other disease-specific medication (n)		Antifibrotic 44 (27.3)	Biological 14 (21.5) ICS 59 (90.8)	ICS 30 (50.0)	Targeted 30 (65.2) CT and/or IT 10 (21.7)	

*Values are displayed as number (%), mean ±SD, or median [interquartile range]. Subgroup 'other ILD' includes interstitial pneumonia with auto-immune features, desquamative interstitial pneumonia, vasculitis, unclassifiable ILD, asbestosis, respiratory bronchiolitis-ILD, drug induced ILD, sarcoidosis, granulomatous-lymphocytic ILD. If available, lung function values post-bronchodilator are displayed. CPFE = combined pulmonary fibrosis and emphysema; CT = chemotherapy; CTD = connective tissue disease; DLCOc = diffusing capacity for carbon monoxide corrected for hemoglobin level; FEV1 = forced expiratory volume in the first second; FVC = forced vital capacity; GOLD = Global Initiative for Chronic Obstructive Lung Disease; HP = hypersensitivity pneumonitis; ICS = inhaled corticosteroid; ILD = interstitial lung disease; iNSIP = idiopathic non-specific interstitial pneumonia; IPF = idiopathic pulmonary fibrosis; IT = immunotherapy; NSCLC = non-small cell lung cancer; py = pack years; %pred = percent of predicted value, calculated based on sex, age and height. ~never smokers (n=99) excluded. * n=7 missing values. ** n=39 missing values. # n=12 missing values. ## n=7 missing values. ^In case of prednisone: dosage ≥10 mg.*

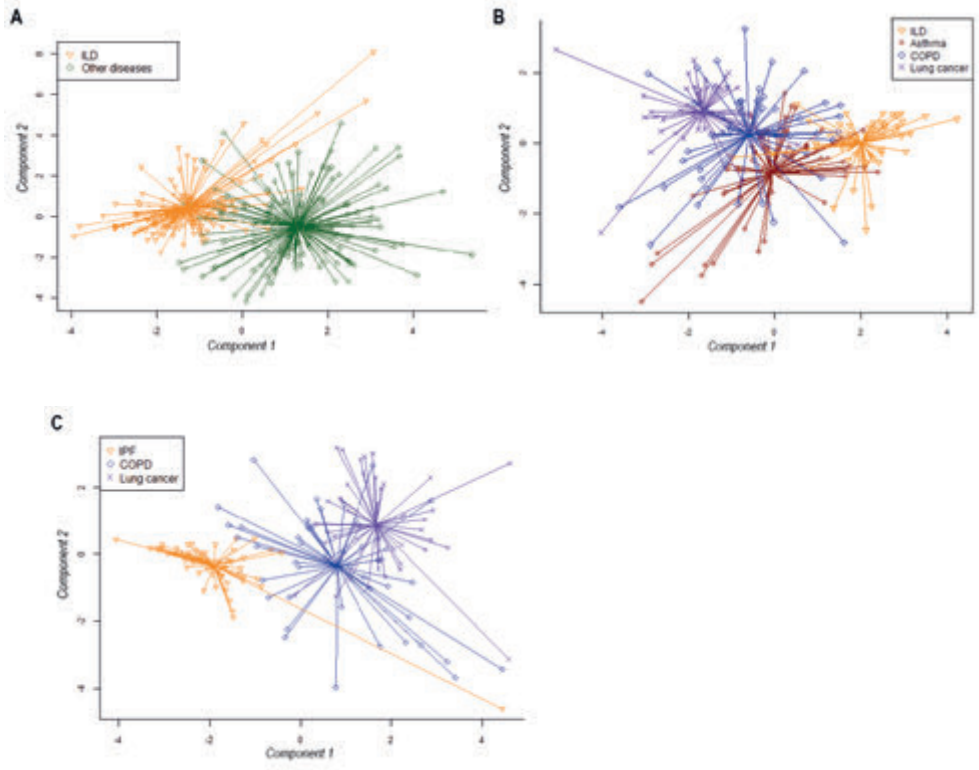


Figure 2: Comparison of breath profiles between patients with ILD and other respiratory diseases.
A. Scatterplot of patients with ILD (n=161) versus other respiratory diagnoses (i.e. asthma, COPD, and lung cancer; n=161). B. Scatterplot of patients with ILD (n=55) versus asthma (n=65) versus COPD (n=50) versus lung cancer (n=46). C. Scatterplot of patients with IPF (n=61) versus COPD (n=50) versus lung cancer (n=46). Each dot represents one patient. Component 1 and 2 are principal components resulting from partial least squares analysis. COPD = chronic obstructive pulmonary disease; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis.

Table 2: Results of breath analysis between patient groups.

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
ILD	108	**Asthma – COPD – Lung Cancer**	108	Training	0.97	*0.95-0.99*	0.93	0.93	0.93	0.93	0.93
	53		53	Test	0.99	*0.97-1.00*	0.89	1.00	0.94	1.00	0.90
ILD	37	**Asthma**	44	Training	0.99	*0.97-1.00*	0.91	1.00	0.95	1.00	0.90
	18		21	Test	1.00	*1.00-1.00*	1.00	1.00	1.00	1.00	1.00
ILD	37	**COPD**	34	Training	0.97	*0.97-1.00*	1.00	0.86	0.93	0.87	1.00
	18		16	Test	0.96	*0.90-1.00*	0.94	0.89	0.91	0.88	0.91
ILD	37	**Lung Cancer**	31	Training	1.00	*1.00-1.00*	1.00	1.00	1.00	1.00	1.00
	18		15	Test	0.98	*0.94-1.00*	0.89	1.00	0.94	1.00	0.88
COPD	34	**Lung Cancer**	31	Training	0.88	*0.79-0.97*	0.88	0.87	0.88	0.88	0.87
	16		15	Test	0.97	*0.90-1.00*	1.00	0.93	0.97	0.94	1.00
COPD	34	**Asthma**	44	Training	0.92	*0.85-0.98*	0.95	0.76	0.87	0.84	0.93
	16		21	Test	0.90	*0.79-1.00*	0.86	0.88	0.86	0.90	0.82
IPF	41	**COPD**	34	Training	0.88	*0.80-0.96*	0.71	0.98	0.85	0.96	0.80
	20		16	Test	0.93	*0.86-1.00*	0.75	0.95	0.86	0.92	0.83
IPF	41	**Lung Cancer**	31	Training	0.91	*0.85-0.98*	0.98	0.68	0.85	0.80	0.95
	20		15	Test	0.93	*0.82-1.00*	1.00	0.87	0.94	0.91	1.00

Results based on 2 principal components. AUC = area under the curve; CI = confidence interval; COPD = chronic obstructive pulmonary disease; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis; NPV = negative predictive value; PPV = positive predictive value.

Predicted probability scores of individual patients

To illustrate how eNose might perform in future clinical practice, the probability of having an ILD was predicted based on eNose breath data of one hundred patients previously diagnosed with ILD. For example, a predicted probability of 88% means that the breath profile of this individual patient fits for 88% with the ILD breath profile. This might help physicians in clinical decision making. **Figure 3** shows the distribution of all individual probability scores in a density plot.

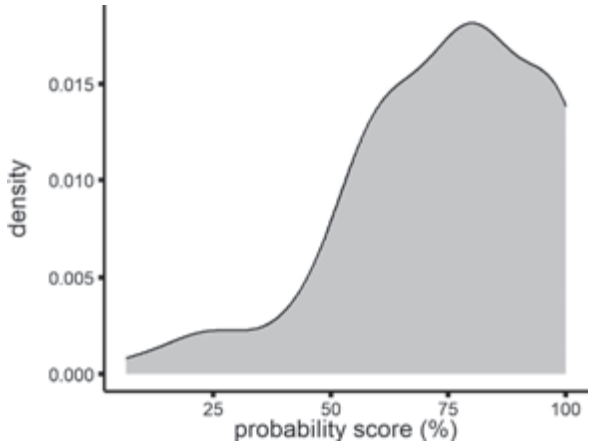


Figure 3: Predicted probability scores of individual patients with ILD based on breath data
Density plot shows the distribution of the predicted individual probability scores of a random sample of 100 unseen dataset of patients with ILD. Probability score is based on their breath profile and the trained PLS-DA model. The density (i.e. relative likelihood) is displayed on the y-axis and the individual probability scores on the x-axis.

Subgroup results

An additional analysis on the influence of smoking is displayed in **Table 3**. eNose technology perfomed equally in the subgroup of ever smokers compared to the results of the full cohort. Moreover, breath profiles were not influenced by sex, or medical center. Current smokers seem to have slightly different breath profiles than former smokers.

Table 3: Results of breath analysis in subgroups.

Group 1		Group 2	n=	Data-set	AUC	95% CI	Speci-ficity	Sensi-tivity	Accu-racy	NPV	PPV
ILD (ever smoking)	75	**Asthma – COPD – Lung Cancer (ever smoking)**	72	Training	0.99	*0.99-1.00*	0.99	0.96	0.97	0.96	0.99
	37		36	Test	0.94	*0.89-0.99*	0.86	0.95	0.90	0.94	0.88
Never smoker	96	**Ever smoker**	223		0.66	*0.60-0.73*					
Current smoker	30	**Former smoker**	193		0.80	*0.73-0.87*					
Male sex	168	**Female sex**	154		0.67	*0.61-0.73*					
Hospital EMC	254	**Hospital FGV**	73		0.64	*0.53-0.74*					

Results of the cohort that is split in a training and test set are based on 2 principal components; results of the unsplit cohort are based on 1 principal component. ^Includes asthma and chronic obstructive pulmonary disease patients only, as the number of patients with interstitial lung disease and lung cancer were too small in medical center FGV. AUC = area under the curve; CI = confidence interval; COPD = chronic obstructive pulmonary disease; EMC = Erasmus Medical Center; FGV = Franciscus Gasthuis&Vlietland; ILD = interstitial lung disease; NPV = negative predictive value; PPV = positive predictive value.

Discussion

Patients with ILD can be distinguished accurately from those with other respiratory diseases using eNose technology, shown in large training and test cohorts of patients with different disease stages and treatments. Moreover, the separation of breath profiles of patients with ILD compared to asthma, COPD or lung cancer individually was highly accurate, independent of age or sex. These results show the potential of using an eNose for detection of ILD non-invasively. If these findings are confirmed in a asymptomatic or early ILD patient cohort, screening or early detection might be possible.

Our results align with previously published results on the performance of eNose technology in differentiating ILD from COPD [11, 12]. Dragonieri et al. compared IPF with COPD and found an AUC of 0.85 in a test cohort, with active smokers being excluded [12]. The study of Krauss et al. aimed to differentiate individual ILDs, but patients with COPD were included as a control group [11]. Comparing CTD-ILD versus COPD resulted in an AUC of 0.85, and cryptogenic organizing pneumonia versus COPD in an AUC of 0.77. Other ILDs were not reported. Moreover, only patients with COPD GOLD stage III-IV were included, and results were not validated in a test set. Although direct comparison of results is difficult as both studies used another eNose device and selected patients with specific ILD diagnoses, all published results emphasize the potential of the overall concept of eNose technology for ILD. To our knowledge, studies that compare ILD with lung cancer or asthma have not been published until date.

No studies have been published on early detection of ILD using an eNose, except for two studies that focus on pneumoconiosis screening in high risk groups [15, 16]. Although these were pilot studies, they found high accuracies when comparing people with and without pneumoconiosis. Recently, studies on lung cancer screening have become available. A prospective study in patients with COPD showed that patients that developed lung cancer had a different breath profile already two years before the diagnosis of lung cancer compared to patients that did not develop lung cancer [25]. Moreover, De Kort et al. published a validation study on the performance of eNose technology for lung cancer screening [26]. They included patients suspected of lung cancer prior to tissue biopsy. In this robust study, the presence of lung cancer could be predicted using an eNose with an AUC of 0.79 in the validation cohort. This performance increased to an AUC 0.86 when known clinical risk factors where added in the model. These studies illustrate the promise of incorporating eNose results in risk models for early detection of respiratory diseases.

Interestingly, in our study we also found an acucurate seperation between patientsin different clinically heterogenous subgroups with smoking-related diagnoses (IPF, COPD and lung cancer). The diagnostic workup of patients with unexplained respiratory symptoms and differentiation between various diagnoses is complex, especially in patients with similar clinical characteristics. Moreover, pulmonary function tests often do not show abnormalities in early disease. Thus, we believe that eNose technology could be of added value to raise early suspicion for ILD and improve referral and adequate diagnosis in both primary and secondary care.

Several limitations of our study should be named. First, we chose only one classification algorithm for data analysis. PLS-DA is an accepted method for classification of groups, but several methods should be compared in validation studies [27, 28]. Second, our study lacks an external validation cohort We minimized the risk for model overfitting by splitting our dataset in a separate training and test set, but an external cohort is necessary to confirm the model performance. Besides, in our study cohort, the prevalence of ILD is much higher than would be expected in a real-life cohort of patients with unexplained respiratory symptoms. In a real-world setting, negative predictive value for ILD would therefore likely be higher, and positive predictive values lower. Lastly, the included cohort might not be representative for the overall population for which a clinical test for early disease detection is most beneficial; i.e. the patients visiting a physician with new or unexplained respiratory symptoms. The majority of the study cohort consisted of prevalent patients, of whom many already used disease-modifying treatment and had advanced disease stage. However, eNose technology achieved high accuracies despite the cohort heterogeneity in terms of treatment, stage and disease severity, indicating

the suitability for application in real-world populations. Nevertheless, we should include patients with suspected and early respiratory diseases from primary health care centers and community sites in future multicenter external validation studies.

Conclusion

eNose technology can be used to distinguish patients with ILD from patients with other respiratory diseases. This technology has high potential as an easily accessible point-of-care medical test for accurate identification of patients with ILD, and could facilitate earlier diagnosis and referral of patients suspected of ILD.

References

1. Soriano JB, Kendrick PJ, Paulson KR, Gupta V, Abrams EM, Adedoyin RA, et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet Respiratory Medicine. 2020 2020/06/01/;8(6):585-96.
2. Hoyer N, Prior TS, Bendstrup E, Wilcke T, Shaker SB. Risk factors for diagnostic delay in idiopathic pulmonary fibrosis. Respir Res. 2019 May 24;20(1):103.
3. Pritchard D, Adegunsoye A, Lafond E, Pugashetti JV, DiGeronimo R, Boctor N, et al. Diagnostic test interpretation and referral delay in patients with interstitial lung disease. Respir Res. 2019 Nov 12;20(1):253.
4. Spagnolo P, Ryerson CJ, Putman R, Oldham J, Salisbury M, Sverzellati N, et al. Early diagnosis of fibrotic interstitial lung disease: challenges and opportunities. The Lancet Respiratory Medicine. 2021 2021/09/01/;9(9):1065-76.
5. Lamas DJ, Kawut SM, Bagiella E, Philip N, Arcasoy SM, Lederer DJ. Delayed access and survival in idiopathic pulmonary fibrosis: a cohort study. Am J Respir Crit Care Med. 2011 Oct 1;184(7):842-7.
6. van der Sar IG, Jones S, Clarke DL, Bonella F, Fourrier JM, Lewandowska K, et al. Patient Reported Experiences and Delays During the Diagnostic Pathway for Pulmonary Fibrosis: A Multinational European Survey. Front Med (Lausanne). 2021;8:711194.
7. Hoffman TW, van Es HW, Biesma DH, Grutters JC. Potential interstitial lung abnormalities on chest X-rays prior to symptoms of idiopathic pulmonary fibrosis. BMC Pulm Med. 2022 Aug 30;22(1):329.
8. van der Sar IG, Wijbenga N, Nakshbandi G, Aerts JGJV, Manintveld OC, Wijsenbeek MS, et al. The smell of lung disease: a review of the current status of electronic nose technology. Respiratory Research. 2021 2021/09/17;22(1):246.
9. van der Sar IG, Wijsenbeek MS, Moor CC. Exhaled breath analysis in interstitial lung disease. Current Opinion in Pulmonary Medicine. 2023.
10. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratú P, Resta O, et al. An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. Respiratory Medicine. 2013 2013/07/01/;107(7):1073-78.
11. Krauss E, Haberer J, Maurer O, Barreto G, Drakopanagiotakis F, Degen M, et al. Exploring the Ability of Electronic Nose Technology to Recognize Interstitial Lung Diseases (ILD) by Non-Invasive Breath Screening of Exhaled Volatile Compounds (VOC): A Pilot Study from the European IPF Registry (eurIPFreg) and Biobank. J Clin Med. 2019 Oct 16;8(10).
12. Dragonieri S, Scioscia G, Quaranta VN, Carratu P, Venuti MP, Falcone M, et al. Exhaled volatile organic compounds analysis by e-nose can detect idiopathic pulmonary fibrosis. J Breath Res. 2020 Jul 21;14(4):047101.
13. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts J, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).
14. van der Sar IG, Moor CC, Oppenheimer JC, Luijendijk ML, van Daele PLA, Maitland-van der Zee AH, et al. Diagnostic Performance of Electronic Nose Technology in Sarcoidosis. Chest. 2022 Mar;161(3):738-47.
15. Yang HY, Peng HY, Chang CJ, Chen PC. Diagnostic accuracy of breath tests for pneumoconiosis using an electronic nose. J Breath Res. 2017 Nov 29;12(1):016001.

16. Xuan W, Zheng L, Bunes BR, Crane N, Zhou F, Zang L. Engineering solutions to breath tests based on an e-nose system for silicosis screening and early detection in miners. Journal of Breath Research. 2022 2022/4;16(3):036001.

17. Travis WD, Costabel U, Hansell DM, King Jr TE, Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. American Journal of Respiratory and Critical Care Medicine. 2013;188(6):733-48.

18. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2018;198(5):e44-e68.

19. Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1;205(9):e18-e47.

20. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention. Available from: www.ginasthma.org

21. Global Initiative for Chronic Obstructive Lung Disease. Available from www.goldcopd.org, published in Deer Park, IL, USA.

22. de Vries R, Brinkman P, van der Schee MP, Fens N, Dijkers E, Bootsma SK, et al. Integration of electronic nose technology with spirometry: validation of a new approach for exhaled breath analysis. J Breath Res. 2015 Oct 15;9(4):046001.

23. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. Eur Respir J. 2018 Jan;51(1).

24. Kumar P, Bhatnagar R, Gaur K, Bhatnagar A. Classification of Imbalanced Data:Review of Methods and Applications. IOP Conference Series: Materials Science and Engineering. 2021 2021/03/01;1099(1):012077.

25. de Vries R, Farzan N, Fabius T, De Jongh FHC, Jak PMC, Haarman EG, et al. Prospective Detection of Early Lung Cancer in Patients With COPD in Regular Care by Electronic Nose Analysis of Exhaled Breath. Chest. 2023 2023/05/19/.

26. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh FH, van Putten JWG, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2023 Mar;163(3):697-706.

27. Leopold JH, Bos LD, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. J Breath Res. 2015 Dec 15;9(4):046002.

28. van der Sar IG, van Jaarsveld N, Spiekerman IA, Toxopeus FJ, Langens QL, Wijsenbeek MS, et al. Evaluation of different classification methods using electronic nose data to diagnose sarcoidosis. J Breath Res. 2023 Aug 29;17(4).

ADDITIONAL FILE 1

Outlier analysis

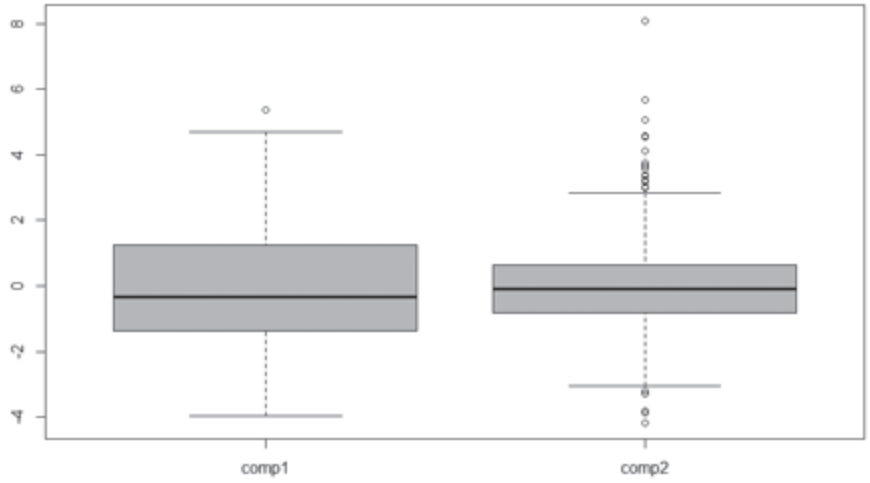


Figure S1: Box-and-whisker plot of the first principal component resulting from breath profile comparison of patients with ILD and other chronic respiratory diagnosis.
Principal component 1 (comp1) and 2 (comp2) result from the partial least squares analysis of breath profiles comparison of patients with ILD and other respiratory diseases. Outliers are marked as dots (n=25).

Table S1: Results of breath analysis comparison without outliers.

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
ILD	101	**Asthma – COPD – Lung cancer**	98	Training	0.97	*0.95-0.99*	0.91	0.95	0.93	0.95	0.91
	50		48	Test	0.98	*0.96-1.00*	0.94	0.94	0.94	0.94	0.94

Results based on 2 principal components. AUC = area under the curve; CI = confidence interval; COPD = chronic obstructive pulmonary disease; ILD = interstitial lung disease.

CHAPTER 5

Validation of electronic nose technology as diagnostic tool for different fibrotic interstitial lung diseases – preliminary report

Unpublished

Iris G. van der Sar[1], Leda Yazbeck[2], Markus Polke[3], Michael Kreuter[4], Philip L. Molyneaux[2], Marlies S. Wijsenbeek[1#], Catharina C. Moor[1#]

[1]Erasmus Medical Center, Rotterdam (the Netherlands); [2]Imperial College, London (United Kingdom); [3]University of Heidelberg, Heidelberg (Germany); [4]Mainz University Medical Center and of Pulmonary, Critical Care & Sleep Medicine, Mainz (Germany); [#]Authors share last authorship

Abstract

Introduction

Diagnosing fibrotic interstitial lung diseases (ILDs) requires chest computed tomography scan, multiple investigations including occasional invasive procedures, followed by a multidisciplinary team (MDT) discussion. Diagnostic and treatment delays are common. Previous single-centre studies showed that profiling of exhaled volatile organic compounds using electronic nose (eNose) sensor technology has potential to identify ILDs quickly and noninvasively. We aimed to validate eNose technology to differentiate various types of fibrotic ILD in an international multicentre cohort.

Methods

Patients with six predefined fibrotic ILD MDT diagnoses that often cause diagnostic dilemmas (idiopathic pulmonary fibrosis, fibrotic hypersensitivity pneumonitis, connective tissue disease related ILD, idiopathic non-specific interstitial pneumonia, interstitial pneumonia with autoimmune features, and unclassifiable ILD) were included in a prospective longitudinal study in three international ILD expert centres. An eNose (SpiroNose®) was used for exhaled breath analysis. Baseline data, split in training and test sets, were analysed with partial least squares discriminant and receiver operating characteristic analyses.

Results

372 patients (36.3% female) were included of whom 40.3% (n=150) had idiopathic pulmonary fibrosis (IPF). Differentiating breath profiles of patients with IPF and other fibrotic ILDs resulted in an area under the curve (AUC) of 0.94 (95%CI 0.91-0.97) in the training and 0.92 (0.87-0.98) in the test set. Moreover, individual ILDs could be discriminated with AUCs ranging 0.91-0.95 in the test sets.

Discussion

This international study shows that eNose technology differentiates breath profiles from patients with various fibrotic ILDs. eNose has the potential to help reducing diagnostic delay in patients with pulmonary fibrosis.

Introduction

Pulmonary fibrosis may manifest across various types of interstitial lung disease (ILD) [1]. In most of these diseases dynamic combinations of inflammatory and fibrotic mechanisms lead to tissue remodelling and deposition of extra cellular matrix, which may set-off self-perpetuating formation of fibrosis [2]. These inflammatory and fibrotic changes impair gas exchange in the lungs and results in generic respiratory symptoms like dyspnoea, reduced exercise tolerance, fatigue, and dry cough. The disease course of patients with a fibrotic ILD is highly variable, ranging from relatively stable to progressive pulmonary fibrosis; however, many patients eventually die from respiratory failure. Available treatments (i.e. anti-inflammatory or antifibrotic agents) may prevent or slow down the formation of fibrosis and (partially) reverse inflammatory changes. This highlights the need for timely diagnosis and treatment to limit the progression of symptoms and lung function decline.

High resolution chest computed tomography (CT) scan is central in the diagnosis of fibrotic ILDs [1]. Clinical and chest CT features of patients with different fibrotic ILDs often overlap and no single conclusive medical test to diagnose an individual ILD exists. Therefore, establishing the specific underlying ILD diagnosis and optimal treatment strategy require a multidisciplinary approach with often multiple investigations, including invasive procedures like bronchoscopies and biopsies. Despite this careful and extensive approach, some uncertainty on diagnosis or treatment strategy often remains for patients with a suspected fibrotic ILD.

Electronic nose (eNose) technology is a potential novel medical test and increasingly studied for diagnosing and monitoring various pulmonary diseases [3]. This sensor-based device analyses volatile organic compounds (VOCs) present in the human exhaled breath. These VOCs originate from various pathophysiological and metabolic processes in the human body and from external factors. We investigated in previous single-centre studies the ability of an eNose to detect various types of ILD and reported high accuracies [4-6].

In the international multicentre ILDnose study, we aim to validate eNose technology to differentiate various types of fibrotic ILDs, with a focus on idiopathic pulmonary fibrosis (IPF) versus other fibrotic ILDs, connective tissue disease related ILD (CTD-ILD), and fibrotic hypersensitivity pneumonitis (fHP). The current manuscript is a preliminary report.

Methods

Study design and population

We conducted a multicentre prospective observational trial in four ILD expert centres: Erasmus University Medical Center (EMC) Rotterdam, Thoraxklinik Heidelberg (TKH), Hôpital Louis-Pradel Lyon, and Royal Brompton Hospital (RBH) London. This manuscript reports on baseline visit results collected until December 2023.

Patients with pulmonary fibrosis and six pre-defined ILD multidisciplinary team (MDT) diagnoses that often cause diagnostic dilemmas were eligible for inclusion (IPF, fHP, CTD-ILD, idiopathic non-specific interstitial pneumonia (iNSIP), interstitial pneumonia with autoimmune features (IPAF), and unclassifiable ILD (U-ILD)). A high resolution CT scan was required to confirm the presence of PF, defined as reticular abnormalities with traction bronchiectasis or bronchiolectasis, with or without honeycombing, as determined by an experienced thoracic radiologist. Patients were classified as being 'incident' if the ILD diagnosis was established ≤6 months prior to inclusion, otherwise as 'prevalent'. Exclusion criteria were recent alcohol intake (<8 hours) or active respiratory infection according to their treating physician.

The ILDnose study is registered at Clinicaltrials.gov (identifier NCT04680832). The study was conducted in accordance with the amended Declaration of Helsinki. All participants signed informed consent before participating. The medical ethics committees of all participating centres approved the study protocol (MEC-2020-0655).

Data collection

An eNose device called SpiroNose® (Breathomix, Leiden, the Netherlands) was used for exhaled breath data collection. This validated eNose device contains seven different metal oxide semiconductor sensors present in various arrays on the inside and outside of the device [7, 8]. Participants were instructed to perform five tidal breaths, followed by an inhalation to total lung capacity, a 5 second breath hold, and a slow expiration. Measurements were performed in duplicate. Subsequently, participants completed a short survey (e.g., smoking history, respiratory symptoms, and recent medication, food or drink intake). Data were stored and processed in a secured, certified online database and data processing platform (BreathBase) [8].

Additional data were collected from medical files if available (e.g., patient demographics, medical history, medication use, and recent diagnostic test results from spirometry, chest CT scan reports, pathology assessments and blood samples) and stored in an online secured database (Castor).

Data analysis

Breath profiles of several patient groups were compared:

- Patients with IPF versus patients with another fibrotic ILD diagnosis (fHP, CTD-ILD, iNSIP, IPAF, and U-ILD);
- Patients with IPF versus fHP, IPF versus CTD-ILD, fHP versus CTD-ILD.

For the first analysis, datasets were split randomly in a training and test set (2:1) for independent validation of results. Second, for external validation, comparison of IPF versus the other fibrotic ILDs was performed in a training set (patient cohort of EMC), and was validated using the cohorts of TKH and RBH (**Figure 1**).

Figure 1: Data analysis groups.

To test for factors potentially influencing breath data, breath profile comparison (IPF versus other fibrotic ILDs) was repeated in several subgroups: males, females, former smokers, never smokers, antifibrotic and immunosuppressant use.

Descriptive statistics were used to analyse baseline data, including $\chi 2$, Student's t, and Mann Whitney tests to compare groups. We displayed normally distributed data as mean values (±standard deviation) and non-normally distributed data as median values (interquartile range). P-values of <0.05 were considered statistically significant. R version 4.3.2 for Windows with mixOmics version 6.26.0 package was used for analysis.

Data pre-processing

Data were retrieved from the BreathBase platform and underwent pre-processing prior to analysis. This processing encompassed identifying the best breath manoeuvre, adjusting for ambient air, scaling data to the stable sensor, and reducing inter-array differences. For statistical evaluation, we used the peak value and the ratio of the peak value to the breath hold of each sensor. The peak value from the stable sensor was omitted, leaving 13 data points per measurement (i.e. the breath profile) labelled with clinical patient data. Measurements of insufficient quality caused by incorrect breathing techniques or unstable ambient conditions were excluded.

Data classification

eNose sensor data classification and comparison between patient groups was conducted using partial least squares discriminant analysis (PLS-DA). PLS-DA reduces dimensionality of data and results in various principal components, i.e. weighted combinations of sensor values. To assess the discriminative ability of eNose, the first two components were used for visualising breath data and calculating area under the curve (AUC) using receiver operating characteristics analysis. The 95% confidence interval (CI), sensitivity, specificity, accuracy, negative predictive value (NPV), and positive predictive value (PPV) were derived from this analysis.

We conducted additional analysis to test whether a regression model of eNose and clinical data could improve diagnostic performance compared to eNose data only. First, we performed a binomial logistic regression with 10-fold cross-validation using the first two principal components that resulted from PLS-DA of the total cohort of patients with IPF versus other fibrotic ILDs. Second, we repeated the logistic regression and created a generalised linear model. Model included PLS-DA components and clinical parameters that are considered increasing the clinical likelihood for IPF compared to another ILD diagnosis: male sex (against female), >60 years of age (against <50 years), former or current smoker (against never), and restrictive lung function (i.e. forced vital capacity (FVC) <70% of predicted; against others) [9]. The significance of each parameter's contribution to model improvement was evaluated.

Results

In total, 372 patients were included in three centres (EMC, TKH, RBH) between November 2020 and December 2023. Overall median age was 71 years [63-77] and 36.3% was female (n=135). Most frequent diagnoses were IPF (n=150, 40.3%) and CTD-ILD (n=91, 24.5%). Between groups, significant differences in smoking history and

diffusion capacity (DLCOc) were found. Only a minority were incident patients (n=82, 28.3%). See **Table 1** for all baseline characteristics. Baseline characteristics divided by diagnosis groups is presented in **Supplementary data A**.

Table 1: Baseline characteristics.

	Overall (n=372)	**EMC (n=240)**	**RBH (n=82)**	**TKH (n=50)**	***p-value***
Female sex	135 (36.3)	81 (33.8)	39 (47.6)	15 (30.0)	*0.05*
Age (years)	71 [63-77]	72 [65-77]	70 [58-77]	68.0 [63-73]	*0.08*
Smoking history					*<0.01*
Never	134 (36.0)	67 (27.9)	45 (54.9)	22 (44.0)	
Former	229 (61.6)	166 (69.2)	36 (43.9)	27 (54.0)	
Current	9 (2.4)	7 (2.9)	1 (1.2)	1 (2.0)	
Pack years (years)	19 [7-30]	20 [8-30]	7 [0-30]	22 [10-37]	*0.14*
FVC (%predicted) *	81 [65-92]	81 [65-92]	83 [70-93]	78 [61-88]	*0.24*
DLCOc (%predicted) ^	48 [37-58]	50 [40-61]	42 [34-52]	45 [37-57]	*0.01*
ILD diagnosis					-
IPF	150 (40.3)	91 (37.9)	35 (42.7)	24 (48.0)	
CTD-ILD	91 (24.5)	57 (23.8)	24 (29.3)	10 (20.0)	
fHP	58 (15.6)	30 (12.5)	19 (23.2)	9 (18.0)	
U-ILD	34 (9.1)	30 (12.5)	-	4 (8.0)	
iNSIP	28 (7.5)	24 (10.0)	2 (2.4)	2 (4.0)	
IPAF	11 (3.0)	8 (3.3)	2 (2.4)	1 (2.0)	
Incident ~	82 (28.3)	74 (30.8)	-	8 (16.0)	-
Antifibrotic medication use	91 (24.5)	68 (23.3)	4 (4.9)	19 (38.0)	*<0.01*
Immunosuppressive medication use	109 (29.3)	92 (38.3)	4 (4.9)	13 (26.0)	*<0.01*
(Probable) UIP pattern on chest CT scan	165 (44.4)	104 (43.4)	40 (48.8)	21 (42.0)	*0.65*

*Values are displayed as number (%), mean ± SD, or median [interquartile range]. *n=11 missing data. ^n=29 missing data. ~ Refers to an established ILD diagnosis ≤6 months prior to inclusion; n=84 missing data (RBH only). CT = computed tomography; CTD = connective tissue disease; DLCOc = diffusing capacity for carbon monoxide corrected for haemoglobin level; EMC = Erasmus Medical Center; fHP = fibrotic hypersensitivity pneumonitis; FVC = forced vital capacity; ILD = interstitial lung disease; iNSIP = idiopathic non-specific interstitial pneumonia; IPF = idiopathic pulmonary fibrosis; IPAF = interstitial pneumonia with autoimmune features; RBH = Royal Brompton Hospital; TKH = Thoraxklinik Heidelberg; U-ILD = unclassifiable ILD; UIP = usual interstitial pneumonia.*

Breath profile comparison

First we compared breath profiles of patients with an IPF (n=150) versus another fibrotic ILD diagnosis (n=222) in randomly divided training and test sets (**Figure 2**). This resulted in an AUC of 0.94 (95%CI 0.91-0.97) in the training sets (n=100 vs. n=148) and an AUC of 0.92 (0.87-0.98) in the independent test sets (n=50 vs. n=74). Training the model with patients recruited at EMC (IPF n=91 vs. other ILDs n=149) resulted in an AUC of 0.94 (0.90-1.00). External validation of this model with patients from TKH and RB (IPF n=59 vs. other ILDs n=73) resulted in an AUC of 0.91 (0.86-1.00).

Additionally, breath profile comparison of selected ILD diagnoses of the full patient cohort resulted in test sets in an AUC of 0.95 (0.90-1.00) for IPF versus CTD-ILD, AUC of 0.91 (0.85-0.98) for IPF versus fHP, and AUC of 0.91 (0.83-0.99) for fHP versus CTD-ILD (**Figure 3**). All results, including corresponding specificity, sensitivity, accuracy, NPV, and PPV are displayed in **Table 2**.

Subgroup analyses

Additional subgroup analyses showed no important influence of sex, smoking status, antifibrotic or immunosuppressive drug use on the discriminative ability of eNose for IPF versus other fibrotic ILDs. Patients with a probable or definite UIP pattern on chest CT scan could be differentiated from those having another pattern (AUC 0.77 (0.72-0.82)). Scatter plots and performance results can be found in **Supplementary data B** (**Figure S1**, **Table S2**).

Regression model

A logistic regression model to distinguish IPF from the other ILDs, using eNose data only resulted in an AUC of 0.93 (95% CI) 0.90-0.96 and accuracy of 0.87. Adding clinical parameters as input variables minimally improved the model's performance and showed an AUC value of 0.94 (95% CI 0.92-0.97) and accuracy of 0.89. Age and sex category had a significant effect on the model outcome. Full results can be found in **Supplementary data C**.

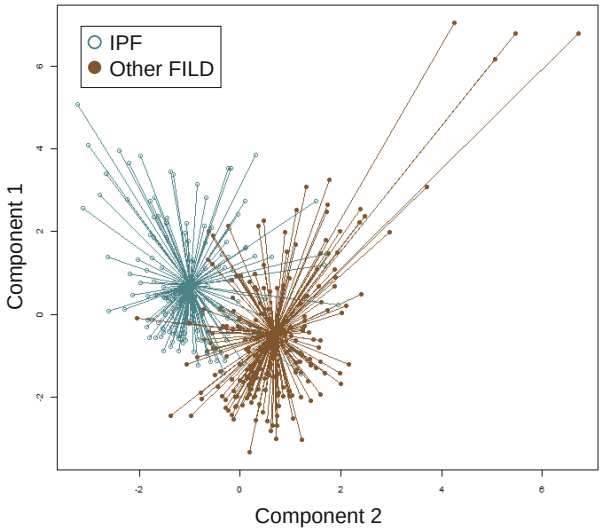


Figure 2: Breath profile comparison IPF versus other fibrotic ILDs.
Scatterplot of individual breath profiles of patients with an IPF versus another fibrotic ILD diagnosis. Each dot in the plot represents one patient. Component 1 and 2 are the first two principal components resulting from sparse partial least squares discriminant analysis. (F)ILD = (fibrotic) interstitial lung disease; IPF = idiopathic pulmonary fibrosis.

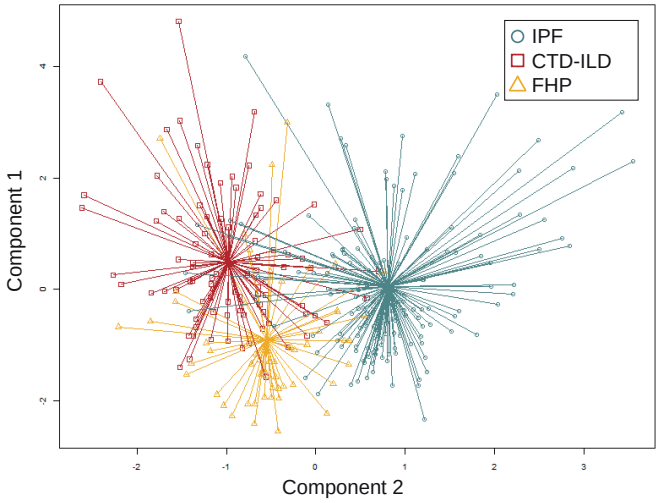


Figure 3: Breath profile comparison specific fibrotic ILDs.
Scatterplot of individual breath profiles of patients with an IPF, fHP and CTD-ILD diagnosis. Each dot in the plot represents one patient. Component 1 and 2 are the first two principal components resulting from sparse partial least squares discriminant analysis. fHP = fibrotic hypersensitivity pneumonitis; ILD = ILD; CTD = connective tissue disease; IPF = idiopathic pulmonary fibrosis.

Table 2: Results breath profile comparison in patients with fibrotic ILD.

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
IPF	100	**Other FILD**	148	Training	0.94	0.91-0.97	0.82	0.98	0.92	0.97	0.89
	50		74	Test	0.92	0.87-0.98	0.82	0.95	0.90	0.91	0.89
IPF	91	**Other FILD**	149	Training EMC	0.94	0.90-1.00	0.90	0.92	0.91	0.87	0.94
	59		73	Validation RBH+TKH	0.91	0.86-1.00	0.85	0.85	0.85	0.82	0.87
IPF	100	**CTD-ILD**	61	Training	0.97	0.94-1.00	0.95	0.93	0.94	0.89	0.97
	50		30	Test	0.95	0.90-1.00	0.90	0.96	0.94	0.93	0.94
IPF	100	**fHP**	39	Training	0.95	0.90-0.99	0.92	0.91	0.91	0.80	0.97
	50		19	Test	0.91	0.85-0.98	0.95	0.74	0.80	0.58	0.97
fHP	39	**CTD-ILD**	61	Training	0.89	0.82-0.97	0.87	0.87	0.87	0.81	0.91
	19		30	Test	0.91	0.83-0.99	0.90	0.77	0.82	0.71	0.92

Results based on 2 principal components resulting from partial least squares discriminant analysis in training sets. Training and test datasets are created by a random split (2:1) of each analysed cohort. AUC = area under the curve; fHP = fibrotic hypersensitivity pneumonitis; CI = confidence interval; CTD-ILD = connective tissue disease related interstitial lung disease; EMC = Erasmus Medical Center; FILD = fibrotic interstitial lung disease; IPF = idiopathic pulmonary fibrosis; NPV = negative predictive value; PPV = positive predictive value; RBH = Royal Brompton Hospital; TKH = Thoraxklinik Heidelberg.

Discussion

eNose technology can distinguish various types of fibrotic ILD. Preliminary results of the ILDnose study showed accurate separation of patients with IPF and other fibrotic ILDs, and patients with various specific fibrotic ILD diagnoses using an eNose. Outcomes were externally validated in an international cohort from three ILD expert centres.

Confirmation of the diagnostic performance for specific ILDs was awaited, following several mostly single-centre studies showing high performance for differentiating ILD from other diseases and healthy controls by using an eNose [10]. Compared to results in the first pilot study in 2021, the current results of IPF versus other ILDs are slightly better. Moor et al. reported an AUC of 0.87 (95% CI 0.77–0.96) in the test set, and we now showed an AUC of 0.92 (95% CI 0.87-0.98) using similar analysis and eNose type. The smaller CI might result from a larger sample size and more homogeneous population, since the previous study also included patients without fibrosis. Moreover, results suggest that multicentre data increase model robustness. Interestingly, additional analyses showed that the regression model with clinical and eNose data input did not significantly improve the model's performance for differentiating IPF from other fibrotic ILDs, compared to eNose data only. Another exploratory analysis suggest that different chest CT patterns is driving breath profile composition, regardless of diagnostic label. Longitudinal data on chest CT scans and eNose measurements including central review of CT scans should further explore the reliability and value of this finding.

International multicentre eNose studies are lacking in most fields of respiratory medicine. Yet, in lung cancer, one study used international multicentre data for validating cancer detection in suspected patients [11]. They found an AUC of 0.83 in the training and 0.79 in the validation group. A second multicentre study compared breath profiles of patients with asthma and chronic obstructive pulmonary disease (COPD) in several national centres. Reported accuracies were 95-97% in the training and 88-90% in the validation cohorts [12, 13]. Worth knowing, analysis type and cohort composition somewhat differed between the training and validation report. This international prospective longitudinal study is the first in ILD and for several reasons essential in the development of eNose applications for clinical practice. International data allow for reliable testing for the influence of diet, environment, living area, or other external factors on breath profiles. Besides, the presented results confirmed previous data that smoking habits, sex and ILD medication use seem not to affect eNose performance significantly [5, 6].

Multicentre data can also be used to evaluate the effect of using similar eNose devices in different research settings. Sensors always vary marginally between devices and change slightly over time (i.e. sensor drift) requiring periodically sensor validation. Data pre-processing should correct for these factors as well as for ambient room air, which is part of the standardised workflow of the currently used device [8]. Presented outcomes confirm that these corrections result in reliable uniform data across centres and countries.

Limitations

Some limitations apply to the current study. First, the accepted gold standard for ILD diagnosis, multidisciplinary team consensus, is used for training eNose models but includes inherent challenges. Despite clinical guidelines, multidisciplinary approach and individual diagnostic conclusions vary between centres [14]. Agreement between MDTs across the world ranges from poor to good, and is generally best for IPF diagnosis [15]. To ensure reliable, high-quality data input for model training, we selected highly experienced centres from the European Reference Network for ILD.

Another limitation is that the presented results are based on incomplete cohorts. For this preliminary report, we only analysed data of the ILD diagnosis subgroups that reached the aimed sample size at each of the three including centres. The complete baseline data are expected soon. These will enable the final analysis of all pre-defined diagnoses to confirm the validated diagnostic performance of this eNose. Also, the effect of treatment and disease severity (based on e.g. extent of fibrosis or pulmonary function) can be assessed. Exploratory unsupervised analysis of the full dataset will reveal what drives individual breath profiles: MDT diagnosis category or disease characteristics. Subsequently, follow-up data will allow evaluation of the performance of eNose as a predictive biomarker for disease course, response to therapy and other sub aims of the ILDnose study.

Future

Once the ILDnose study is completed and outcomes confirm that individual fibrotic ILDs have distinct breath profiles, more steps are warranted to collect the required evidence for development and approval of a clinically applicable diagnostic ILD tool. First, a diagnostic model needs to be designed. The best performing classification method has to be selected using the ILDnose dataset [16]. Then, in a multicentre international clinical trial, the designed model should be evaluated by comparing MDT and eNose model outcomes in newly recruited patients with various diagnoses,

aetiologies and stages of fibrotic ILD in secondary and tertiary health care centres. Important for this trial is the structural registration of the likelihood of MDT consensus ILD diagnoses as 'confident' (>90% confidence), 'provisional' (51-69% or 70-89%) or 'unclassifiable' (<50%) [17]. Furthermore, new studies ideally include a retrospective review of patient subset from each including centre to assess MDT agreement.

Once eNose is approved as diagnostic test, we believe that embedding eNose results in the MDT discussion will improve the diagnostic speed and confidence, in particular in non-expert centres, and might prevent invasive additional testing like lung biopsies in patients suspected of ILD. Ultimately, besides diagnostic classification, eNose might serve as a biomarker for underlying disease pathophysiology, chest CT scan pattern and recommend an optimal treatment strategy.

Conclusion

These preliminary data of the ILDnose study show that fibrotic ILDs can be highly accurately distinguished using an eNose, confirmed in an international validation cohort. These unique findings encourage development and evaluation of diagnostic models based on eNose data. Once approved as point-of-care medical test, eNose could facilitate a higher diagnostic confidence of individual MDT diagnoses non-invasively. This will improve patient care by enabling faster and more accurate diagnosis, leading to better treatment strategies.

References

1. Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1;205(9):e18-e47.
2. Wijsenbeek MS, Suzuki A, Maher TM. Interstitial lung diseases. The Lancet. 2022 2022/09/03/;400(10354):769-86.
3. van der Sar IG, Wijbenga N, Nakshbandi G, Aerts JGJV, Manintveld OC, Wijsenbeek MS, et al. The smell of lung disease: a review of the current status of electronic nose technology. Respiratory Research. 2021 2021/09/17;22(1):246.
4. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).
5. van der Sar IG, Moor CC, Oppenheimer JC, Luijendijk ML, van Daele PLA, Maitland-van der Zee AH, et al. Diagnostic Performance of Electronic Nose Technology in Sarcoidosis. Chest. 2022 Mar;161(3):738-47.
6. van der Sar IG, Wijsenbeek MS, Braunstahl GJ, Loekabino JO, Dingemans AC, In 't Veen J, et al. Differentiating interstitial lung diseases from other respiratory diseases using electronic nose technology. Respir Res. 2023 Nov 6;24(1):271.
7. de Vries R, Brinkman P, van der Schee MP, Fens N, Dijkers E, Bootsma SK, et al. Integration of electronic nose technology with spirometry: validation of a new approach for exhaled breath analysis. J Breath Res. 2015 Oct 15;9(4):046001.
8. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. Eur Respir J. 2018 Jan;51(1).
9. Cottin V, Tomassetti S, Valenzuela C, Walsh SLF, Antoniou KM, Bonella F, et al. Integrating Clinical Probability into the Diagnostic Approach to Idiopathic Pulmonary Fibrosis: An International Working Group Perspective. American Journal of Respiratory and Critical Care Medicine. 2022 2022/08/01;206(3):247-59.
10. van der Sar IG, Wijsenbeek MS, Moor CC. Exhaled breath analysis in interstitial lung disease. Current Opinion in Pulmonary Medicine. 2023.
11. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh FH, van Putten JWG, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2023 Mar;163(3):697-706.
12. Niki F, Aeilko HZ, Marc PvdS, Selma BdN, Erica D, Albert CR, et al. Exhaled Breath Profiling Enables Discrimination of Chronic Obstructive Pulmonary Disease and Asthma. American Journal of Respiratory and Critical Care Medicine. 2009;180(11):1076-82.
13. Fens N, Roldaan AC, van der Schee MP, Boksem RJ, Zwinderman AH, Bel EH, et al. External validation of exhaled breath profiling using an electronic nose in the discrimination of asthma with fixed airways obstruction and chronic obstructive pulmonary disease. Clinical & Experimental Allergy. 2011;41(10):1371-78.
14. Cottin V, Martinez FJ, Smith V, Walsh SLF. Multidisciplinary teams in the clinical care of fibrotic interstitial lung disease: current perspectives. European Respiratory Review. 2022;31(165):220003.

15. Walsh SLF, Wells AU, Desai SR, Poletti V, Piciucchi S, Dubini A, et al. Multicentre evaluation of multidisciplinary team meeting agreement on diagnosis in diffuse parenchymal lung disease: a case-cohort study. Lancet Respir Med. 2016 Jul;4(7):557-65.

16. van der Sar IG, van Jaarsveld N, Spiekerman IA, Toxopeus FJ, Langens QL, Wijsenbeek MS, et al. Evaluation of different classification methods using electronic nose data to diagnose sarcoidosis. J Breath Res. 2023 Aug 29;17(4).

17. Ryerson CJ, Corte TJ, Lee JS, Richeldi L, Walsh SLF, Myers JL, et al. A Standardized Diagnostic Ontology for Fibrotic Interstitial Lung Disease. An International Working Group Perspective. Am J Respir Crit Care Med. 2017 Nov 15;196(10):1249-54.

SUPPLEMENTARY DATA A

Baseline characteristics per diagnosis group

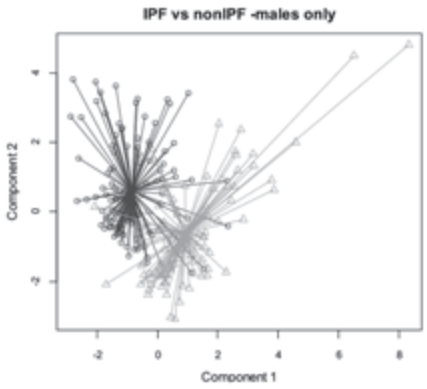
Table S1: Baseline characteristics divided by diagnosis

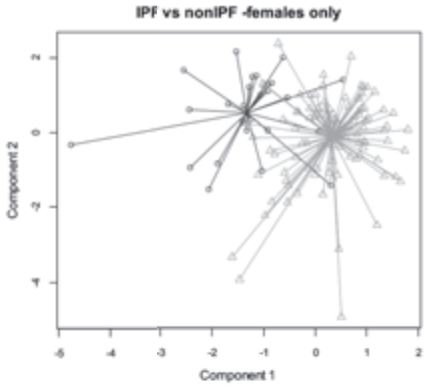
	IPF (n=150)	CTD-ILD (n=91)	FHP (n=58)	U-ILD (n=34)	iNSIP (n=28)	IPAF (n=11)
Female sex	26 (17.3)	55 (60.4)	31 (53.4)	9 (26.5)	9 (32.1)	5 (45.5)
Age (years)	73.0 [67.0, 77.0]	65.0 [55.5, 74.0]	68.50 [60.3, 75.0]	73.0 [66.0, 79.5]	74.0 [68.8, 77.3]	68.0 [64.0, 78.0]
Smoking history						
Never	46 (30.7)	43 (47.3)	23 (39.7)	8 (23.5)	10 (35.7)	4 (36.4)
Former	100 (66.7)	45 (49.5)	34 (58.6)	25 (73.5)	18 (64.3)	7 (63.6)
Current	4 (2.7)	3 (3.3)	1 (1.7)	1 (2.9)	-	-
Pack years (years)	20 [8, 32]	16 [67, 30]	10 [3, 28]	18 [9, 44]	25 [21, 38]	10 [6, 18]
FVC (%predicted) *	83 [70, 94]	77 [60, 90]	80 [60, 90]	79 [67, 90]	89 [76, 99]	69 [60, 89]
DLCOc (%predicted) ^	49 [40, 59]	47 [36, 56]	47 [31, 57]	48 [37, 55]	55 [45, 68]	48 [38, 60]
Including center						
TKH	24 (16.0)	10 (11.0)	9 (15.5)	4 (11.8)	2 (7.1)	1 (9.1)
RBH	35 (23.3)	24 (26.4)	19 (32.8)	-	2 (7.1)	2 (18.2)
EMC	91 (60.7)	57 (62.6)	30 (51.7)	30 (88.2)	24 (85.7)	8 (72.7)
Incident ~	26 (22.6)	19 (28.4)	10 (25.6)	15 (44.1)	9 (34.6)	3 (33.3)
Antifibrotic medication use	56 (37.3)	16 (17.6)	10 (17.2)	5 (14.7)	3 (10.7)	1 (9.1)
Immunosuppressive medication use	10 (6.7)	47 (51.6)	25 (43.1)	13 (38.2)	8 (28.6)	6 (54.6)
(Probable) UIP pattern on chest CT scan	122 (81.3)	17 (18.7)	11 (19.0)	8 (23.5)	3 (10.7)	4 (36.4)

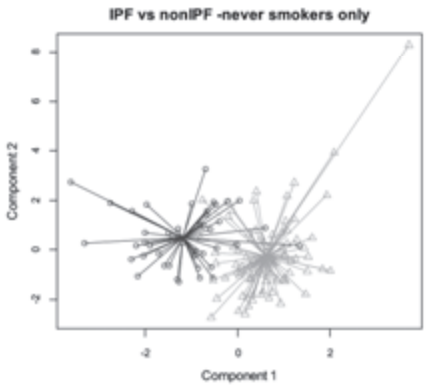
*Values are displayed as number (%), mean ± SD, or median [interquartile range]. *n=11 missing data. ^n=29 missing data. ~ Refers to an established ILD diagnosis ≤6 months prior to inclusion; n=84 missing data (RBH only). CT = computed tomography; CTD = connective tissue disease; DLCOc = diffusing capacity for carbon monoxide corrected for haemoglobin level; EMC = Erasmus Medical Center; fHP = fibrotic hypersensitivity pneumonitis; FVC = forced vital capacity; ILD = interstitial lung disease; iNSIP = idiopathic non-specific interstitial pneumonia; IPF = idiopathic pulmonary fibrosis; IPAF = interstitial pneumonia with autoimmune features; RBH = Royal Brompton Hospital; TKH = Thoraxklinik Heidelberg; U-ILD = unclassifiable ILD; UIP = usual interstitial pneumonia.*

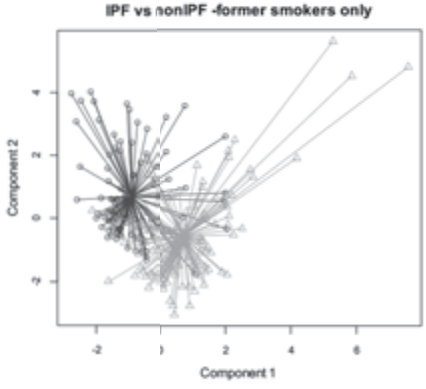
SUPPLEMENTARY DATA B

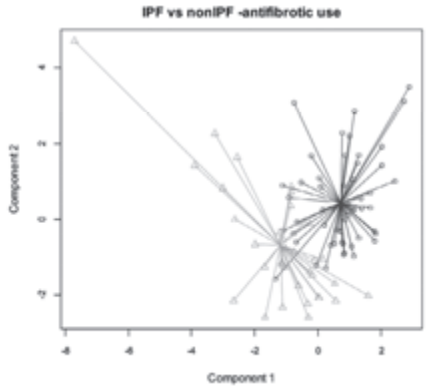
Subgroup analysis

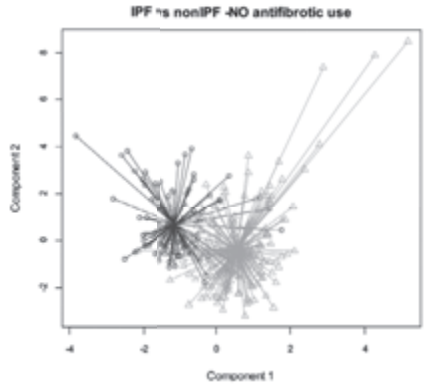












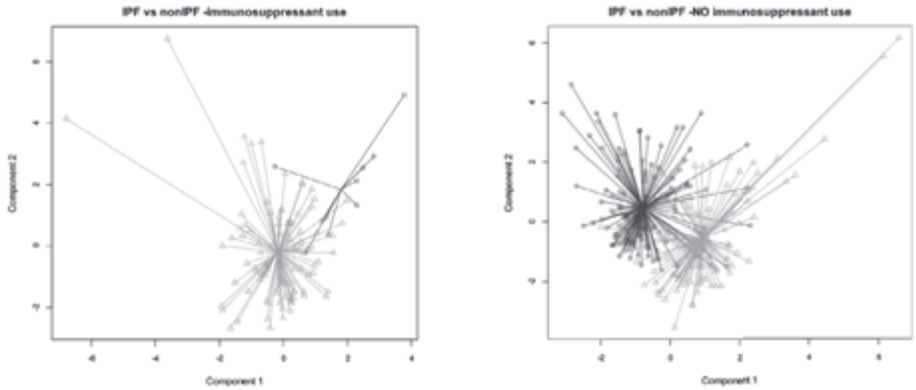


Figure S1: Breath profile comparison IPF versus other fibrotic ILDs in selected subgroups for assessment of influencing factors.
Scatterplot of individual breath profiles of patients with an IPF (purple) versus another fibrotic ILD diagnosis (grey) in subgroups based on gender, smoking history or medication use. Each dot in the plot represents one patient. Component 1 and 2 are the first two principal components resulting from sparse partial least squares discriminant analysis. ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis.

Table S2: Results breath profile comparison of patients with fibrotic ILD in selected subgroups

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
IPF	100	**Other FILD**	129	Former smokers	0,91	0,87-0,95	0,79	0,93	0,87	0,90	0,85
IPF	46	**Other FILD**	88	Never smokers	0,96	0,92-1,00	0,93	0,92	0,93	0,86	0,96
IPF	124	**Other FILD**	113	Males	0,93	0,87-0,97	0,81	0,97	0,89	0,97	0,82
IPF	26	**Other FILD**	109	Females	0,93	0,86-1,00	0,88	0,93	0,92	0,74	0,97
IPF	56	**Other FILD**	35	Antifibrotic use	0,95	0,92-0,99	0,82	1,00	0,89	1,00	0,78
IPF	94	**Other FILD**	187	No antifibrotic use	0,94	0,91-0,97	0,88	0,91	0,90	0,84	0,94
IPF	10	**Other FILD**	99	Immunosuppressant use	0,95	0,90-1,00	0,90	0,89	0,89	0,45	0,99
IPF	140	**Other FILD**	123	No immunosuppressant use	0,93	0,89-0,96	0,84	0,93	0,88	0,94	0,83
(p)UIP	165	**No (p)UIP**	207	Overall cohort	0,77	0,72-0,82	0,80	0,67	0,74	0,75	0,73

Results based on 2 principal components resulting from partial least squares discriminant analysis. AUC = area under the curve; CI = confidence interval; (F)ILD = (fibrotic) interstitial lung disease; IPF = idiopathic pulmonary fibrosis; NPV = negative predictive value; PPV = positive predictive value; (p)UIP = (probable) usual interstitial pneumonia (i.e. pattern on chest computed tomography scan).

SUPPLEMENTARY DATA C

Regression model results

Table S3: Coefficients of a logistic regression model to diagnose patients with IPF or another fibrotic ILD from eNose data.

Predictor variable	Estimated coefficient	Standard error	*p-value*
Component 1	2.08	0.22	*<0.05*
Component 2	-0.83	0.14	*<0.05*

Results of a generalized linear model based on logistic regression with 10-fold cross-validation. Input variables are the first two principal components resulting from partial least squares discriminant analysis of eNose sensor data from IPF (n=150) vs. other fibrotic ILDs (n=222). Both components have significant effects on the model outcome. Model resulted in a area under the curve of 0.93 (95%CI 0.90-0.96), accuracy 0.87 and kappa 0.73 for diagnosing IPF. CI = confidence interval; eNose = electronic nose; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis.

Table S4: Coefficients of a logistic regression model to diagnose patients with IPF or another fibrotic ILD from eNose data and clinical parameters.

Predictor variable	Estimated coefficient	Standard error	*p-value*
Component 1	2.06	0.23	*<0.05*
Component 2	-0.88	0.15	*<0.05*
Age	-1.18	0.50	*<0.05*
Sex	1.20	0.39	*<0.05*
Smoking history	0.00	0.38	*0.10*
FVC value	0.65	0.36	*0.05*

Results of a generalized linear model based on logistic regression with 10-fold cross-validation. Input variables are relevant clinical parameters and the first two principal components resulting from partial least squares discriminant analysis of eNose sensor data from IPF (n=150) vs. other fibrotic ILDs (n=222). Patients' clinical parameters were categorized as age ≤60 (n=74) or >60 years old (n=298), male (n=237) or female sex (n=135), ever (n=238) or never smoking (n=134), FVC <70% (n=174) or ≥70% of predicted (n=198). Both components, age and sex have significant effects on the model outcome. Model resulted in a area under the curve of 0.94 (95% CI 0.92-0.97), accuracy 0.89 and kappa 0.77 for diagnosing IPF. CI = confidence interval; eNose = electronic nose; FVC = forced vital capacity; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis.

CHAPTER 6

Diagnostic performance of electronic nose technology in sarcoidosis

Chest. 2022 Mar;161(3):738-747.

Iris G. van der Sar[1], Catharina C. Moor[1], Judith C. Oppenheimer[1], Megan L. Luijendijk[1], Paul L.A. van Daele[1], Anke-Hilse Maitland-van der Zee[2], Paul Brinkman[2], Marlies S. Wijsenbeek[1]

Erasmus Medical Center, Rotterdam (the Netherlands); [2]Amsterdam University Medical Center, Amsterdam (the Netherlands)

Abstract

Background

Diagnosing sarcoidosis can be challenging, and a non-invasive diagnostic method is lacking. The electronic nose (eNose) profiles volatile organic compounds in exhaled breath, and has potential as a point-of-care diagnostic tool.

Research question

Can we use eNose technology to distinguish accurately between sarcoidosis, interstitial lung disease (ILD) and healthy controls, and between sarcoidosis subgroups?

Study Design and Methods

In this cross-sectional study, exhaled breath of patients with sarcoidosis, ILD, and healthy controls was analyzed using an eNose (SpiroNose). Clinical characteristics were collected from medical files. Partial least square discriminant and ROC analysis was applied to a training and independent validation cohort.

Results

We included 252 patients with sarcoidosis, 317 with ILD and 48 healthy controls. In the validation cohorts, eNose distinguished sarcoidosis from controls with an AUC of 1.00, and pulmonary sarcoidosis from other ILD (AUC 0.87 (0.82-0.93)) and hypersensitivity pneumonitis (AUC 0.88 (0.75-1.00)). Exhaled breath of sarcoidosis patients with and without pulmonary involvement, pulmonary fibrosis, multiple organ involvement, pathology supported diagnosis, and immunosuppressive treatment showed no distinctive differences. Breath profiles differed between patients with a slightly and highly elevated soluble interleukin-2 receptor level (median cut off 772.0 U/mL; AUC 0.78 (0.64-0.92)).

Interpretation

Patients with sarcoidosis can be distinguished from ILD and healthy controls using eNose technology, indicating that this may facilitate accurate diagnosis in the future. Further research is warranted to understand the value of eNose in monitoring sarcoidosis activity.

Keywords

Breath test; diagnostic tool; electronic nose; interstitial lung disease; sarcoidosis

Take-home Points

Study Question: Can eNose technology be used to distinguish accurately between sarcoidosis, ILD, and healthy control subjects, and between sarcoidosis subgroups?

Results: In a study cohort of 252 patients with sarcoidosis, 317 with ILD, and 48 healthy control subjects, eNose accurately distinguished sarcoidosis from control subjects (AUC, 1.00 in the validation cohort), and pulmonary sarcoidosis from other ILD (AUC, 0.87; 95% CI, 0.82-0.93 in the validation cohort).

Interpretation: Patients with sarcoidosis can be distinguished from ILD and healthy control subjects by using eNose technology, indicating that this method may facilitate accurate diagnosis in the future.

Introduction

Sarcoidosis is a granulomatous inflammatory disease without a known cause that can affect roughly any organ. The lungs are involved in the vast majority of patients (89-99%) [1]. Diagnosis can be challenging because no standardized diagnostic procedure exists. The three major criteria for diagnosis are compatible clinical features, pathology tissue assessment, and exclusion of other granulomatous diagnoses [2].

Due to the heterogeneity of sarcoidosis, disease course and treatment outcomes are difficult to predict. Severity of symptoms, organs affected, disease progression, and treatment response vary widely between individuals [3, 1]. In clinical practice, patients may be divided in limited disease (i.e. involution or stable) and (potentially) progressive disease with threat to organ function [4].

Current serum biomarkers for diagnosing, monitoring or predicting disease course of sarcoidosis lack validity and/or reliability [5]. Despite that, the serum level of soluble interleukin-2 receptor (sIL-2R) is often used in clinical practice as a follow-up marker for disease activity [6]. sIL-2R also correlates with inflammatory activity on positron emission tomography (PET) scan [5]. The sIL-2R value is not specific for a sarcoidosis diagnosis and not available worldwide.

Breath biomarkers are increasingly studied in respiratory diseases, as exhaled volatile organic compounds (VOCs) reflect pathophysiological processes in the human body [7, 8]. Techniques such as gas chromatography and mass spectrometry can be used to identify individual VOCs, but are time-consuming and complex. To the best of our knowledge, three studies identified individual VOCs in sarcoidosis using these techniques [9]. However, VOC identification lacked reproducibility in external validation cohorts [10]. It is more likely that analysis of a profile of VOCs (a 'breathprint') using electronic nose (eNose) technology will be of added value in clinical practice. This breath analysis tool is quick, easier, and cheaper than GC-MS analysis [11, 8]. eNose devices contain multiple gas-sensors that react to a broad range of VOCs [12]. An eNose creates an individual breathprint after pooling and processing sensor deflections.

Until now, only one small pilot study evaluated the potential of eNose technology to detect sarcoidosis [13]. A cohort of 11 untreated sarcoidosis patients could be distinguished from 25 healthy controls. Thus, further research in larger patient groups is warranted to confirm these promising results.

The aim of this study was to evaluate the reliability and validity of exhaled breath analysis using eNose technology to differentiate between sarcoidosis, healthy controls,

and interstitial lung disease (ILD). Moreover, we aimed to evaluate whether breathprint data could distinguish between subgroups of sarcoidosis patients based on clinical characteristics.

Study design and methods

Study design and population

This single-center cross-sectional study was performed in the Erasmus Medical Center (Rotterdam, the Netherlands) between August 2019 and March 2021. Outpatients with an established diagnosis of sarcoidosis according to the ATS/ERS/WASOG criteria or ILD according to the American Thoracic Society/European Respiratory Society criteria were eligible for inclusion [14-16, 2]. Data of a subset of patients in this study, was also used in a previous publication by Moor et al [17]. Healthy controls were recruited among healthcare staff of the Erasmus Medical Center. Subjects in the control group had a negative history of respiratory diseases and did not use pulmonary medication. The study was conducted in accordance with the amended Declaration of Helsinki. Patients and control subjects with pulmonary infection were excluded. All participants signed informed consent before participating. The medical ethics committee approved the study protocol (MEC-2019-0230).

Data collection

The SpiroNose (Breathomix, Leiden, the Netherlands) was used for exhaled breath analysis. The SpiroNose is a validated eNose device containing seven different metal-oxide semiconductor sensors [18, 19]. Measurements were performed as described previously [17]. Participants were instructed to perform five tidal breaths, followed by an inhalation to total lung capacity, a five-second breath hold and a slow expiration. Data was stored and processed in a secured certified online database and data processing platform (BreathBase) [19].

Participants completed a short questionnaire including ethnicity, smoking, recent food or drink intake, inhaler use and signs of pulmonary infection. Information on patient characteristics, medical history, medication use, and most recent available diagnostic test results (e.g. spirometry, chest imaging, pathologic assessment, blood samples) were collected from medical files. If available, the most recent chest HRCT scan was evaluated for the presence of pulmonary fibrosis. Patients were classified as having pulmonary fibrosis when reticulations with traction bronchiectasis were present on HRCT scan as reviewed by an experienced thoracic radiologist. Clinical subgroups were defined depending on organ involvement, presence of pulmonary fibrosis, current

immunosuppressive treatment, availability of histology for diagnosis, and sIL-2R level. To explore if breathprints correlate with disease activity, the sIL-2R level was used as marker for activity. In our laboratory, a sIL-2R value of ≤ 550 U/ml is considered normal. The median value of elevated sIL-2R levels was used as a cut-off to define the lower and upper 50% group.

Data analysis

Sensor data resulting from the measurements were extracted from the database. Prior to statistical analysis, eNose sensor signals were processed. Sensor signals were corrected for ambient air, peak values were normalized to the most stable sensor and inter-array differences were reduced [18, 19]. Sensor peak values and ratios between peak value and breath hold were both used for analysis. The sensor data of each patient was labelled with the patient and disease characteristics. Partial least square discriminant analysis (PLS-DA) was used for analyzing sensor data. This method reduces the dimensionality of data and results in a set of multivariate components. Each PLS-DA component is a weighted combination of the original sensor variables. The first two components explain the greatest variance of sensor data. PLS-DA component 1 and 2 were therefore used for comparing data between diagnosis groups. Component 1 was used for analysis within the sarcoidosis groups to avoid overfitting the model. For linear regression analysis, PLS-DA component 1 was used. Results from the PLS-DA analyses were visualized as scatterplots with component 1 on the x-axis and 2 on the y-axis. Each dot represents one patient and the center of the dot cloud represents the mean value of the components. After applying a generalized linear model prediction method to the PLS-DA component 1 and 2, receiver operating characteristic (ROC) analysis was performed using the odds (a value between 0 and 1) that a patient does belong to either of the groups based on the sensor data. The area under the curve (AUC) values and corresponding 95% confidence intervals were derived from that analysis. Additionally, sensitivity, specificity, accuracy, and negative and positive predictive values were calculated. Additional background information on sensor data processing and analysis is provided in the text and **e-Figures 1 to 5** of **e-Appendix 1**.

Before analysis, diagnosis groups were randomly divided in a training and independent validation set (2:1), following recommendations for metabolomics experiments [20]. The PLS-DA components 1 and 2 derived from the training set were applied to the independent validation set to validate the results. For analysis within the sarcoidosis cohort, subgroups were not split in a training and validation set. Descriptive statistics were used to analyze baseline data. Normally distributed data are displayed as mean values with standard deviation and non-normally distributed data as median with

interquartile range. Between-group comparisons were done using Chi-squared tests, Kruskal–Wallis tests, and Mann Whitney tests. Analyses were performed using R version 4.0.3 for Mac OS X GUI (PBC, Boston, MA, USA) using the mixOmics package version 6.14.0 and ggpubr package version 0.4.0.

Results

In total, 569 outpatients were included: 252 with sarcoidosis and 317 with ILD. 48 healthy controls were included. The ILD cohort comprised patients with IPF (n=124), connective tissue disease related ILD (n=64), hypersensitivity pneumonitis (HP; n=50) and other ILD (n=79). Baseline characteristics of the study groups are presented in **Table 1** and **2**. Patients with ILD were older than patients with sarcoidosis and healthy controls ($p<0.05$). Patients with sarcoidosis had a higher diffusion capacity for carbon monoxide and forced vital capacity compared to patients with ILD ($p<0.05$).

Table 1: Baseline characteristics.

	Sarcoidosis (n=252)	ILD (n=317)	Healthy controls (n=48)
Age (years)	53.1 ±11.4[a]	70.0 (62.0-76.0)[a]	36.5 (27.0-48.3)[a]
Male	134 (53.2)[b]	195 (61.5)[b]	15 (31.3)
BMI (kg/m^2)	27.1 (24.7-30.6)[b]	26.3 (24.2-29.4)[b]	22.6 (20.7-24.5)
Smoking status[a]			
- Never smoker	154 (61.1)	90 (28.4)	37 (77.1)
- Former smoker	83 (32.9)	217 (68.5)	7 (14.6)
- Current smoker	15 (6.0)	10 (3.2)	4 (8.3)
FVC (% of predicted)	89.0 (78.0-98.0)[c]	78.8 ±20.0	-
DLCO (% of predicted)	78.5 (63.0-89.0)[c]	50.2 ±15.4	-

Values are displayed as number (%), mean ±SD or median (interquartile range). BMI = body mass index; DLCO = diffusion capacity for carbon monoxide; FVC = forced vital capacity; ILD = interstitial lung disease; SD = standard deviation. [a]Significantly different between subgroups. [b]Significantly different from healthy controls. [c]Significantly different from ILD.

Table 2: Distribution of diagnoses in ILD cohort (n=317).

Type of ILD	Number (%)
Idiopathic pulmonary fibrosis	124 (39.1)
Connective tissue disease related ILD	64 (20.2)
Hypersensitivity pneumonitis	50 (15.8)
Idiopathic nonspecific interstitial pneumonia	20 (6.3)
Interstitial pneumonia with autoimmune features	14 (4.4)
Combined pulmonary fibrosis and emphysema	10 (3.2)
(Cryptogenic) organizing pneumonia	9 (2.8)
Unclassifiable	8 (2.5)
Granulomatosis with polyangiitis	4 (1.3)
Respiratory bronchiolitis ILD	4 (1.3)
Asbestosis	3 (0.9)
Desquamative interstitial pneumonia	3 (0.9)
Drug-induced ILD	2 (0.6)
Other	2 (0.6)

ILD = interstitial lung disease

Sarcoidosis versus healthy controls

Patients with sarcoidosis and healthy controls were divided in a training (n=168 sarcoidosis, n=32 controls) and validation set (n=84 sarcoidosis, n=16 controls; **Figure 1**). Differentiation between patients and controls resulted in an AUC of 1.00 in both training and validation set. Corresponding sensitivity, specificity and accuracy are displayed in **Table 3**.

When comparing patients with pulmonary involvement (n=224) to control subjects, similar results were found in both training set (n=150 sarcoidosis, n=32 controls, AUC 1.00) and validation set (n=74 sarcoidosis, n=16 controls, AUC 1.00). Sarcoidosis patients treated with immunosuppressive medication (training n=81, validation n=40) could also be differentiated from healthy controls (training n=32, validation n=16) with an AUC of 1.00 in both sets.

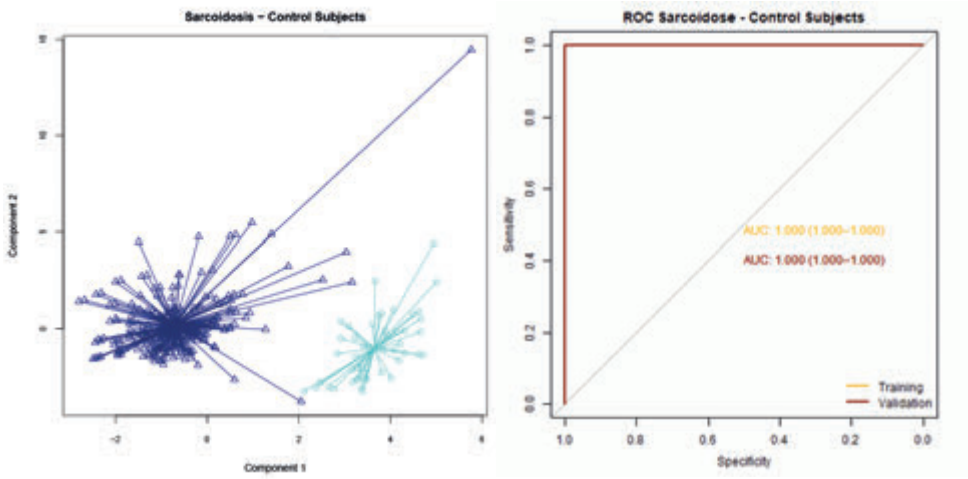


Figure 1: eNose data of patients with sarcoidosis and healthy controls.
A. Scatterplot of eNose data of PLS-DA component 1 and 2 for full data set (n=252 sarcoidosis, n=48 control subjects). Each data point represents one patient; the center of the dot cloud represents the mean value of the components. Blue = patient with sarcoidosis, turquoise = control subject. B. ROC curves for training and validation set. AUC = area under the curve; eNose = electronic nose; PLS-DA = partial least square discriminant analysis; ROC = receiver operating characteristic.

Pulmonary sarcoidosis versus ILD

eNose data of sarcoidosis patients with pulmonary involvement (n=224) were compared to patients with ILD (n=317; **Figure 2** and **Table 3**). This resulted in an AUC of 0.90 (0.87-0.94) in the training set (n=150 sarcoidosis, n=212 ILD) and an AUC of 0.87 (0.92-0.93) in the validation set (n=74 sarcoidosis, n=105 ILD).

The comparison between pulmonary sarcoidosis and HP yielded an AUC of 0.95 (0.90-0.99) in the training set (n=150 sarcoidosis, n=34 HP), and an AUC of 0.88 (0.75-1.00) in the validation set (n=74 sarcoidosis, n=16 HP) (**Figure 3**).

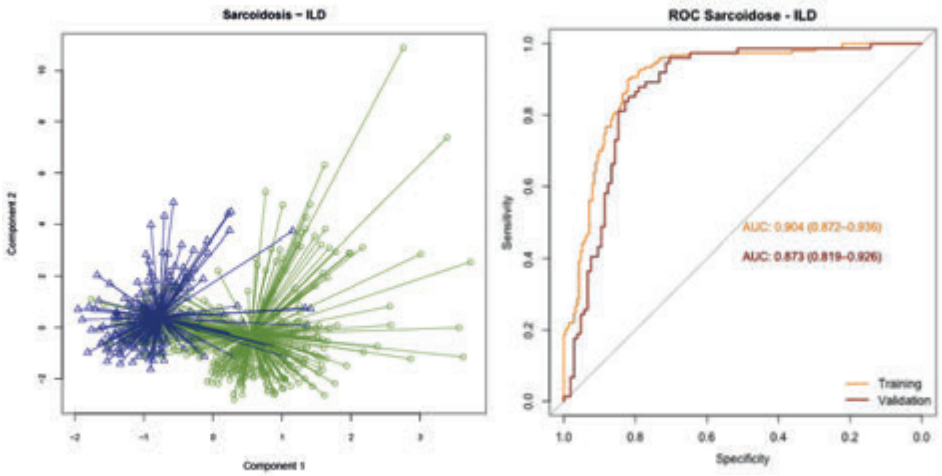


Figure 2: eNose data of patients with pulmonary sarcoidosis and interstitial lung disease.
A. Scatterplot of eNose data of PLS-DA component 1 and 2 for full data set (n=224 sarcoidosis, n=317 ILD). Each data point represents one patient; the center of the dot cloud represents the mean value of the components. Blue = patient with sarcoidosis, green = patient with ILD. B. ROC curves for training and validation set. AUC = area under the curve; eNose = electronic nose; ILD = interstitial lung disease; PLS-DA = partial least square discriminant analysis; ROC = receiver operating characteristic.

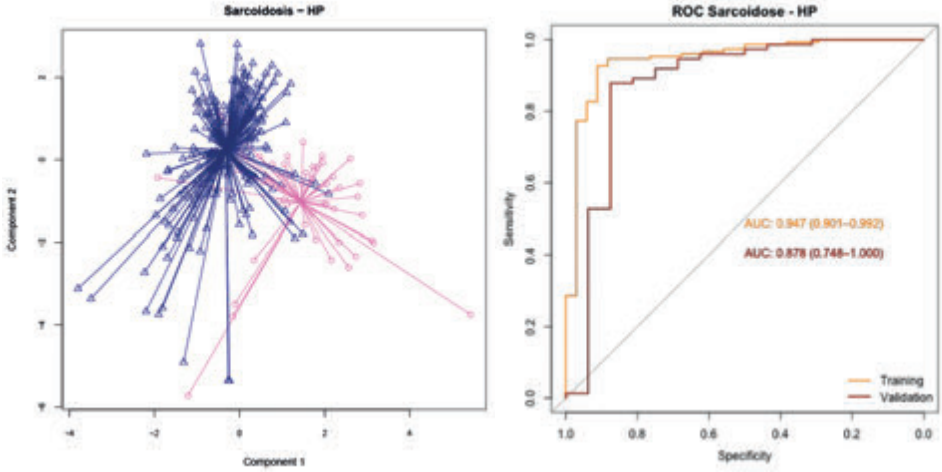


Figure 3: eNose data of patients with pulmonary sarcoidosis and hypersensitivity pneumonitis.
A. Scatterplot of eNose data of PLS-DA component 1 and 2 for full data set (n=224 sarcoidosis, n=50 HP). Each data point represents one patient; the center of the dot cloud represents the mean value of the components. Blue = patient with sarcoidosis, pink = patient with HP. B. ROC curves for training and validation set. AUC = area under the curve; eNose = electronic nose; HP = hypersensitivity pneumonitis; PLS-DA = partial least square discriminant analysis; ROC = receiver operating characteristic.

Table 3: Diagnostic performance of eNose technology.

Group 1	n=	Group 2	n=	Data set	AUC (CI 95%)	Sensitivity	Specificity	Accuracy	NPV	PPV
Sarcoidosis	168	**Healthy controls**	32	Training	1.00 (1.00-1.00)	100%	100%	100%	100%	100%
	84		*16*	*Validation*	*1.00 (1.00-1.00)*	*100%*	*100%*	*100%*	*100%*	*100%*
Sarcoidosis (pulmonary)	150	**Healthy controls**	32	Training	1.00 (1.00-1.00)	100%	100%	100%	100%	100%
	74		*16*	*Validation*	*1.00 (1.00-1.00)*	*100%*	*100%*	*100%*	*100%*	*100%*
Sarcoidosis (treated)	81	**Healthy controls**	32	Training	1.00 (1.00-1.00)	100%	100%	100%	100%	100%
	40		*16*	*Validation*	*1.00 (1.00-1.00)*	*100%*	*100%*	*100%*	*100%*	*100%*
Sarcoidosis (pulmonary)	150	**ILD**	212	Training	0.90 (0.87-0.94)	90.0%	82.1%	85.4%	92.1%	78.0%
	74		*105*	*Validation*	*0.87 (0.82-0.93)*	*85.1%*	*81.9%*	*83.2%*	*88.7%*	*76.8%*
Sarcoidosis (pulmonary)	150	**HP**	34	Training	0.95 (0.90-0.99)	92.7%	91.2%	92.4%	73.8%	97.9%
	74		*16*	*Validation*	*0.88 (0.75-1.00)*	*87.8%*	*87.5%*	*87.8%*	*60.9%*	*97.0%*

Results of the validation set are in italic. AUC = area under the curve; CI = confidence interval; HP = hypersensitivity pneumonitis; ILD = interstitial lung disease.

Sarcoidosis

Additional clinical characteristics of the sarcoidosis cohort are described in **Table 4**. The comparison of breathprints between sarcoidosis subgroups resulted in AUCs ranging from 0.55 to 0.64 (**Table 5**). The presence or absence of pulmonary involvement, and pulmonary fibrosis in particular, multiple organ involvement, pathology supported diagnosis or immunosuppressive treatment did not influence patients' breathprint, as all 95% confidence intervals are close to 0.5.

sIL-2R level was available in 132 patients. eNose data did not distinguish patients with normal sIL-2R levels from elevated levels (cut-off 550 U/mL). In patients with elevated sIL-2R levels (n=43), the median was 772.0 U/mL. In this group, differences in breathprint were found between the lower and upper 50% (AUC 0.78; 0.64-0.92, n=21 lower 50%, n=22 upper 50%). Explorative regression analysis did not show a correlation between breathprint and sIL-2R levels. Additional subgroup analyses showed that smoking status, age and gender did not influence the outcomes. The results of these analyses are shown in **e-Figure 6 to 21** of **e-Appendix 2**.

Table 4: Sarcoidosis patient characteristics.

Sarcoidosis cohort	Number of patients
Self-reported ethnicity	252 (100)
European/Caucasian	170 (67.5)
South and Latin American	59 (23.4)
Asian	11 (4.4)
Northern African	7 (2.8)
Sub-Saharan African	5 (2.0)
Time from diagnosis	252 (100)
Time (months)	68.0 (28.3-139.0)
Diagnosis supported by pathology	188 (74.6)
Numbers of organs involved	252 (100)
1 organ	24 (9.5)
>1 organ	228 (90.5)
Pulmonary involvement	224 (88.9)
Pulmonary fibrosis	52 (23.2)
No pulmonary fibrosis	148 (66.1)
Fibrosis unknown[a]	24 (10.7)
Extrapulmonary involvement	250 (99.2)
Lymph nodes	232 (92.8)
Skin	48 (19.2)

Sarcoidosis cohort	Number of patients
Eyes	46 (18.4)
Muscle/joints	30 (12.0)
Cardiac	21 (8.4)
Small fiber neuropathy	11 (4.4)
Central nervous system	6 (2.4)
Other organs	50 (20.0)
Current immunosuppressive treatment[b]	121 (48.0)
Corticosteroids	70 (57.9)
Methotrexate	70 (57.9)
TNF inhibitors	19 (15.7)
Azathioprine	8 (6.6)
Mycophenolate mofetil	2 (1.7)
Rituximab	1 (0.8)
No current immunosuppressive treatment	131 (52.0)
sIL-2R results[c]	132 (52.4)
Level (U/mL)	458.0 (325.5-625.8)
Normal sIL-2R (≤ 550 U/mL)	89 (35.3)
Level (U/mL)	383.0 (297.0-458.0)
Elevated sIL-2R (> 550 U/mL)	43 (17.1)
Level (U/mL)	772.0 (632.5-1289.5)

Values are displayed as number (%) or median (interquartile range). Percentages calculated of subgroup total. [a]No HRCT available. [b]Some patients used a combination of different medications. [c]sIL-2R level was not available for 120 (47.6%) sarcoidosis patients. HRCT = high resolution computed tomography; sIL-2R = soluble interleukin-2 receptor; TNF = tumor necrosis factor.

Table 5: Diagnostic performance of eNose in sarcoidosis subgroups.

Group 1	n=	Group 2	n=	AUC (CI 95%)
Disease characteristics				
Pulmonary involvement	224	No pulmonary involvement	28	0.64 (0.54-0.73)
Pulmonary fibrosis	52	No pulmonary fibrosis	148	0.59 (0.51-0.68)
1 organ involved	24	>1 organ involved	228	0.64 (0.53-0.76)
Immunosuppressive treatment	121	No immunosuppressive treatment	131	0.55 (0.48-0.62)
Pathology supported	188	No pathology	64	0.61 (0.52-0.69)
sIL-2R level				
Normal	89	Elevated	43	0.61 (0.51-0.71)
Elevated lower 50%	21	Elevated upper 50%	22	0.78 (0.64-0.92)

AUC = area under the curve; CI = confidence interval; sIL-2R = soluble interleukin-2 receptor.

Discussion

This study evaluated the diagnostic performance of eNose technology in a large cohort of patients with sarcoidosis. The eNose accurately differentiated between patients with sarcoidosis and healthy controls with an AUC of 1.00. Breathprints of patients with ILD, and HP in particular, could also be adequately distinguished from pulmonary sarcoidosis. These findings were confirmed in a validation cohort. Within sarcoidosis, breathprints of patient subgroups were similar, except for those with elevated sIL-2R levels.

The accuracy of eNose technology to differentiate sarcoidosis from controls was significantly better than in the only previous study assessing eNose technology in sarcoidosis. Dragonieri et al. reported a cross-validated accuracy of 83.3% to distinguish sarcoidosis from healthy controls, while in the current study the accuracy was 100% [13]. Moreover, Dragonieri et al. did not find a difference in breathprint between treated sarcoidosis patients and healthy controls. The difference between the studies might be explained by the much smaller cohort size in the study of Dragonieri, as well as the use of a different eNose device.

Interestingly, in our cohort, breathprints were similar in sarcoidosis subgroups. A specific signal originating from the disease itself seems to dominate the patients' breathprints, despite clinical heterogeneity [4]. The finding that breathprints of patients with and without pulmonary fibrosis were not significantly different, implies an influence of inflammation on exhaled VOCs. This is supported by increasing evidence from studies on different breath analysis techniques in other diseases [21]. In this study, we also showed that the eNose could separate sarcoidosis patients with high and low inflammatory activity, based on sIL-2R levels, and might serve as a new marker for inflammatory activity. However, no correlation between breathprints and sIL-2R levels was found. This could be due to a relatively small number of patients with an available sIL-2R level in our cohort, of which the majority had only slightly elevated levels (median 772.0 U/mL). More extensive follow-up studies with successive within-patient measurements will lead to a better understanding of the influence of disease activity and treatment on breathprints, and the relation with sIL-2R levels and inflammatory activity on PET scans. According to a longitudinal study in asthmatic subjects with unsupervised clustering of eNose data, it might be possible to identify changes in inflammatory activity or immunosuppressive treatment [22].

In clinical practice, it can be challenging to establish a diagnosis of sarcoidosis, and in particular to differentiate between other granulomatous diseases, such as HP [23]. Notably, our results showed that sarcoidosis could be accurately separated from HP. A limitation of the current study was the absence of patients with granulomatous

diseases such as tuberculosis and sarcoid-like reactions, due to the low prevalence of these diseases. Previous studies did show that tuberculosis can be accurately differentiated from healthy controls and from patients with suspected tuberculosis using an eNose [24, 25]. eNose technology therefore holds the potential to guide multidisciplinary team discussions in patients with a granulomatous disease. Future studies should assess the value of the eNose in differentiating between a broader range of granulomatous entities. Especially in areas with limited access to diagnostic procedures and/or a high prevalence of tuberculosis, eNose might be of added value as an easy accessible and accurate point of care tool in clinical practice.

The new sarcoidosis diagnostic guideline states that histopathology is not always needed to establish the diagnosis if all other findings are consistent with sarcoidosis [2]. In the current study, breathprints of patients with and without a diagnosis confirmed by tissue sampling did not differ, which supports the recommendations in the guideline. This finding emphasizes the potential of eNose technology as an accurate diagnostic tool for sarcoidosis, without the need for invasive tissue sampling.

Strengths of this study are its large sample size and real-world population, including patients with comorbidities or medication use. Additionally, we validated the results obtained from the training set in an independent validation cohort. A limitation is that our dataset contains some missing data. sIL-2R values were not available for all patients, which might influence the outcome and strength of the analysis. Hence, further studies to extend and confirm these results are warranted. Moreover, the compared groups are not matched regarding certain baseline variables such as gender, smoking status and age. However, additional subgroup analyses did not show an effect of these variables on results. Lastly, the results of our single center study still need to be confirmed and validated by external patient cohorts in a multicenter multinational study [26]. External validation, design of a diagnostic algorithm and test cohorts are required steps before implementation of the SpiroNose as a diagnostic tool can be realized (**Figure 4**).

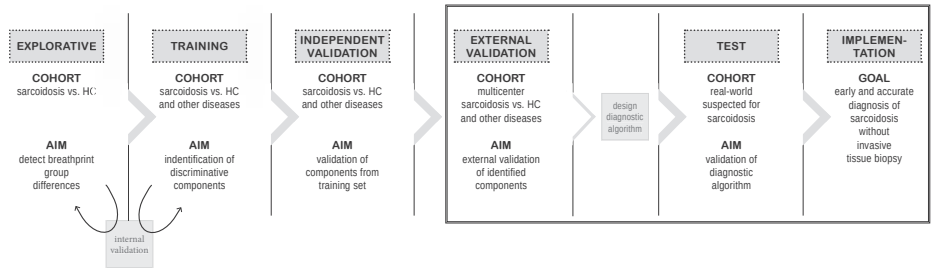


Figure 4: Development steps of eNose technology towards a diagnostic tool for sarcoidosis.
In the current study, data analysis of a training and independent validation cohort have been performed. Research steps in the rectangle box are still required before the SpiroNose could be used as a diagnostic tool in suspected sarcoidosis patients. HC = healthy controls.

Interpretation

This study shows a reliable and accurate differentiation of patients with sarcoidosis from patients with ILD and healthy controls, based on eNose data. The results confirm the potential of eNose technology as a non-invasive diagnostic tool to obtain an early, accurate sarcoidosis diagnosis and reduce the number of invasive diagnostic procedures in the diagnostic trajectory. This encourages further research in external cohorts of patients with sarcoidosis to validate the diagnostic properties of eNose technology (**Figure 4**).

Within sarcoidosis, breathprints were similar between subgroups, except for patients with high inflammatory activity. This emphasizes the potential value of eNose technology in monitoring disease activity. Longitudinal studies need to explore its ability to monitor disease activity.

References

1. Grunewald J, Grutters JC, Arkema EV, Saketkoo LA, Moller DR, Müller-Quernheim J. Sarcoidosis. Nat Rev Dis Primers. 2019 Jul 4;5(1):45.
2. Crouser ED, Maier LA, Wilson KC, Bonham CA, Morgenthau AS, Patterson KC, et al. Diagnosis and Detection of Sarcoidosis. An Official American Thoracic Society Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2020;201(8):e26-e51.
3. Judson MA. The Clinical Features of Sarcoidosis: A Comprehensive Review. Clin Rev Allergy Immunol. 2015 Aug;49(1):63-78.
4. Pereira CA, Dornfeld MC, Baughman R, Judson MA. Clinical phenotypes in sarcoidosis. Curr Opin Pulm Med. 2014 Sep;20(5):496-502.
5. Ramos-Casals M, Retamozo S, Sisó-Almirall A, Pérez-Alvarez R, Pallarés L, Brito-Zerón P. Clinically-useful serum biomarkers for diagnosis and prognosis of sarcoidosis. Expert Rev Clin Immunol. 2019 Apr;15(4):391-405.
6. Schimmelpennink MC, Quanjel M, Vorselaars A, Wiertz I, Veltkamp M, Van Moorsel C, et al. Value of serum soluble interleukin-2 receptor as a diagnostic and predictive biomarker in sarcoidosis. Expert Rev Respir Med. 2020 Jul;14(7):749-56.
7. van de Kant KDG, van der Sande LJTM, Jöbsis Q, van Schayck OCP, Dompeling E. Clinical use of exhaled volatile organic compounds in pulmonary diseases: a systematic review. Respiratory Research. 2012 2012/12/21;13(1):117.
8. van der Schee MP, Paff T, Brinkman P, van Aalderen WMC, Haarman EG, Sterk PJ. Breathomics in lung disease. Chest. 2015 Jan;147(1):224-31.
9. Terrington DL, Hayton C, Peel A, Fowler SJ, Fraser W, Wilson AM. The role of measuring exhaled breath biomarkers in sarcoidosis: a systematic review. J Breath Res. 2019 Jun 24;13(3):036015.
10. Fijten RRR, Smolinska A, Drent M, Dallinga JW, Mostard R, Pachen DM, et al. The necessity of external validation in exhaled breath research: a case study of sarcoidosis. J Breath Res. 2017 Nov 29;12(1):016004.
11. Fens N, van der Schee MP, Brinkman P, Sterk PJ. Exhaled breath analysis by electronic nose in airways disease. Established issues and key questions. Clinical & Experimental Allergy. 2013;43(7):705-15.
12. Wilson AD, Baietto M. Advances in electronic-nose technologies developed for biomedical applications. Sensors (Basel). 2011;11(1):1105-76.
13. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratú P, Resta O, et al. An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. Respir Med. 2013 Jul;107(7):1073-8.
14. Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. Am J Respir Crit Care Med. 1999 Aug;160(2):736-55.
15. Travis WD, Costabel U, Hansell DM, Jr. TEK, Lynch DA, Andrew GN, et al. An Official American Thoracic Society/European Respiratory Society Statement: Update of the International Multidisciplinary Classification of the Idiopathic Interstitial Pneumonias. American Journal of Respiratory and Critical Care Medicine. 2013;188(6):733-48.

16. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2018;198(5):e44-e68.

17. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee A-H, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. European Respiratory Journal. 2021;57(1):2002042.

18. de Vries R, Brinkman P, van der Schee MP, Fens N, Dijkers E, Bootsma SK, et al. Integration of electronic nose technology with spirometry: validation of a new approach for exhaled breath analysis. J Breath Res. 2015 Oct 15;9(4):046001.

19. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. European Respiratory Journal. 2018;51(1):1701817.

20. Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics. 2006 2006/12/01;2(4):171-96.

21. Azim A, Barber C, Dennison P, Riley J, Howarth P. Exhaled volatile organic compounds in adult asthma: a systematic review. European Respiratory Journal. 2019;54(3):1900056.

22. Brinkman P, Wagener AH, Hekking P-P, Bansal AT, Maitland-van der Zee A-H, Wang Y, et al. Identification and prospective stability of electronic nose (eNose)–derived inflammatory phenotypes in patients with severe asthma. Journal of Allergy and Clinical Immunology. 2019 2019/05/01/;143(5):1811-20.e7.

23. Buendia-Roldan I, Aguilar-Duran H, Johannson KA, Selman M. Comparing the Performance of Two Recommended Criteria for Establishing a Diagnosis for Hypersensitivity Pneumonitis. Am J Respir Crit Care Med. 2021 Jul 7.

24. Bruins M, Rahim Z, Bos A, van de Sande WW, Endtz HP, van Belkum A. Diagnosis of active tuberculosis by e-nose analysis of exhaled air. Tuberculosis (Edinb). 2013 Mar;93(2):232-8.

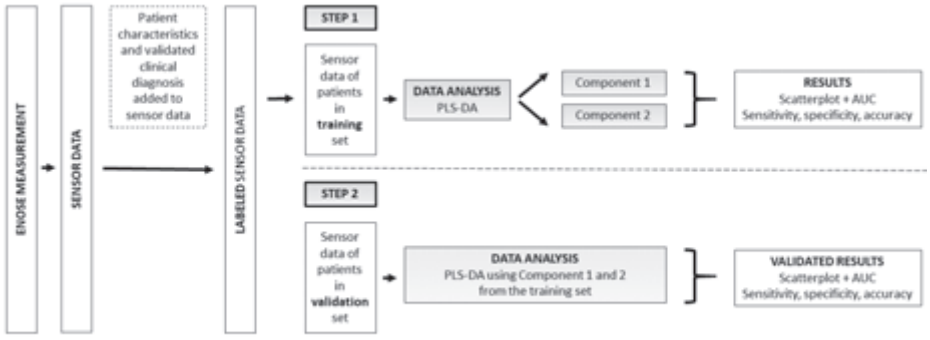
25. Saktiawati AMI, Stienstra Y, Subronto YW, Rintiswati N, Sumardi, Gerritsen J-W, et al. Sensitivity and specificity of an electronic nose in diagnosing pulmonary tuberculosis among patients with suspected tuberculosis. PLOS ONE. 2019;14(6):e0217963.

26. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Medicine. 2015 2015/01/06;13(1):1.

E-APPENDIX 1

eNose data collection, processing and analysis

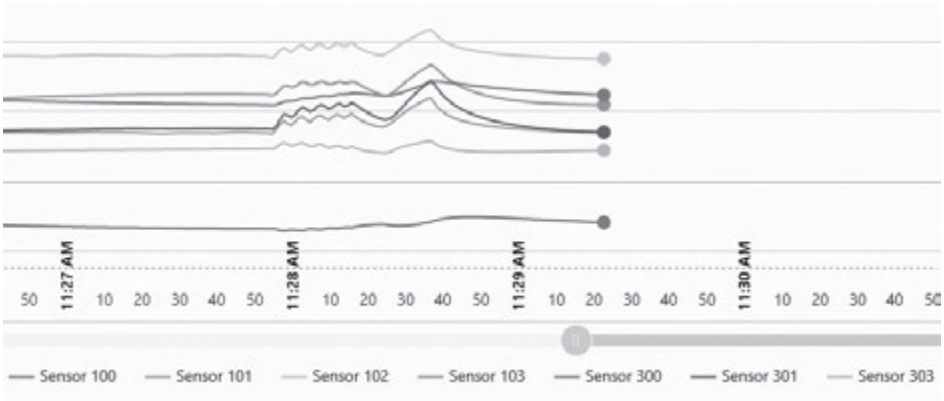
e-Figure 1 provides a summary of the whole process of eNose data collection and analysis. Steps are further explained in the text of this e-Appendix.



e-Figure 1: Summary of eNose data collection and analysis in a separate training and validation dataset, explaining the methods used in the current paper.

Data collection

When performing an eNose measurement, the patients exhales into the eNose; the sensor deviations as a response to the exhaled breath are displayed as seven signal lines in the online platform. An example is displayed in e-Figure 2.



e-Figure 2: Example of sensor signals appearing in real-time during a eNose measurement using the BreathBase platform connected to the SpiroNose.
The first five small deviations indicate five tidal breaths, the consecutive large deviation indicates the maximal inhalation (descending line), breath hold (lowest point) and exhalation (ascending line) until the sensor peak value (highest point).

Data processing

After all eNose measurements are performed, raw sensor data are exported combined with clinical patient data. Raw sensor data are processed as follows: sensor signals are corrected for ambient air, peak values are normalized to sensor 2 and inter-array differences are reduced [1, 2]. Examples of the sensor data are displayed in e-Figure 3. The sensor data contain sensor peak values and peak to breath hold (BH) ratios. Peak values are the peaks of the sensor deflections, displayed as the peak of the lines in e-Figure 2. As sensor deflections are a sensor-to-sensor ratio (normalized to sensor 2, which has the value 1.00), they have no absolute values and no unit. The sensor data shown in e-Figure 3 are used as input for data analyses in statistical software programs.

CV	CW	CX	CY	CZ	DA	DB	DC	DD	DE	DF	DG	DH	DI
Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	S1/BH	S2/BH	S3/BH	S4/BH	S5/BH	S6/BH	S7/BH
0.97	1.00	0.59	1.36	0.70	1.44	0.29	0.16	0.14	0.06	0.10	0.29	0.10	0.20
1.72	1.00	0.78	1.26	0.64	1.27	0.16	0.15	0.04	0.04	0.05	0.16	0.06	0.65
0.68	1.00	0.61	1.50	0.45	1.52	0.35	0.15	0.09	0.10	0.15	0.28	0.13	0.04
1.69	1.00	0.64	1.23	1.01	1.31	0.29	0.20	0.06	0.03	0.05	0.20	0.06	0.29
0.96	1.00	0.59	1.33	0.68	1.45	0.29	0.15	0.10	0.08	0.09	0.29	0.11	0.21
1.68	1.00	1.20	1.24	0.75	1.29	0.16	0.15	0.09	0.06	0.16	0.20	0.15	0.41

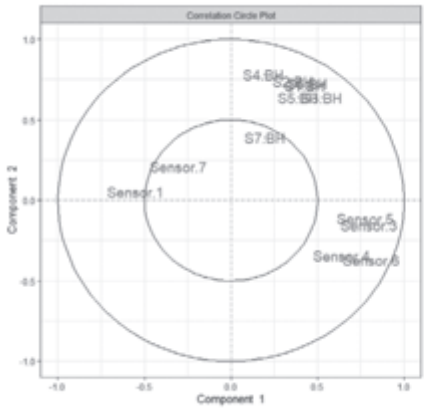
e-Figure 3: Example of processed sensor variables that serves as input for statistical analyses.
Sensor x = peak sensor value; Sx/BH = ratio of the peak sensor value divided to the breath hold value.

Data analysis

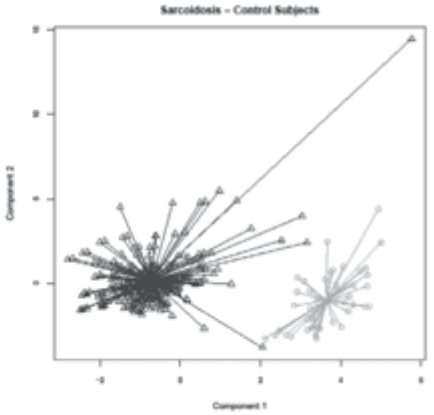
Analyses are performed with R version 4.0.3 for Mac OS X GUI (PBC, Boston, MA, USA) using the mixOmics package version 6.14.0 [3, 4]. In this paper, we use a supervised machine learning technique called partial least square discriminant analysis (PLS-DA) to examine group differences. As this is a supervised analysis, the sensor data are labelled by the investigator before the analysis with diagnosis and other clinical information of the patient.

PLS-DA is a frequently used validated machine learning method that objectively reduces the dimensionality of data. The model reduces the data derived from all sensors to different individual components. In our analyses we have chosen to use two components, because the first two components resulting from an PLS-DA analysis explain the variance of data best. A component is a combination of the weighted values of all sensors. For all individual diseases (with other exhaled VOCs), the components will be different, as different sensors have higher or lower discriminative values to specific exhaled VOCs. So for each new analysis, components 1 and 2 are created based on the most discriminative sensor values. In the analysis of sarcoidosis versus healthy controls for example, the sensors 3, 4, 5 and 6 contribute the most to component 1, followed by sensor 1, etcetera. This example is visualized in e-Figure 4 with a correlation circle plot.

The calculated components are used to create scatterplots (e-Figure 5) and the receiver operating characteristic (ROC) curves to show the ability to differentiate between two diagnosis groups. In a scatterplot, component 1 is displayed on the x-axis and 2 on the y-axis. Each dot represents one patient and the center of the dot cloud represents the mean value of the components per patient group. Before calculating the ROC curve, we first apply a generalized linear model prediction method to the two PLS-DA components. ROC analysis is consequently performed using the odds (value between 0 and 1) that a patient belongs to either of the groups. AUC values with corresponding 95% confidence intervals, sensitivity, specificity, accuracy, negative predictive values and positive predictive values are derived from that analysis.



e-Figure 4: Example of a correlation circle, showing the correlation between sensor variables and PLS-DA components; each point corresponds to a sensor variable.



e-Figure 5: Example of a scatterplot of two groups of data with component 1 and 2 on the x- and y-axis; each point corresponds to a patient.

Training and validation

In this paper, the data is randomly divided in a training and validation set (2:1) using the 'sample' function in R. PLS-DA was first performed on the training set, which is displayed in e-Figure 1 as 'Step 1'. This analysis resulted in two components. The analysis is repeated in the validation set, with use of the two components derived from the training set ('Step 2'). In this way, the components are validated in a new set of labelled sensor data.

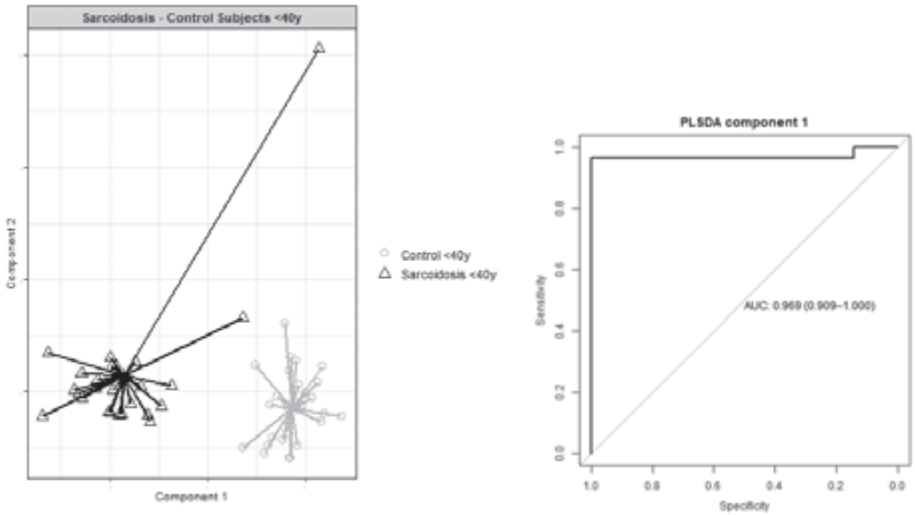
References

1. de Vries R, Brinkman P, van der Schee MP, Fens N, Dijkers E, Bootsma SK, et al. Integration of electronic nose technology with spirometry: validation of a new approach for exhaled breath analysis. J Breath Res. 2015 Oct 15;9(4):046001.
2. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. European Respiratory Journal. 2018;51(1):1701817.
3. Barker M, Rayens W. Partial least squares for discrimination. Journal of Chemometrics. 2003;17(3):166-73.
4. Rohart F, Gautier B, Singh A, KA LC. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput Biol. 2017 Nov;13(11):e1005752.

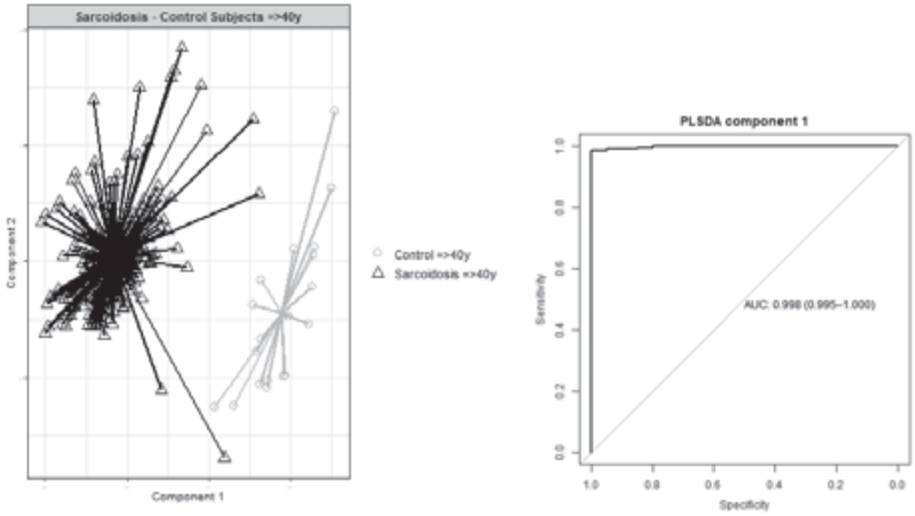
E-APPENDIX 2

Additional subgroup analyses

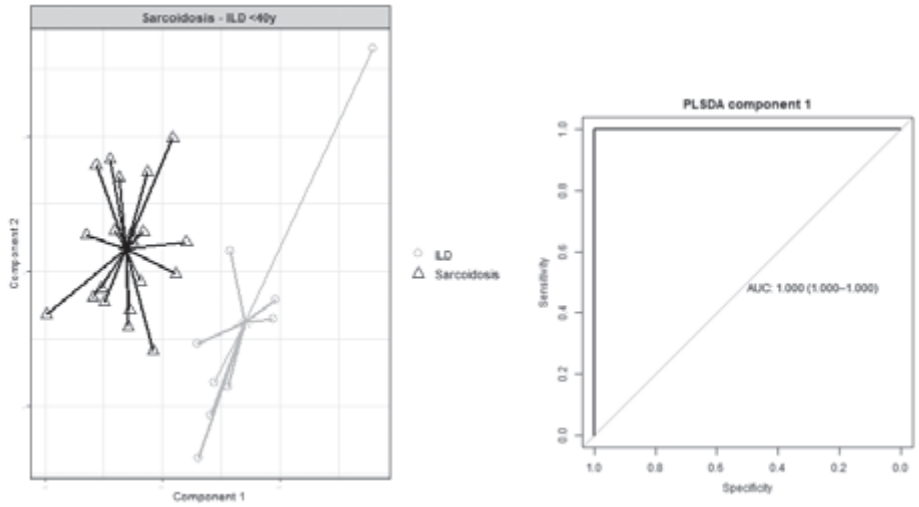
Influence of age on breathprints



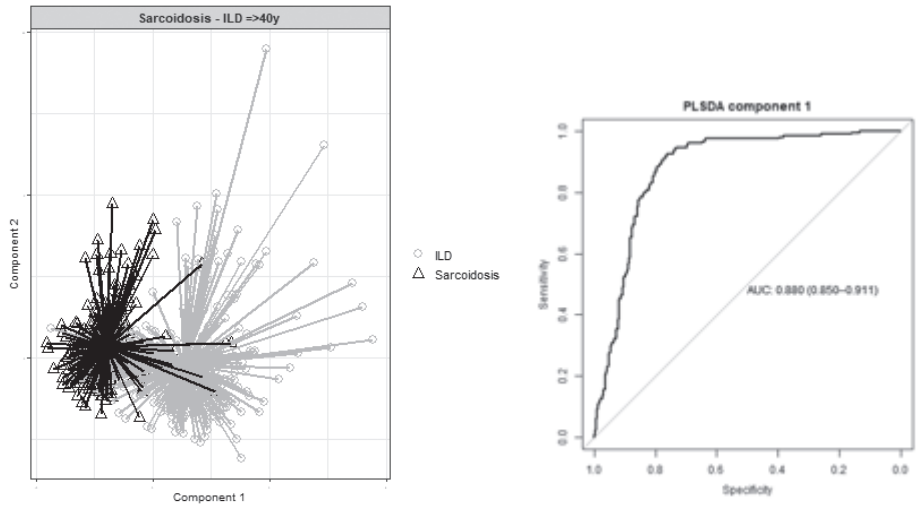
e-Figure 6: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients (n=28) and healthy controls (n=28) below 40 years.



e-Figure 7: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients (n=224) and healthy controls (n=20) above or equal to 40 years.

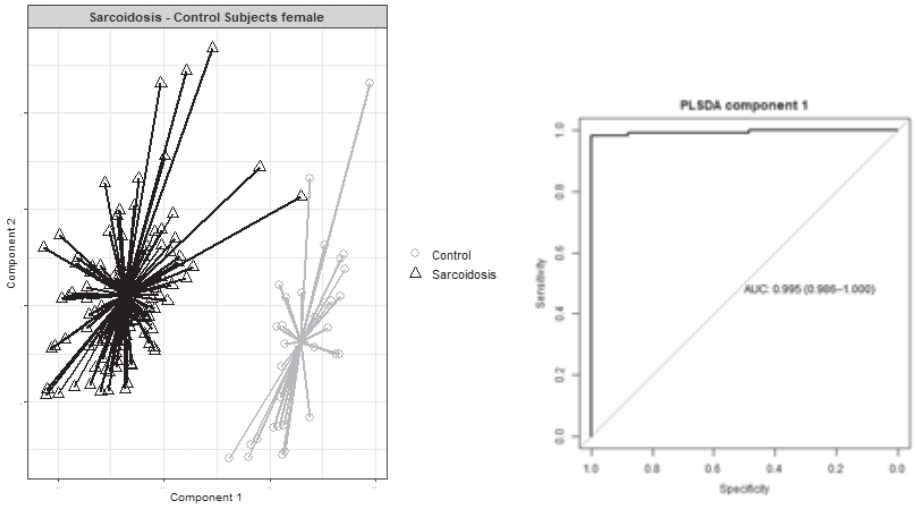


e-Figure 8: Scatterplot and corresponding ROC curve of eNose data from pulmonary sarcoidosis patients (n=20) and ILD patients (n=9) below 40 years.

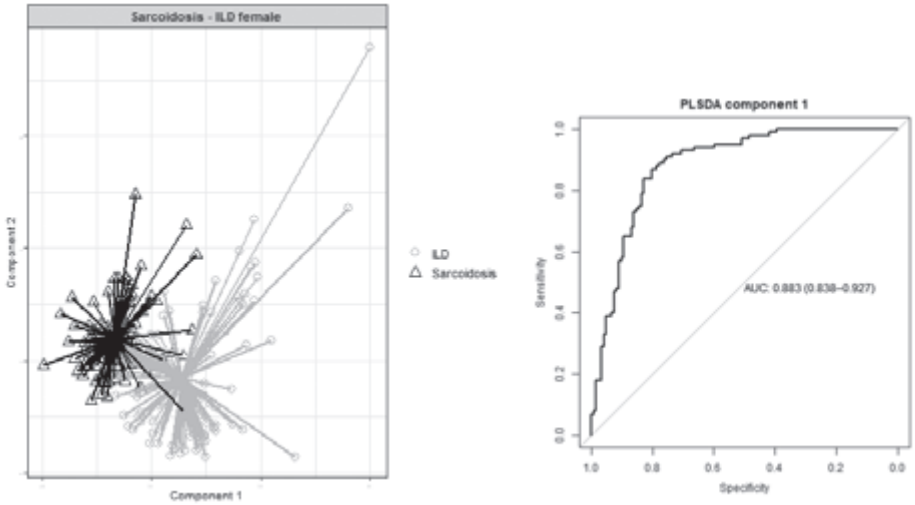


e-Figure 9: Scatterplot and corresponding ROC curve of eNose data from pulmonary sarcoidosis patients (n=204) and ILD patients (n=308) above or equal to 40 years.

Influence of gender on breathprints

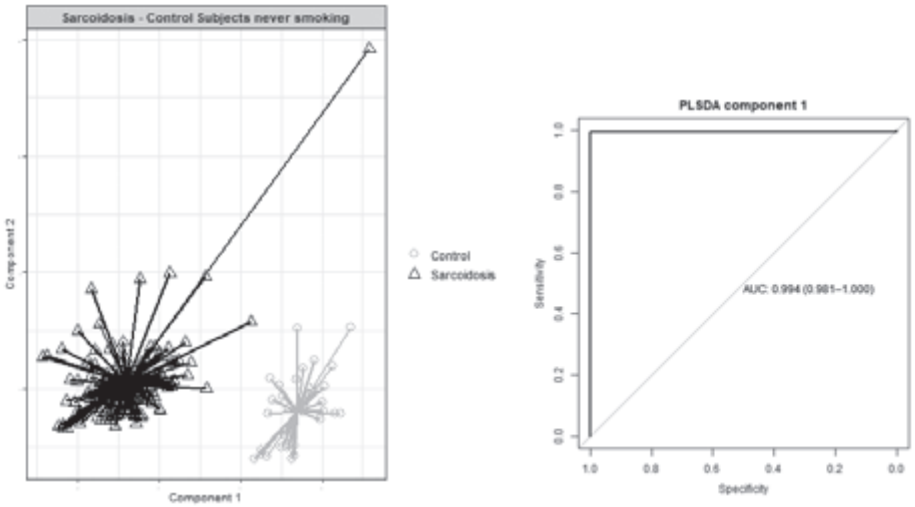


e-Figure 10: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients (n=118) and healthy controls (n=33), female patients only.

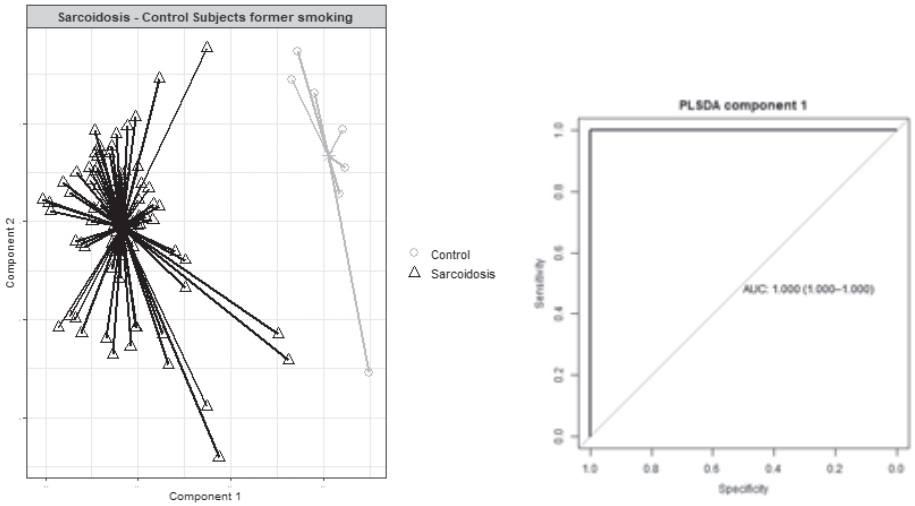


e-Figure 11: Scatterplot and corresponding ROC curve of eNose data from pulmonary sarcoidosis patients (n=100) and ILD patients (n=122), female patients only.

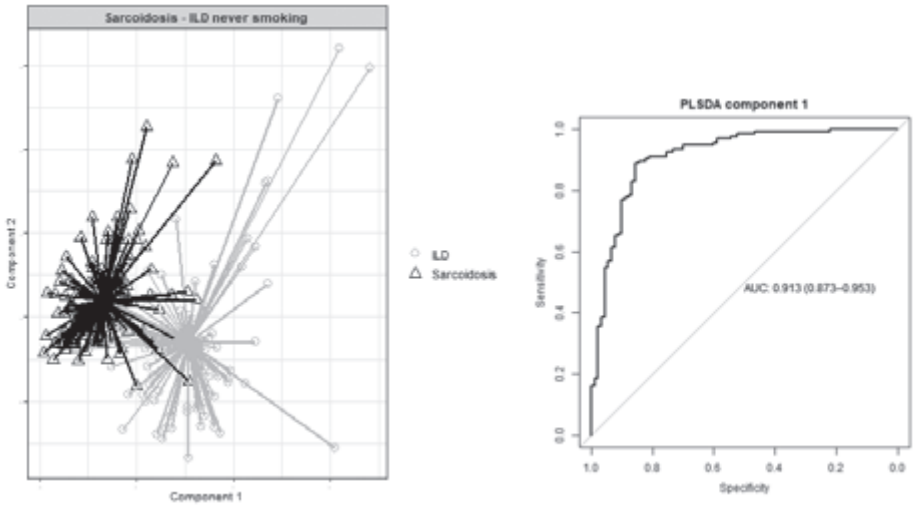
Influence of smoking status on breathprints



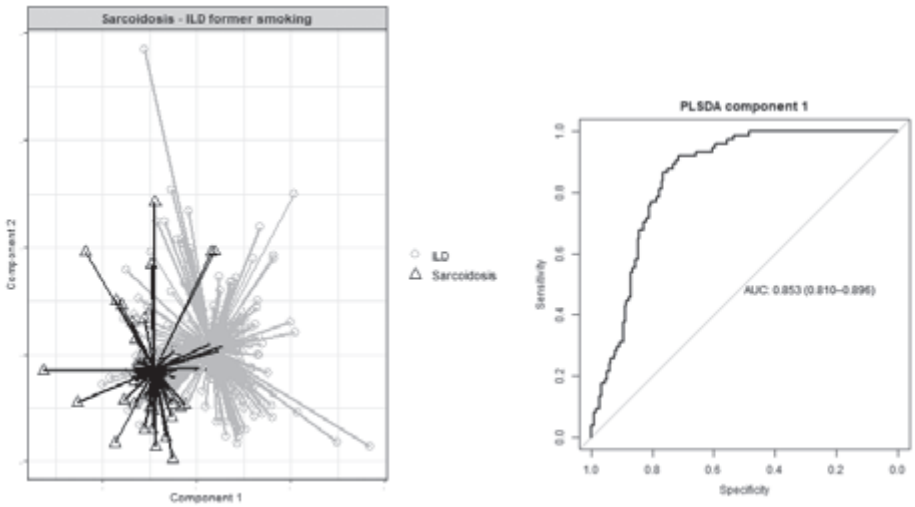
e-Figure 12: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients (n=154) and healthy controls (n=37), never smoking.



e-Figure 13: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients (n=83) and healthy controls (n=7), former smoking.

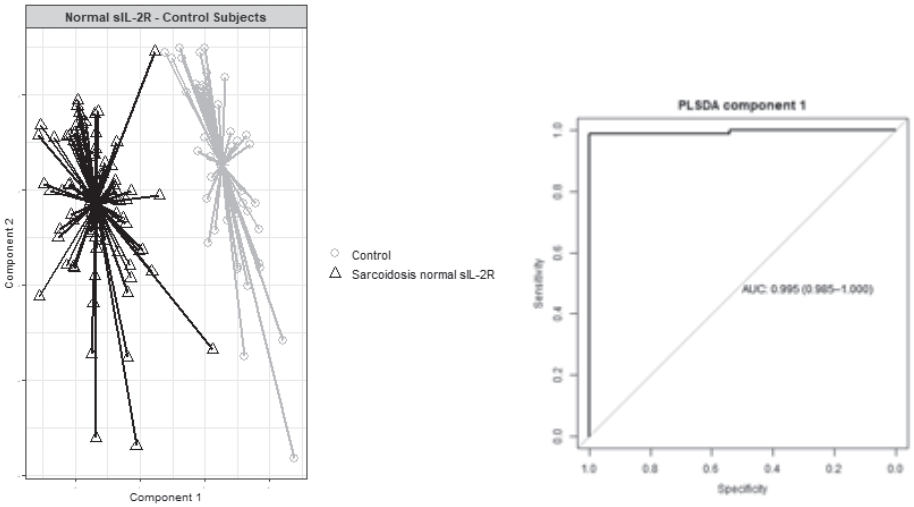


e-Figure 14: Scatterplot and corresponding ROC curve of eNose data from pulmonary sarcoidosis patients (n=135) and ILD patients (n=90), never smoking.

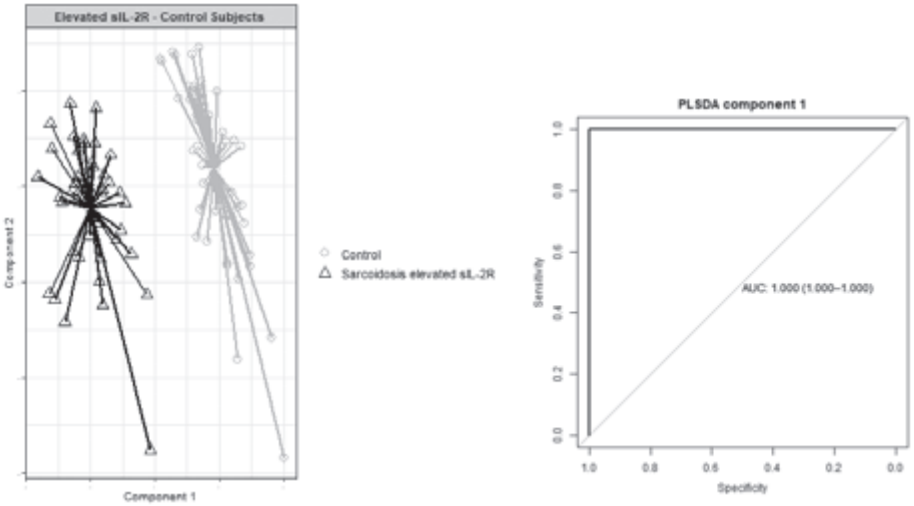


e-Figure 15: Scatterplot and corresponding ROC curve of eNose data from pulmonary sarcoidosis patients (n=74) and ILD patients (n=217), former smoking.

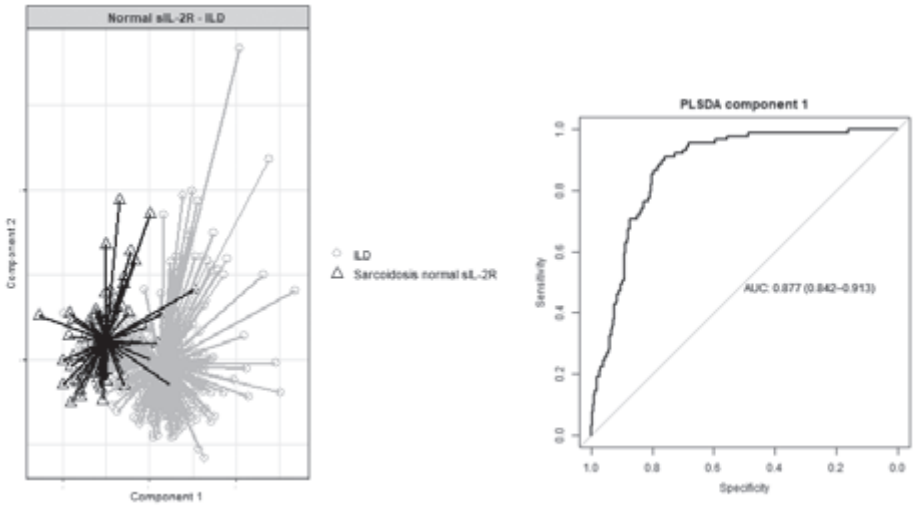
Influence of sIL-2R level on breathprints



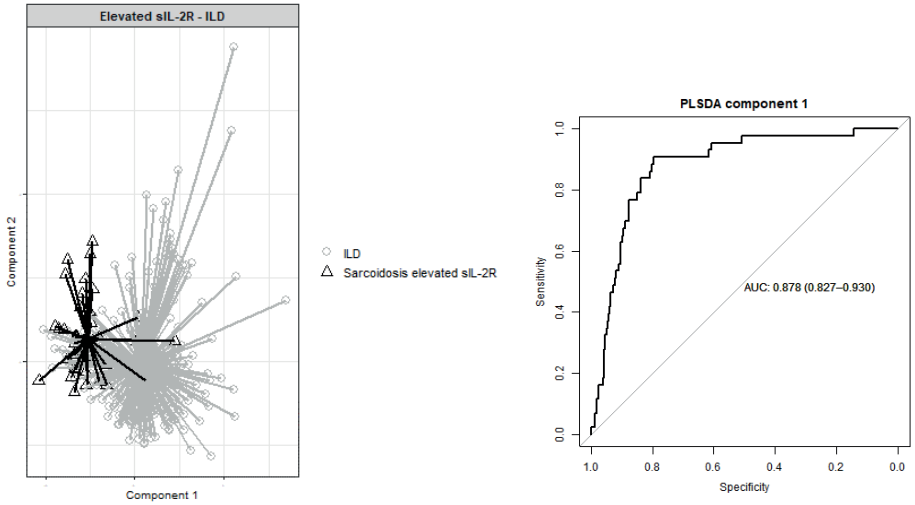
e-Figure 16: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients with normal sIL-2R level (n=89) and healthy controls (n=48).



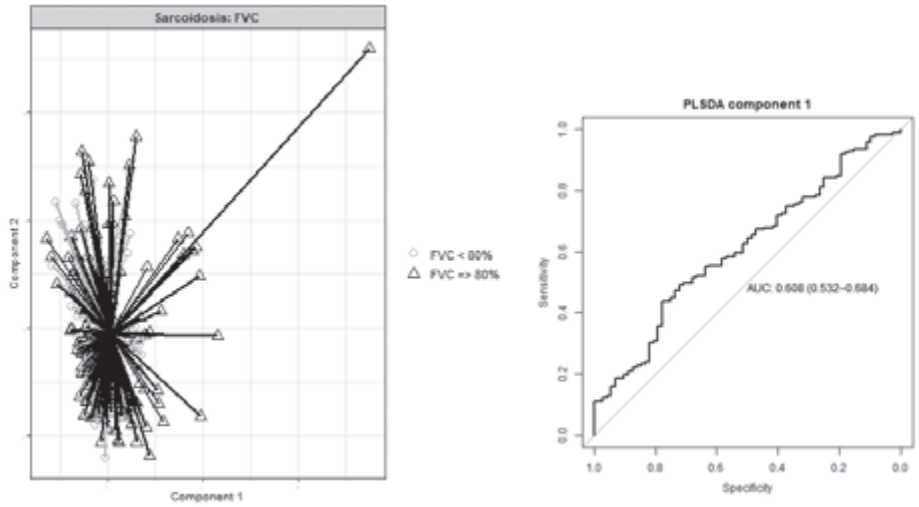
e-Figure 17: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients with elevated sIL-2R level (n=43) and healthy controls (n=48).



e-Figure 18: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients with normal sIL-2R level (n=89) and ILD patients (n=317).

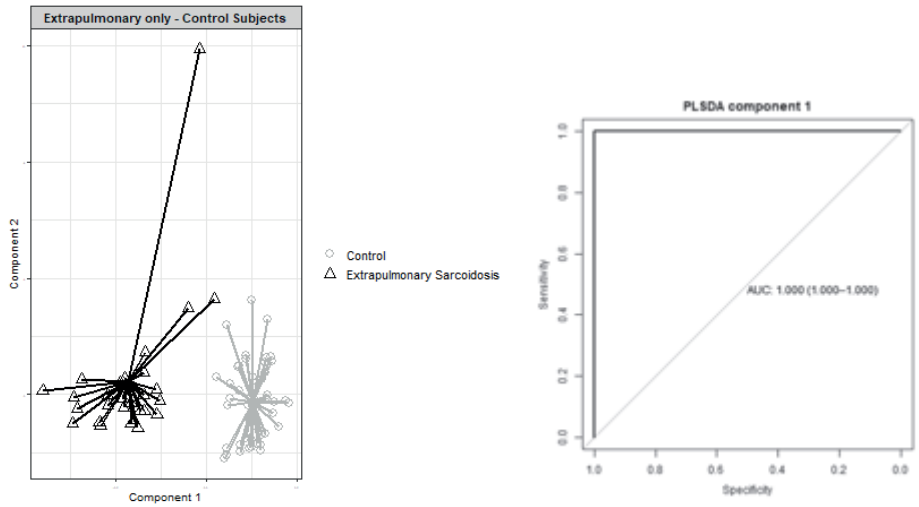


e-Figure 19: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients with elevated sIL-2R level (n=43) and ILD patients (n=317).



e-Figure 20: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients divided by lung function: normal (≥80%, n=171) and abnormal forced vital capacity (<80%, n=72).

Influence of sarcoidosis organ involvement on breathprints



e-Figure 21: Scatterplot and corresponding ROC curve of eNose data from sarcoidosis patients with extrapulmonary disease only (n=28) and healthy controls (n=48).

CHAPTER 7

Evaluation of different classification methods using electronic nose data to diagnose sarcoidosis

J Breath Res. 2023 Aug 29;17(4).

Iris G. van der Sar[1], Nynke van Jaarsveld[1,2,3], Imme A. Spiekerman[1,2,3], Floor J. Toxopeus[1,2,3], Quint L. Langens[1,2,3], Marlies S. Wijsenbeek[1], Justin Dauwels[3#], Catharina C. Moor[1#]

[1]Erasmus Medical Center, Rotterdam (the Netherlands); [2]Leiden University Medical Center, Leiden (the Netherlands); authors contributed equally; [3]Delft University of Technology, Delft (the Netherlands); [#]Authors share last authorship.

Abstract

Introduction

Electronic nose (eNose) technology is an emerging diagnostic application, using artificial intelligence (AI) to classify human breath patterns. These patterns can be used to diagnose medical conditions. Sarcoidosis is an often difficult to diagnose disease, as no standard procedure or conclusive test exists. An accurate diagnostic model based on eNose data could therefore be helpful in clinical decision-making.

Aim

The aim of this paper is to evaluate the performance of various dimensionality reduction methods and classifiers in order to design an accurate diagnostic model for sarcoidosis.

Methods

Various methods of dimensionality reduction and multiple hyperparameter optimised classifiers were tested and cross-validated on a dataset of patients with pulmonary sarcoidosis (n=224) and other interstitial lung disease (n=317). Best performing methods were selected to create a model to diagnose patients with sarcoidosis. Nested cross-validation was applied to calculate the overall diagnostic performance.

Results

A classification model with feature selection and random forest classifier showed the highest accuracy. The overall diagnostic performance resulted in an accuracy of 87.1% and area-under-the-curve of 91.2%.

Conclusion

After comparing different dimensionality reduction methods and classifiers, a highly accurate model to diagnose a patient with sarcoidosis using eNose data was created. The random forest classifier and feature selection showed the best performance. The presented systematic approach could also be applied to other eNose datasets to compare methods and select the optimal diagnostic model.

Keywords

Electronic nose, breath analysis, classification model, diagnostic test, sarcoidosis, interstitial lung disease, volatile organic compounds

Introduction

New applications of artificial intelligence (AI) in pulmonary medicine have been increasingly studied and published over the last years. However, no applications have yet been approved for use in clinical practice. Investigated applications range from automatic interpretation of pulmonary function tests and chest computed tomography scans, to predicting disease exacerbations using home monitoring data [1]. AI models sometimes achieve the accuracy level of human experts [2]. Therefore, it is likely that AI will support clinical decision making in the near future.

Electronic nose (eNose) technology is one of the upcoming new technologies for clinical practice that uses AI. An eNose device analyses exhaled breath in real-time, using multiple cross-reactive sensors with different sensitivities. By using classification models to categorize generated sensor data, the eNose device has the potential to be used as non-invasive diagnostic tool. Hence, different eNose devices and clinical applications are currently studied in the field of pulmonary medicine [3].

Interstitial lung diseases (ILDs) comprise a large group of heterogeneous rare individual diseases that affect the interstitium of the lungs. Patients usually present with non-specific symptoms, and disease course and response to therapy widely varies. Sarcoidosis, a form of ILD, is a multisystem granulomatous disease with lung involvement occurring in 89-99% of patients [4]. In the current guidelines, three main criteria are proposed to diagnose sarcoidosis: a compatible clinical presentation, the finding of nonnecrotizing granulomatous inflammation in tissue samples, and the exclusion of alternative causes of granulomatous disease [5]. However, no objective measures exist to judge whether these criteria are satisfied. Consequently, the established consensus diagnosis always contains a certain margin of uncertainty for each individual, despite multiple diagnostic test, often including invasive tissue biopsy. Therefore, accurate, non-invasive and fast diagnostic modalities are highly needed.

Studies that tested performance of eNose technology as a diagnostic tool for ILD show accuracies varying from 49 to 100% [6-12]. The large spread might be explained by differences in study design and eNose devices. Moreover, these studies used different classifiers to analyse the sensor data: Neural Networks (NN), Canonical Discriminant Analysis (CDA), K-Nearest Neighbour (KNN), Linear Discriminant Analysis (LDA), Partial Least Squares Discriminant Analysis (PLS-DA), Random Forest (RF), Support Vector Machines (SVM), and Extreme Gradient Boosting (XGBoost). We previously showed that PLS-DA accurately distinguished sarcoidosis from other forms of ILD, but we did not evaluate the performance of different classifiers or models [10, 11].

In the field of machine learning, various models are usually compared before selecting a final machine learning model [13]. This might also be a good approach for clinical eNose research, as performance might differ per dataset and classification model [14]. Until now, only two eNose studies in ILD evaluated multiple models. They showed fair and comparable model performance on training datasets, but performance in test and validation sets varied [7, 12].

The main aim of this paper is to evaluate the performance of various dimensionality reduction methods and classifiers to design the most accurate diagnostic model for sarcoidosis.

Methods

Dataset and materials

The used dataset includes eNose sensor and clinical data of patients with pulmonary sarcoidosis (n=224) and patients with other ILDs (n=317) from the Erasmus Medical Centre (Rotterdam, the Netherlands). Clinical characteristics have been published previously [11]. We collected exhaled breath data using the SpiroNose (Breathomix, Leiden, Netherlands) which is connected to an online secured platform and database called BreathBase. Breath manoeuvres were performed in duplicate. Each manoeuvre included five tidal breaths, followed by a maximal inhalation to vital capacity, five-second breath hold and slow maximal exhalation leading to a sensor peak value. During the measurements, a mouthpiece with bacterial filter (Pulmosafe 3, Lemon Medical GmbH, Hammelburg, Germany) and a nose clamp were used. The investigator checks the quality of each measurement in real-time during the breath manoeuvre by inspecting the sensor deviation curves that appear in BreathBase. The investigator can provide the patient feedback for the second manoeuvre if necessary. Specifications of the device and manoeuvres have previously been published and are specified in **Supplementary data A** [15]. Sensor characteristics, system verification procedures, and conditions and contra-indications for using the device are also described in **Supplementary data A**.

The SpiroNose contains seven different metal oxide semiconductor sensors, present in duplicate on the inside and outside of the device. After data pre-processing (including scaling and correction for ambient air), both the sensor peak value and peak to breath-hold ratio are extracted from each sensor signal, leading to 14 sensor values per patient. The peak value of sensor 2 is set to a constant value and is used for scaling of the other sensor values. The peak value of sensor 2 does not serve as an input variable. The data processing has been described previously [15]. **Figure S1** and **S2** in

Supplementary data B shows some examples of sensor diagrams and corresponding input variables.

Clinical characteristics were obtained from medical files and patient questionnaires. The study protocol was approved by the local ethical committee of Erasmus Medical Center (MEC-2019-0230). Analyses were conducted in Matlab (version R2021b), Statistics and Machine Learning Toolbox [16]. The final script to generate the results of this paper was run in June 2022. The full Matlab scripts are freely available on request.

Model design and testing

Based on previously published eNose studies and compatibilities of Matlab, classifiers k-NN, LDA, NN, RF, and SVM were selected for evaluation of their binary classification performance using eNose sensor data of patients with sarcoidosis and ILD. The overall process of model design and evaluation consisted of several consecutive steps:

1. Testing several methods of dimensionality reduction to select the best performing method to train the model (Fig. 1A);
2. Training and testing several hyperparameter optimised classifiers using 10-fold cross-validation to select the most accurate classifier to train the model (Fig. 1B);
3. Validating the overall diagnostic performance of the model using nested cross-validation (Fig. 1C);
4. Applying the trained final model on random patients to show the individual diagnostic probability;
5. Assessing the sufficiency of dataset size by calculating model accuracies on increasing sample size proportions.

Dimensionality reduction

First, the dimensionality of the dataset was reduced using feature selection or feature extraction, and this was compared to using no dimensionality reduction. The input variables (i.e. features) were the 13 peak sensor and peak to breath-hold values per eNose measurement of a patient. All three methods were tested on 80% of the data using 10-fold cross-validation (**Figure 1A**). The method with the highest cross-validated accuracy (CVA) was implemented in training the final model. The dimensionality reduction cross-validation was performed once, as the outcome of dimensionality reduction depends on the dataset itself, not on the classifier [17, 18].

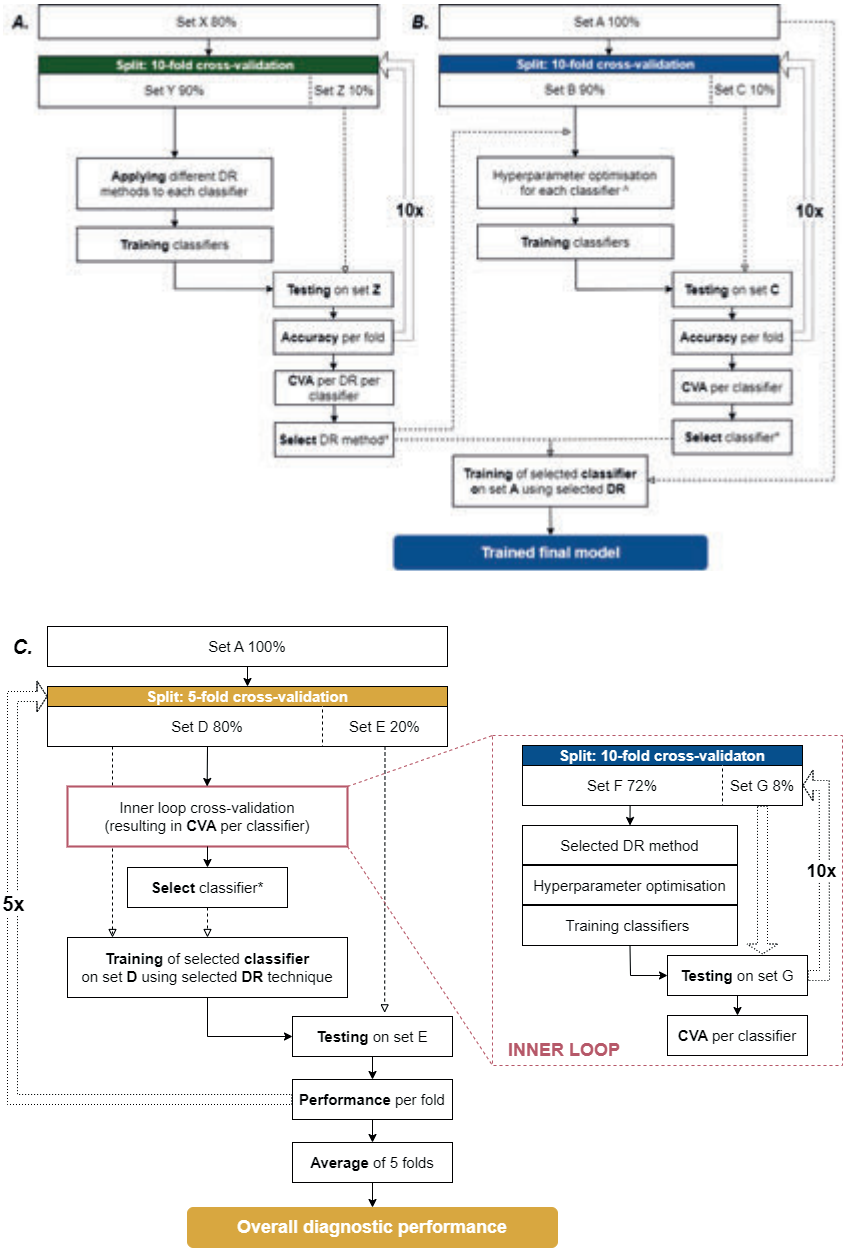


Figure 1: Summary of model design and evaluation.
*A. Testing several methods of dimensionality reduction to select the best performing method. B. Training and testing different classifiers to select the best performing method and train the final model. C. Validating the overall diagnostic performance of the model using nested cross-validation. Cross-validation in the inner loop is similar to the cross-validation in B. *The classifier or DR technique resulting in the best CVA is selected. ^Using Bayesian optimisation. CVA = cross-validated accuracy; DR = dimensionality reduction*

Feature extraction was performed using Principal Component Analysis (PCA) with Matlab's function pca [19]. PCA results in a set of multivariate components, where each component is a combination of the original 13 sensor values. The first PCA component explains the greatest variance of the data and the last PCA component the least. To determine which components to include, percentage of variability thresholds of ≥ 90%, 95% or 99% were compared. The singular value decomposition algorithm was selected within Matlab's pca function.

Feature selection was performed with Matlab's fscchi2 function [20]. This function was used to calculate the weight of each feature by taking the negative logarithm of the p-value resulting from a chi-squared test. This weight represents the extent to which a single feature influences the outcome of the model; a higher score indicates more influence. Following the feature weights calculation, various weight thresholds were tested to select a certain number of contributing features. Three thresholds that resulted in five up to ten contributing features were eventually tested.

Hyperparameter optimisation

In each fold, hyperparameter optimisation was executed while training each classifier. This was done by setting the option OptimizeHyperparameters in each classifier to 'auto'. This led to 2 to 4 parameters being optimised per classifier. The type of parameters depended on which classifier was being trained. Specifications of the optimizations can be found in **Supplementary data C**. Using Bayesian optimisation, the 5-fold cross-validated loss per set of hyperparameters was calculated over 30 iterations [21]. The set with the minimal cross-validation loss was selected.

The RF method required several other parameters to be defined in the function fitcensemble [22]. In Matlab, the type of learner method was set as 'decision tree' and the aggregation method as 'bag'. Bootstrap aggregation (i.e. bagging) reduces the variance of weak learners such as RF. This specification cannot be combined with the OptimizeHyperparameters option. Thus, hyperparameter optimisation was performed separately using the Bayesopt function in the same manner as for the other classifiers [21].

Model training, testing and selecting

To select the most accurate classifier, 10-fold cross-validation was performed on the full dataset (set A) for each classifier using the selected dimensionality reduction method (section 2.2.1). The data splits for 10-fold cross-validation were made using the function cv-partition [23]. Nine folds formed the training set (set B) and the other fold the test set (set C). The CVA per classifier was calculated as the average of the accuracies of the ten folds and included a range (i.e. minimum and maximum accuracy of the folds). The classifier with the highest CVA was selected and trained on set A. Hyperparameter optimisation was executed anew. For SVM, the selected kernel type was 'linear'. This resulted in the final trained model to classify patients based on eNose data (**Figure 1B**).

Diagnostic performance calculation

The overall diagnostic performance of this final model was determined by repeating the initial 10-fold cross-validation within a 5-fold cross-validation (i.e. nested cross-validation) including all five classifiers (k-NN, LDA, NN, RF, and SVM) (**Figure 1C**). Nested cross-validation leads to less bias than single-loop cross-validation used for training the final model (section 2.2.3) as the results do not depend on a single data split [24].

To execute this validation method, the full dataset (set A) was split into five folds resulting in four folds representing 80% (set D) and one representing 20% of the data (set E), the so-called outer loop. Set D was used for inner loop 10-fold cross-validation and therefore divided into set F and G (**Figure 1C**). These sets F and G underwent the exact same process as the initial training and test sets B and C (**Figure 1B**).

Each of the five folds resulted in a best performing classifier, and this classifier was subsequently trained on set D and tested on set E to calculate the diagnostic performance (accuracy, specificity, sensitivity and AUC values) of that fold using Matlab's function confusionmat. The accuracy was calculated as (true negatives + true positives) / (true negatives + true positives + false negatives + true positives), specificity as true negatives / (true negatives + false positives), and sensitivity as true positives / (true positives + false negatives). Finally, the overall diagnostic performance of the model was the average of these values of the five folds.

Classifying individual patients

To simulate the model's ability in diagnosing a 'new' patient, the trained final model was applied to sensor data from random patients from set A. The model predicted for an individual patient the class it belongs to (sarcoidosis or ILD), including the

probability of this prediction and the time it took to complete the prediction. A higher probability means a higher likelihood of the prediction being correct for this individual. The probability was calculated by multiplying the prior probability with multivariate normal density and expressed as an percentage [25].

Evaluation size dataset

In order to evaluate whether the final model would benefit by training on more data or if sample size was sufficient, the model's accuracy was calculated for smaller training dataset sizes. The final trained model was repeatedly trained on an increasing proportion of data to calculate the corresponding accuracy.
The entire dataset was first split into a new training (90%) and test set (10% of data). The model was trained using 1 up to 100% of the training data, each attempt increasing with 1%. The corresponding accuracy was tested using the full test set. Training and testing was repeated 20 times per proportion of training data, resulting in an average accuracy per proportion used.

Results

Dimensionality reduction

CVA values resulting from the five different classifiers after applying 'feature selection', 'feature extraction' (i.e. PCA) and 'no dimensionality reduction' are shown in **Figure 2**. A feature selection weight threshold set on 1 resulted in 10 features, 3 in 6, and 5 in 5 features. A threshold of 1 resulted in the highest CVA in four out of five classifiers. Therefore, this method was chosen for dimensionality reduction to implement in the final model. The weights of each feature are shown in **Figure 3**. The peak value of sensor 3 and 5, and peak to breath-hold ratio of sensor 2 did not reach the optimal threshold of 1, and were excluded for training the final model.

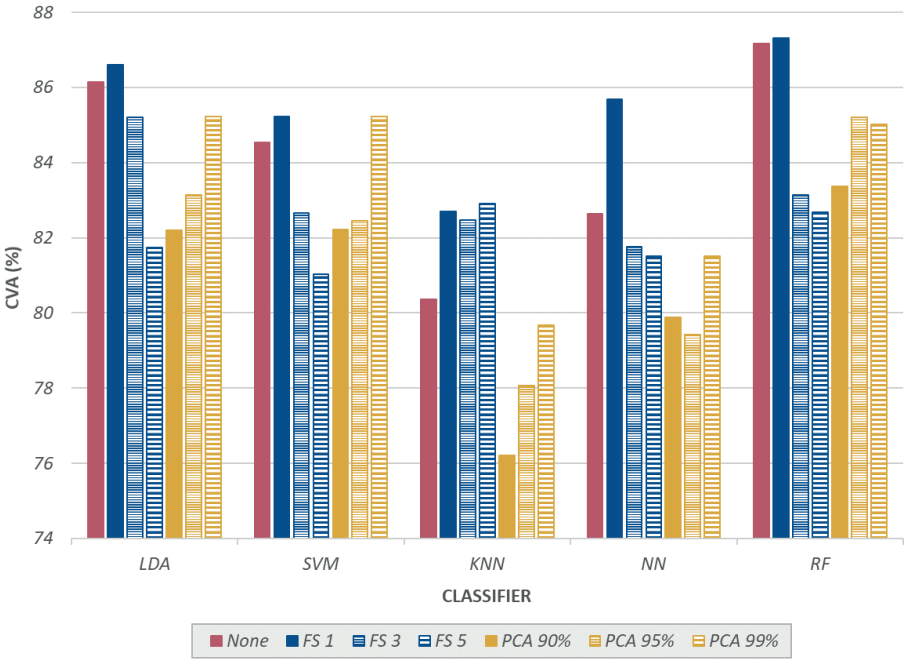


Figure 2: Cross-validated accuracy per classifier when applying different dimensionality reduction method, including three different thresholds for FS (1, 3 and 5) and PCA (90, 95 and 99%).
CVA = cross-validated accuracy; FS = feature selection; KNN = K-nearest neighbour; LDA = linear discriminant analysis; NN = neural networks; PCA = principle component analysis; RF = random forest; SVM = support vector machines.

Model training, testing and selecting

After applying feature selection with threshold of 1, CVAs for all classifiers were calculated separately. RF showed the highest CVA of 87.6% with a range of 79.6-96.3% (**Figure 4**) and was therefore selected as classifier for the final model.

Hyperparameter optimisation resulted in 100 learning cycles and 23 bins, which were used to train the random forest classifier and the final model. All data subsets in the ten folds had approximately the same class distribution of ILD and sarcoidosis as the complete dataset.

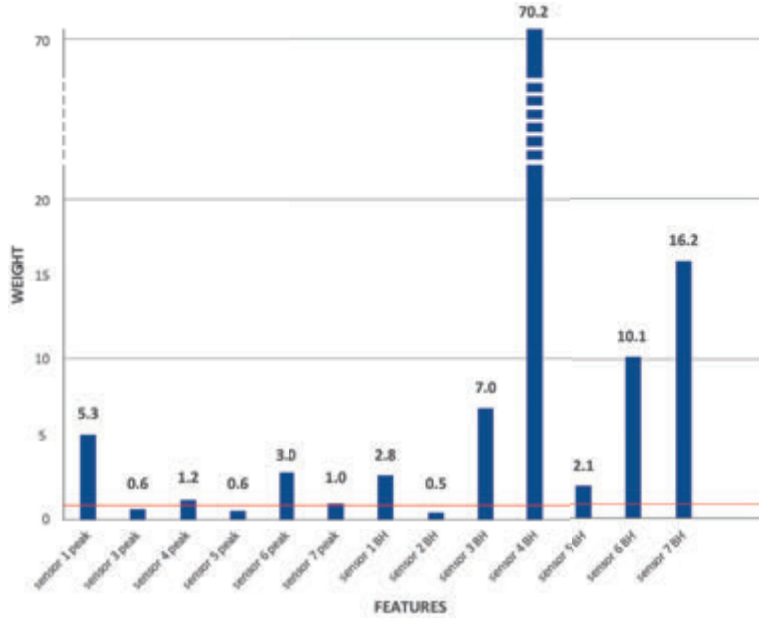


Figure 3: Weight per feature of the trained final model.
Features include peak values and peak to breath-hold ratios. Weights represent the extent to which a sensor value influences the response variable of the model. Weight = -log(p-value per feature). The red line illustrates the weight of 1 used as threshold for feature selection. BH = breath hold.

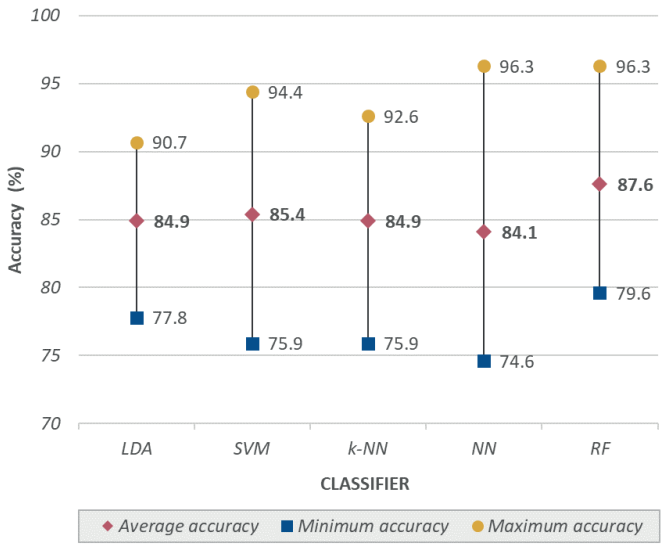


Figure 4: Comparison of the minimum, maximum, and the average accuracy of the ten folds, displayed per classifier calculated by 10-fold cross validation.
K-NN = K-nearest neighbour; LDA = linear discriminant analysis; NN = neural networks; RF = random forest; SVM = support vector machines.

Diagnostic performance calculation

The best performing classifier and the corresponding diagnostic performance values resulting from each fold of the nested cross-validation are shown in **Table 1**. RF performed best in three out of five folds. The CVA resulting from the five folds was 87.1% ranging from 80.7 to 92.6%. The average sensitivity was 91.4% (range 86.4-96.6%) and specificity 82.2% (range 74.0% - 90.5%). The AUC of the receiver operating characteristic curves varied from 83.7-96.8% with an average of 91.2%. The accuracy for each five classifiers of all five folds and the receiver operating characteristic curve resulting from each fold can be found in **Supplementary data D** (**Table S3** and **Figure S3**).

Table 1: Overall diagnostic performance of the final model displayed as the average accuracy (i.e. CVA), sensitivity, specificity and AUC of the five folds.

	Classifier	Accuracy (%)	Sensitivity	Specificity	AUC (%)
Fold 1	RF	90.7	90.9	90.5	93.9
Fold 2	SVM	80.7	86.4	74.0	83.7
Fold 3	RF	86.1	88.7	82.6	93.3
Fold 4	SVM	85.2	94.3	76.4	88.1
Fold 5	RF	92.6	96.6	87.8	96.8
Average	-	*87.1*	*91.4*	*82.2*	*91.2*
95% CI		*84.29, 89.91*	*88.99, 93.81*	*78.63, 85.77*	*90.76, 91.64*

The best performing classifier per fold was selected based on the highest accuracy. AUC = area under the curve; CI = confidence interval; CVA = cross-validated accuracy; RF = random forest; SVM = support vector machines.

Evaluation size dataset

Increasing the training dataset from 80 to 100% resulted in 0.7% accuracy improvement (87.5 to 88.2%), indicating that the model is likely trained on sufficient data. The model's accuracy when training with a smaller dataset size is shown in **Figure S4** in **Supplementary data E**.

Classifying individual patients

The model's output for each individual patient includes a diagnosis and diagnostic probability based on eNose data. An example of the model's output of 10 randomly selected individuals from the full dataset is shown in **Table 2**.

Table 2: Example of the diagnostic model's output of 10 randomly selected patients including the probability of the assigned class and the time needed to classify.

	Diagnosis	Probability (%)	Prediction time (s)
Patient 1	ILD	94	0.11
Patient 2	Sarcoidosis	93	0.09
Patient 3	ILD	89	0.09
Patient 4	ILD	88	0.13
Patient 5	Sarcoidosis	97	0.08
Patient 6	Sarcoidosis	97	0.07
Patient 7	ILD	86	0.07
Patient 8	Sarcoidosis	85	0.08
Patient 9	ILD	84	0.06
Patient 10	Sarcoidosis	95	0.06

All patients were classified correctly. ILD = interstitial lung disease; s = seconds

Discussion

In this paper, we evaluated multiple classification methods to design a highly accurate model using eNose data for diagnosing patients with pulmonary sarcoidosis within a group of patients with ILD. Different dimensionality reduction methods and classifiers were trained, tested and compared systematically. Feature selection and RF resulted in the highest diagnostic performance compared to the other methods assessed and were trained to create a final diagnostic model. Diagnostic performance resulted in a CVA of 87.1%. The presented approach for comparing different dimensionality reduction methods and classifiers to design a diagnostic eNose model has not been described previously. A strength of the designed model is the ability to show a specific diagnostic probability for an individual patient, which will facilitate translation of eNose technology into clinical practice.

When starting to design a diagnostic model for a certain condition using eNose data, the most important factor that determines model performance is whether the selected condition can be detected in exhaled breath accurately. A proof-of-concept study should clarify this first before designing a diagnostic model, like we performed for sarcoidosis previously using the PLS-DA classifier [11]. In the current comparative

analysis of classifiers, RF turned out to be the best performing classifier for this dataset. RF has been used previously to classify various medical conditions using eNose data [14, 3]. In general, the majority of eNose papers focus on a single analysis method to classify patients supervised without a clear rationale for the selected method. In this paper, we show a systematic comparative approach to justify the choice for a certain analysis method.

Although RF showed the highest accuracy, differences between classifiers were small and all showed good accuracies. When designing a model for clinical applications, also other factors besides performance have to be considered, such as speed of the model, visualization, and outcome parameters [26]. Our trained final model shows a diagnosis within 1 second for an individual patient including a diagnostic certainty. The latter is important for clinician to interpret the eNose results correctly when using this test in clinical practice.

Before implementing the eNose as a diagnostic tool for sarcoidosis in clinical practice, the current model needs to be trained and tested on an independent heterogeneous multicentre cohort including patients with various related conditions, with respiratory complaints without a diagnosis, and healthy controls matched by possible confounders (e.g. age and sex), to confirm the models robustness and to prevent overfitting [14]. Additionally, analysis of unlabelled patient data need to confirm the hypothesis of this diagnostic tool. This is in particular important for a sarcoidosis cohort due to several reasons. First, patients from different healthcare settings should be included, not only from ILD and sarcoidosis expert centres like the current cohort, as patients' characteristics and diagnostic certainty might differ. Second, given the lack of clear objective diagnostic criteria for sarcoidosis and ILD, the reached consensus diagnosis always includes some uncertainty. It is inevitable that training data of the current dataset are not 100% accurate. Moreover, the time between a patient's diagnosis and eNose measurement varied. Additionally, class frequencies are assumed a realistic representation of prior probabilities, which might vary in other care settings. Lastly, most patients have received or were receiving therapy, which could have influenced the eNose measurements. Nevertheless, previous analyses of this sarcoidosis cohort suggested that the extent of disease activity and treatment does not significantly affect the accuracy of eNose results [11].

When looking at potential clinical applications of diagnostic AI tools, including eNose technology, it is unlikely that AI will fully replace clinical decision making, as both clinicians and AI systems have unique strengths. It is well recognised that humans outperform machines in detection, perception, improvisation, long-term memory,

induction, and judgement, and machines outperform humans in response speed and precision, repetition, short-term memory, deductive reasoning, and handling complex operations [27]. Thus, especially the use of AI combined with clinical decision-making is likely to be of added value. This accounts in particular for diseases without a conclusive diagnostic test, such as sarcoidosis, where pattern recognition is of great importance.

Another prerequisite for a fruitful implementation of eNose technology in clinical practice is trust of clinicians in the capability of the technology [28]. In the current paper, we aim to provide insights to clinicians with regard to data processing, model design and performance. This will build trust in eNose technology and encourage correct interpretation of the model output. Essential for correct model output interpretation and integration in clinical decision-making is the individual diagnostic probability score provided in the current paper. Besides, clinicians should know on what data the model is trained to identify the correct patients for applying the model to.

Several limitations of the developed model and proposed method should be addressed. The PLS-DA classifier that was used for analyses in the previous proof-of-concept paper on the same sarcoidosis cohort, which led to an accuracy of 83.2% in the validation set, is not evaluated in the current paper. The way PLS-DA reduces and classifies data is substantially different from the other selected classifiers and less commonly used in machine learning [29]. Besides, PLS-DA is not supported by a compatible Matlab package. Moreover, some of the classifiers presented in this paper achieve better accuracy than 83.2%. However, for proof-of-concept studies to explore whether eNose technology is able to distinguish certain patient groups there is no need to compare multiple classifiers and PLS-DA is a reliable method to use [29].

The calculated threshold for feature selection was based on an independent 10-fold cross-validation (**Figure 1A**). Preferably, threshold optimisation would have been included in the 10-fold cross-validation when each classifier was tested and train, to select the most relevant features. This was not executed due to computational limits.

The current results cannot yet be used in clinical practice due to the lack of external validation of the model. Due to the rarity of the disease and the small number of specialized treatment centres, an external patient cohort is difficult to create. To generate robust results and avoid overfitting of the model using the available data, the nested cross-validation was performed as an extra step in testing the model following recommendations from Cawley and Talbot [24]. Another possible source of bias is the absence of data from patients suspected of pulmonary sarcoidosis or ILD.

Conclusion

Evaluation of various classification methods resulted in an accurate diagnostic model for sarcoidosis based on exhaled breath eNose data. To design this model, frequently used dimensionality reduction methods and classifiers were assessed and compared systematically by rigorous procedures such as nested cross-validation. For the current eNose dataset, a model based on feature selection followed by RF yield the best results. The proposed strategy to design and evaluate a diagnostic model can serve as an example for other researchers and is applicable to other eNose datasets.

The outcome of the model includes a specific diagnostic probability for an individual patient, which will facilitate translation into clinical practice. After optimising the model with a multicentre training dataset and validating the developed model with eNose data of patients with suspected pulmonary sarcoidosis, eNose models might be integrated in clinical decision making in order to facilitate a fast, accurate and non-invasive diagnosis.

References

1. Mekov E, Miravitlles M, Petkov R. Artificial intelligence and machine learning in respiratory medicine. Expert Review of Respiratory Medicine. 2020 2020/06/02;14(6):559-64.
2. Topalovic M, Das N, Burgel P-R, Daenen M, Derom E, Haenebalcke C, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. European Respiratory Journal. 2019;53(4):1801660.
3. van der Sar IG, Wijbenga N, Nakshbandi G, Aerts JGJV, Manintveld OC, Wijsenbeek MS, et al. The smell of lung disease: a review of the current status of electronic nose technology. Respiratory Research. 2021 2021/09/17;22(1):246.
4. Grunewald J, Grutters JC, Arkema EV, Saketkoo LA, Moller DR, Müller-Quernheim J. Sarcoidosis. Nature Reviews Disease Primers. 2019 2019/07/04;5(1):45.
5. Crouser ED, Maier LA, Wilson KC, Bonham CA, Morgenthau AS, Patterson KC, et al. Diagnosis and Detection of Sarcoidosis. An Official American Thoracic Society Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2020;201(8):e26-e51.
6. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratú P, Resta O, et al. An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. Respiratory Medicine. 2013 2013/07/01/;107(7):1073-78.
7. Yang H-Y, Peng H-Y, Chang C-J, Chen P-C. Diagnostic accuracy of breath tests for pneumoconiosis using an electronic nose. Journal of Breath Research. 2017 2017/11/29;12(1):016001.
8. Krauss E, Haberer J, Maurer O, Barreto G, Drakopanagiotakis F, Degen M, et al. Exploring the Ability of Electronic Nose Technology to Recognize Interstitial Lung Diseases (ILD) by Non-Invasive Breath Screening of Exhaled Volatile Compounds (VOC): A Pilot Study from the European IPF Registry (eurIPFreg) and Biobank. Journal of Clinical Medicine. 2019;8(10):1698.
9. Dragonieri S, Scioscia G, Quaranta VN, Carratu P, Venuti MP, Falcone M, et al. Exhaled volatile organic compounds analysis by e-nose can detect idiopathic pulmonary fibrosis. Journal of Breath Research. 2020 2020/07/21;14(4):047101.
10. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. European Respiratory Journal. 2021;57(1):2002042.
11. van der Sar IG, Moor CC, Oppenheimer JC, Luijendijk ML, van Daele PLA, Maitland-van der Zee AH, et al. Diagnostic Performance of Electronic Nose Technology in Sarcoidosis. Chest. 2022 2022/03/01/;161(3):738-47.
12. Xuan W, Zheng L, Bunes BR, Crane N, Zhou F, Zang L. Engineering solutions to breath tests based on an e-nose system for silicosis screening and early detection in miners. Journal of Breath Research. 2022 2022/04/07;16(3):036001.
13. Statnikov A, Aliferis CF, Tsamardinos I, Hardin D, Levy S. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics. 2004;21(5):631-43.
14. Leopold JH, Bos LD, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. J Breath Res. 2015 Dec 15;9(4):046002.

15. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. European Respiratory Journal. 2018;51(1):1701817.

16. Statistics and machine learning toolbox. MathWorks.

17. Jolliffe I. Principal component analysis. 2nd ed. Springer-Verlag; 2002.

18. Sánchez-Maroño N, Alonso-Betanzos A, Tombilla-Sanromán M. Filter Methods for Feature Selection – A Comparative Study. 2007. p. 178-87.

19. PCA - Principal component analysis of raw data. MathWorks.

20. Fscchi2 - Univariate feature ranking for classification using chi-square tests. Mathworks.

21. Bayesopt - Select optimal machine learning hyperparameters using Bayesian optimization. MathWorks.

22. Fitcensemble - Fit ensemble of learners for classification. MathWorks.

23. cvpartition - Partition data for cross-validation. MathWorks.

24. Cawley G, Talbot N. On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. Journal of Machine Learning Research. 2010 2010;11:2079-107.

25. predict - Predict labels using discriminant analysis classification model. MathWorks.

26. Gromski PS, Correa E, Vaughan AA, Wedge DC, Turner ML, Goodacre R. A comparison of different chemometrics approaches for the robust classification of electronic nose data. Analytical and Bioanalytical Chemistry. 2014 2014;406:7581-90.

27. Fitts PM. Human engineering for an effective air-navigation and traffic-control system. 1951.

28. Asan O, Bayrak AE, Choudhury A. Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. J Med Internet Res. 2020;22(6):e15154.

29. Lee LC, Liong C-Y, Jemain AA. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. Analyst. 2018;143(15):3526-39.

SUPPLEMENTARY DATA A

Quality control and conditions of use of the SpiroNose®

(Information is copied from the Clinical Investigators Brochure v.1.0 with permission of the manufacturer Breathomix)

System verification

At least once per month the performance of the SpiroNose sensors must be verified. This is done via a custom-designed Verification Tool. This Verification Tool can be connected to the front of the SpiroNose to allow the quality control (QC) gas to flow through it.

During the verification, a quality gas is flushed through 5 tubes of the SpiroNose using the verification tool. The gas (composition mentioned below) is supplied in gas cannisters. Each gas canister can only be used once.

Composition: 0.30% Acetylene (C2H2), 0.30% Methane (CH4), 0.30% Carbon monoxide (CO) 0.31%, Oxygen (O2) 20.9%, rest Nitrogen (N2).

Maintenance

System and software maintenance is performed by Breathomix.

Conditions and contra-indications for using the SpiroNose®

Certain contraindications apply to the breath analysis using the SpiroNose. Study participants should not perform a breath measurement using the SpiroNose if they:

- Suffer from sever shortness of breath;
- Suffer from a pneumothorax;
- Had recent chest surgery;
- Suffer from nausea of are vomiting;
- Have sever wounds in or on the mouth;
- Have consumed alcohol <8 hours.

Breath Measurement

- Breath analysis using the SpiroNose should be performed in a separate, well-ventilated room. A stable environment will benefit the performance of the measurements.
- Breath analysis using the SpiroNose must always be carried out with accessories (mouthpiece, nose clamp and bacterial/viral filter) provided by Breathomix from

Lemon Medical, Germany (Pulmosafe V3/2). Failure to do so may influence the SpiroNose sensor data.
- The mouthpieces are single participant use only. Further re-use could influence the SpiroNose sensor data and could increase the risk of cross-contamination.
- The mouthpiece should be disposed after use, in accordance with local waste disposal guidance.

SpiroNose

- The use of alcohol disinfection agents in the same room as the SpiroNose should be avoided. The alcohol will influence the SpiroNose sensors which will subsequently influence the measurement data.
- The housing of the SpiroNose must not be cleaned with alcohol-containing agents as the vapors of these agents will damage the sensors.
- The outside of the SpiroNose should only be disinfected with hydrogen peroxide 3% (provided by Breathomix) sprayed on a paper towel.
- Hydrogen Peroxide 3% should not be directly sprayed on the SpiroNose.
- The inside of the SpiroNose should not be cleaned.
- Under no circumstances should the instrument be immersed or splashed with liquid.

It is important that the SpiroNose and the Gateway are always connected to a power source, so that the sensors can stabilize to the environment. This way the life span of the sensors will be extended.

Sensor specifications of the SpiroNose®

(Information is copied from the Clinical Investigators Brochure v.1.0 with permission of the manufacturer Breathomix)

Table S1: Characteristics of the sensors used in the SpiroNose®.

Sensor type	High sensitivity to	Range
TGS 2600	Hydrogen, Ethanol	1-30 ppm
TGS 2602	VOCs, ammonia, H2S	1-30 ppm
TGS 2603	Trimethylamine, methylmercaptan	1-30 ppm
TGS 2610-C00	Butane, Propane	500-10.000 ppm
TGS 2611-C00	Methane	500-10.000 ppm
TGS 2612-D00	Methane, Propane, Butane	500-1.000 ppm
TGS 2620	Alcohol, solvent vapors	50-5.000 ppm

The breathing maneuver

(Information is copied from the BreathBase Research Manual rev 3.0 with permission of the manufacturer Breathomix)

Steps of breath maneuver

To ensure high quality measurement data, it is very important that the breathing maneuver is performed correctly. Therefore, follow the following steps carefully:

1. The patient breathes normally through the device 5 times.
2. The patient inhales as deeply as possible and hold the breath for 5 seconds.
3. The patient exhales as slowly and as completely as possible.
4. The patient breathes in deeply once more, releases the mouthpiece and steps away from the device.

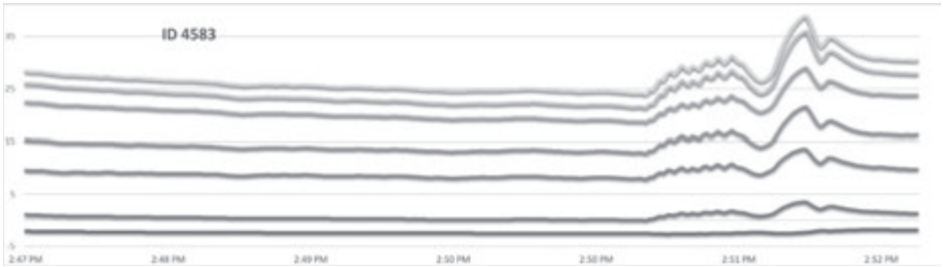
It is very important that the breathing maneuver is performed correctly so that a reliable analysis of the signals can take place. Each step in the measurement maneuver is important for a reliable breath measurement and the importance of the steps is described below:

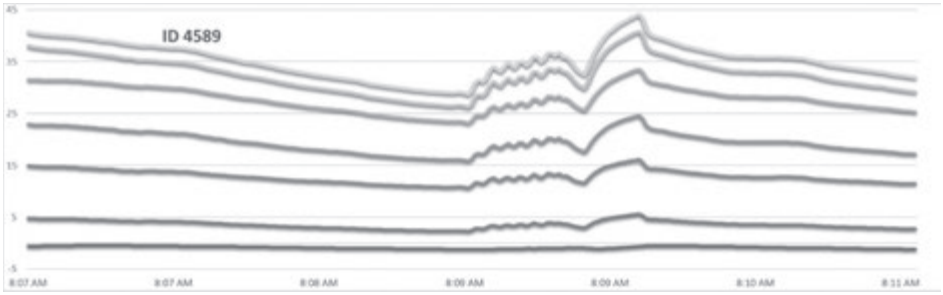
1. The 5 tidal breaths allow the patient to get used to breathing through the mouthpiece and SpiroNose.
2. The patient inhales as deeply as possible and holds it for 5 seconds, in order to have enough air for the slow maximum exhalation and to allow the VOCs in the lungs to mix well with the inhaled air.
3. The patient exhales as slowly and as maximally as possible to exhale all VOCs in the SpiroNose. This must be done slowly because the SpiroNose generates 1 datapoint per second.
4. The patient takes a final deep breath as they need air after the slow maximum exhalation, and this allows ambient air to re-enter the SpiroNose causing the sensor values to return to a stable value more quickly. This creates a peak which is necessary for the analysis.

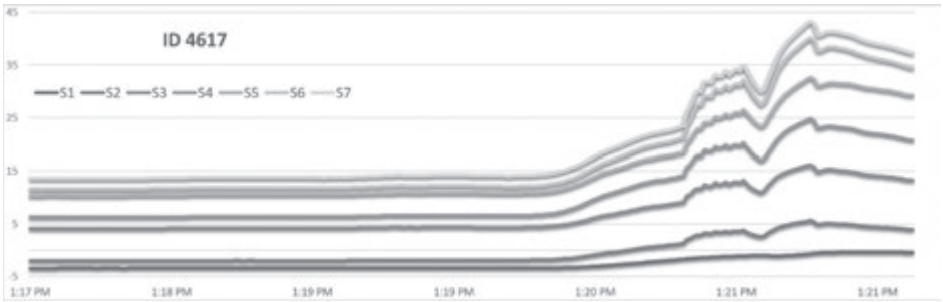
NOTE: No alcohol should be consumed up to 8 hours before the test. This also applies to mouthwash with alcohol.
WARNING: The disposable kit (mouthpiece, bacterial & virus filter, and nose clip) should be disposed after use, in accordance with local waste disposal guidance.
WARNING: Do not skip steps during the measurement and its preparation. This may result in low quality measurements.
WARNING: Always clean the SpiroNose after each measurement with H2O2 3% (Reinigungsmiddel FAR-2020-11, Bloc Medical).
WARNING: Under no circumstances should the SpiroNose be disinfected with alcohol.
WARNING: Under no circumstances should the SpiroNose be submerged or sprinkled with water.

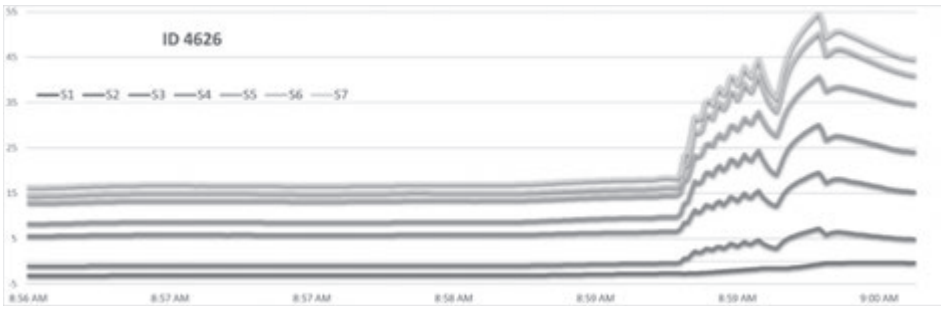
SUPPLEMENTARY DATA B

Examples of raw sensor diagrams resulting from the SpiroNose sensors during a breath manoeuvre and corresponding sensor values after data pre-processing



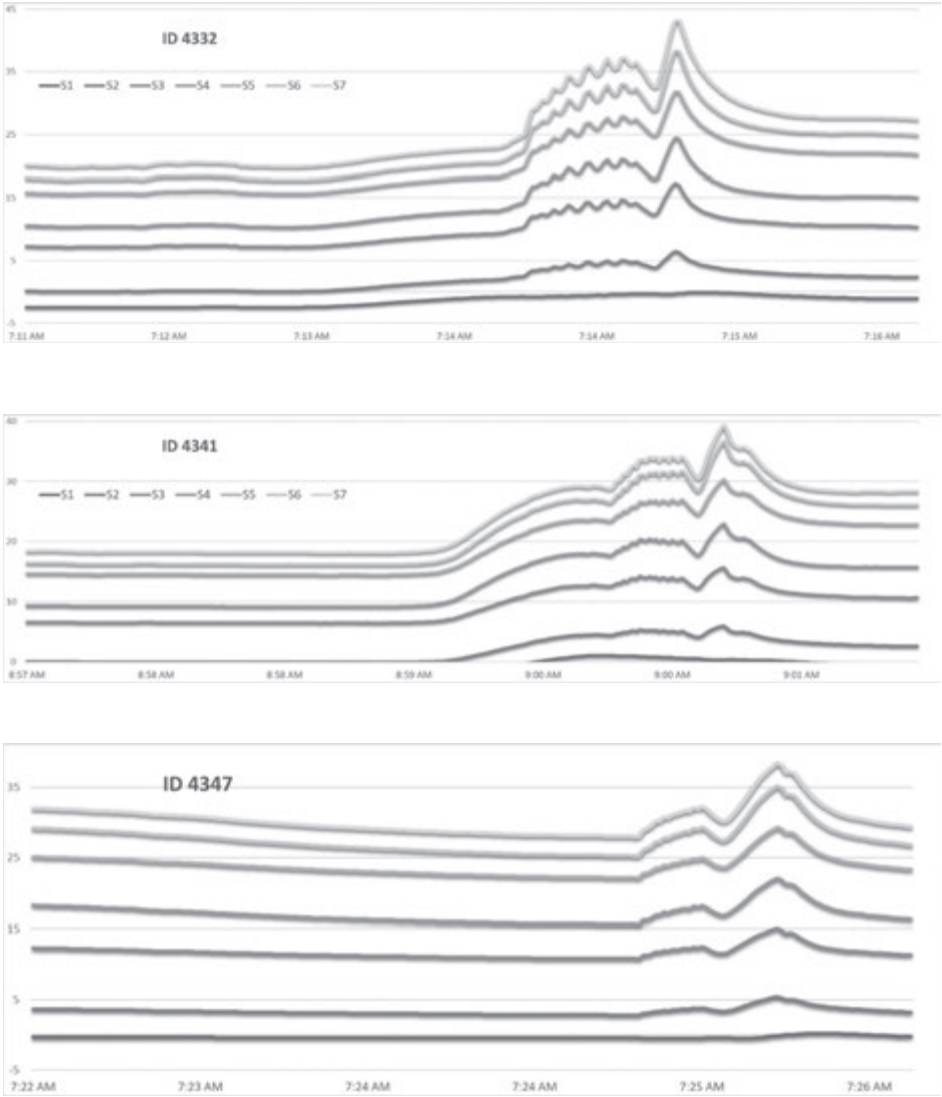


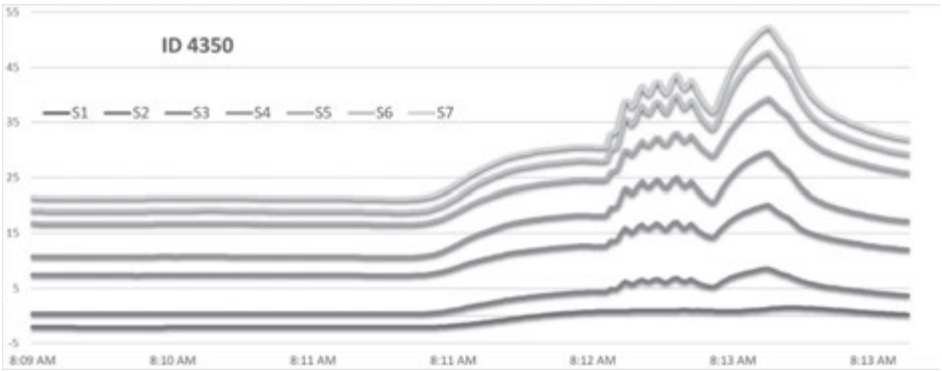




ID	S1	S2	S3	S4	S5	S6	S7	S1/BH	S2/BH	S3/BH	S4/BH	S5/BH	S6/BH	S7/BH
4583	1.43	1	0.81	1.04	1.08	1.21	0.53	0.50	0.46	0.35	0.60	0.16	0.47	0.11
4589	1.37	1	0.70	1.14	0.80	1.37	0.24	0.25	0.10	-0.06	0.27	0.39	0.14	0.11
4617	1.38	1	0.84	0.82	0.23	1.00	0.59	0.20	0.18	0.03	0.35	0.42	0.19	0.24
4626	0.37	1	1.00	1.33	1.13	1.37	0.42	0.38	0.34	0.42	0.47	0.59	0.42	0.31

Figure S1: Sensor diagram from patients with sarcoidosis and corresponding sensor values after data pre-processing that serve as input variables.
Diagrams exported from the BreathBase® platform. S = sensor; BH = breath hold.





ID	S1	S2	S3	S4	S5	S6	S7	S1/BH	S2/BH	S3/BH	S4/BH	S5/BH	S6/BH	S7/BH
4332	0.99	1	0.74	1.24	0.96	1.39	0.43	0.10	0.11	0.02	0.37	0.02	0.22	0.11
4341	1.07	1	0.74	1.21	0.87	1.18	0.51	0.00	0.01	0.21	0.40	0.16	0.12	0.15
4347	1.53	1	0.73	1.26	0.56	1.15	0.38	0.10	0.06	0.07	0.31	0.04	0.22	0.06
4350	0.84	1	0.75	1.32	1.10	1.33	0.49	0.01	0.01	0.18	0.43	0.08	0.06	0.20

Figure S2: Sensor diagrams from patients with another interstitial lung disease (idiopathic pulmonary fibrosis) and corresponding sensor values after data pre-processing that serve as input variables.

Diagrams reconstructed using raw unprocessed data similar to the lay-out of the BreathBase® platform. S = sensor; BH = breath hold.

SUPPLEMENTARY DATA C

Specification hyperparameter optimization options

Table S2: Specification of applied hyperparameter optimization options per classifier.

SVM	KNN	LDA	NN	RF
BoxConstraint	NumNeighbors	Delta	Activations	Number of learning cycles: range 1-100
KernelScale	Distance	Gamma	Standardize	Number of bins: range 2-50
			Lambda	
			LayerSizes	

Automated optimization as defined by Matlab function was used if available [1-5]. For the RF classifier, cycles and bins were applied using the Bayesian optimization function [6]. KNN = K-Nearest Neighbour; LDA = Linear Discriminant Analysis; NN = neural networks; RF = random forest; SVM = support vector machines.

References

1. fitcsvm - Train support vector machine (SVM) classifier for one-class and binary classification. MathWorks.
2. fitcknn - Fit k-nearest neighbor classifier. MathWorks.
3. fitcdiscr - Fit discriminant analysis classifier. MathWorks.
4. fitcnet - Train neural network classification model. MathWorks.
5. fitcensemble - Fit ensemble of learners for classification. MathWorks.
6. bayesopt - Select optimal machine learning hyperparameters using Bayesian optimization. MathWorks.

SUPPLEMENTARY DATA D

Table S3: Accuracy for each classifier of each fold.

CVA 5 folds	LDA	SVM	KNN	NN	RF
Fold 1	0.845	0.857	0.843	0.859	**0.864**
Fold 2	0.857	**0.870**	0.854	0.852	0.861
Fold 3	0.861	0.861	0.854	0.859	**0.866**
Fold 4	0.843	**0.855**	0.843	0.852	0.852
Fold 5	0.836	0.843	0.829	0.822	**0.843**

Best performing per fold is displayed in bold. CVA = cross-validated accuracy; KNN = K-Nearest Neighbour; LDA = Linear Discriminant Analysis; NN = neural networks; RF = random forest; SVM = support vector machines.

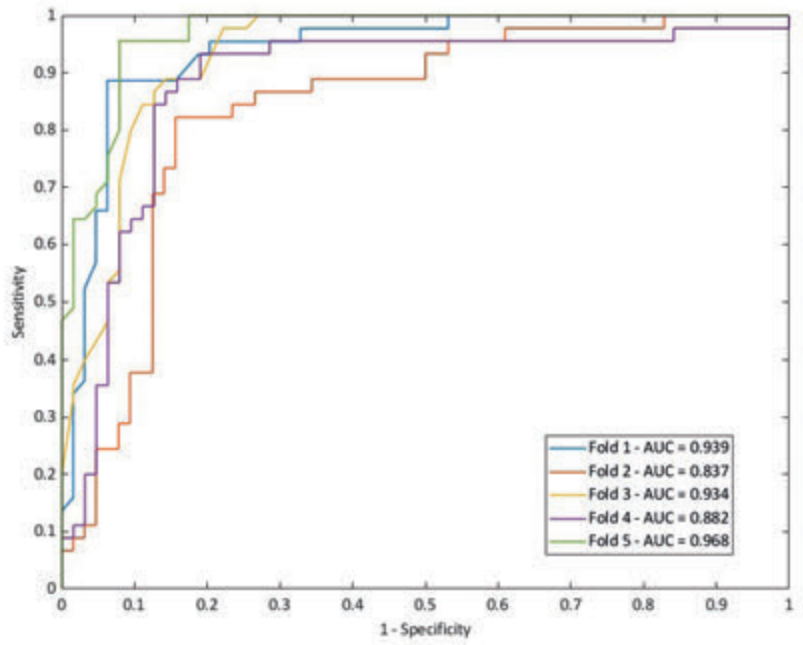


Figure S3: Receiver operating characteristic curves of the final model for each fold.
AUC = area under the curve.

SUPPLEMENTARY DATA E

Evaluation size dataset

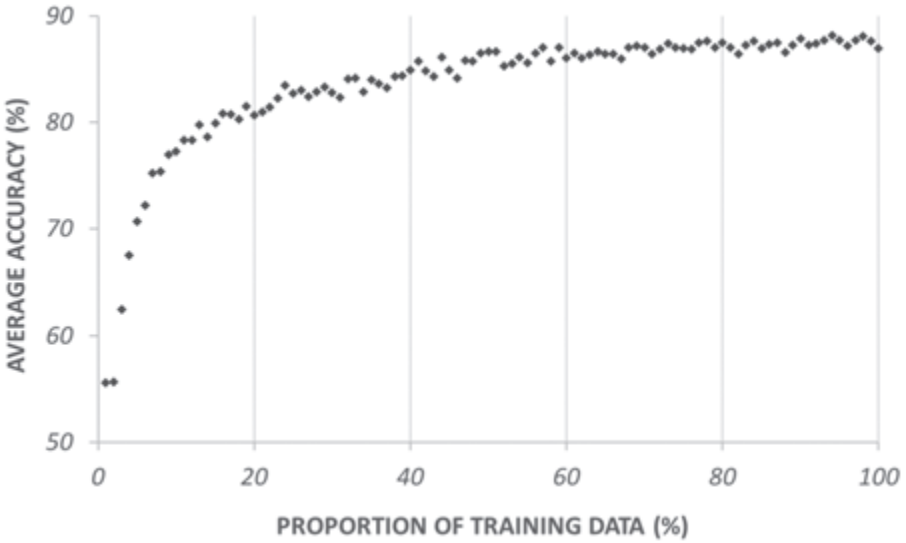


Figure S4: The average accuracy of the final trained model using the random forest classifier plotted against the proportion of data used for training the model to evaluate the sufficiency of the dataset size.
Full training set consisted of 90% of total data amount and the test set of 10%. The average accuracy resulted from repeating training and testing 20 times per proportion of training data.

CHAPTER 8

Classifying interstitial lung disease: omics are in the air

Am J Respir Crit Care Med. 2024 May 3.

Iris G. van der Sar, Catharina C. Moor, Marlies S. Wijsenbeek

Erasmus Medical Center, Rotterdam (the Netherlands)

Correspondence letter

Reply to: Huang Y, Ma SF, Oldham JM, Adegunsoye A, Zhu D, Murray S, Kim JS, Bonham C, Strickland E, Linderholm AL, Lee CT, Paul T, Mannem H, Maher TM, Molyneaux PL, Strek ME, Martinez FJ, Noth I. Machine Learning of Plasma Proteomics Classifies Diagnosis of Interstitial Lung Disease. Am J Respir Crit Care Med. 2024 Feb 29.

Keywords

Interstitial lung disease, proteomics, breathomics, diagnostics, artificial intelligence

To the editor

We currently stand at the beginning of an era where artificial intelligence will likely become integrated into medical practice and health care on a broad scale. In the scientific quest for novel diagnostic tests and biomarkers, vast amounts of data are collected. Data originating from specific biological sources is known as omics data and may include genes (i.e. genomics), serum proteins (i.e. proteomics) or breath molecules (i.e. breathomics). Machine learning algorithms are often used for analyzing these extensive datasets, revealing patterns in data that exceed human capabilities. This approach is particularly useful to extract insightful information for better diagnosis of complex rare diseases such as interstitial lung diseases (ILDs).

In this context, we find Huang and colleagues' report on serum plasma proteomics for diagnosing patients with pulmonary fibrosis especially interesting [1]. They aimed to design a proteomic classifier for differentiating patients with idiopathic pulmonary fibrosis (IPF) and connective tissue disease related ILD (CTD-ILD) by including an impressively large cohort (n=1247 IPF; n=352 CTD-ILD) representing 42 clinical centers across the USA. The developed classifier for differentiation consisted of 37 proteins. Subsequently, the accuracy of four algorithms for classifying patients with IPF or CTD-ILD were calculated. All algorithms yielded high accuracies (77.3-82.5%) in an independent dataset. Considering the high performance of the classifier and non-invasive nature of blood sampling, integrating proteomics analysis in medical practice holds promise. However, the current analysis process is still too elaborative and should be converted into an easily applicable and affordable method to enable integration in medical practice.

We would like to challenge Huang and colleagues that exhaled breath data resulting from eNose analysis (i.e. breathomics data) has greater potential to serve as an easily integrated diagnostic tool for ILD. An eNose using multiple sensors to analyze the >2000 types of volatile organic compounds present in breath. The technology can provide individual results in real-time and comes with low costs. Besides, the procedure is patient-friendly, as testing takes less than two minutes and involves slow breathing without forced maneuvers. Opposite to invasive procedures like tissue biopsies, an eNose test can be repeated during the disease course allowing to monitor or reclassify an ILD diagnosis. Importantly, studies reported high performance for differentiating types of ILD, and pulmonary fibrosis in particular. Two single-center studies from 2019 and 2021 compared IPF and CTD-ILD cohorts, similar to Huang and colleagues, and reported area under the curve values for classifying breath profiles using different type of eNoses ranged from 0.84 to 0.96 [2, 3].

The examples of proteomics and breathomics research show that the analysis of high volume omics data analyzed with machine learning algorithms can support diagnosing rare diseases. One could envision that combining results from the plasma proteomics classifier and an eNose breathomics profile can lead to improved diagnostic confidence of ILD multidisciplinary team discussions and limit the need for invasive biopsies. Given the limited number of expertise centers for ILD and potential differences in ILDs across the world, international collaboration is essential to guarantee the collection of sufficient data for robust models. Thus, we need to start sharing omics data in secured and constantly updated cloud-connected databases to facilitate the development and use of high-quality clinically applicable diagnostic models.

As omics are in the air, we look forward to a bright future in which we utilize biological data to its full potential to enable accurate, fast and non-invasive diagnostic trajectories for most patients with ILD worldwide.

References

1. Huang Y, Ma SF, Oldham JM, Adegunsoye A, Zhu D, Murray S, et al. Machine Learning of Plasma Proteomics Classifies Diagnosis of Interstitial Lung Disease. Am J Respir Crit Care Med. 2024 Feb 29.
2. Krauss E, Haberer J, Maurer O, Barreto G, Drakopanagiotakis F, Degen M, et al. Exploring the Ability of Electronic Nose Technology to Recognize Interstitial Lung Diseases (ILD) by Non-Invasive Breath Screening of Exhaled Volatile Compounds (VOC): A Pilot Study from the European IPF Registry (eurIPFreg) and Biobank. J Clin Med. 2019 Oct 16;8(10).
3. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).

PART III

ENOSE FOR SCREENING

CHAPTER 9

Detection of systemic sclerosis-associated interstitial lung disease by exhaled breath analysis using electronic nose technology

Am J Respir Crit Care Med. 2024 Jun 6. (Research Letter)

Emiel R. Marges[1#], Iris G. van der Sar[2#], Jeska K. de Vries-Bouwstra[1], Tom W.J. Huizinga[1], Paul L.A. van Daele[2], Marlies S. Wijsenbeek[2], Catharina C. Moor[2#], J.J. Miranda Geelhoed[1#]

[1]Leiden University Medical Center, Leiden (the Netherlands); [2]Erasmus Medical Center, Rotterdam (the Netherlands); [#]Authors share first/last authorship

Abstract

Introduction

Interstitial lung disease (ILD) remains the leading cause of death in patients with systemic sclerosis (SSc). Although several risk factors of developing SSc-ILD are known, timely detection remains a major challenge. The analysis of volatile organic compounds in exhaled breath using a sensor-based electronic nose (eNose) device is increasingly studied and shows promising results in detecting ILD. We investigated whether eNose technology can detect SSc-ILD amongst patients with SSc and other types of ILD.

Methods

We conducted a cross-sectional multicenter study in patients with SSc. Breath analysis using an eNose (SpiroNose®) was performed. eNose data of patients with SSc-ILD were compared with SSc without ILD and with other types of ILD. Cohorts were split in a training and test set to internally validate results. eNose data were analyzed using partial least square discriminant and receiver operating characteristics analysis.

Measurements and Main results

223 patients with SSc were included of whom 110 had ILD. eNose distinguished patients with SSc-ILD from those without ILD with an AUC of 0.84 (0.75-0.94) in the test set. Comparison of SSc-ILD with other types of ILD (n=300) yielded an AUC of 0.84 (0.76-0.91) in the test set; AUC of SSc-ILD versus other CTD-ILD (n=47) was 0.74 (0.60-0.88).

Conclusion

eNose technology can accurately differentiate patients with SSc-ILD from those without ILD, and can discriminate SSc-ILD from other types of ILD. eNose breath analysis is a promising non-invasive screening tool for SSc-ILD.

Introduction

Systemic sclerosis (SSc) is a rare autoimmune connective tissue disease (CTD) characterized by diffuse fibrosis and vasculopathy of multiple organ systems [1]. Despite improvements in survival over the last three decades due to earlier diagnosis and treatment, SSc has a high mortality rate (cumulative survival at 10 years of 62.5%) [2, 3]. Systemic sclerosis-associated interstitial lung disease (SSc-ILD) has a prevalence of 30-50% in patients with SSc and is the leading cause of death [4-6]. SSc-ILD is considered to be the result of intra-alveolar inflammation of epithelium, initiating an inflammatory cascade ultimately leading to interstitial fibrosis formation [7, 8]. When progressive, pulmonary fibrosis leads to increasing symptoms, lung function decline and respiratory failure [4-6]. Early detection of ILD might prevent lung function decline and improve survival due to timely treatment, but no uniform screening guidelines are available [9].

Presence of anti-topoisomerase (anti Scl-70) antibodies and diffuse skin involvement have been identified as risk factors for development of SSc-ILD [10-12]. Moreover, elevated inflammatory markers, extensive skin involvement, and SSc development within 18 months after first non-Raynaud phenomenon have been suggested as independent predictors of early forced vital capacity (FVC) decline in SSc patients [13, 14]. However, none of these factors are unequivocally associated with development or progression of ILD in patients with SSc.

The gold standard to detect SSc-ILD is a chest high resolution computed tomography (HRCT) scan, which is widely used for ILD screening at time of SSc diagnosis. Yearly re-screening in high-risk patients should be considered after an initial negative screening [15]. Since HRCT scans come with relatively high costs and potential detrimental effects of cumulative radiation exposure, a non-invasive, cheap, and accurate screening tool for ILD is warranted to facilitate early detection and monitoring of SSc-ILD.

Over the past decades, the analysis of volatile organic compounds (VOCs) in exhaled breath is increasingly studied as a source of biomarkers to detect respiratory diseases. Two breath analysis techniques are mainly used in research. A lab-based method using gas chromatography and mass spectrometry (GC-MS) can identify individual VOCs [16-18]. A second method for VOC analysis is a sensor-based technique called electronic nose (eNose) technology. This technology detects a wide range of VOCs to generate breath profiles. After training a model with large datasets, pattern recognition algorithms can assign individual breath profiles to a specific diagnosis group. An eNose has potential as point-of-care diagnostic test in clinical practice, since it is non-invasive and can provide real-time results. Multiple studies, using different devices, already

demonstrated its capability to accurately identify respiratory diseases in individual patients, including ILD, sarcoidosis and post-corona virus disease lung disease [19-30].

In this study, we evaluated whether exhaled breath analysis using eNose technology can differentiate between SSc patients with and without ILD, and whether known risk factors for developing SSc-ILD influence the breath profile. Additionally, we evaluated the discriminative value of an eNose for differentiating patients with SSc-ILD from other ILDs, including CTD-ILD.

Methods

Study design

In this cross-sectional multicenter study at the Leiden University Medical Center (LUMC; the Netherlands) and Erasmus Medical Center Rotterdam (EMC; the Netherlands), we included patients with a confirmed diagnosis of SSc according to the current diagnostic guideline between December 2020 and March 2023 [31]. In patients with SSc-ILD, diagnosis was established by a multidisciplinary team [32]. A thoracic HRCT scan and pulmonary function test had to be available. LUMC patients were recruited during their annual visit as part of the prospective cohort study CCISS (combined care in systemic sclerosis) [33]. EMC patients were recruited during regular outpatient clinic visits. All patients were included irrespective of disease duration and treatment. Patients were excluded if they had an ILD diagnosis other than SSc-ILD, an active malignancy, or current respiratory infection. Other contra-indications were alcohol consumption <8 hours, since sensor signals are very sensitive to alcohol. A previously included cohort with other ILD diagnoses was used as a control group [28].

All patients signed written informed consent. The study was approved by medical ethics committees in both participating sites and conducted in accordance with the amended Declaration of Helsinki.

Data collection

We used the SpiroNose® (Breathomix, Leiden, the Netherlands) for eNose breath analysis. The device contains multiple cross-reactive metal oxide sensors on both the inside and the outside of the device, to enable correction for ambient air. Sensor signals are stored on the online secured and certified BreathBase® platform. Patients were instructed to perform a breath maneuver twice, using a standardized procedure of five tidal breaths, a deep inhalation, a 5-second breath-hold, and a slow maximal exhalation. Technological aspects of the device and data collection are described

in more detail elsewhere [34, 35]. Patients were asked about their last food or drink intake, smoke, and inhalation medication use. Additional information was extracted from the patient's medical file (e.g. demographics, medical history, smoking status, current medication use). Results from the most recent thoracic HRCT scan, pulmonary function test and laboratory tests were collected. HRCT scans were reviewed for the presence of interstitial lung abnormalities by experienced thoracic radiologists.

Data analysis

After downloading eNose sensor responses from the BreathBase® platform, data were pre-processed. This included normalization of the data, correction for ambient air, and reduction of inter-array differences. For each patient, the best measurement was selected to calculate the peak value and the peak to breath hold ratio, which resulted in 13 values per patient. Breath data were labeled with clinical characteristics of the patient. An overview of data collection and processing including visual examples can be found in a previously published paper [28].

To answer the main study aims, we compared eNose sensor data of patients with SSc and diagnosed with SSc-ILD versus those without ILD. Second, data of patients with SSc-ILD were compared to patients with other types of ILD and to patients with another type of CTD-ILD only. For these main analyses, we split the dataset in a training and test set to avoid overfitting (ratio 2:1) as recommended for metabolomics research [36]. We performed partial least square discriminant analysis (PLS-DA) and receiver operating characteristics to assess between-group differences in breath data using the two first principal components (PCs) resulting from the PLS-DA. Each principal component is a weighted combination of the processed sensor values. Results were visualized as scatterplots with PC1 on the x-axis and PC2 on the y-axis. Area under the curve (AUC) values with 95% confidence intervals (CI), sensitivity, specificity, accuracy, and negative and positive predictive values (NPV and PPV) were derived from ROC analyses. The main analyses were repeated without outlying values of PC1 and PC2 to assess influence of outliers on the separation of groups. Outliers were defined as measurements with PC1 or PC2 outside the upper and lower limits of a box-and-whisker plot. Limits were calculated as quartile 1 and 3 +/- 1.5 * interquartile range.

Additionally, the impact of factors that could potentially influence breath profiles such as sex, smoking status, age, use of immunosuppressive drugs, disease duration (early or late SSc diagnosis, according to date of SSc diagnosis with a cut-off of 24 months) and SSc-ILD severity (mild or severe, using FVC 70% of predicted as cut-off [37]) was assessed within the full dataset. Subsequently, in order to evaluate the impact

of certain patient features on disease detection by eNose, we repeated the analysis of SSc-ILD versus SSc without ILD in subgroups of patients with characteristics that are associated with ILD development (anti Scl-70 positivity, male sex, diffuse SSc), and with early or late SSc diagnosis. For the additional analyses, we did not split the dataset and used the first PC only.

Descriptive statistics were used to analyze baseline data, including χ^2, Student's t, and Mann Whitney tests between groups. We displayed normally distributed data as mean values (± standard deviation) and non-normally distributed data as median values (interquartile range). R version 4.2.1 for Windows with mixOmics version 6.20.0 package was used for the analyses.

Results

Baseline characteristics

In total, 223 patients with SSc were included, of whom 110 had confirmed diagnosis of SSc-ILD (**Table 1**). Median age of patients was 60.0 years, and 74.7% were female with a higher proportion in the group without ILD (p=0.01). No significant difference was observed in smoking status between the two groups. FVC and diffusion capacity (DLCOc) were lower in patients with ILD compared to patients without ILD. The majority of patients had limited SSc, with a predominance in the group without ILD. The cohort of other types of ILD consisted of 300 patients, of which 47 patients were diagnosed with CTD-ILD other than SSc including Sjögren's syndrome, anti-synthetase syndrome, undifferentiated/mixed CTD, systemic lupus erythematosus and/or rheumatoid arthritis [28]. Add+-250itional baseline characteristics can be found in **Supplementary Data A** (**Table S1**).

Main results

The eNose distinguished patients with SSc-ILD from those without ILD with an AUC of 0.79 (95% CI 0.72-0.87) in the training set and an AUC of 0.84 (0.75-0.94) in the test set (**Figure 1A**). Comparing eNose data of patients with SSc-ILD to other types of ILD, resulted in an AUC of 0.87 (0.81-0.92) in the training set and an AUC of 0.84 (0.76-0.91) in the test set.

SSc-ILD breath profiles were separated from other CTD-ILDs with an AUC of 0.88 (0.80-0.95) in the training and AUC of 0.74 (0.60-0.88) in the test set (**Figure 1B**). Corresponding specificity, sensitivity, accuracy, NPV, and PPV can be found in **Table 2**.

Table 1: Baseline characteristics SSc cohort.

	All patients with SSc	Without ILD	SSc-ILD	*p-value*
Subjects	223	113	110	
Females	165 (74.7) *	93 (82.3)	72 (66.7)	*0.01*
Age (years)	60.0 [51.0, 69.0] *	57.0 [50.0, 67.0]	63.50 [54.0, 71.0]	*<0.01*
Smoking status	*			*0.33*
Never	89 (40.8)	45 (40.2)	44 (41.5)	
Former	108 (49.5)	53 (47.3)	55 (51.9)	
Current	21 (9.6)	14 (12.5)	7 (6.6)	
FVC (%pred)	93.7 (20.0) *	100.6 (17.4)	86.5 (20.1)	*<0.01*
DLCOc (%pred)	69.3 (20.9) *	77.3 (18.9)	61.1 (19.6)	*<0.01*
Pulmonary fibrosis	98 (43.9)	0 (0.0)	98 (89.1)	*<0.01*
Limited SSc	172 (77.1)	99 (87.6)	73 (66.4)	*<0.01*
Established PAH	23 (10.3)	12 (10.6)	11 (10.0)	*1.00*
Anti Scl-70 +	53 (24.2) *	10 (8.8)	43 (40.6)	*<0.01*
Anti CENP +	86 (40.2) **	72 (63.7)	14 (13.9)	<0.01
Anti RNA polymerase III +	10 (4.9) ***	5 (4.4)	5 (5.4)	0.99
Antifibrotic drug use	9 (4.0)	0 (0.0)	9 (8.2)	0.01
Immunosuppressive drug use	78 (35.0)	18 (15.9)	60 (54.5)	<0.01
Pulmonary comorbidity	5 (2.2)	4 (3.5)	1 (0.9)	0.38
Early SSc	39 (17.5)	17 (15.0)	22 (20.0)	0.43

*Values are displayed as number (%), mean ±SD, or median [interquartile range]. Early SSc is defined as time since SSc diagnosis <24 months. CENP = centromeric proteins; DLCOc = diffusing capacity for carbon monoxide corrected for hemoglobin level; FVC = forced vital capacity; ILD = interstitial lung disease; PAH = pulmonary artery hypertension; RNA = ribonucleic acid; Scl-70 = topoisomerase I; SSc = systemic sclerosis; %pred = percent of predicted value, calculated based on sex, age and height. * n=1-5 values missing. ** n=9 values missing. *** n=18 values missing.*

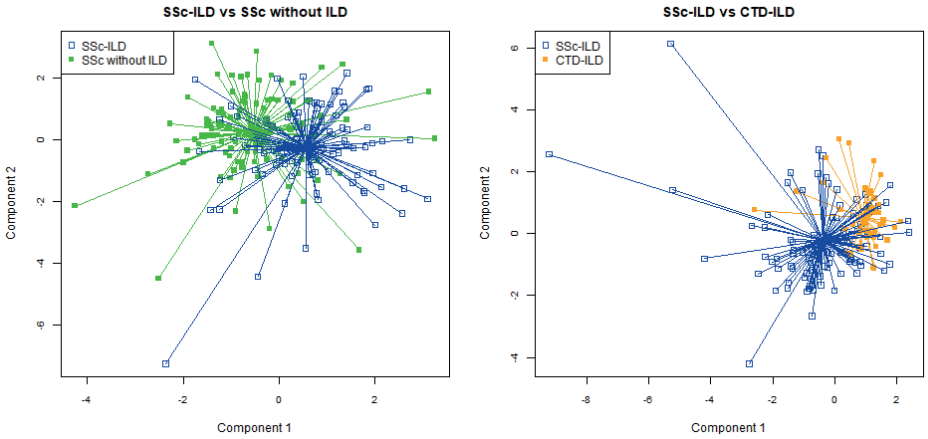


Figure 1: Comparison of breath profiles between patients with SSc, SSc-ILD and other CTD-ILDs.
A. Scatterplot of patients with SSc-ILD (n=110) versus patients with SSc without ILD (n=113). B. Scatterplot of patients with SSc-ILD versus patients with other CTD-ILDs (n=47). X- and Y-axis represent first 2 principal components. CTD = connective tissue disease; ILD = interstitial lung disease; SSc = systemic sclerosis.

There were 17 patients with outlying values of PC1 or PC2 (11 in the comparison of SSc and SSc-ILD, 6 in the comparison of SSc-ILD and CTD-ILD). Removal of outliers had no significant effect on the main results (see **Figure S1-2** and **Table S2-3** in **Supplementary Data B**).

Subgroup analysis

Additional analysis revealed that sex, smoking status, age, use of immunosuppressive drugs, SSc duration and ILD severity did not notably influence breath profiles according to AUC values with confidence intervals close to 0.5 (**Supplementary Data C**, **Table S4**).

Subgroup analyses in those patients with potential risk factors for development of SSc-ILD showed no clear impact on the accuracy for SSc-ILD detection as AUC values remained high (0.80-0.88; **Table 3**). Disease duration of SSc did not influence the ability of eNose for differentiating SSc patients with and without ILD either, indicated by AUC values of 0.86 (0.72-0.99) for disease duration <24 months and 0.79 (0.73-0.86) for ≥24 months.

Table 2: Results of eNose breath analysis between patients with SSc, SSc-ILD and other ILDs.

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
SSc-ILD	74	**SSc without ILD**	76	Training	0.79	*0.72-0.87*	0.83	0.73	0.78	0.76	0.81
	36		37	Test	0.84	*0.75-0.94*	0.78	0.83	0.81	0.83	0.79
SSc-ILD	74	**Other ILDs**	200	Training	0.87	*0.81-0.92*	0.77	0.85	0.79	0.93	0.58
	36		100	Test	0.84	*0.76-0.91*	0.76	0.81	0.77	0.92	0.55
SSc-ILD	74	**Other CTD-ILDs**	32	Training	0.88	*0.80-0.95*	0.88	0.84	0.85	0.70	0.94
	36		15	Test	0.74	*0.60-0.88*	0.80	0.69	0.73	0.52	0.89

Results based on 2 principal components. AUC = area under the curve; CI = confidence interval; CTD = connective tissue disease; ILD = interstitial lung disease; NPV = negative predictive value; PPV = positive predictive value; SSc = systemic sclerosis.

Table 3: Results of eNose breath analysis between SSc and SSc-ILD in SSc subgroups.

Subgroup	n=	SSc-ILD	n=	No ILD	n=	*AUC*	*95% CI*
Anti Scl-70 +	53		43		10	0.88	*0.79-0.98*
Males	56		36		20	0.80	*0.67-0.93*
Diffuse SSc	51		37		14	0.82	*0.69-0.95*
Early SSc	39		22		17	0.86	*0.72-0.99*
Late SSc	184		88		96	0.79	*0.73-0.86*

Results based on 1 principal component. Cut-off for early or late SSc diagnosis is 24 months since SSc diagnosis. AUC = area under the curve; CI = confidence interval; ILD = interstitial lung disease; Scl-70 = topoisomerase I; SSc = systemic sclerosis.

Discussion

Early detection of ILD remains a major challenge in patients with SSc, but has important therapeutic and prognostic complications. This multicenter cross-sectional study shows that eNose technology has high potential for detection of ILD in patients with SSc. The eNose accurately discriminates between breath profiles of SSc patients with or without ILD, and between SSc-ILD and other types of ILD. Interestingly, the accuracy of an eNose breath test was similar within subgroups of patients with risk factors for ILD development.

To our knowledge, this is the first study showing a difference between eNose breath profiles in patients with SSc with and without ILD. This difference could not be explained by patient-related factors, such as use of immunosuppressives, age, smoking habits, sex, and disease duration or severity, according to our results. An explanation for this finding could be that difference in the alveolar inflammation results in distinct VOC patterns. Both inflammatory neutrophils and monocyte-derived macrophages in the lung promote proinflammatory cytokines and chemokines when chronic inflammation is present [38, 39]. In later stages, mesenchymal cells, fibroblasts and tissue macrophages are predominantly present in lung tissue [40]. Macrophages are capable of producing all kinds of substances like nitric oxide, indoleamine-pyrrole 2,3-dioxygenase, arginase-1, reactive oxygen species and matrix metalloproteases, besides cytokines like IL-6. Some of these substances might cause the variation in breath profiles [41, 42].

A previous study using GC-MS compared exhaled breath from patients with SSc and healthy subjects, and showed 16 discriminative VOCs [43]. In contrast to the eNose, GC-MS detects single VOCs and could elucidate the pathophysiologic process of ILD development. However, the GC-MS findings have not been validated and there was no comparison between SSc with and without ILD. Moreover, GC-MS is less feasible as diagnostic tool, as it is a time consuming and complex technique, which is difficult to replicate.

Also in subgroups with risk factors for ILD, the eNose could accurately identify the patients with SSc-ILD. This supports the use of eNose technology as a screening tool for SSc-ILD, probably in a model combined with clinical risk factors. Recently, a model including eNose values and clinical parameters has been tested for screening of non-small cell lung cancer [44]. This model resulted in a higher sensitivity and negative predictive value, compared to models with breath or clinical data only. Future studies in SSc should clarify whether adding clinical parameters to our current model will improve performance as well.

Besides the differences in breathprint in SSc-ILD versus SSc without ILD, we also found a distinctive breath profile in patients with SSc-ILD compared to other forms of ILD, including other types of CTD-ILD. This indicates that eNose might be of added value in the diagnostic work-up of patients with suspected ILD, especially if the disease is difficult to classify. A possible explanation for the differences in breath profiles can be that serum biomarker profiles are known to be different in patients with e.g. rheumatoid arthritis-related ILD and SSc-ILD [45, 46]. These differences reflect disease specific inflammatory pathways and might influence the VOC composition in those patients.

Our study has several weaknesses. First, this is a cross-sectional study including only prevalent patients with SSc, of which the majority received treatment. Second, some subgroups were too small to split into a training and test cohort. Third, patients are included in the same area of one country. Therefore, future international prospective follow-up studies with larger subgroups including suspected or newly diagnosed SSc patients cohorts are needed to confirm and externally validate our results to investigate the potential of eNose as screening tool more comprehensively and develop prediction models combining breath and clinical data.

In conclusion, our study shows that eNose technology can accurately differentiate patients with SSc and ILD from those without ILD, and SSc-ILD from other types of ILD including CTD-ILD. Thus, eNose analysis is a promising non-invasive tool for detection of ILD in SSc.

References

1. Denton CP, Khanna D. Systemic sclerosis. Lancet. 2017 Oct 7;390(10103):1685-99.
2. Nihtyanova SI, Tang EC, Coghlan JG, Wells AU, Black CM, Denton CP. Improved survival in systemic sclerosis is associated with better ascertainment of internal organ disease: a retrospective cohort study. QJM. 2010 Feb;103(2):109-15.
3. Rubio-Rivas M, Royo C, Simeon CP, Corbella X, Fonollosa V. Mortality and survival in systemic sclerosis: systematic review and meta-analysis. Semin Arthritis Rheum. 2014 Oct;44(2):208-19.
4. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972-2002. Ann Rheum Dis. 2007 Jul;66(7):940-4.
5. Tyndall AJ, Bannert B, Vonk M, Airo P, Cozzi F, Carreira PE, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis. 2010 Oct;69(10):1809-15.
6. Wangkaew S, Euathrongchit J, Wattanawittawas P, Kasitanon N, Louthrenoo W. Incidence and predictors of interstitial lung disease (ILD) in Thai patients with early systemic sclerosis: Inception cohort study. Mod Rheumatol. 2016 Jul;26(4):588-93.
7. Khanna D, Tseng CH, Farmani N, Steen V, Furst DE, Clements PJ, et al. Clinical course of lung physiology in patients with scleroderma and interstitial lung disease: analysis of the Scleroderma Lung Study Placebo Group. Arthritis Rheum. 2011 Oct;63(10):3078-85.
8. Ho YY, Lagares D, Tager AM, Kapoor M. Fibrosis--a lethal component of systemic sclerosis. Nat Rev Rheumatol. 2014 Jul;10(7):390-402.
9. Lamas DJ, Kawut SM, Bagiella E, Philip N, Arcasoy SM, Lederer DJ. Delayed access and survival in idiopathic pulmonary fibrosis: a cohort study. Am J Respir Crit Care Med. 2011 Oct 1;184(7):842-7.
10. Assassi S, Sharif R, Lasky RE, McNearney TA, Estrada YMRM, Draeger H, et al. Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. Arthritis Res Ther. 2010;12(5):R166.
11. Jaeger VK, Wirz EG, Allanore Y, Rossbach P, Riemekasten G, Hachulla E, et al. Incidences and Risk Factors of Organ Manifestations in the Early Course of Systemic Sclerosis: A Longitudinal EUSTAR Study. PLoS One. 2016;11(10):e0163894.
12. Wu W, Jordan S, Graf N, de Oliveira Pena J, Curram J, Allanore Y, et al. Progressive skin fibrosis is associated with a decline in lung function and worse survival in patients with diffuse cutaneous systemic sclerosis in the European Scleroderma Trials and Research (EUSTAR) cohort. Ann Rheum Dis. 2019 May;78(5):648-56.
13. Ghuman A, Khanna D, Lin CJF, Furst DE, Raghu G, Martinez FJ, et al. Prognostic and predictive markers of systemic sclerosis-interstitial lung disease in a clinical trial and long-term observational cohort. Rheumatology (Oxford). 2023 May 25.
14. Khanna D, Maher TM, Volkmann ER, Allanore Y, Smith V, Assassi S, et al. Effect of nintedanib in patients with systemic sclerosis-associated interstitial lung disease and risk factors for rapid progression. RMD Open. 2023 Feb;9(1).
15. American College of Rheumatology (ACR) Guideline for the Screening and Monitoring of Interstitial Lung Disease in People with Systemic Autoimmune Rheumatic Disease. 2023.
16. Phillips M, Cataneo RN, Saunders C, Hope P, Schmitt P, Wai J. Volatile biomarkers in the breath of women with breast cancer. J Breath Res. 2010 Jun;4(2):026003.

17. Song G, Qin T, Liu H, Xu GB, Pan YY, Xiong FX, et al. Quantitative breath analysis of volatile organic compounds of lung cancer patients. Lung Cancer. 2010 Feb;67(2):227-31.

18. Syslova K, Kacer P, Kuzma M, Pankracova A, Fenclova Z, Vlckova S, et al. LC-ESI-MS/MS method for oxidative stress multimarker screening in the exhaled breath condensate of asbestosis/silicosis patients. J Breath Res. 2010 Mar;4(1):017104.

19. Westhoff M, Litterst P, Freitag L, Baumbach JI. Ion mobility spectrometry in the diagnosis of sarcoidosis: results of a feasibility study. J Physiol Pharmacol. 2007 Nov;58 Suppl 5(Pt 2):739-51.

20. Jalali M, Zare Sakhvidi MJ, Bahrami A, Berijani N, Mahjub H. Oxidative Stress Biomarkers in Exhaled Breath of Workers Exposed to Crystalline Silica Dust by SPME-GC-MS. J Res Health Sci. 2016 Summer;16(3):153-61.

21. Fijten RRR, Smolinska A, Drent M, Dallinga JW, Mostard R, Pachen DM, et al. The necessity of external validation in exhaled breath research: a case study of sarcoidosis. J Breath Res. 2017 Nov 29;12(1):016004.

22. Yamada YI, Yamada G, Otsuka M, Nishikiori H, Ikeda K, Umeda Y, et al. Volatile Organic Compounds in Exhaled Breath of Idiopathic Pulmonary Fibrosis for Discrimination from Healthy Subjects. Lung. 2017 Apr;195(2):247-54.

23. Yang HY, Shie RH, Chang CJ, Chen PC. Development of breath test for pneumoconiosis: a case-control study. Respir Res. 2017 Oct 17;18(1):178.

24. Guiot J, Zanella D, Stefanuto PH, Giltay L, Henket M, Guissard F, et al. Exhaled Volatile Organic Compounds are Able to Diagnose Systemic Sclerosis. European Respiratory Journal. 2020 2020-09-07 00:00:00;56.

25. Hayton CT, D. Ahmed W. Exhaled volatile organic compounds in idiopathic pulmonary fibrosis and disease progression. European Respiratory Journal. 2020 2020-09-07 00:00:00;56.

26. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).

27. Plantier L, Smolinska A, Fijten R, Flamant M, Dallinga J, Mercadier JJ, et al. The use of exhaled air analysis in discriminating interstitial lung diseases: a pilot study. Respir Res. 2022 Jan 20;23(1):12.

28. van der Sar IG, Moor CC, Oppenheimer JC, Luijendijk ML, van Daele PLA, Maitland-van der Zee AH, et al. Diagnostic Performance of Electronic Nose Technology in Sarcoidosis. Chest. 2022 Mar;161(3):738-47.

29. van der Sar IG, Wijsenbeek MS, Moor CC. Exhaled breath analysis in interstitial lung disease. Curr Opin Pulm Med. 2023 Sep 1;29(5):443-50.

30. van Raaij BFM, Veltman JD, Hameete JF, Stoger JL, Geelhoed JJM. Diagnostic performance of eNose technology in COVID-19 patients after hospitalization. BMC Pulm Med. 2023 Apr 20;23(1):134.

31. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. Ann Rheum Dis. 2013 Nov;72(11):1747-55.

32. Flaherty KR, King TE, Jr., Raghu G, Lynch JP, 3rd, Colby TV, Travis WD, et al. Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? Am J Respir Crit Care Med. 2004 Oct 15;170(8):904-10.

33. Liem SIE, Ciaffi J, van Leeuwen NM, Boonstra M, Ahmed S, Beaart-van de Voorde LJJ, et al. Step forward in early recognition of systemic sclerosis: data from the Leiden CCISS cohort. RMD Open. 2023 Apr;9(2).

34. de Vries R, Brinkman P, van der Schee MP, Fens N, Dijkers E, Bootsma SK, et al. Integration of electronic nose technology with spirometry: validation of a new approach for exhaled breath analysis. J Breath Res. 2015 Oct 15;9(4):046001.

35. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. Eur Respir J. 2018 Jan;51(1).

36. Broadhurst DKD. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics. 2006;2(4):171-96.

37. Goh NS, Desai SR, Veeraraghavan S, Hansell DM, Copley SJ, Maher TM, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. Am J Respir Crit Care Med. 2008 Jun 1;177(11):1248-54.

38. Trombetta AC, Soldano S, Contini P, Tomatis V, Ruaro B, Paolino S, et al. A circulating cell population showing both M1 and M2 monocyte/macrophage surface markers characterizes systemic sclerosis patients with lung involvement. Respir Res. 2018 Sep 24;19(1):186.

39. Lescoat A, Lecureur V, Varga J. Contribution of monocytes and macrophages to the pathogenesis of systemic sclerosis: recent insights and therapeutic implications. Curr Opin Rheumatol. 2021 Nov 1;33(6):463-70.

40. Valenzi E, Bulik M, Tabib T, Morse C, Sembrat J, Trejo Bittar H, et al. Single-cell analysis reveals fibroblast heterogeneity and myofibroblasts in systemic sclerosis-associated interstitial lung disease. Ann Rheum Dis. 2019 Oct;78(10):1379-87.

41. Chen W. IDO: more than an enzyme. Nat Immunol. 2011 Aug 18;12(9):809-11.

42. Rath M, Muller I, Kropf P, Closs EI, Munder M. Metabolism via Arginase or Nitric Oxide Synthase: Two Competing Arginine Pathways in Macrophages. Front Immunol. 2014;5:532.

43. Zanella D, Guiot J, Stefanuto PH, Giltay L, Henket M, Guissard F, et al. Breathomics to diagnose systemic sclerosis using thermal desorption and comprehensive two-dimensional gas chromatography high-resolution time-of-flight mass spectrometry. Anal Bioanal Chem. 2021 Jun;413(14):3813-22.

44. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh FH, van Putten JWG, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2023 Mar;163(3):697-706.

45. Bonhomme O, Andre B, Gester F, de Seny D, Moermans C, Struman I, et al. Biomarkers in systemic sclerosis-associated interstitial lung disease: review of the literature. Rheumatology (Oxford). 2019 Sep 1;58(9):1534-46.

46. Van Kalsbeek D, Brooks R, Shaver D, Ebel A, Hershberger D, Schmidt C, et al. Peripheral Blood Biomarkers for Rheumatoid Arthritis-Associated Interstitial Lung Disease: A Systematic Review. ACR Open Rheumatol. 2023 Apr;5(4):201-26.

SUPPLEMENTARY DATA A

Additional baseline characteristics

Table S1: Baseline characteristics ILD cohort.

	Patients with ILD
Subjects	300
Females	107 (35.7)
Age (years)	71.00 [63.0, 76.0]
Smoking status	
Never	81 (27.0)
Former	209 (69.7)
Current	10 (3.3)
FVC (%pred)	78.68 (20.3) *
DLCOc (%pred)	50.30 (15.7) **
Diagnosis	
Idiopathic pulmonary fibrosis	124 (41.3)
Chronic hypersensitivity pneumonitis	50 (16.7)
Connective tissue disease-related ILD	47 (15.7)
Idiopathic NSIP	20 (6.7)
IPAF	15 (5.0)
CPFE	10 (3.3)
Unclassifiable	9 (3.0)
Cryptogenic organizing pneumonia	9 (3.0)
Respiratory bronchiolitis-ILD	4 (1.3)
Vasculitis	4 (1.3)
Desquamative interstitial pneumonia	3 (1.0)
Asbestosis	3 (1.0)
Drug induced ILD	2 (0.7)

*Values are displayed as number (%), mean ±SD, or median [interquartile range]. CPFE = combined pulmonary fibrosis and emphysema; DLCOc = diffusing capacity for carbon monoxide corrected for hemoglobin level; FVC = forced vital capacity; ILD = interstitial lung disease; IPAF = interstitial pneumonia with autoimmune features; NSIP = nonspecific interstitial pneumonia; %pred = percent of predicted value, calculated based on sex, age and height. * n=4 values missing. ** n=18 values missing.*

SUPPLEMENTARY DATA B

Outlier detection results

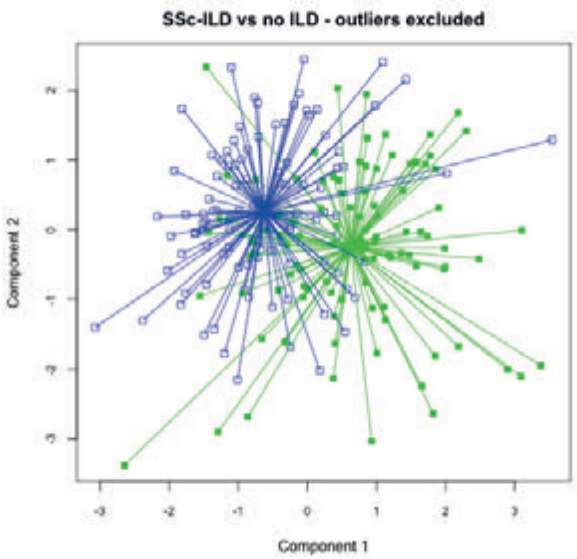


Figure S1: Scatterplot of patients with SSc-ILD versus patients with SSc without ILD excluding outliers. *Outliers n=11, resulting in SSc-ILD n=106 and SSc without ILD n=106. X- and Y-axis represent first 2 principal components. ILD = interstitial lung disease; SSc = systemic sclerosis.*

Table S2: Results of eNose breath analysis between patients with SSc with and without ILD excluding outliers.

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
SSc-ILD	71	**SSc without ILD**	71	Training	0.85	*0.78-0.92*	0.89	0.73	0.81	0.77	0.87
	35		35	Test	0.77	*0.66-0.88*	0.77	0.69	0.73	0.71	0.75

Outliers n=11, resulting in SSc-ILD n=106 and SSc without ILD n=106. Results based on 2 principal components. AUC = area under the curve; CI = confidence interval; ILD = interstitial lung disease; NPV = negative predictive value; PPV = positive predictive value; SSc = systemic sclerosis.

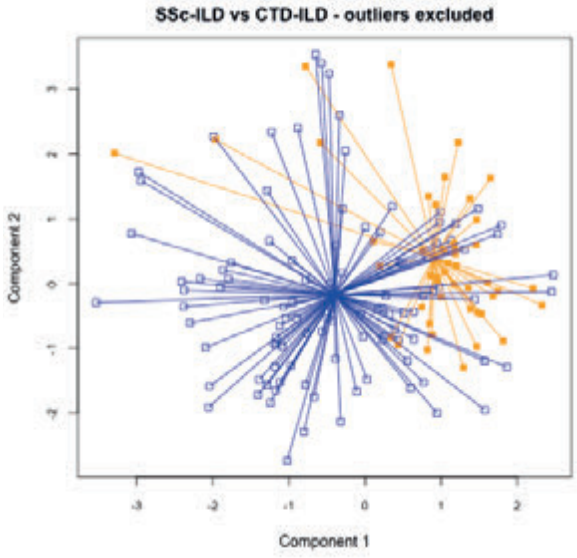


Figure S2: Scatterplot of patients with SSc-ILD versus patients with other CTD-ILDs excluding outliers. *Outliers n=6, resulting in SSc-ILD n=105 and CTD-ILD n=46. X- and Y-axis represent first 2 principal components. CTD = connective tissue disease; ILD = interstitial lung disease; SSc = systemic sclerosis.*

Table S3: Results of eNose breath analysis between patients with SSc with ILD and other CTD-ILD excluding outliers.

Group 1	n=	Group 2	n=	Dataset	AUC	95% CI	Specificity	Sensitivity	Accuracy	NPV	PPV
SSc-ILD	70	**CTD-ILD**	31	Training	0.87	*0.80-0.94*	0.90	0.76	0.80	0.62	0.95
	35		15	Test	0.71	*0.55-0.88*	0.73	0.74	0.74	0.55	0.87

Outliers n=6, resulting in SSc-ILD n=105 and CTD-ILD n=46. Results based on 2 principal components. AUC = area under the curve; CI = confidence interval; CTD = connective tissue disease; ILD = interstitial lung disease; NPV = negative predictive value; PPV = positive predictive value; SSc = systemic sclerosis.

SUPPLEMENTARY DATA C

Potential influencing factors

Table S4: Results of eNose breath analysis between patient subgroups with potential influencing factors of breath profiles.

Factor	Group 1	n=	Group 2	n=	Cohort	AUC	95% CI
Sex	**Male**	56	**Female**	165	All SSc	0.66	0.57-0.74
Smoking status	**Never**	89	**Ever**	129	All SSc	0.57	0.50-0.65
Age	**<60 years**	115	**>60 years**	106	All SSc	0.66	0.59-0.73
Immunosuppressive drug use	**Yes**	78	**No**	178	All SSc	0.64	0.56-0.71
SSc duration	**Early**	39	**Late**	184	All SSc	0.62	0.52-0.72
ILD severity	**Mild**	89	**Severe**	20	SSc-ILD	0.71	0.58-0.84

Results based on 1 principal component. Age cut-off is based on median age of included SSc cohort. Cut-off for early or late SSc diagnosis is 24 months since SSc diagnosis. AUC = area under the curve; CI = confidence interval; ILD = interstitial lung disease; SSc = systemic sclerosis.

CHAPTER 10

Detection of drug-induced interstitial lung disease caused by cancer treatment using electronic nose exhaled breath analysis

Ann Am Thorac Soc. 2024 Mar 25.

Iris G. van der Sar, Marlies S. Wijsenbeek, Daphne W. Dumoulin, Agnes Jager, Astrid A.M. van der Veldt, Mariska J.P. Rossius, Anne-Marie C. Dingemans, Catharina C. Moor

Erasmus Medical Center, Rotterdam (the Netherlands)

Introduction

The introduction of novel cancer treatments such as immune checkpoint inhibitors, tyrosine kinase inhibitors and antibody-drug conjugates have improved survival of patients with cancer. However, many cancer treatments are associated with potentially life-threatening adverse events, including drug-induced interstitial lung disease (DIILD), also known as drug-related pneumonitis. Currently, chest CT scans are used for screening of DIILD in high-risk patients, but are often inconclusive and entail radiation exposure [1].

Incidence of DIILD in patients receiving cancer treatment ranges from 1.12-4.77% in EGFR inhibitor and 1.14-6.25% in ALK inhibitor use [2]. Meta-analyses report incidence rates for all-grade DIILD of 2.7% in PD1 inhibitor and for severe DIILD 1.7% in mTOR inhibitor use [3, 4]. Notably, clinical trials with specific antibody-drug conjugates reported incidence rates up to 26% [1]. Median time to onset is several months, ranging from days to years [5].

Diagnosing DIILD can be challenging, due to overlapping features with pulmonary infections, progressive malignancy, radiation pneumonitis, lymphangitis carcinomatosa or metastases. Therefore, a multidisciplinary approach is warranted for diagnosis and treatment [6]. High-dose corticosteroids and interrupting suspected causative drug are main aspects of treatment. Despite this, mortality rates remain high in severe DIILD [6].

A non-invasive, rapid, accurate test to facilitate frequent monitoring, timely diagnosis, and adequate treatment is needed. Volatile organic compounds (VOCs) in exhaled breath are products from various metabolic and pathophysiologic processes in the human body, and diffuse from the blood stream to the lungs or are produced within the lungs. VOCs can be measured in real-time using an electronic nose (eNose). We previously showed that patients with ILD have a specific VOC breath profile, different from healthy controls and other respiratory diseases [7, 8]. Here, we investigated whether eNose technology can differentiate patients with cancer who have DIILD, and those without DIILD.

Keywords
Breath tests, electronic nose, interstitial lung diseases, antineoplastic agents, early diagnosis

Materials and Methods

Patients with pathology-proven cancer diagnosis and suspected DIILD caused by cancer treatment, were included October 2021-November 2023. A separate control cohort was recruited among patients with similar cancer diagnosis and treatment, without ILD present on CT scan. Patients with a pulmonary infection or alcohol consumption ≤8 hours were excluded.

A multidisciplinary team consisting of pulmonologists with expertise in ILD and oncology, and a thoracic radiologist discussed all cases. Combining baseline characteristics, CT scan results and clinical course led to a DIILD diagnostic likelihood (≤50, 51-69, 70-89, ≥90%) and Common Terminology Criteria for Adverse Events grade for each case [9, 10]. Patients were excluded if DIILD likelihood was ≤50% or another diagnosis was more likely. The date that the treating physician raised suspicion of DIILD in the patient file was set as start of DIILD.

The SpiroNose® (Breathomix) was used for breath analysis [11]. After signing informed consent, patients performed a breath maneuver and answered questions (including recent food intake, medication use, smoking history). Breath maneuvers consisted of five tidal breaths, followed by maximum inspiration, five second breath-hold and maximum exhalation. Technical details of the device can be found elsewhere [11]. Clinical data were collected from files and stored in secured databases (BreathBase®, Castor®).

Consequently, breath data were pre-processed and low-quality measurements excluded. Labelled breath data were analyzed using sparse partial least squares discriminant analysis including cross-validation. This method first reduces the dimensionality of data and results in several principal components that are a weighted combination of eNose sensor data. The first component represents the highest variation of data to discriminate between cohorts, and was therefore used for classification of patients with DIILD versus patients without and receiver operating characteristic analysis. Sub analyses were conducted to indicate potential influence of corticosteroid use. R version 4.2.1 for Windows with mixOmics package version 6.20.0 was used for analyses.

Results

25 patients suspected of DIILD were included; one low-quality measurement and four with a more likely alternative diagnosis were excluded. Subsequently, 20 matched control patients were recruited. DIILD occurred at median 2.8 [1.5, 6.1] months after start of cause-related cancer treatment, and 11 (55%) patients were treated with corticosteroids for DIILD at inclusion. Patients' characteristics are described in **Table 1**.

Table 1: Baseline characteristics included patients.

	DIILD n=20	CONTROL n=20
Female sex	9 (45)	9 (45)
Age (years)	65.3 ±11.7	67.6 ±9.4
Smoking history		
Never	3 (15)	7 (35)
Former	17 (85)	13 (65)
Cancer diagnosis		
Non-small cell lung cancer	8 (40)	8 (40)
Urogenital	5 (25)	5 (25)
Melanoma	4 (20)	4 (20)
Mesothelioma	3 (15)	3 (15)
Cancer therapy		
Nivolumab (ipilimumab)	8 (40)	8 (40)
Pembrolizumab, pemetrexed, carboplatin	4 (20)	4 (20)
Sotorasib	2 (10)	2 (10)
Capmatinib	1 (5)	1 (5)
Dabrafenib, trametinib	1 (5)	1 (5)
Gemcitabine (carboplatin)	1 (5)	1 (5)
GEN1046-04 (pembrolizumab)	1 (5)	1 (5)
Osimertinib	1 (5)	1 (5)
Taxane	1 (5)	1 (5)
Time from start cancer treatment to inclusion (months)	4.9 [2.2,10.1]	3.8 [11.4,20.1]
Time from start cancer treatment to DIILD (months)	2.8 [1.5,6.1]	-
Current corticosteroid use for DIILD	11 (55)	-
MDT likelihood DIILD diagnosis		
51-69%	2 (10)	-
71-89%	8 (40)	-
≥90%	10 (50)	-
CTCAE grade DIILD		
I-II	8 (40)	-
III-IV	9 (45)	-
V	3 (15)	-

Values are displayed as number (%), mean ± SD, or median [interquartile range]. Treatment within brackets was not received by one or more patients in either of the groups. CTCAE = Common Terminology Criteria for Adverse Events (version 5.0); DIILD = drug-induced interstitial lung disease; MDT = multidisciplinary team.

Comparing breath data of the patients with and without DIILD resulted in an area under the curve of 0.81 (95%CI 0.67-0.95, **Figure 1**) with corresponding sensitivity and specificity of 0.75.

Repeating the analysis in patients with DIILD using of corticosteroids versus patients without DIILD, resulted in a similar area under the curve of 0.80 (0.63-0.96). Patients with DIILD without corticosteroid use versus patients without DIILD increased the area under the curve value to 0.87 (0.74-1.00).

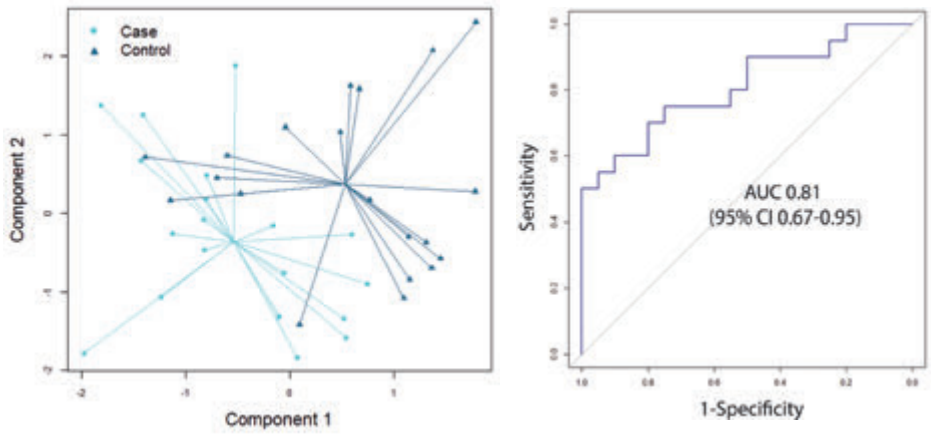


Figure 1: Breath profile comparison.
A. Scatterplot of individual breath profiles of patients with cancer diagnosed with DIILD (case) versus patients without DIILD (control). Each dot in the scatterplot represents one patient. Component 1 and 2 are the first two principal components resulting from sparse partial least squares discriminant analysis. B. Receiver operating characteristics curve, based on the first component. AUC = area under the curve; CI = confidence interval; DIILD = drug-induced interstitial lung disease.

Discussion

This proof-of-concept study showed that eNose technology can distinguish patients with and without DIILD in a cohort of patients with various types of cancer and treatment. This important first step confirms the potential of exhaled breath analysis for detecting pulmonary toxicity caused by cancer treatment.

This is the first study on breath biomarkers in DIILD. Other studies investigated serum proteins and whole genome sequencing, but findings were not validated [12-14]. Studies in patients with lung cancer and radiation pneumonitis identified serum

amyloid A and KL6 as potentially predictive [15, 16]. These proteins have been previously found to be elevated in ILDs, and are not specific for pulmonary toxicity due to cancer treatment. Also, in other ILDs no diagnostic biomarker is validated for routine use in clinic [17]. It seems unlikely that a single marker exists for these complex, heterogeneous pulmonary condition. Sets of biomarkers, e.g. a breath profile, might have more potential as valid test as it includes more disease-specific information.

Our results suggest that DIILD has a different breath profile compared to patients without DIILD. Further research needs to validate this in prospective follow-up studies. Projects should focus on the potential of eNose for screening purposes in patients with increased DIILD risk, due to cancer treatment or patient-related factors [6]. Using eNose for ruling out DIILD has most benefits in these high-risk patients to prevent extra radiation exposure. This is especially relevant in younger patients using cancer medication for a long period. Besides, accessible screening can facilitate earlier DIILD treatment. This may increase the number of patients with reversible pulmonary damage and possibilities to continue cancer treatment.

A limitation of our study is the sample size. We applied cross-validation to avoid overfitting and present robust results. Our study also comprises a heterogeneous population including patients with various malignancies, cancer treatments and corticosteroids dose. The high performance of eNose for detecting DIILD (AUC of 0.81) despite this heterogeneity, suggests that breath profiles do not seem to be driven by these patient-related factors. Which is also confirmed by sub analyses in patients with or without corticosteroid use. This supports the hypothesis that eNose can detect DIILD, and breath profiles do not seem to be driven by malignancy, cancer treatment or corticosteroids. Another limitation is the majority of patients having high grade DIILD, indicating possible selection bias. This is not a significant concern for this proof-of-concept study, but should be taken into account for future studies.

eNose technology seems to differentiate patients with and without DIILD, and has therefore potential as a novel point-of-care test for screening and monitoring DIILD in patients with cancer and current treatment. Our findings encourage validation of eNose for identifying DIILD in prospective multicenter studies, in the current era of expanding access to pulmonary toxic cancer treatments.

References

1. Swain SM, Nishino M, Lancaster LH, Li BT, Nicholson AG, Bartholmai BJ, et al. Multidisciplinary clinical guidance on trastuzumab deruxtecan (T-DXd)-related interstitial lung disease/pneumonitis-Focus on proactive monitoring, diagnosis, and management. Cancer Treat Rev. 2022 May;106:102378.
2. Johkoh T, Lee KS, Nishino M, Travis WD, Ryu JH, Lee HY, et al. Chest CT Diagnosis and Clinical Management of Drug-related Pneumonitis in Patients Receiving Molecular Targeting Agents and Immune Checkpoint Inhibitors: A Position Paper from the Fleischner Society. Radiology. 2021 Mar;298(3):550-66.
3. Nishino M, Giobbie-Hurder A, Hatabu H, Ramaiya NH, Hodi FS. Incidence of Programmed Cell Death 1 Inhibitor-Related Pneumonitis in Patients With Advanced Cancer: A Systematic Review and Meta-analysis. JAMA Oncol. 2016 Dec 1;2(12):1607-16.
4. Zhang X, Ran YG, Wang KJ. Risk of mTOR inhibitors induced severe pneumonitis in cancer patients: a meta-analysis of randomized controlled trials. Future Oncol. 2016 Jun;12(12):1529-39.
5. Spagnolo P, Bonniaud P, Rossi G, Sverzellati N, Cottin V. Drug-induced interstitial lung disease. Eur Respir J. 2022 Oct;60(4).
6. Conte P, Ascierto PA, Patelli G, Danesi R, Vanzulli A, Sandomenico F, et al. Drug-induced interstitial lung disease during cancer therapies: expert opinion on diagnosis and treatment. ESMO Open. 2022 Apr;7(2):100404.
7. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).
8. van der Sar IG, Wijsenbeek MS, Braunstahl GJ, Loekabino JO, Dingemans AC, In 't Veen J, et al. Differentiating interstitial lung diseases from other respiratory diseases using electronic nose technology. Respir Res. 2023 Nov 6;24(1):271.
9. National Cancer Institute Common terminology criteria for adverse events (CTCAE) version 5.0.
10. Ryerson CJ, Corte TJ, Lee JS, Richeldi L, Walsh SLF, Myers JL, et al. A Standardized Diagnostic Ontology for Fibrotic Interstitial Lung Disease. An International Working Group Perspective. Am J Respir Crit Care Med. 2017 Nov 15;196(10):1249-54.
11. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. Eur Respir J. 2018 Jan;51(1).
12. Atagi S, Katakami N, Yoshioka H, Fukuoka M, Kudoh S, Ogiwara A, et al. Nested case control study of proteomic biomarkers for interstitial lung disease in Japanese patients with non-small-cell lung cancer treated with erlotinib: a multicenter phase IV study (JO21661). Clin Lung Cancer. 2013 Jul;14(4):407-17.
13. Nakamura K, Kato M, Shukuya T, Mori K, Sekimoto Y, Ihara H, et al. Surfactant protein-D predicts prognosis of interstitial lung disease induced by anticancer agents in advanced lung cancer: a case control study. BMC Cancer. 2017 May 2;17(1):302.
14. Udagawa C, Horinouchi H, Shiraishi K, Kohno T, Okusaka T, Ueno H, et al. Whole genome sequencing to identify predictive markers for the risk of drug-induced interstitial lung disease. PLoS One. 2019;14(10):e0223371.

15. Goto K, Kodama T, Sekine I, Kakinuma R, Kubota K, Hojo F, et al. Serum levels of KL-6 are useful biomarkers for severe radiation pneumonitis. Lung Cancer. 2001 Oct;34(1):141-8.

16. Wang YS, Chang HJ, Chang YC, Huang SC, Ko HL, Chang CC, et al. Serum amyloid a as a predictive marker for radiation pneumonitis in lung cancer patients. Int J Radiat Oncol Biol Phys. 2013 Mar 1;85(3):791-7.

17. Inoue Y, Kaner RJ, Guiot J, Maher TM, Tomassetti S, Moiseev S, et al. Diagnostic and Prognostic Biomarkers for Chronic Fibrosing Interstitial Lung Diseases With a Progressive Phenotype. Chest. 2020 Aug;158(2):646-59.

PART IV

GAS CHROMATOGRAPHY-MASS SPECTROMETRY

CHAPTER 11

Gas chromatography -mass spectrometry exhaled breath analysis for phenotyping interstitial lung disease

Submitted.

Iris G. van der Sar[1#], Iris A. Simons[2#], Roxanne van Duren[1], Esther J. Nossent[2], Marlies S. Wijsenbeek[1], Jan Willem Duitman[2], Arlette E. Odink[1], Lilian J. Meijboom[2], Catharina C. Moor[1#], Paul Brinkman[2#]

[1]Erasmus Medical Center, Rotterdam (the Netherlands); [2]Amsterdam University Medical Center, Amsterdam (the Netherlands); [#]Authors share first/last authorship

Abstract

Background

Interstitial lung disease (ILD) encompasses pulmonary disorders characterised by varying degrees of inflammation and/or fibrosis. The presence and extent of these pulmonary abnormalities on (high-resolution) computed tomography (CT) have consequences for diagnosis and treatment; however, inter-observer assessment varies. Analysis of exhaled volatile organic compounds (VOCs) through gas chromatography-mass spectrometry (GC-MS) offers a noninvasive approach to biomarker discovery and pathophysiology understanding. Our study aims to explore the ability of GC-MS-driven exhaled breath analysis to differentiate ILD patients with predominant fibrotic, inflammatory, or a combination of fibrotic and inflammatory pulmonary abnormalities in a training and an external validation cohort.

Methods

In a multicentre cross-sectional study, patients diagnosed with ILD were recruited. After central review of chest CT scans by independent radiologists, patients were categorised as fibrotic, inflammatory or mixed phenotype group based on the percentage of chest CT scan abnormalities. Breath samples were collected and analysed via GC-MS. Significantly different VOC fragments between groups were selected and used to differentiate groups in the training cohort with sparse partial least squares discriminant analysis. Analyses were validated with patients from an external cohort.

Results

53 patients were included, 21 patients in the fibrotic, 14 in the inflammatory and 18 in the mixed phenotype group. AUCs for discrimination between groups ranged from 0.83-0.95 in training cohorts. An attempt to confirm these findings in our external validation cohort resulted in AUCs of 0.57-0.63.

Conclusions

This study shows that GC-MS driven exhaled breath analysis towards differentiation of ILD phenotypes is challenging. Current findings emphasise the importance of predefined validation steps during the process of biomarker discovery.

Keywords

Exhaled breath analysis, volatile organic compounds, phenotype, interstitial lung disease, validation study

Background

Interstitial lung disease (ILD) encompasses a heterogeneous group of pulmonary disorders characterised by varying degrees of inflammation and/or fibrosis within the pulmonary interstitium [1]. ILD can be detected on chest computed tomography (CT) scans, but chest CT assessment is subject to inter-observer variation [2] and is not conclusive on ILD diagnosis. No single test for diagnosing ILD exists. Therefore, a comprehensive multidisciplinary approach, integrating radiological, histopathological and clinical examination results is essential for establishing a consensus diagnosis and formulating effective therapeutic strategies.

In clinical practice, individual therapeutic strategies are often guided by chest CT in the absence of better biomarkers but the choice and timing of therapy remain challenging, in particular for non-IPF pulmonary fibrosis. Immunosuppressive agents are often prescribed for inflammation-dominant disease, while pulmonary fibrosis is treated with antifibrotic therapy [3]. The progressive pulmonary fibrosis (PPF) phenotype is a treatment indication for antifibrotic therapy, but guidance on the use of immunosuppression is lacking. There is limited data that BAL lymphocytosis may predict response to immunosuppression, but this is not validated and requires invasive procedures. Therefore, there is a high need for phenotype biomarkers to guide treatment strategies.

The analysis of exhaled volatile organic compounds (VOCs) is a potential source of novel biomarkers for diagnosing and monitoring respiratory diseases, including ILD [4]. VOCs reflect both physiological and pathophysiological processes. The exhaled mixture of VOCs leads to an individual breath profile and can be analysed noninvasively by two main techniques. Electronic nose (eNose) technology uses multiple chemical sensors to generate sensor response patterns. It has the potential as a point-of-care clinical test with real-time results. Previous results show different eNose breath profiles for patients with Idiopathic pulmonary fibrosis (IPF) and other ILDs and therefore may hold clinical promise [5]. A limitation of this technology is the inability of VOC identification preventing to unravel the origin of breath profile differences between patient groups. A second technique often used for exhaled VOC analysis is gas chromatography-mass spectrometry (GC-MS) and does enable the identification of VOCs based on chromatographic peaks. This may offer insights into specific pathophysiological processes and associated VOCs. So far, two biomarker discovery studies reported that VOCs measured by GC-MS significantly differed in IPF and connective tissue disease (CTD) associated ILD, or healthy controls [6, 7]. In line, another study compared sarcoidosis with healthy persons and found significant differences in exhaled VOCs, but external validation appeared hard [8].

Our study aims to evaluate whether ILD patients with a predominant fibrotic, inflammatory, or combined fibrotic and inflammatory phenotype can be differentiated using GC-MS-driven exhaled breath analysis in a training and as additional step in an external validation cohort.

Methods

Study Design & Population

In this multicentre cross-sectional observational cohort study, we recruited patients from the outpatient pulmonary clinics of two ILD expert centres in the Netherlands, the Erasmus University Medical Centre Rotterdam (Erasmus MC; training cohort) and the Amsterdam University Medical Centre (Amsterdam UMC; validation cohort) between September 2022 and September 2023.

Patients with an ILD diagnosis established in the multidisciplinary team (consisting of pulmonologists, radiologists and pathologists specialized in ILD) meeting (MDT) at the treating medical centre according to the current guidelines [1, 9-11] were included if they met the criteria of one of the following groups:

- *Fibrotic Phenotype:* IPF diagnosis with a fibrosis extent ≥10% and inflammatory irregularities <10% on chest CT scan.
- *Inflammatory Phenotype:* A non-IPF ILD diagnosis characterised by inflammation such as CTD-ILD or sarcoidosis, with a fibrosis extent <10% and inflammatory irregularities ≥10% on chest CT scan.
- *Mixed Phenotype:* A non-IPF ILD diagnosis with a fibrosis extent ≥10% and inflammatory irregularities ≥10% on chest CT scan.

Exclusion criteria were anti-inflammatory or antifibrotic treatment, no available chest (high-resolution) CT scan with thin slices within six months prior to inclusion, the presence of another pulmonary disease, malignancy, or current respiratory infection, and alcohol consumption within 8 hours before breath measurement.

After inclusion, there was a central review of all chest CTs by experienced thoracic radiologists from both centres (AO and LM), including discussion till consensus in case of discrepancies. A patient was excluded if no consensus was reached about phenotype group classification. Pulmonary fibrosis was defined as the presence of reticulation and traction bronchiectasis with or without honeycombing [9], and inflammatory changes were defined as ground glass opacities, consolidations, and/or diffuse nodular abnormalities [12].

Group Comparison

With the Erasmus MC patients serving as training cohort and patients from the Amsterdam UMC as external validation cohort, breath data from the following groups were compared: patients with a fibrotic vs inflammatory phenotype, a fibrotic vs mixed phenotype, and an inflammatory vs mixed phenotype. Additionally, we performed the same analyses but excluded the sarcoidosis patients to assess whether this heterogeneous patient category influences results.

Data Collection

Breath samples were collected during the patient's outpatient appointment. Each patient performed a single breath manoeuvre through a portable breath sampling setup to collect exhaled breath for GC-MS analysis (see **Figure 1**). This setup consisted of a T-piece with a one-way valve (Directional valve 1954000, Intersurgical Ltd, Berkshire, UK; ISO 9001:2015, ISO 13485:2016, ISO 14001:2015 certified and EC certified) connected to a facemask equipped with an anti-viral and antibacterial filter (ClearLiteTM, anaesthetic face mask 7292001, Intersurgical Ltd, Berkshire, UK; ISO 9001:2015, ISO 13485:2016, ISO 14001:2015 and EC certified). A carbon filter (Carbon Filter N06575001L, Honeywell Safety Products Ltd, UK; EN14387 and EC certified) was attached to one end of the T-piece to reduce contamination from ambient air gasses.

Figure 1: Example of the portable breath sampling setup used for breath collection.
Credits to the breath research group of the Amsterdam UMC.

Patients were instructed to take ten tidal breaths followed by maximum inhalation and breath hold. During the breath hold, a plastic sampling bag (Mylar 800/PET 45x50cm, Meda-Pak BV, Uithoorn, The Netherlands; ISO 9001:2015 and ISO 14001:2015 certified) was attached to the other end of the T-adapter. Patients were subsequently instructed to exhale maximally into the bag.

To trap VOCs from the collected breath, breath was directed through a stainless-steel thermal desorption tube filled with Tenax GR 60/80 (CAMSCO, Interscience, Breda, The Netherlands) using a handheld air pump (Gastec, Kanagawa, Japan). The pump pulled the breath through the tube for two minutes at a flow rate of 250 ml/min. During this process, exhaled VOCs are trapped on the Tenax. Subsequently, the tubes were securely sealed and stored refrigerated until laboratory analysis. The carbon filter and pump were cleaned for reuse, while the other materials were disposed.

We recorded various clinical parameters from medical files. The parameters included demographics information like ILD diagnosis, and results of the most recent lung function test (forced vital capacity (FVC) and diffusion capacity of the lung for carbon monoxide corrected for haemoglobin (DLCOc)), chest (high-resolution) CT scan, and lung biopsy results if available. In addition, patients answered a short questionnaire regarding potential factors influencing exhaled breath, including recent food, drink, and medication intake within the past two hours, smoking history, and alcohol consumption.

GC-MS Analysis

Collected breath samples were analysed at Amsterdam UMC. First, the tubes were transferred to a thermal desorption unit (Markes TD100, Cincinnati, Ohio, USA) and heated to 250°C for 15 minutes at a flow rate of 30 ml/min to release VOCs from the sorbent tubes. These VOCs were then captured on a cold trap of 25°C and subsequently heated to 300°C for one minute to release them again. The VOCs were injected through a transfer line at 180°C and a flow rate of 1.2 ml/min onto an Inertcap 5MS/Sil gas-chromatography column (30m, ID 0.25mm, film thickness 1µm, 1,4-bis-(dimethylsiloxy)phenylene dimethyl polysiloxam; Restek, Breda, The Netherlands). For five minutes, the oven temperature was maintained at 40°C, after which the temperature was increased at a rate of 10°C per minute until it reached 280°C. The temperature was held isothermal for another five minutes.

Molecules were ionised through electron ionisation (70 eV), and the ions were detected using a quadruple mass-spectrometer (GCMS-GP2010; Shimadzu, Den Bosch, The

Netherlands) within a scan range of 37-300 Da. The peaks corresponding to the ion fragments were grouped according to the retention time.

Data Analysis

Raw data resulting from the GC-MS analysis existing of fragment peaks defined by a specific combination of m/z ratio and retention time, were pre-processed before further analysis. Pre-processing included removing siloxane-related fragments which originate from the GC-MS (e.g column bleed). Subsequently, Wilcoxon rank tests were performed between all phenotype groups to identify significant fragments contributing to breath profile differences. The National Institute of Standards and Technology's mass spectral database (NIST) library matching was used to tentatively identify VOCs based on the significant fragments using Shimadzu GCMSsolution software. In case of uncertainty, the chromatogram peak was analysed using the Automated Mass Spectral Deconvolution and Identification System (http://chemdata.nist.gov/mass-spc/amdis). These steps were independently executed by two authors (IGS and IAS). Discrepancies were discussed between the authors IGS, IAS and PB. A fragment was removed if no consensus was reached, if its identified VOC was considered a contaminant (i.e., derivatives of substances of GC-MS analysis), if its mass exceeded >250 (not considered volatile), or if library matching revealed <80% similarity with the NIST library.

The remaining list of included fragments for each group comparison was used for the classification of breath profiles between the two groups. Supervised data analysis using sparse partial least squares discriminant analysis (sPLS-DA) was performed. sPLS-DA first reduces the dimensionality of the data, resulting in multiple components per measurement. Each component is a weighted combination of most discriminatory fragments. Components 1 and 2 were used to create scatter plots, and component 1 for receiver operating characteristic analyses to calculate the area under the curve (AUC) values.

To test whether other clinical characteristics than the described phenotypes can be associated with the obtained exhaled VOCs, unsupervised analyses were conducted through principal component analysis and cluster analysis on the full dataset and the dataset excluding sarcoidosis patients. More details can be found in **Supplementary Data C**.

All the statistical analyses were performed by using R statistical software (version 4.3.2, R Foundation for Statistical Computing, Vienna, Austria). A significance level of 0.05 was used to indicate statistical significance.

Results

66 patients were initially included (n=22 fibrotic, n=22 inflammatory, n=22 mixed; of whom half in Erasmus MC and Amsterdam UMC). After central review of chest CT scans, 13 patients were excluded and six were appointed to another group (**Figure 2**). Main reason for exclusion was the absence of ≥10% inflammation. The final cohort for analysis consisted of 53 patients, with 21 patients in the fibrotic, 18 in the inflammatory and 14 in the mixed phenotype group.

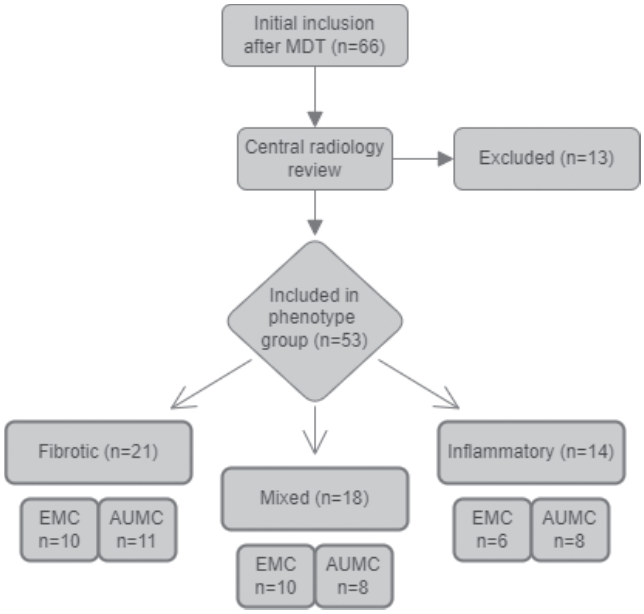


Figure 2: Flowchart for overview of patient inclusion and classification in phenotypes group.
AUMC = Amsterdam University Medical Centre; EMC = Erasmus University Medical Centre Rotterdam.

The total cohort (n=53) had a median age of 69.0 [61.0-76.0] years and 71.7% was male. Only a minority was currently smoking (n=2, 3.8%). All baseline characteristics per phenotype group can be found in **Table 1** and per including centre in **Supplementary Data A** (**Table S1**).

Table 1. Baseline characteristics.

	Fibrotic (n=21)	Inflammatory (n=14)	Mixed (n=18)
Including centre Erasmus MC	10 (47.6)	6 (42.9)	10 (55.6)
Female	4 (19.0)	6 (42.9)	5 (27.8)
Age, years	69.0 [65.0, 76.0]	59.0 [41.0, 68.8]	70.0 [59.5, 75.5]
BMI, kg/m2	26.4 [24.4, 29.3]	24.9 [23.0, 26.5]	27.1 [25.1, 30.6]
Smoking status			
Never	2 (9.5)	7 (50.0)	6 (33.3)
Former	19 (90.5)	5 (35.7)	12 (66.7)
Current	-	2 (14.3)	-
Pack years, years *	10.0 [3.0, 32.0]	6.0 [4.0, 12.0]	24.5 [18.8, 43.3]
Lung function			
FVC, % pred **	83.2 (19.3)	75.7 (16.6)	67.3 (12.1)
DLCOc, % pred **	51.7 (15.4)	61.9 (18.1)	42.1 (17.3)
Histopathology-supported diagnosis	-	4 (28.6)	2 (11.1)
ILD diagnosis			
IPF	21 (100.0)	-	-
Sarcoidosis	-	3 (21.4)	1 (5.6)
CTD-ILD	-	4 (28.6)	-
U-ILD	-	1 (7.1)	8 (44.4)
iNSIP	-	1 (7.1)	4 (22.2)
Other ^	-	5 (35.7)	5 (27.8)

*Values are displayed as the number (%), mean ± SD, or median [interquartile range]. *Former and current smokers only. **Missing data n=2. ^Other diagnoses include cryptogenic organising pneumonia, drug-induced ILD, eosinophilic pneumonia, chronic hypersensitivity pneumonitis, interstitial pneumonia with autoimmune features, and respiratory bronchiolitis-ILD. BMI = body mass index; CTD = connective tissue disease; DLCOc = diffusing capacity of the lungs for carbon monoxide corrected for haemoglobin level; MC = Medical Centre; FVC = forced vital capacity; ILD = interstitial lung disease; iNSIP = idiopathic non-specific interstitial pneumonia; IPF = idiopathic pulmonary fibrosis; U-ILD = unclassifiable ILD.*

Breath Profile Classification

Univariate analysis on the training set resulted in 22 significantly different fragments between the fibrotic and inflammatory phenotype. After identification of the fragments, removal of contaminant or unidentified VOCs resulted in 14 included fragments for sPLS-DA, which represented 13 unique VOCs. Similar analysis between fibrotic and mixed phenotypes resulted in 30 different fragments, of which 21 were included for sPLS-DA representing 14 unique VOCs. Analysis between inflammatory and mixed phenotypes resulted in 108 fragments, of which 25 were included for sPLS-DA that were all unique (**Additional file B**, **Table S2**).

Figure 3 shows the scatter plots resulting from sPLS-DA between the phenotype groups of the training dataset. Breath profile classification resulted in high mean AUC values (0.83-0.97) in all group comparisons in the training dataset. These values decreased (0.57-0.63) for all group comparisons after applying the trained sPLS-DA model to the validation dataset (**Table 2**). Repeating this analysis without sarcoidosis patients did not improve the validated model performance, as indicated in **Additional file C**.

No influence of including centre or phenotype on breath profiles was found, according to results of additional unsupervised analyses (i.e., PCA). Unsupervised clustering did neither lead to patient clusters with significantly different clinical features. Results are described in **Additional file D**.

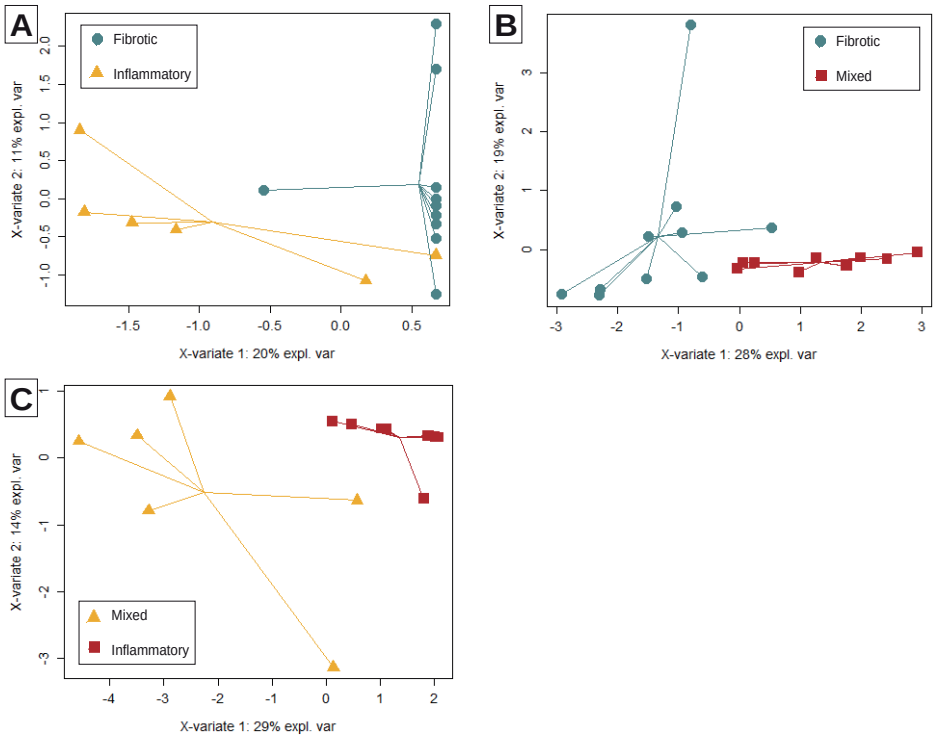


Figure 3. Breath profile comparison training cohort patients by phenotype.
Scatterplots of individual breath profiles classification of patients in the training dataset with a fibrotic and inflammatory (A), fibrotic and mixed (B), or inflammatory and mixed phenotype (C). Each dot in the scatterplot represents one patient. X-variates 1 and 2 are principal components 1 and 2 resulting from sparse partial least squares discriminant analysis.

Table 2. Results of the sPLS-DA training and validation datasets.

Phenotype group 1 (n=)	Phenotype group 2 (n=)	Fragments (n=)	AUC	95% CI	sPLS-DA component no.	Dataset
Fibrotic	**Inflammatory**					
- EMC (10)	- EMC (6)	14	0,83	0,52-1,00	1	Training
- AUMC (11)	*- AUMC (8)*		*0,57*	*0,29-0,84*	*1*	*Validation*
Fibrotic	**Mixed**					
- EMC (10)	- EMC (10)	21	0,97	0,90-1,00	1	Training
- AUMC (11)	*- AUMC (8)*		*0,58*	*0,30-0,86*	*1*	*Validation*
Inflammatory	**Mixed**					
- EMC (6)	- EMC (10)	42	0,95	0,85-1,00	1	Training
- AUMC (8)	*- AUMC (8)*		*0,63*	*0,32-0,93*	*1*	*Validation*

Results derived from 1 principal component resulting from sPLS-DA analyses between groups using the included fragments. AUC = area under the curve; AUMC = Amsterdam University Medical Centre; CI = confidence interval; EMC = Erasmus University Medical Centre Rotterdam; sPLS-DA = sparse partial least squares discriminant analysis..

Discussion

This study indicates that our applied GC-MS driven exhaled breath analysis towards differentiating patients with ILD and fibrotic or inflammatory abnormalities on chest CT scans reveals no reliable biomarker. Results from the training cohort demonstrated accurate discrimination between groups. However, the difficulty to validate these outcomes was evident, reflected by a substantial decline in AUC values in the validation cohorts.

In contrast to exhaled breath analysis using an eNose, GC-MS breath analysis allows the identification of individual VOCs that are contributing to group differences. Despite our efforts to pinpoint specific VOCs associated with fibrotic or inflammatory pathophysiologic processes in lung parenchyma, no clear and unique biomarker for either phenotype group emerged. Our findings in the fibrotic versus inflammatory group analysis parallel those of Plantier et al., who identified alkanes as significant contributors to differentiating IPF from CTD-ILD [7]. Straight-chain alkanes have been found in healthy humans [13] and might originate from oxidative stress [14]. However, identified individual VOCs show no similarity across ILD studies, complicating the discovery of underlying disease mechanisms [8, 6, 7]. This mechanism might explain the presence of alkanes in exhaled breath of patients with ILD, but it remains to be elucidated as to why some alkanes are up and others are downregulated.

Review of two other studies in ILD utilising GC-MS or similar breath analysis techniques show a lack of a validation cohort [6, 7]. An exception is the study of Fijten et al. where a separate training and validation analysis of a sarcoidosis cohort yielded an unexplained drop in model performance, comparable to our results [8]. In other fields of medicine, externally validated exhaled breath studies with positive results have been published. A recent comprehensive study in patients with breast cancer successfully validated a diagnostic model and reached a specificity of 87.7% in external validation cohorts [15]. Although a different analysis method was used (high-pressure photon ionization-time-of-flight mass spectrometry), this example shows that breath analysis results in specific medical conditions can be validated and might lead to new disease insights.

The presented robust multicentre study design with external validation is unique in the field of ILD breath analysis. This study illustrates that validation of results in external cohorts is essential for future studies. Without validation, this study would have presented the opposite wrong conclusion. External validation for studies involving an omics approach, as for instance in exhaled VOC analysis, is highly recommended by the ERS task force [16] and the Institute of Medicine committee [17]. Omics studies often have no pre-defined target and use therefore data-mining approaches, i.e. searching large amounts of data to discover hidden patterns or useful information. This increases the likelihood of coincidental findings that cannot be confirmed in different patient populations and settings [8]. Another strength is that patient inclusion was not only based on MDT-approved ILD diagnosis but that each individual's phenotype was confirmed by central radiologic review. We recommend all future exhaled breath studies to be multicentre with external validation cohorts following robust pre-defined standardised methodologies for patient selection, data collection, processing and analysis.

Several limitations of this study must be acknowledged. First of all, the contribution of external factors to performance discrepancies between the training and test cohorts could not be definitely excluded. Variability between including centres, such as other investigators performing breath collection, seasonal influences, and research environments may have played a role. Importantly, unsupervised analysis revealed no clear clusters based on including centres. Moreover, other potential external factors like inclusion and technical biases, are largely excluded by the stringent study design and standard operating procedures. Second, the heterogeneity of ILD and small sample size might limit the power to identify a biomarker in this study. While acknowledging these constraints, no VOCs were related to the fibrotic or inflammatory pathophysiologic pathway. This suggests that single VOCs representing fibrosis or inflammation on chest CT might not exist. Third, the phenotypes were defined based on chest CT scan characteristics which may not perfectly correlate with fibrotic and

inflammatory extent on a pathophysiological level. This highlights the need for better non-invasive biomarkers in ILD. In the current study, we optimized the homogeneity of included patient groups by central review of CT scans. Future follow-up studies should consider including disease behaviour like treatment response as criteria for defining phenotypes. Last, the identification of VOCs based on raw GC-MS data is complicated and might hamper the generalisability of findings. A single fragment resulting from breath analysis can fit the profile of multiple VOCs. We tried to overcome this and optimised the reliability of our findings by comparing different databases and identifying fragments by independent researchers for consensus. Furthermore, clarifying a VOCs clinical significance or biological source remains challenging. Some metabolic products can be related to the consumption of certain foods and therefore considered a confounding factor. For example, allyl methyl sulphide is related to garlic consumption. Nonetheless, this compound is also reported in studies to be disease-specific [18, 19]. Due to limited sample size, we were not able to test and correct for potential confounders. It is recommended for future studies with larger datasets to investigate confounding factors in more detail. This will increase understanding of specific compounds in terms of the potential influence on results and role in pathophysiologic pathways.

Conclusions

In conclusion, this multicentre GC-MS study with external validation in ILD underlines the importance of result validation when performing biomarker development studies. The current findings in the training cohort could not be validated, despite the elaborative procedures accompanying patient selection and GC-MS breath analysis. This suggests that our applied GC-MS exhaled breath analysis strategy does not distinguish fibrotic and inflammatory ILD phenotypes based on chest CT scan to guide treatment decisions in future clinical practice.

Declarations

The study was exempt from ethics and was conducted in accordance with the amended Declaration of Helsinki. All study participants provided informed consent prior to inclusion.

References

1. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med. 2013 Sep 15;188(6):733-48.
2. Walsh SLF, Wells AU, Desai SR, Poletti V, Piciucchi S, Dubini A, et al. Multicentre evaluation of multidisciplinary team meeting agreement on diagnosis in diffuse parenchymal lung disease: a case-cohort study. Lancet Respir Med. 2016 Jul;4(7):557-65.
3. Wijsenbeek M, Cottin V. Spectrum of Fibrotic Lung Diseases. N Engl J Med. 2020 Sep 3;383(10):958-68.
4. van der Sar IG, Wijsenbeek MS, Moor CC. Exhaled breath analysis in interstitial lung disease. Curr Opin Pulm Med. 2023 Sep 1;29(5):443-50.
5. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).
6. Yamada YI, Yamada G, Otsuka M, Nishikiori H, Ikeda K, Umeda Y, et al. Volatile Organic Compounds in Exhaled Breath of Idiopathic Pulmonary Fibrosis for Discrimination from Healthy Subjects. Lung. 2017 Apr;195(2):247-54.
7. Plantier L, Smolinska A, Fijten R, Flamant M, Dallinga J, Mercadier JJ, et al. The use of exhaled air analysis in discriminating interstitial lung diseases: a pilot study. Respir Res. 2022 Jan 20;23(1):12.
8. Fijten RRR, Smolinska A, Drent M, Dallinga JW, Mostard R, Pachen DM, et al. The necessity of external validation in exhaled breath research: a case study of sarcoidosis. J Breath Res. 2017 Nov 29;12(1):016004.
9. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2018 Sep 1;198(5):e44-e68.
10. Crouser ED, Maier LA, Wilson KC, Bonham CA, Morgenthau AS, Patterson KC, et al. Diagnosis and Detection of Sarcoidosis. An Official American Thoracic Society Clinical Practice Guideline. Am J Respir Crit Care Med. 2020 Apr 15;201(8):e26-e51.
11. Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1;205(9):e18-e47.
12. Nemec SF, Eisenberg RL, Bankier AA. Noninfectious inflammatory lung disease: imaging considerations and clues to differential diagnosis. AJR Am J Roentgenol. 2013 Aug;201(2):278-94.
13. de Lacy Costello B, Amann A, Al-Kateb H, Flynn C, Filipiak W, Khalid T, et al. A review of the volatiles from the healthy human body. J Breath Res. 2014 Mar;8(1):014001.
14. Burk RF, Ludden TM. Exhaled alkanes as indices of in vivo lipid peroxidation. Biochem Pharmacol. 1989 Apr 1;38(7):1029-32.
15. Liu J, Chen H, Li Y, Fang Y, Guo Y, Li S, et al. A novel non-invasive exhaled breath biopsy for the diagnosis and screening of breast cancer. J Hematol Oncol. 2023 Jun 16;16(1):63.

16. Horvath I, Barnes PJ, Loukides S, Sterk PJ, Hogman M, Olin AC, et al. A European Respiratory Society technical standard: exhaled biomarkers in lung disease. Eur Respir J. 2017 Apr;49(4).

17. McShane LM, Cavenagh MM, Lively TG, Eberhard DA, Bigbee WL, Williams PM, et al. Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration. BMC Med. 2013 Oct 17;11:220.

18. Siriwong N, Sukaram T, Tansawat R, Apiparakoon T, Tiyarattanachai T, Marukatat S, et al. Exhaled volatile organic compounds for cholangiocarcinoma diagnosis. Liver Research. 2022 2022/09/01/;6(3):191-97.

19. Sukaram T, Tansawat R, Apiparakoon T, Tiyarattanachai T, Marukatat S, Rerknimitr R, et al. Exhaled volatile organic compounds for diagnosis of hepatocellular carcinoma. Sci Rep. 2022 Mar 29;12(1):5326.

ADDITIONAL FILE A

Baseline table divided by including centre

Table S1. Baseline characteristics per centre.

	EMC (n=26)	AUMC (n=27)	*P-value*
Female	9 (27.3)	10 (30.3)	1.00
Age, years	6 (23.1)	9 (33.3)	0.60
BMI, kg/m2	26.8 [24.9, 30.3]	25.2 [23.6, 28.6]	0.15
Smoking status			0.26
Never	9 (34.6)	6 (22.2)	
Former	17 (65.4)	19 (70.4)	
Current	-	2 (7.4)	
Pack years, years	10.0 [4.0, 25.0]	15.0 [10.0, 34.0]	0.51
Lung function			
FVC, % pred **	77.9 (19.3)	73.7 (15.8)	0.40
DLCOc, % pred **	52.7 (19.7)	49.0 (16.4)	0.47
Pathology-proven diagnosis	2 (7.7)	4 (14.8)	0.70
ILD diagnosis			0.11
IPF	10 (38.5)	11 (40.7)	
Sarcoidosis	2 (7.7)	2 (7.4)	
CTD-ILD	-	4 (14.8)	
U-ILD	3 (11.5)	6 (22.2)	
iNSIP	4 (15.4)	1 (3.7)	
Other ^	7 (26.9)	3 (11.1)	

*Values are displayed as the number (%), mean ± SD, or median [interquartile range]. *Former and current smokers only. **Missing data n=2. ^Other diagnoses include cryptogenic organising pneumonia, drug-induced ILD, eosinophilic pneumonia, chronic hypersensitivity pneumonitis, interstitial pneumonia with autoimmune features, and respiratory bronchiolitis-ILD. AUMC = Amsterdam University Medical Centre; BMI = body mass index; CTD = connective tissue disease; DLCOc = diffusing capacity of the lungs for carbon monoxide corrected for haemoglobin level; EMC = Erasmus Medical Centre; FVC = forced vital capacity; ILD = interstitial lung disease; iNSIP = idiopathic non-specific interstitial pneumonia; IPF = idiopathic pulmonary fibrosis; U-ILD = unclassifiable ILD.*

ADDITIONAL FILE B

Identified compounds

Table S2. Comparison of the identified included VOCs per group.

Fibrotic vs. inflammatory phenotype		Fibrotic vs. mixed phenotype		Inflammatory vs. mixed phenotype	
VOC name	**CAS registry no**	**VOC name**	**CAS registry no**	**VOC name**	**CAS registry no**
1,4-Cyclohexadiene, 1-Methyl-4-(1-methylethyl)	99-85-4	1,3,7-Octatriene, 3,7-dimethyl-	502-99-8	1,3,6-Octatriene, 3,7-dimethyl-, (Z)-	3338-55-4
1-Hexanol	111-27-3	1,4-Cyclohexadiene, 1-Methyl-4-(1-methylethyl)	99-85-4	1,3,7-Octatriene, 3,7-dimethyl-	502-99-8
1-Pentanol, 3-methyl-	589-35-5	1-Dodecene	112-41-4	1-Decanol, 2,2-dimethyl-	2370-15-2
2(5H)-Furanone, 5,5-dimethyl-	20019-64-1	1-Iodo-2-methylundecane	73105-67-6	1-Hexanol	111-27-3
2-Hexanol, acetate	5953-49-1	3-Acetoxydodecane	60826-26-8	1-Hexanol, 2-ethyl-	104-76-7
2-Pentanone	107-87-9	Butane	106-97-8	1-Pentanol, 3-methyl-	589-35-5
Butane, 2,3-dimethyl-	79-29-8	Cyclopropyl carbinol	2516-33-8	2(5H)-Furanone, 5,5-dimethyl-	20019-64-1
Decane	124-18-5	Decane, 3,7-dimethyl-	17312-54-8	2-Pentanone	107-87-9
Decane, 5-methyl-	13151-35-4	Dodecane	112-40-3	2-Pentanone, 4-hydroxy-	4161-60-8
Dodecane	112-40-3	Isosorbide	652-67-5	3,4-Hexanediol, 2,5-dimethyl-	22607-11-0
Furan, tetrahydro-	109-99-9	Nonadecane	629-92-5	3-Pentanol, 2,4-dimethyl-	600-36-2
Methyl 2,2-dimethyl-3-hydroxypropionate	14002-80-3	Octadecane, 1-(ethenyloxy)-	930-02-9	Bicyclo[3.1.1] heptane, 6,6-dimethyl-2-methylene-, (1S)-	18172-67-3
Octadecane, 1-(ethenyloxy)-	930-02-9	Octane	111-65-9	Butane, 2,3-dimethyl-	79-29-8
		Pentanal	110-62-3	Cyclopropyl carbinol	2516-33-8
				Decane, 3,7-dimethyl-	17312-54-8
				Dodecane	112-40-3

Fibrotic vs. inflammatory phenotype		Fibrotic vs. mixed phenotype		Inflammatory vs. mixed phenotype	
VOC name	**CAS registry no**	**VOC name**	**CAS registry no**	**VOC name**	**CAS registry no**
				Hexadecane	544-76-3
				Hexane, 1-(ethenyloxy)-	5363-64-4
				Limonene	138-86-3
				Methyl 2,2-dimethyl-3-hydroxypropionate	14002-80-3
				Octane, 2,7-dimethyl-	1072-16-8
				Sulfide, allyl methyl	10152-76-8
				Tetradecane	629-59-4
				Tridecane	629-50-5
				Undecane	1120-21-4

CAS = Chemical Abstracts Service; VOC = volatile organic compound.

ADDITIONAL FILE C

Results of the sPLS-DA without sarcoidosis

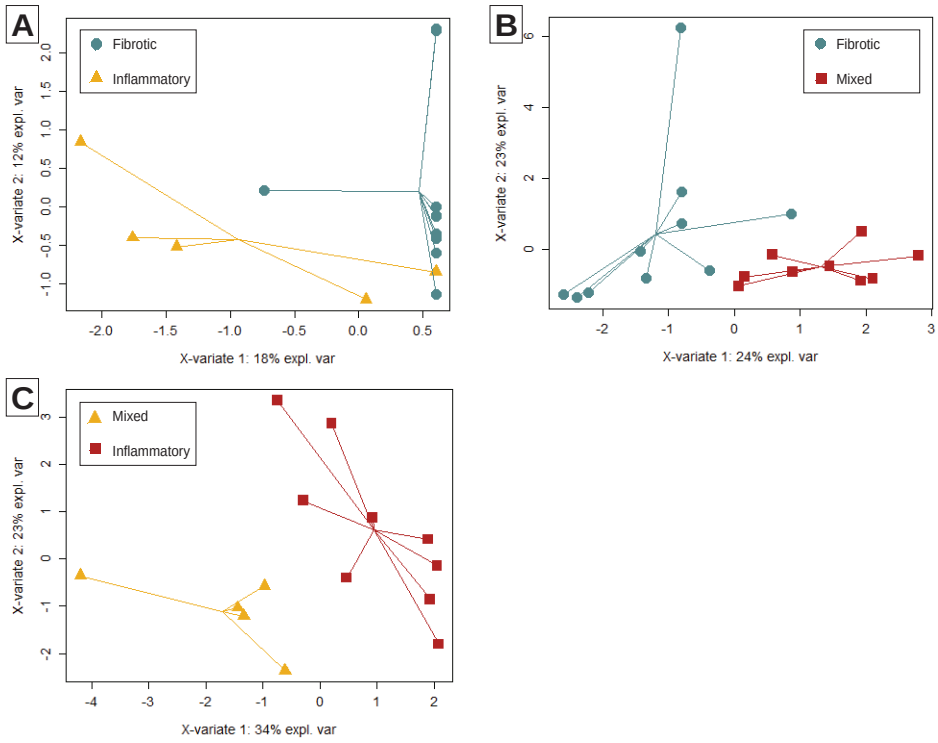


Figure S1: Breath profile comparison training cohort patients by phenotype excluding patients with sarcoidosis.

Scatterplots of individual breath profiles classification of patients in the training dataset with an fibrotic and inflammatory (A), fibrotic and mixed (B), or inflammatory and mixed phenotype (C). Patients with sarcoidosis (n=4) were excluded. Each dot in the scatterplot represents one patient. X-variate 1 and 2 are principal component 1 and 2 resulting from sparse partial least squares discriminant analysis.

Table S3. Results of the sPLS-DA training and validation datasets.

Phenotype group 1 (n=)	Phenotype group 2 (n=)	Fragments (n=)	AUC	95% CI	sPLS-DA component no	Dataset
Fibrotic	**Inflammatory**					
- EMC (10)	- EMC (5)	14	0,87	0,65-1,00	1	Training
- AUMC (11)	*- AUMC (6)*		*0,62*	*0,34-0,90*	*1*	*Validation*
Fibrotic	**Mixed**					
- EMC (10)	- EMC (9)	17	0,97	0,89-1,00	1	Training
- AUMC (11)	*- AUMC (8)*		*0,60*	*0,32-0,88*	*1*	*Validation*
Inflammatory	**Mixed**					
- EMC (5)	- EMC (9)	33	0,98	0,92-1,00	1	Training
- AUMC (6)	*- AUMC (8)*		*0,52*	*0,19-0,85*	*1*	*Validation*

Results derived from 1 principal component resulting from sPLS-DA analyses between groups using the included fragments. AUC = area under the curve; AUMC = Amsterdam University Medical Centre; CI = confidence interval; EMC = Erasmus University Medical Centre Rotterdam; sPLS-DA = sparse partial least squares discriminant analysis.

ADDITIONAL FILE D

Unsupervised analyses

Methods

After breath data pre-processing (see Methods section in main text for more details), we first removed fragments with selected retention times. Retention times were defined based on fragments of which the identified VOC was considered a contaminant (i.e., containing siloxanes) or had a mass exceeding >250 (not volatile at room temperature). The remaining dataset was used for unsupervised data analysis, including a principal component (PCA) and cluster analysis.

PCA was performed with the first two principal components (PCs) using MixOmics R package (version 6.26.0). PC 1 and 2 were used to create scatter plots of the breath data. Subsequently, each breath profile was marked according to the patient's phenotype group or including centre (EMC or AUMC).

For cluster analysis, partitioning around medoids consensus cluster analysis using Euclidean distance was applied to the breath data (R Package ConsensusClusterPlus; version 1.66.0) [1]. The optimal number of stable clusters (range k = 2-10) was determined based on combining visual inspection of consensus matrices and cumulative distribution function (CDF) curves. Comparison of clinical parameters between clusters was performed using one-way ANOVA, Kruskal-Wallis and chi-squared tests.

Results

After removal of selected fragments, the GC-MS dataset of all patient measurements (n=53) consisted of 1529 fragments. The scatterplot (**Figure S2-A**) shows the individual breath profile distribution based on PCA analysis. Patients are indicated per phenotype (B) and including centre (C), without a clear visual separation of the groups. Consensus clustering analysis of the breath data revealed two clusters being most stable, see **Figure S3** for consensus matrices and CDF curves. Clinical characteristics of the clustered patients are displayed in **Table S4**. No significant differences were found between the groups.

Conclusion

Unsupervised breath data analysis did not reveal distinct segregation of breath profiles among the defined phenotype groups. supporting results from the supervised analyses. Patients neither clustered per including centre. suggesting that the outcomes of supervised analyses are not influenced by institution-specific factors. Moreover, the lack of significant differences observed across clusters. does not suggest the presence of unrecognised or novel ILD phenotypes based on the provided clinical characteristics.

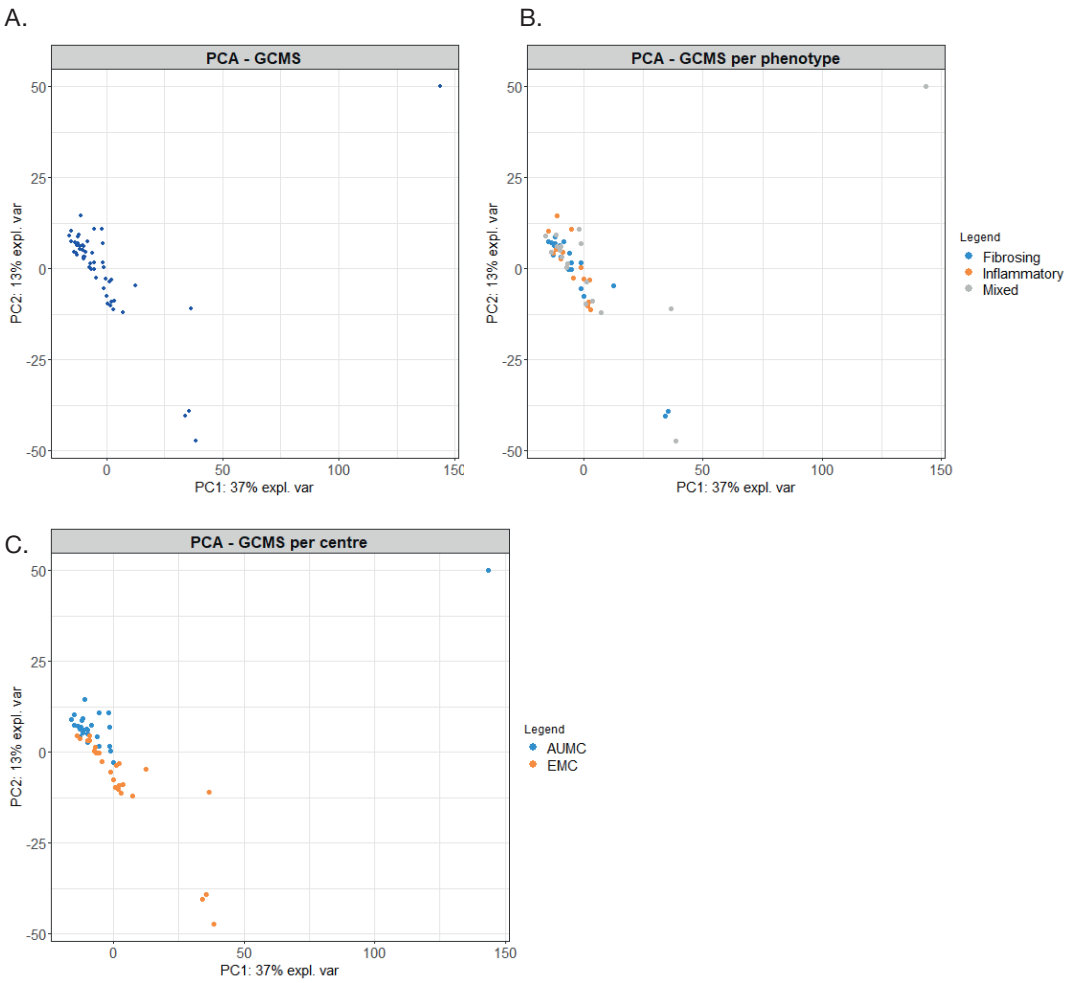


Figure S2: PCA breath profile distribution.
Scatterplots of individual breath profiles of all included patients based on unsupervised analysis (A). Classified by phenotype (B) and including centre (C). Each dot in the scatterplot represents one patient. PC 1 and 2 are the first two principal components resulting from principal component analysis. AUMC = Amsterdam University Medical Centre; EMC = Erasmus Medical Centre; PC = principal component.

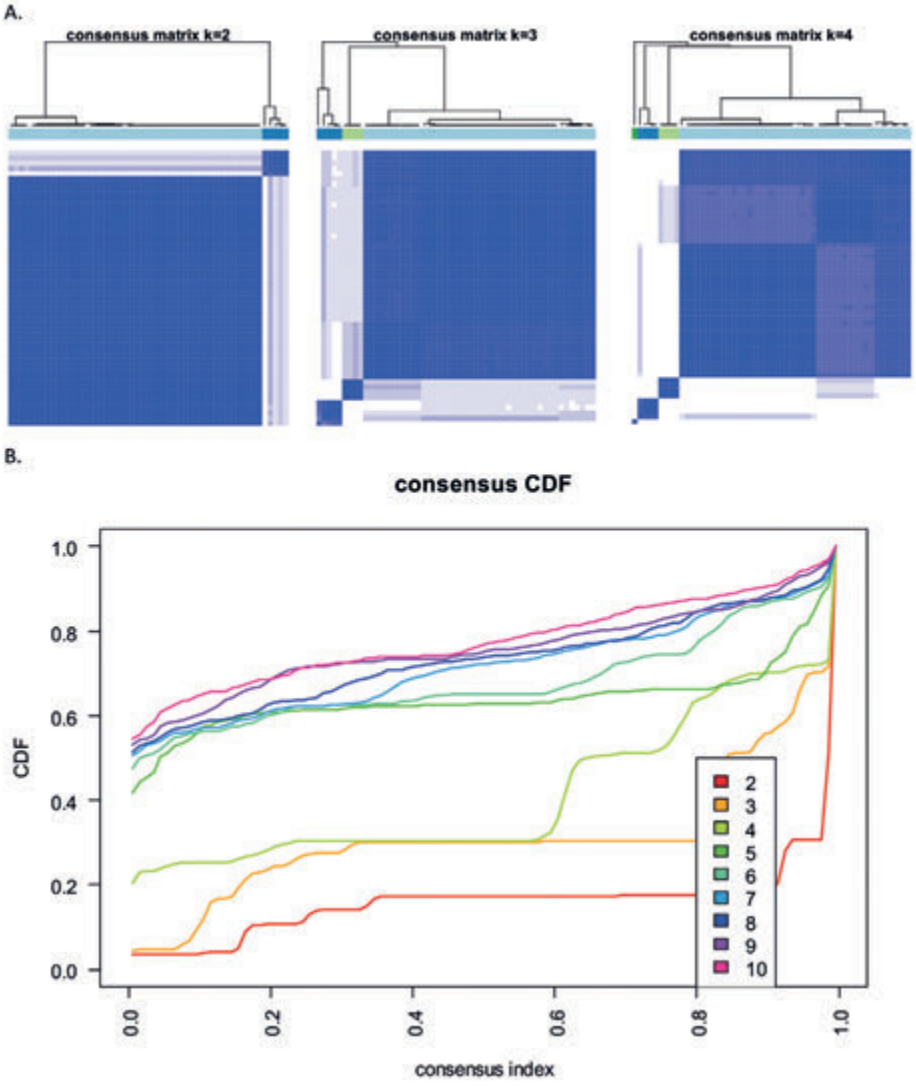


Figure S3: Cluster stability.
(A) Heat map of consensus matrix when k = 2. 3 or 4. (B) CDF curve when k = 2-10. CDF = cumulative distribution function.

Table S4: Clinical characteristics of patient in k=2 clusters.

	Cluster 1 (n=45)	Cluster 2 (n=5)	*p-value*
Including centre EMC	24 (50.0)	2 (40.0)	1.00
Female	13 (27.1)	2 (40.0)	0.93
Age, years	67.3 (12.6)	52.8 (14.7)	0.02
Smoking history			0.13
Never	14 (29.2)	1 (20.0)	
Former	32 (68.8)	3 (60.0)	
Current	1 (2.1)	1 (20.0)	
Phenotype			0.06
Fibrotic	21 (43.8)	-	
Inflammatory	13 (27.1)	1 (20.0)	
Mixed	14 (29.2)	4 (80.0)	
FVC, %pred *	76.1 (18.1)	72.6 (12.2)	0.68
DLCOc, %pred *	50.7 (18.6)	53.5 (9.7)	0.77
ILD diagnosis			0.49
IPF	21 (43.8)	-	
Sarcoidosis	3 (6.2)	1 (20.0)	
CTD-ILD	4 (8.3)	-	
U-ILD	8 (16.7)	1 (20.0)	
iNSIP	3 (6.2)	2 (40.0)	
Other	9 (18.8)	1 (20.0)	

*Values are displayed as the number (%) or mean ± SD. *Missing data n=2. ^Other diagnoses include cryptogenic organising pneumonia. Drug-induced ILD. eosinophilic pneumonia. chronic hypersensitivity pneumonitis. interstitial pneumonia with autoimmune features. and respiratory bronchiolitis-ILD. CTD = connective tissue disease; DLCOc = diffusing capacity of the lungs for carbon monoxide corrected for haemoglobin level; EMC = Erasmus medical centre; FVC = forced vital capacity; ILD = interstitial lung disease; iNSIP = idiopathic non-specific interstitial pneumonia; IPF = idiopathic pulmonary fibrosis; U-ILD = unclassifiable ILD.*

References

1. Monti S, Tamayo P, Mesirov J, Golub T. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. Machine Learning. 2003 2003/07/01;52(1):91-118.

PART V

DISCUSSION

GENERAL DISCUSSION

The main aim of this thesis was to investigate the value of exhaled breath analysis using electronic nose (eNose) technology for detecting interstitial lung disease (ILD), with a focus on the potential to improve the diagnostic process and screen high-risk patients. To answer this aim, we conducted several studies with various cohorts of patients with ILD and controls (**Chapter 4, 6, 9, 10**), and within an ILD cohort (**Chapter 5**). Moreover, we investigated how to use eNose data for developing an optimal diagnostic model (**Chapter 7**). Lastly, the potential of gas chromatography-mass spectrometry (GC-MS) breath analysis for ILD was explored (**Chapter 11**).

Below I discuss the outcomes and implications of this thesis. The first part focusses on the latest results and challenges of exhaled breath analysis in ILD. The second part includes an outlook on eNose technology research and implementation in daily ILD care. To conclude, perspectives on the incorporation of artificial intelligence (AI)-based medical testing are presented.

Exhaled breath analysis in ILD

eNose for diagnosis

The need for a new diagnostic test in ILD is high as the diagnostic process is elaborative and often requires invasive procedures resulting in notable diagnostic delay (**Chapter 3**). Breath analysis using eNose technology has many benefits compared to existing medical tests. The results can be derived non-invasive, quick, relatively cheap, without complication risks, and depending on the device, in real-time.

In several chapters of this thesis we analyzed data from various patient cohorts to demonstrate the performance of eNose in diagnosing ILD. These results confirm the potential of eNose analysis for detection of ILD that was raised by previous single-center studies that primarily compared breath of ILD and healthy controls [1-4]. The results in **Chapter 4** add to this that eNose cannot only be used to differentiate ILD from healthy controls, but also from other respiratory diseases (asthma, chronic obstructive pulmonary disease and lung cancer) with a reported area under the curve (AUC) of 0.99 (95% confidence interval (CI) 0.97–1.00) in the test set. This observation suggests that patients with ILD exhale a distinct volatile organic compound (VOC) pattern, likely indicative of a disease-specific underlying pathobiology. Previously published results from analyses with different eNose devices support this theory, reporting an AUC of 0.85 when comparing idiopathic pulmonary fibrosis (IPF) or connective tissue disease related ILD with chronic obstructive pulmonary disease [2, 3].

Furthermore, patients with various fibrotic ILDs have different breath profiles, as shown in the international prospective study (**Chapter 5**). AUC for differentiating IPF from other fibrotic ILDs was 0.92 (95% CI 0.87-0.98) in the external validation set. Subgroup analyses of earlier studies already suggested unique profiles for individual ILDs; however, they did not validate their results externally [2, 4]. The outcomes from our unique multicenter cohort from different geographical locations and research settings, confirm these previous findings and support algorithm design and validation to enable use and approval as a medical diagnostic test.

The presented differences in breath profiles between various fibrotic ILDs are interesting in the light of the current perception of pathobiology in pulmonary fibrosis. The current conceptual thinking is that initially different inflammatory and fibrotic mechanisms may play a role in the development of an ILD. Though, once fibrosis is progressive, similar mechanisms on a tissue level result in similar disease behavior regardless of diagnosis [5, 6]. Recent revised clinical guidelines therefore recommend prescribing antifibrotic treatment for all patients with a progressive pulmonary fibrosis behavior [7]. In the cohort described in **Chapter 5**, patients were evaluated based on diagnostic labels irrespective of disease stage or behavior. The expected follow-up data from this study which includes multiple eNose measurements and information on disease progression, will be the first in revealing whether disease course or phenotype drives VOC production. These insights will elucidate if eNose testing has not only value for diagnosis, but also for guiding treatment decisions and monitoring disease course in the future.

Our results on sarcoidosis in **Chapter 6** indicate that eNose breath analysis might also serve as a future diagnostic test for this heterogeneous condition. An AUC value of 1.00 was found when comparing breath profiles of patients with healthy controls, regardless of treatment and pulmonary involvement, and AUCs ranging 0.87-0.95 when compared with other ILDs. So far, one other study compared patients with sarcoidosis and healthy controls using a different eNose device. They reported a high discriminative ability for untreated patients versus controls (AUC of 0.83-0.85) [1]. Breath profiles of patients with and without treatment showed a large overlap in the cohort of Dragonieri, similar to the findings in our cohort. Interestingly, they reported accuracies for comparing treated patients versus healthy controls of only 64.4-66.7% in contrast to 100% in our cohort. This shows how exhaled breath study results can differ in similar populations, hypothetically depending on the use of different eNose devices with other sensors and breath collection methods (including breathing technique and materials). However, no studies exist that compare results of multiple devices within one cohort or between similar cohorts. This is complicated by the inability to compare or merge eNose data collected with different devices due to incompatible output data formats.

eNose for screening

Screening for ILD is relevant in selected patient groups that have an increased risk for developing ILD. High-risk groups include, but are not limited to, those with connective tissue disease, high-risk medication or occupational exposure as well as genetic predisposition. Early ILD detection is currently possible using a chest CT scan, but CT scans come with higher costs, limited availability in some countries and increased radiation exposure compared to a point-of-care test like eNose. Alternatively, using pulmonary function testing (PFT) often leads to normal results in early disease, and PFT abnormalities are usually not specific for ILD. To date, two studies in patients at risk for and diagnosed with ILD (i.e. pneumoconiosis) have been published evaluating the eNose screening potential [8, 9]. They report AUC values ranging from 0.77 to 0.94 in groups with various types of pneumoconiosis compared to those without ILD.

Studies described in **Chapter 9 and 10** of this thesis investigated other populations, and showed great potential of eNose for future screening. First of all, we studied a systemic sclerosis (SSc) cohort in **Chapter 9** as many patients with SSc develop ILD during the disease course, indicated by an overall SSc-ILD prevalence of 47% [10]. We showed a high differentiating ability for patients with and without ILD. Second, we aimed to investigate screening for drug induced ILD (DIILD; **Chapter 10**). Many types of drugs contain the risk of causing DIILD but diagnosing DIILD is complicated. Clinical and chest CT scan features may mimic other diagnoses such a infections, cardiac failure or malignancy, and risk factors are largely unknown. eNose showed accurate differentiation of patients with cancer treatment who developed DIILD and those without DIILD. These novel findings open perspectives to new ways of screening and warrant further research, as discussed in the “Future of eNose in ILD” section.

Gas chromatography-mass spectrometry

Other than eNose technology, GC-MS can be used to elucidate the composition of breath profiles on a molecular level. We evaluated the use of this analysis technique to gain more insights in the underlying biological mechanisms of fibrotic and inflammatory processes in the lungs (**Chapter 11**). Understanding these mechanisms would ideally lead to identification of a specific biomarker to guide treatment decisions throughout the disease course. However, GC-MS analysis did not reveal a specific single VOC or group of VOCs that could be related to the pathophysiologic process underlying fibrotic or inflammatory abnormalities on chest computed tomography (CT) scan. We could separate groups with different chest CT scan abnormalities within one center but could not validate these findings in an external cohort, similar to previously published

exhaled breath GC-MS research in sarcoidosis [11]. Additionally, we did not identify the reported VOCs of a GC-MS study in similar patient cohorts of a preceding study [12].

Our study highlights the necessity for external validation of exhaled breath findings. **Figure 1** displays the general hierarchy in study designs with externally validated studies providing the highest quality of evidence. Externally validated positive results are not available for GC-MS studies in ILD to date. In light of our findings, the conclusion drawn by Plantier et al. that patients with IPF and connective tissue disease related ILD can be differentiated by specific identified VOCs should therefore be reconsidered [12].

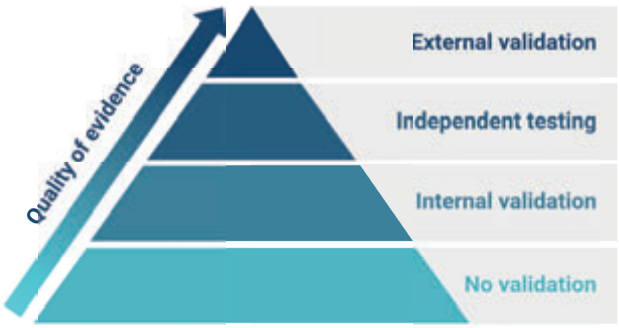


Figure 1: General hierarchy of study design according to acquired scientific evidence.
High quality of evidence obtained from a robust study design is essential to establish validity of omics study findings, e.g. electronic nose studies. Non-validated studies result in the weakest evidence and are most frequently conducted, whereas externally validated studies result in the strongest evidence and are less frequently conducted.

Challenges

In context of the results of the current thesis and previous studies, some challenges of exhaled breath research in ILD must be addressed.

One of the issues in eNose research is the unknown origin and type of VOCs constructing breath profiles. Various GC-MS studies have tried to unravel the ILD VOC profile, but no disease-specific VOCs have been externally validated (reviewed in **Chapter 2**). The lack of successful validation of single VOCs might not be surprising, as the pathophysiology of ILDs is complex, heterogeneous and not fully understood. The opposite accounts for generally approved breath biomarkers like the carbon urea breath test for Helicobacter Pylori infections that do have a clearly understood underlying biological mechanism [13]. One of the potential underlying mechanisms of

disease-related VOC production in patients with ILD is oxidative stress. This damages or alters human cells and is suggested to contribute to the origin of many diseases. Products of oxidative stress-induced peroxidation of membrane lipids can used as (breath) biomarkers. However, oxidative stress is not a disease-specific process nor clarifies the found differences between individual ILDs. Besides, oxidative stress is part of the normal process of human aging [14]. So it is likely that multiple biological mechanisms involved in ILDs are reflected by the breath profile.

Although the exact composition of ILD breath profiles is not unraveled yet, this does not hamper the development of an useful test. Pattern recognition in large eNose datasets by algorithms can be compared to the gut feeling of medical professionals. Professionals need sufficient experience to develop a reliable intuition to recognize repeating patterns in daily medical practice. For algorithms, most important is extensive validation of results if the exact origin is unknown. Therefore, our results from the externally validated cohort confirming that eNose can identify ILDs highly accurate is more important than understanding the exact composition of breath profiles (**Chapter 5**). The high volume 'breathomics' data captured by eNose sensors seem to better represent the complex disease mechanism of ILD than a single biomarker, since many attempts to discover a biomarker for ILD have not led to validated results to date.

Another important topic is whether VOC profiles change throughout the disease course. A hypothesis often implicitly adopted from cross-sectional studies and not questioned, is that a breath profile represents the disease regardless of severity, stage or activity. However, to date is unknown if breath composition is similar for early and late ILD, or if it changes throughout the disease course. This question is currently being investigated in the longitudinal international ILDnose study. Understanding effects of time and disease behavior on breath profiles will provide grounds for the potential of screening with an eNose but also to monitor or predict treatment response or disease course. Furthermore, knowledge of individual breath profile changes will improve rationale for patient selection in breath research.

A third point of debate in breath research, is whether and to what extent external factors influence the breath profile, and whether or when to correct for any of those factors. Several eNose studies tried to answer these questions for specific potential influencing factors.

To date, most evidence is collected on the influence of smoking, but results are non-validated nor conclusive. One study compared eNose breath profiles of healthy controls that never or ever smoked. They did not find significant differences [15].

Another study assessed the influence of smoking within-patient, by performing an eNose measurements before and after smoking. This resulted in significant changes in breath profiles with a larger effect at 30 and 60 minutes after smoking, compared to 5 minutes [16]. Authors suggest larger effects of post-cigarette inflammation than of tobacco odorants, but did not compare to non-smoking persons.

The effect of recent food intake or diet is largely unknown. One study conducted repeated eNose measurements, at baseline and 10 minutes, 1 hour, and 2 hours after food intake. Breath profiles differed at 10 minutes and 1 hour after intake compared to baseline with a discriminative accuracy of 65%, but results have not been validated [17].

For other factors like age and gender no contribution to eNose breath profile composition has been found [18]. Studies with successive eNose measurements were performed in order to examine the role of the circadian rhythm or menstrual cycle but were not conclusive [19, 20]. No associations between breath change over time and biological variations were found. Similarly, no associations were found in a study examining the link between physical activity and acetone levels in breath [21].

Based on the cited studies, all conducted in healthy control cohorts without external validation, it is yet unclear whether external factors significantly influence VOC results and disease detection, and if, when and how correction for factors is needed. An additional challenge arises from the subjective interpretation by investigators, given the lack of consensus on cut-off values for significant influence. Furthermore, it is crucial to acknowledge that minor alterations in VOC composition resulting from external factors level out disease-related VOCs when a sufficient number of patients are included in the analysis. Hence, eNose studies should not focus on correcting individual breath profiles. Collecting data of possible external influencing factors to conduct post-hoc analysis on population basis provides sufficient evidence applicable to the investigated cohort. This approach also enables the use of eNose tests in real-world clinical setting, generating reliable results regardless of patients' lifestyle.

The studies presented in this thesis followed the suggested post-hoc approach for several parameters. In the cohort of **Chapter 4** with patients with various lung diseases, smoking status potentially influenced breath profiles (current versus former smoking AUC 0.80), but accuracy for differentiating the diseases did not drop in subgroups with similar smoking history. In other studies, no notable impact of the analyzed external factors was noted. Importantly, results from our international cohorts that include patients with a varying dietary habits and environmental conditions, suggest that external factors do not affect disease detection by eNose (**Chapter 5**).

A last interesting point in light of the discovery of new ILD biomarkers, is that multidisciplinary team (MDT) discussions are the gold standard for diagnosis. The gold standard typically serves as the benchmark for evaluating performance of a new diagnostic test, but no sensitivity and specificity of MDT discussions is known. Besides, in training eNose models the diagnostic label serves as input for classifying patients. However, quality of MDT discussion is dependent on the local expertise, and an MDT cannot always establish a confident diagnosis. Different MDTs also reach a different conclusion on the most likely ILD diagnosis in over 40% of cases [22]. Another complicating factor can be the development of new or progressive symptoms during a patient's follow-up after an initial diagnosis, that may increase the likelihood of alternative ILD diagnoses. The varying quality and insecurities inherent to MDT conclusions highlight the importance of training eNose algorithms with data from multiple high-quality MDTs. Therefore, we selected acknowledged ILD expert centers only as including sites for the ILDnose study to guarantee highly reliable data.

To conclude, although challenges are inherent to eNose research and test development in ILD, most of the addressed challenges can be conquered. In the following paragraph 'Future of eNose in ILD' solutions and suggestions to work towards clinical application are suggested.

Future of eNose in ILD

Considering the positive results presented in this thesis and the non-invasive quick nature of the eNose test, it seems likely that the future diagnostic trajectory of patients with ILD will include eNose assessment. Until then, several important developmental steps in novel research projects need to be taken. These research steps are discussed below for different test aims (e.g. diagnosing, screening, monitoring, phenotyping). **Figure 2** provides an overview of the collected and warranted evidence for these aims in ILD. The position and benefits of eNose technology in future clinical practice are discussed afterwards.

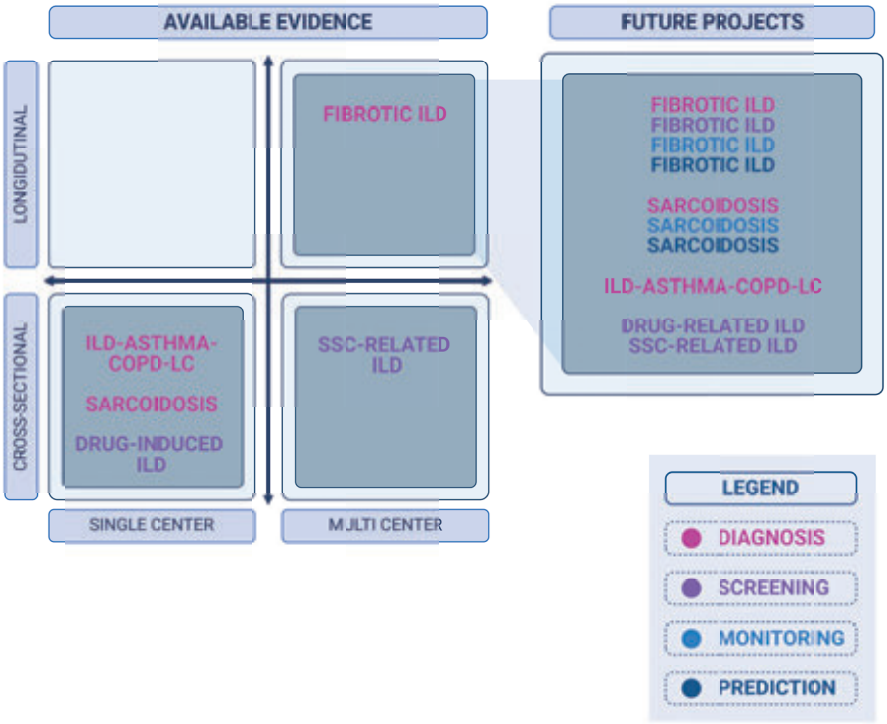


Figure 2: Available evidence and future projects regarding different applications of eNose in ILD.
A. Available evidence: representing studies published in this thesis; B. Future projects: needed and/or anticipated studies following conducted studies to validate the available evidence to work towards clinical application. COPD = chronic obstructive pulmonary disease; eNose = electronic nose; ILD = interstitial lung disease; LC = lung cancer.

Research

Most robust evidence so far is collected in specialized hospitals for the aim of diagnosing individual ILDs (**Figure 2**). Once all data of the international validation study are collected (**Chapter 5**) and validated accuracies remain high, the data can be used to design a diagnostic model for individual fibrotic ILDs. This requires collaboration with engineers to design and compare the performance of various algorithms, like the example in **Chapter 7** with sarcoidosis data. Subsequently, the model has to be evaluated in a separate real-world patient cohort. That study should recruit undiagnosed patients suspected of ILD to compare their MDT diagnosis with the most likely diagnosis resulting from the eNose test. After revealing the eNose result to the MDT, the team is requested to indicate whether the diagnosis or diagnostic likelihood changes. Additionally, in cases where a bronchoscopy and/or transbronchial biopsy was recommended, the MDT should state if this procedure would still be advised

when including the eNose test outcome. The number of potentially prevented invasive procedures quantitatively indicates the reduction of costs and complication risks.

For investigating the broader aim of diagnosing multiple pulmonary diagnoses using one eNose model in first or second line health care, new studies should be designed that recruit any patient with respiratory symptoms and a suspected pulmonary condition. This approach can possibly confirm the results of **Chapter 4** showing that breath profiles of patients with ILD were different from those with other respiratory illnesses recruited in expertise hospitals. Subsequently, this study will enable the design of diagnostic models. Importantly, pre-test probabilities of various diagnoses in the different health care settings have to be taken into account to ensure reliable model outcomes. Subsequently, testing these models in real-world cohorts will establish the added value of eNose testing.

Presented data for exploring ILD screening in specific subgroups (**Chapter 9 and 10**) were not internationally nor longitudinally collected. Results of these proof-of-concept studies showed high accuracy of the eNose in discriminating the groups with and without ILD. Further research should confirm the performance of eNose as a screening tool by prospective inclusion and follow-up of populations at risk from start of diagnosis (e.g. SSc) or exposure (e.g. high-risk drug).

Furthermore, apart from diagnostic and screening applications, using eNose for prediction or monitoring disease progression and treatment response would be of great value for supporting clinical decisions and personalized medicine. Currently, no generally accepted and implemented single biological test or clinical parameter are available to predict treatment response or disease progression. Prospective longitudinal databases like those generated in the ILDnose study, are awaited to assess whether eNose correlates with these parameters. If results will show a correlation, the aim of using eNose for monitoring and prediction can be further investigated.

Lastly, the study in **Chapter 11** could not relate GC-MS-based VOC profiles with clinical phenotypes based on chest CT scan features (i.e. fibrotic or inflammatory). Generically, a role for GC-MS breath analysis in clinical daily practice is not expected considering difficulties in validating results and the time-consuming nature of analyses. Nevertheless, GC-MS studies can theoretically facilitate optimization and specification of eNose sensors for ILD by improving the understanding of disease-specific VOC production.

Daily clinical care

In future clinical practice, once the performance of data models is validated, eNose can seize several unmet needs for ILD patient care due to multiple benefits compared to other available tests. Available tests, like PFT and chest CT require more advanced trained personnel, more time, and more expensive materials compared to an eNose. Whereas shortage of health care staff, money, time and resources threatens the availability and quality of care across the world, eNose fits the desired sustainable health care system of the future. Furthermore, eNose technology has benefits for patients comfort compared to currently used medical tests. CT scans can be burdensome for dyspneic patients and require radiation exposure. Pulmonary function tests are exhausting and it can be impossible to achieve reliable results in patients with advanced pulmonary diseases.

Though, before eNose systems can get implemented in practice, it is important to consider how to present the eNose results and at what time point of the patient journey the test should be introduced. The best and most useful manner to present outcomes of a diagnostic eNose test for patients suspected of ILD, is as likelihoods for various differential diagnoses to enable incorporation in the current MDT discussions. eNose likelihoods for several relevant diagnoses can be calculated based on the percentage of similarity between the breath profile of a patient and the average profile for those diagnoses, like we presented in **Chapter 4 and 7**. Binary results will probably not be of added value in clinical practice, seeing the complexity and overlap of individual ILDs, limitations of the current gold standard, and trust of physicians. Depending on the clinical setting, physicians should be able to request likelihoods of several selected pulmonary disease categories like ILD, obstructive or malignant disease (e.g. at a general practitioner's clinic to stimulate adequate referral if indicated) or individual ILDs (e.g. at specialized ILD clinics to serve as input for ILD MDT discussions to obtain higher diagnostic confidence).

Lastly, after implementing eNose testing, maintaining the model quality will need continuous attention. To ensure this quality, we should strive to build a collective eNose database with the ILD research community to enable model updates with new patient data.

Artificial intelligence in ILD

In today's era numerous technological innovations and data applications emerge that use AI to support people's daily lives. Similarly, the development of AI applications is rapidly expanding in clinical medicine [23]. Writing reports, assisting in surgical procedures, generating warnings for change in patients' conditions or informing patients are only a few of the numerous imaginable or already used applications [24]. eNose breath analysis is one of the technologies using AI algorithms for omics data interpretation. Below, the use of omics data for ILD, other than breathomics, is discussed and several examples of investigated AI-based tests. Additionally, important factors to consider when introducing AI-based tests to medical professionals are discussed. Various terms regarding AI used in this paragraph are explained in **Table 1**.

Table 1: Terms and definitions for basic understanding data analysis in the field of artificial intelligence.

Term	Definition
Artificial intelligence	A broad range of technologies allowing computers to simulate human's intelligence processes (e.g. learning, reasoning).
Algorithm	A calculation or set of rules solving a specific problem or performing a particular task. Used in computer science including artificial intelligence applications (e.g. automated processes, decision making and data analyses).
Machine learning	Applications enabling systems to learn and improve their performance autonomously based on available data. A subtype of artificial intelligence.
Supervised learning	An algorithm training models with labeled input data, aiming to learn the model to predict or classify a certain outcome in new unseen data. A type of machine learning approach.
Unsupervised learning	An algorithm training models with unlabeled input data, aiming discover patterns or relations in the input data to generate new hypotheses and insights. A type of machine learning approach.
Big data	Large and complex datasets requiring artificial intelligence for analysis. Often, big data are high dimensional, i.e. datasets in which the number of features exceeds the number of observations.
Omics data	Large datasets containing information from a biological source like genes (i.e. genomics), proteins (i.e. proteomics) or breath (i.e. breathomics). Often analyzed with a data-mining approach using supervised or unsupervised learning. A type of big data.
Dimensionality reduction	Methods for transforming data from high-dimensional to low-dimensional. The transformed data represent the original data. Often used in analysis of big data.

Adjusted from [25-27].

Omics

Omics studies use a big amount of data from a particular biological source that require intelligent multiple-step data analysis to retrieve significant information for answering a research question, e.g. the breathomics studies using eNose sensor data presented in this thesis [28]. An omics approach is particularly popular for discovery of new biomarkers or disease phenotypes, as it entails a data-mining method enabling supervised classification or unsupervised clustering (**Figure 3**). Besides breath, various biological sources like serum, BAL fluid and genes have been examined in the quest for new ILD biomarkers [29]. All sources likely contain slightly different yet valuable information regarding biomarker discovery, stimulating the collection and integration of multiple omics data types. International collaboration will enable the optimal use of both existing and upcoming data, eventually leading to a non-invasive diagnostic and personalized approach for patients with ILD (**Chapter 8**).

A challenge in omics study designs is a correct sample size or power calculation, due to several factors inherent to the nature of omics data. Omics data usually contain a large number of variables that are exposed to dimensionality reduction. Moreover, data are more likely to be non-normally distributed, interactions between variables are complex, and multiple tests are conducted [30]. Only a minority of omics studies report a power calculation, using different methods either pre or post hoc [31]. Peer-reviewers should encourage publication of power calculations, or a sample size rationale if calculation is not possible. Mainly because including the correct number of participants results in valid outcomes with less patient burden, and time, money and resources spend on trials. A variety of statistical methods and machine learning algorithms can help to overcome or limit these challenges leading to new insights.

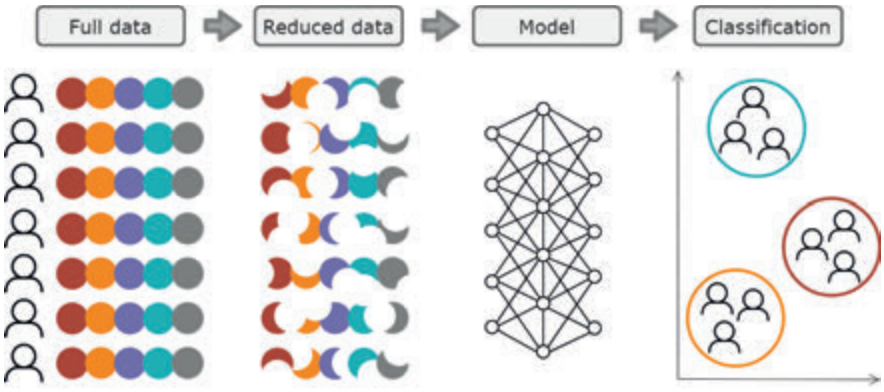


Figure 3: Schematic overview of omics data processing using data reduction and modeling to group patients.

Novel tests

Radiology is the most swiftly growing medical field in terms of development and research of AI algorithms and the number of U.S. Food and Drug Administration approved applications [32]. In ILD, algorithms are developed that use radiomics data resulting from chest CT scans to establish a fibrosis pattern (e.g. UIP) or an IPF diagnosis, and predict forced vital capacity values [33]. Also, automated assessment of chest CTs for phenotyping, prognostication and therapeutic response have been investigated [25]. Other studies combined chest CT, PFT and demographic data to predict disease progression in patients with pulmonary fibrosis [34]. Despite these efforts, experts state that ILD applications for scoring chest CT scans are not yet recommended for use in clinical practice [35]. The lack of clear definitions of clinical outcomes, validation studies, and evidence on clinical benefits of algorithms over standard of care hamper implementation [25].

The genomic classifier is another investigated AI-based tool in ILD. This classifier uses a genomics approach to detect histopathologic UIP patterns based on expressed genes on lung biopsy tissue specimen. The specificity for detecting histopathologic UIP patterns is high [36]. However, low sensitivity and absence of large-scale studies still prevent using a genomic classifier to support MDT diagnostic decisions. Importantly, using a genomic classifier did not impact progression-free survival or pulmonary function decline [37]. Another study evaluates deep learning for supporting ILD pathologist [38]. They report an AUC of 0.86 for predicting UIP patterns in a validation cohort. Although these methods might increase accuracy of tissue diagnosis, they will not reduce the number of non-invasive procedures in the diagnostic trajectory, neither improve clinical outcomes. The potential benefit for patients and patient management is therefore limited.

Furthermore, the promise of proteomics data obtained from serum samples is recently shown by Huang and colleagues [39]. A developed classifier of 37 proteins achieved high accuracies (77.3-82.5%) for differentiating fibrotic connective tissue disease related ILD and IPF in a large cohort. This diagnostic biomarker is non-invasive but the analysis method needs to be improved to achieve a cheap and easily accessible applicable test.

Other studied applications of AI that do not use biological or omics data can support physicians in diagnosing ILD in other ways. For example, the recognition of lung sounds using digital auscultation, automated medical interviews, and interpretation of pulmonary function tests have been studied [26, 40]. Except for diagnostic support, AI can also be implemented for improving preventive and personalized medicine by

automated prediction of disease course and risk factor assessment. Lastly, clinical trials can benefit from AI when using algorithms to assist in analyzing large datasets, retrieving data from electronic medical records or interpreting results from home monitoring devices [26].

Implementation

Integration of AI-based tests in medical practice can improve healthcare for both patients and medical professionals. As shown in various chapters of this thesis for breath analysis, proving the validity of diagnostic AI models requires careful study design to collect sufficient and robust evidence. Moreover, professionals need to trust and interpret results correctly [41]. Park et al. have proposed study phases 0-IV for evaluation of AI applications, mimicking the standardized study phases for approval of new drugs [42]. Medical research communities should consider adopting such an approach to have a standard for sufficient high-quality evidence to guide discussions regarding approval of novel AI-based tests. Several characteristics of AI-based tests like eNose breath analysis, and differences with conventional medical tests (e.g. blood sampling or radiologic imaging) will be explained here, including how to overcome potential barriers for acceptance and implementation.

A major difference with conventional medical tests is that AI-based tests are, in general, designed and validated to provide an explicit answer to a specific question (e.g. Does this breath profile indicate a diagnosis of IPF? What is the most likely diagnosis?). An algorithm instead of a medical professional, interprets the input data and answers the question. Second, algorithms are exceptionally good in discovering patterns in large datasets. Factors that a medical professional would not think of or is not capable of can be related. Mainly, because an algorithm is able to process large amounts of data rapidly. Second, an algorithm does not take standard clinical patterns, correlations or causality into account that are considered by professionals in daily practice. This is demonstrated in a trial that aimed to predict all-cause long-term mortality using deep learning model based on single chest X-rays. The algorithm selected obvious features like size of the heart, prominent pulmonary vasculature or sternotomy wires, but also specific shadows possibly reflecting body habitus as important features for longevity [43].

Moreover, an AI-based test is often perceived as a ‘black box’. This term is easily used if potential users find an electronic system difficult to understand, but does not mean that all AI processes are utterly opaque or unexplainable. In general, lack of transparency can be experienced due to the absence of specific skills or education, or due to the complexity of the system [44]. Obviously, opacity can also be intentional to withhold

information from other parties but is typically more common in commercial than in scientific endeavors. In regard to eNose testing, all three forms of opacity can contribute to some extent to the black box feeling of physicians. This can be improved by educating physicians and using simple transparent algorithms. Though, another contributing factor to the black box perception in eNose testing that is harder to tackle, is the inability to explain the exact type and origin of VOCs that create a particular breath profile.

Another difference with conventional medical tests, is that no normal range or reference value exist for patterns recognized by algorithms, e.g. eNose breath profiles. A paper in mass spectrometry breath analysis aimed to define the 'healthy' breath profile and identified 48 components, but results are not validated yet [45]. eNose studies with this aim have not been published. To date, healthy volunteers have only served as control groups or to assess influence of external factors or time on breath. Device manufacturers could consider defining a device-specific healthy breath profile that can serve as reference for future studies.

Lastly, the population in which a test can be applied can be restricted for AI-base tests according to the cohort used for model training. Most conventional medical tests can be used for all patients; only reference values might differ (e.g. lower limit of hemoglobin concentration varies according to sex, PFT reference values are adjusted for BMI and gender) [46, 47]. This might not apply to AI-based tests, as an algorithm's output depends on the input data, i.e. training population. For example, a skin melanoma algorithm, which is trained on white-skinned people only, cannot be applied to individuals with black skins. Among many other issues, the mentioned differences might hamper the acceptance and understanding of AI-based test results by physicians [48]. More transparency in AI decisions, also called explainable AI, can improve acceptance and understanding [49]. One way to explain AI decisions is 'post-hoc' transparency including providing information on individual algorithmic decisions on a technical or data level [49]. For eNose this could entail adding individual sensor response values of individual measurements. However, this type of transparency may confuses the user more seeing the large amount of data, no direct correlation with exhaled VOCs and complex parameter tuning done by algorithms [50]. A better way to explain AI is by means of 'institutional explanation'. Institutional explanation will answer the question why one can rely on the decisions made by the algorithm by showing how the system is designed, which data are used for training the algorithm, how is bias avoided, etcetera [49]. For eNose tests, information on the applicable patient population and clinical setting should be included in test reports. This likely increases trust and correct interpretation of the outcomes.

Apart from technical and clinical validation, and test opacity and acceptance by professional, several ethical and legal aspects of AI have to be discussed before tests can be incorporated [51]. Is a doctor responsible for wrong decisions made by algorithms? Is all patient data allowed to be used to improve self-learning algorithms? How to avoid overdiagnosis? What should be included in screening reports when eNose can reliably detect various pulmonary disease including cancer?
Discussing and improving the mentioned issues by collaborating with experts from various fields preferably lead to consensus documents that will guide research and enable implementation of AI-based tests in future clinics. Because most important, apart from the challenges presented, AI-based tests and other applications offers numerous opportunities to improve patient care and support medical professionals [23].

Conclusion

To conclude, the encouraging results of eNose exhaled breath analysis in ILD presented in this thesis show multiple potential applications for diagnosing, screening, monitoring and phenotyping patients. The collected data to date allow the design of diagnostic models which should be tested in new robust studies. Collaborative efforts of experts from various fields will enable the development of a technically and medically approved and accepted eNose breath tests that will improve ILD care throughout the disease course.

References

1. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratú P, Resta O, et al. An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. Respir Med. 2013 Jul;107(7):1073-8.
2. Krauss E, Haberer J, Maurer O, Barreto G, Drakopanagiotakis F, Degen M, et al. Exploring the Ability of Electronic Nose Technology to Recognize Interstitial Lung Diseases (ILD) by Non-Invasive Breath Screening of Exhaled Volatile Compounds (VOC): A Pilot Study from the European IPF Registry (eurIPFreg) and Biobank. J Clin Med. 2019 Oct 16;8(10).
3. Dragonieri S, Scioscia G, Quaranta VN, Carratu P, Venuti MP, Falcone M, et al. Exhaled volatile organic compounds analysis by e-nose can detect idiopathic pulmonary fibrosis. J Breath Res. 2020 Jul 21;14(4):047101.
4. Moor CC, Oppenheimer JC, Nakshbandi G, Aerts JGJV, Brinkman P, Maitland-van der Zee AH, et al. Exhaled breath analysis by use of eNose technology: a novel diagnostic tool for interstitial lung disease. Eur Respir J. 2021 Jan;57(1).
5. Don CR, Bell PD, Joseph AH. Fibrosis — A Common Pathway to Organ Injury and Failure. New England Journal of Medicine. 2015;372(12):1138-49.
6. Wijsenbeek M, Cottin V. Spectrum of Fibrotic Lung Diseases. N Engl J Med. 2020 Sep 3;383(10):958-68.
7. Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1;205(9):e18-e47.
8. Yang HY, Peng HY, Chang CJ, Chen PC. Diagnostic accuracy of breath tests for pneumoconiosis using an electronic nose. J Breath Res. 2017 Nov 29;12(1):016001.
9. Xuan W, Zheng L, Bunes BR, Crane N, Zhou F, Zang L. Engineering solutions to breath tests based on an e-nose system for silicosis screening and early detection in miners. J Breath Res. 2022 Apr 7;16(3).
10. Joy GM, Arbiv OA, Wong CK, Lok SD, Adderley NA, Dobosz KM, et al. Prevalence, imaging patterns and risk factors of interstitial lung disease in connective tissue disease: a systematic review and meta-analysis. Eur Respir Rev. 2023 Mar 31;32(167).
11. Fijten RRR, Smolinska A, Drent M, Dallinga JW, Mostard R, Pachen DM, et al. The necessity of external validation in exhaled breath research: a case study of sarcoidosis. J Breath Res. 2017 Nov 29;12(1):016004.
12. Plantier L, Smolinska A, Fijten R, Flamant M, Dallinga J, Mercadier JJ, et al. The use of exhaled air analysis in discriminating interstitial lung diseases: a pilot study. Respir Res. 2022 Jan 20;23(1):12.
13. Sabbagh P, Mohammadnia-Afrouzi M, Javanian M, Babazadeh A, Koppolu V, Vasigala VR, et al. Diagnostic methods for Helicobacter pylori infection: ideals, options, and limitations. European Journal of Clinical Microbiology & Infectious Diseases. 2019 2019/01/01;38(1):55-66.
14. Dalle-Donne I, Rossi R, Colombo R, Giustarini D, Milzani A. Biomarkers of Oxidative Damage in Human Disease. Clinical Chemistry. 2006;52(4):601-23.
15. Principe S, van Bragt J, Longo C, de Vries R, Sterk PJ, Scichilone N, et al. The Influence of Smoking Status on Exhaled Breath Profiles in Asthma and COPD Patients. Molecules. 2021 Mar 4;26(5).

16. Dragonieri S, Quaranta VN, Buonamico E, Battisti C, Ranieri T, Carratu P, et al. Short-Term Effect of Cigarette Smoke on Exhaled Volatile Organic Compounds Profile Analyzed by an Electronic Nose. Biosensors (Basel). 2022 Jul 13;12(7).

17. Dragonieri S, Quaranta VN, Portacci A, Ahroud M, Di Marco M, Ranieri T, et al. Effect of Food Intake on Exhaled Volatile Organic Compounds Profile Analyzed by an Electronic Nose. Molecules. 2023 Jul 30;28(15).

18. Dragonieri S, Quaranta VN, Carratu P, Ranieri T, Resta O. Influence of age and gender on the profile of exhaled volatile organic compounds analyzed by an electronic nose. J Bras Pneumol. 2016 Apr;42(2):143-5.

19. Dragonieri S, Quaranta VN, Carratu P, Ranieri T, Resta O. The ovarian cycle may influence the exhaled volatile organic compound profile analyzed by an electronic nose. J Breath Res. 2018 Mar 1;12(2):021002.

20. Liotino V, Dragonieri S, Quaranta VN, Carratu P, Ranieri T, Resta O. Influence of circadian rhythm on exhaled breath profiling by electronic nose. J Biol Regul Homeost Agents. 2018 Sep-Oct;32(5):1261-65.

21. Kim MJ, Hong SH, Cho W, Park D-H, Lee E-B, Song Y, et al. Breath Acetone Measurement-Based Prediction of Exercise-Induced Energy and Substrate Expenditure. Sensors. 2020;20(23):6878.

22. De Sadeleer LJ, Meert C, Yserbyt J, Slabbynck H, Verschakelen JA, Verbeken EK, et al. Diagnostic Ability of a Dynamic Multidisciplinary Discussion in Interstitial Lung Diseases: A Retrospective Observational Study of 938 Cases. Chest. 2018 2018/06/01/;153(6):1416-23.

23. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. N Engl J Med. 2023 Mar 30;388(13):1201-08.

24. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, et al. Foundation models for generalist medical artificial intelligence. Nature. 2023 2023/04/01;616(7956):259-65.

25. Barnes H, Humphries SM, George PM, Assayag D, Glaspole I, Mackintosh JA, et al. Machine learning in radiology: the new frontier in interstitial lung diseases. The Lancet Digital Health. 2023 2023/01/01/;5(1):e41-e50.

26. Handa T. The potential role of artificial intelligence in the clinical practice of interstitial lung disease. Respiratory Investigation. 2023 2023/11/01/;61(6):702-10.

27. Pammi M, Aghaeepour N, Neu J. Multiomics, artificial intelligence, and precision medicine in perinatology. Pediatric Research. 2023 2023/01/01;93(2):308-15.

28. Newton CA, Herzog EL. Molecular Markers and the Promise of Precision Medicine for Interstitial Lung Disease. Clin Chest Med. 2021 Jun;42(2):357-64.

29. Inoue Y, Kaner RJ, Guiot J, Maher TM, Tomassetti S, Moiseev S, et al. Diagnostic and Prognostic Biomarkers for Chronic Fibrosing Interstitial Lung Diseases With a Progressive Phenotype. Chest. 2020 Aug;158(2):646-59.

30. Krassowski M, Das V, Sahu SK, Misra BB. State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing. Frontiers in Genetics. 2020 2020-December-10;11.

31. Balki I, Amirabadi A, Levman J, Martel AL, Emersic Z, Meden B, et al. Sample-Size Determination Methodologies for Machine Learning in Medical Imaging Research: A Systematic Review. Canadian Association of Radiologists Journal. 2019;70(4):344-53.

32. FDA. Artificial intelligence and machine learning (AI/ML)–enabled medical devices.: FDA.

33. Dack E, Christe A, Fontanellaz M, Brigato L, Heverhagen JT, Peters AA, et al. Artificial Intelligence and Interstitial Lung Disease: Diagnosis and Prognosis. Invest Radiol. 2023 Aug 1;58(8):602-09.

34. Walsh SLF, Mackintosh JA, Calandriello L, Silva M, Sverzellati N, Larici AR, et al. Deep Learning-based Outcome Prediction in Progressive Fibrotic Lung Disease Using High-Resolution Computed Tomography. Am J Respir Crit Care Med. 2022 Oct 1;206(7):883-91.

35. Rajan SK, Cottin V, Dhar R, Danoff S, Flaherty KR, Brown KK, et al. Progressive pulmonary fibrosis: an expert group consensus statement. Eur Respir J. 2023 Mar;61(3).

36. Kheir F, Uribe Becerra JP, Bissell B, Ghazipura M, Herman D, Hon SM, et al. Use of a Genomic Classifier in Patients with Interstitial Lung Disease: A Systematic Review and Meta-Analysis. Ann Am Thorac Soc. 2022 May;19(5):827-32.

37. Chaudhary S, Weigt SS, Ribeiro Neto ML, Benn BS, Pugashetti JV, Keith R, et al. Interstitial lung disease progression after genomic usual interstitial pneumonia testing. Eur Respir J. 2023 Apr;61(4).

38. Uegami W, Bychkov A, Ozasa M, Uehara K, Kataoka K, Johkoh T, et al. MIXTURE of human expertise and deep learning-developing an explainable model for predicting pathological diagnosis and survival in patients with interstitial lung disease. Mod Pathol. 2022 Aug;35(8):1083-91.

39. Huang Y, Ma SF, Oldham JM, Adegunsoye A, Zhu D, Murray S, et al. Machine Learning of Plasma Proteomics Classifies Diagnosis of Interstitial Lung Disease. Am J Respir Crit Care Med. 2024 Feb 29.

40. Siebert JN, Hartley MA, Courvoisier DS, Salamin M, Robotham L, Doenz J, et al. Deep learning diagnostic and severity-stratification for interstitial lung diseases and chronic obstructive pulmonary disease in digital lung auscultations and ultrasonography: clinical protocol for an observational case-control study. BMC Pulm Med. 2023 Jun 2;23(1):191.

41. van de Sande D, Van Genderen ME, Smit JM, Huiskens J, Visser JJ, Veen RER, et al. Developing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter. BMJ Health Care Inform. 2022 Feb;29(1).

42. Park Y, Jackson GP, Foreman MA, Gruen D, Hu J, Das AK. Evaluating artificial intelligence in medicine: phases of clinical research. JAMIA Open. 2020 Oct;3(3):326-31.

43. Lu MT, Ivanov A, Mayrhofer T, Hosny A, Aerts H, Hoffmann U. Deep Learning to Assess Long-term Mortality From Chest Radiographs. JAMA Netw Open. 2019 Jul 3;2(7):e197416.

44. Wadden JJ. Defining the undefinable: the black box problem in healthcare artificial intelligence. J Med Ethics. 2021 Jul 21.

45. Sasiene ZJ, LeBrun ES, Schaller E, Mach PM, Taylor R, Candelaria L, et al. Real-time breath analysis towards a healthy human breath profile. J Breath Res. 2024 Jan 23;18(2).

46. Beutler E, Waalen J. The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? Blood. 2006 Mar 1;107(5):1747-50.

47. Bhakta NR, McGowan A, Ramsey KA, Borg B, Kivastik J, Knight SL, et al. European Respiratory Society/American Thoracic Society technical statement: standardisation of the measurement of lung volumes, 2023 update. Eur Respir J. 2023 Oct;62(4).

48. Sullivan E. Understanding from Machine Learning Models. The British Journal for the Philosophy of Science. 2020 2022/03/01;73(1):109-33.

49. Theunissen M, Browning J. Putting explainable AI in context: institutional explanations for medical AI. Ethics Inf Technol. 2022;24(2):23.

50. Zednik C. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. Philosophy & Technology. 2021 2021/06/01;34(2):265-88.

51. Morley J, Floridi L. An ethically mindful approach to AI for health care. The Lancet. 2020 2020/01/25/;395(10220):254-55.

PART VI

APPENDICES

SUMMARY

Part I: Introduction

The group of interstitial lung disease (ILD) includes >200 different lung disorders affecting the lung interstitium. Patients present with generic respiratory symptoms like dyspnea on exertion, cough and fatigue. Most ILDs are rare and idiopathic of origin. Besides, disease course varies widely between and within ILDs, but disease is often chronic with progressive symptoms. ILD is usually suspected on an high-resolution chest computed tomography (CT) scan, but a CT scan is not specific to establish an exact ILD diagnosis. After multiple additional tests, the ILD diagnosis is determined in a multidisciplinary team discussion with experts from various fields of medicine, including but not limited to, pulmonologists, radiologists, pathologists and rheumatologists. Treatment options include anti-inflammatory and antifibrotic drugs. Determining the optimal treatment strategy depends on multiple patient-related factors but is often challenging and response to treatment is difficult to predict.

Exhaled breath contains various compounds including volatile organic compounds (VOCs) that originate from various metabolic and pathophysiologic processes within the body or from direct or indirect external influences. The total of VOCs that a person exhales, is called a breath profile and represents the person's health status. Over the past decades the analysis of VOCs is studied using two main techniques, gas chromatography-mass spectrometry (GC-MS) and electronic nose (eNose) sensor technology. The latter has potential as point-of-care medical test seeing its quick, non-invasive, cheap and accessible nature. **Chapter 1** describes technical details of eNose technology in general and provides an overview of characteristics of available devices. Moreover, the available evidence collected for diagnosing and monitoring respiratory diseases using eNose technology are reviewed. In **Chapter 2** focuses specifically on ILD and reviews the published evidence from exhaled breath studies using eNose or GC-MS technology. eNose seems to have the highest potential for point-of-care medical testing in ILD populations and is therefore the main aim of this thesis.

Chapter 3 describes results from a patient survey revealing patient-related and healthcare-related factors that contribute to delays throughout the diagnostic journey. Longer time from symptom onset to a final ILD diagnosis and subsequent treatment is associated with worse patient outcomes. An accurate single non-invasive diagnostic test or biomarker, like VOCs in exhaled breath, might help reducing the diagnostic delay.

Part II of this thesis, contains several original research papers that aimed to further investigate the diagnostic accuracy of an eNose for various ILDs. **Part III** focuses on the potential of eNose technology as screening tool in patients at high-risk for developing ILD. In **Part IV**, we explored the value of breath analysis using GC-MS for ILD phenotyping.

Part II: Diagnosis

Chapter 4 reports results from a cross-sectional study in patients with any type ILD and other pulmonary diseases like asthma, chronic obstructive pulmonary disease (COPD), and lung cancer from two hospitals. The aim was to evaluate the discriminating ability of eNose technology for breath profiles of patients with ILD from those with asthma, COPD, and lung cancer. Results showed a high accuracy for distinguishing the patient groups in both the training and test cohorts. Also individual patient groups of ILD, asthma, COPD and lung cancer could be separated accurately. This indicates a potential future role for using an eNose as medical test for identification of ILD amongst patients with respiratory symptoms. Ideally, eNose will facilitate earlier referral of patients with a high probability of ILD by showing physicians the probability scores for various lung diseases after a single eNose measurement.

In **Chapter 5** preliminary baseline results of a longitudinal international study for external validation of eNose performance in patients with fibrotic ILD are presented. The ILDnose study is currently conducted in four international ILD expertise centers. The preliminary results that resulted from three centers showed high discriminative values for patients with idiopathic pulmonary fibrosis versus patients with another fibrotic ILD. Results from the training set could be validated in the test and external validation sets, indicating robust reliable results. Full baseline data are awaited and can be used to build diagnostic algorithms for various ILD diagnoses. This will hopefully reduce the need for invasive diagnostic procedures when diagnostic confidence of multidisciplinary team conclusions increases. Moreover, longitudinal data should reveal the value of an eNose in monitoring and predicting disease.

Chapter 6 evaluates whether sarcoidosis has a distinct eNose breath profile compared to healthy persons and patients with other ILDs. Sarcoidosis is a multisystem disease and represents specific entity within the group of ILDs. The majority of patients have pulmonary involvement, which can mimic other forms of ILD. eNose perfectly discriminated sarcoidosis from healthy persons in both a training and test set. Also breath profiles of patients with pulmonary involvement could be separated accurately from patients with other ILDs. Moreover, the influence of immunosuppressive treatment on the detection of sarcoidosis was investigated, but treatment did not appear to significantly impact detection. Interestingly, patients with higher levels of soluble interleukin-2 receptor, possibly indicating active inflammation in sarcoidosis, had different breath profiles that patients with lower levels. The results indicate that eNose may facilitate a fast and accurate sarcoidosis diagnosis in the future. New multicenter studies are now warranted to validate results externally and longitudinal studies for understanding the value of eNose in monitoring disease activity.

Chapter 7 reports the design and performance of a diagnostic eNose model for pulmonary sarcoidosis. We aimed to evaluate various dimensionality reduction methods and classifiers in order to design an accurate diagnostic model with eNose data. After multiple steps including dimensionality reduction, hyperparameter optimization and cross-validation, the random forest classifier resulted in the highest overall diagnostic performance for pulmonary sarcoidosis among patients with other ILDs. This chapter shows an example of a robust method for model design which can be applied to other eNose datasets as well.

Chapter 8 contains a reply to a recently published paper on a developed serum proteomics classifier for diagnosing fibrotic ILDs. Our correspondence letter highlights the potential of and contains perspectives on using various omics data sources, like proteomics and breathomics, for enhance the diagnostic trajectory of ILD.

Part III: Screening

Chapter 9 explores whether an eNose can be used to screen for systemic sclerosis (SSc) related ILD, which is an important cause of morbidity and mortality in patients with SSc. A multicenter study was performed to answer this question. This resulted in a high accuracy for differentiating eNose breath profiles of patients with SSc without ILD and patients with SSc-ILD. Breath profiles also differed from other types of connective tissue related ILD. These results indicate the potential of eNose as a screening test for SSc-ILD, which should be further explored in longitudinal studies.

A proof-of-concept study in **Chapter 10** investigates the potential of screening in a population at risk for developing drug-induced ILD: patients with cancer and antineoplastic treatment. Breath profiles of patients that were diagnosed with drug-induced ILD confirmed by an multidisciplinary team discussion, were compared to patients without ILD. eNose technology could distinguish these groups with a high area under the curve value, indicating that eNose might detect pulmonary toxicity caused by cancer treatment. To confirm the ability of an eNose to serve as screening test, longitudinal studies in patients with cancer are warranted.

Part IV: Gas chromatography-mass spectrometry

Chapter 11 evaluates results of breath analysis in patients with ILD using GC-MS technology. The patient cohorts were included in two hospitals, of which one served as a training and one as an external validation cohort. We aimed to unravel the VOC differences between patients with fibrotic or inflammatory interstitial abnormalities on

chest CT scan, or a combination of both. Identifying VOCs that are specific for these three phenotypes groups, might guide future treatment decisions. However, results of the training cohort could not be validated in the other cohort. This suggests no role for GC-MS exhaled breath analysis in differentiating ILD phenotypes and underlines the need of external validation of results in biomarker studies.

Part V: Discussion

The general discussion discusses the relevance, future implications and limitations of the results published in this thesis.

SAMENVATTING

Deel I: Introductie

De groep interstitiële longziekten (ILD) omvat >200 verschillende longaandoeningen die het long interstitium (i.e. de ruimte tussen de longblaasjes en de bloedvaten) aantasten. Patiënten presenteren zich met algemene respiratoire symptomen zoals kortademigheid bij inspanning, hoesten en vermoeidheid. De meeste ILD's zijn zeldzaam met een onduidelijke oorsprong. Bovendien varieert het ziektebeloop sterk tussen en binnen verschillende ILD's, maar de ziekte is vaak chronisch met toenemende klachten. ILD wordt meestal vermoed op een computer tomografie (CT) scan van de borstkas, maar de CT-scan is niet specifiek om een exacte diagnose van ILD vast te stellen. Na meerdere aanvullende onderzoeken wordt de meest waarschijnlijke ILD diagnose vastgesteld in een multidisciplinaire teambespreking met deskundigen uit verschillende medische disciplines, onder andere longartsen, radiologen, pathologen en reumatologen. Behandelingsopties omvatten afweeronderdrukkende en fibrose remmende medicatie. Het bepalen van de optimale behandelstrategie hangt af van meerdere patiënt gerelateerde factoren, maar is vaak een uitdaging en de respons op behandeling is moeilijk te voorspellen.

Uitgeademde lucht bevat verschillende componenten, waaronder vluchtige organische stoffen (VOC's) die afkomstig zijn van verschillende metabolische en pathofysiologische processen in het lichaam of van (in)directe invloeden van buitenaf. Het totaal aan VOC's dat een persoon uitademt, wordt een ademprofiel genoemd en is een afspiegeling van de gezondheidsstatus van het lichaam. In de afgelopen decennia is de analyse van VOC's bestudeerd met behulp van twee technieken, gaschromatografie-massaspectrometrie (GC-MS) en elektronische neus (*eNose*) sensortechnologie. De laatste heeft potentieel als medische *point-of-care* test (i.e. sneltest) gezien de snelle, niet-invasieve, goedkope en toegankelijke aard. **Hoofdstuk 1** beschrijft de technische details van *eNose*-technologie in het algemeen en geeft een overzicht van de kenmerken van beschikbare apparaten. Bovendien wordt een overzicht gegeven van het beschikbare bewijsmateriaal dat is verzameld voor de diagnose en het monitoren van longziekten met behulp van *eNose*-technologie. **Hoofdstuk 2** richt zich specifiek op ILD en geeft een overzicht van de gepubliceerde onderzoeken naar uitademingslucht analyse met behulp van *eNose*- of GC-MS-technologie. De *eNose* lijkt het grootste potentieel te hebben als medische sneltest bij ILD en is daarom het hoofddoel van dit proefschrift.

Deel II: Diagnose

Hoofdstuk 4 rapporteert de resultaten van een cross-sectioneel onderzoek verricht in twee ziekenhuizen bij patiënten met een vorm van ILD en andere longziekten zoals astma, chronisch obstructieve longziekte (COPD) en longkanker. Het doel was om de nauwkeurigheid van *eNose*-technologie te evalueren met betrekking tot het onderscheiden van ademprofielen van patiënten met ILD en die van patiënten met astma, COPD en longkanker. De resultaten toonden een hoog onderscheidend vermogen in zowel de trainings- als testcohorten. Ook de individuele patiëntengroepen ILD, astma, COPD en longkanker konden nauwkeurig worden onderscheiden. Dit wijst op een mogelijke toekomstige rol voor het gebruik van een *eNose* als medische test voor de identificatie van ILD onder patiënten met longklachten. Idealiter zal de *eNose* het juist doorverwijzen van patiënten met een hoge verdenking op ILD versnellen doordat artsen waarschijnlijkheidsscores voor verschillende longziekten te zien krijgen na een enkele *eNose*-meting.

In **hoofdstuk 5** worden de voorlopige resultaten gepresenteerd van een longitudinaal internationaal onderzoek ter externe validatie van de prestaties van *eNose* bij patiënten met longfibrose. Het '*ILDnose'* onderzoek wordt momenteel uitgevoerd in vier internationale ILD-expertisecentra. De voorlopige resultaten van drie centra lieten hoge discriminerende waarden zien voor patiënten met idiopathische longfibrose versus patiënten met een ander type longfibrose. Resultaten van de trainingsset konden worden gevalideerd in de test- en externe validatiesets, wat duidt op robuuste betrouwbare resultaten. De volledige resultaten worden binnenkort verwacht en kunnen worden gebruikt om diagnostische algoritmen te bouwen voor verschillende ILD-diagnoses. Hopelijk vermindert dat de noodzaak tot invasieve diagnostische procedures door een hogere diagnostische zekerheid in multidisciplinaire teambesprekingen. Daarnaast zullen longitudinale resultaten de waarde van een *eNose* bij het monitoren en voorspellen van ziektebeloop aantonen.

Hoofdstuk 6 evalueert of sarcoïdose een ander *eNose*-ademprofiel heeft in vergelijking met gezonde personen en patiënten met een andere ILD. Sarcoïdose is een multisysteemziekte en vormt een specifieke entiteit binnen de groep ILD's. De meerderheid van de patiënten heeft longbetrokkenheid, die kan lijken op andere vormen van ILD. *eNose* onderscheidde sarcoïdose perfect van gezonde personen in zowel de trainings- als testset. Ook ademprofielen van sarcoïdose patiënten met longbetrokkenheid konden nauwkeurig worden gescheiden van patiënten met andere vorm van ILD. Daarnaast werd de invloed van het gebruik van afweeronderdrukkende medicatie op de detectie van sarcoïdose onderzocht, maar dit lijkt geen belangrijke invloed te hebben. Interessant genoeg hadden patiënten met hoge interleukine-2

receptor niveaus, wat kan wijzen op hoge ontstekingsactiviteit, andere ademprofielen dan patiënten met lage niveaus. De resultaten geven aan dat *eNose* in de toekomst een snelle en nauwkeurige diagnose van sarcoïdose kan ondersteunen. Nieuwe multicenteronderzoeken zijn nu eerst nodig om de genoemde resultaten extern te valideren en longitudinale onderzoeken om de rol van *eNose* bij het monitoren van ziekteactiviteit beter te begrijpen.

Hoofdstuk 7 beschrijft het ontwerpen en de prestaties van een diagnostisch *eNose*-model voor pulmonale sarcoïdose. We evalueerden verschillende manieren van dimensionaliteitsreductie en classificatie om een nauwkeurig diagnostisch model te bouwen op basis van *eNose*-gegevens. Na meerdere stappen waaronder dimensionaliteitsreductie, hyperparameteroptimalisatie en kruisvalidatie, toonde de *random forest* classificatiemethode de hoogste nauwkeurigheid voor het diagnosticeren van pulmonale sarcoïdose onder patiënten met andere vormen van ILD. Dit hoofdstuk toont een voorbeeld van een robuuste methode voor modelontwerp die ook kan worden toegepast op andere *eNose*-datasets.

Hoofdstuk 8 bevat een antwoord op een recent gepubliceerd artikel over een ontwikkelde serum *proteomics classifier* voor het diagnosticeren van fibrotische ILDs. Onze schriftelijke reactie beschrijft het potentieel van en geeft perspectieven op het gebruik van verschillende *omics* bronnen, zoals *proteomics* en *breathomics*, voor het verbeteren van het diagnostische traject van ILD.

Deel III: Screening

In **hoofdstuk 9** hebben we onderzocht of een *eNose* kan worden gebruikt om te screenen op systemische sclerose (SSc) geassocieerde ILD, een belangrijke oorzaak van morbiditeit en mortaliteit bij patiënten met SSc. Er werd een multicenteronderzoek uitgevoerd om deze vraag te beantwoorden. Dit resulteerde in een hoge nauwkeurigheid voor het onderscheiden van *eNose*-ademprofielen van patiënten met SSc zonder ILD en patiënten met SSc-ILD. Ademprofielen verschilden ook van mensen met andere typen auto-immuunziekten geassocieerde ILD. Deze resultaten wijzen op het potentieel van *eNose* als screeningtest voor SSc-ILD, dat verder moet worden onderzocht in longitudinale studies.

De *proof-of-concept*-studie beschreven in **hoofdstuk 10** onderzoekt de mogelijkheid van screening in een populatie met een verhoogd risico op het ontwikkelen van medicatie geïnduceerde ILD: patiënten met kanker en antineoplastische behandeling. Ademprofielen van patiënten met de diagnose medicatie geïnduceerde ILD, bevestigd

in een multidisciplinaire teambespreking, werden vergeleken met patiënten zonder ILD. *eNose*-technologie kon deze groepen onderscheiden met een hoge *area under the curve* waarde, wat aangeeft dat een *eNose* pulmonale toxiciteit veroorzaakt door kankerbehandeling zou kunnen detecteren. Om het vermogen van een *eNose* als screeningstest te bevestigen, zijn longitudinale studies bij patiënten met kanker nodig.

Deel IV: Gaschromatografie-massaspectrometrie

Hoofdstuk 11 evalueert de resultaten van ademanalyse bij patiënten met ILD die is uitgevoerd met behulp van GC-MS-technologie. Patiënten werden geïncludeerd in twee ziekenhuizen, waarvan één ziekenhuis diende als trainingscohort en één als extern validatiecohort. We wilden de verschillen in VOC's achterhalen tussen patiënten met longfibrose, -inflammatie of een combinatie van beide op een CT-scan van de borstkas. Het identificeren van VOC's die specifiek zijn voor deze drie fenotypegroepen, zou een leidraad kunnen zijn voor toekomstige behandelingsbeslissingen. De resultaten van het trainingscohort konden echter niet worden gevalideerd in het andere cohort. Dit suggereert geen rol voor GC-MS-analyse van uitgeademde lucht bij het onderscheiden van ILD fenotypen en onderstreept de noodzaak van externe validatie van resultaten in onderzoeken naar biologische markers.

Deel V: Algemene discussie

De algemene discussie bespreekt de relevantie, implicaties en beperkingen van de in dit proefschrift gepubliceerde bevindingen.

PHD PORTFOLIO

PhD student

Naam	I.G. van der Sar
University	Erasmus University Rotterdam
Department	Respiratory Medicine
Research School	Molecular Medicine
Period	October 2020 – April 2024
Promotor	Prof.dr. M.S. Wijsenbeek
Co-promotor	Dr. C.C. Moor

Summary of PhD training and teaching activities

Courses, seminars and workshops	**Year**	**Workload (ECTS)**
Workshop Endnote	2020	0.2
Biomedical Writing (MolMed)	2020	2.5
eBROK® course	2020	1.5
Getting started with Castor EDC	2020	0.1
Workshop on Indesign (MolMed)	2021	0.15
Workshop on Photoshop/Illustrator (MolMed)	2021	0.3
Scientific integrity	2021	0.3
Basic introduction course on SPSS (MolMed)	2021	1.0
Basic course on R (MolMed)	2021	1.8
The Personal Leadership & Communication for PhD students and Post Docs (MolMed)	2021	1.0
ILD course Davos	2022	1.0
Planning and evaluation of screening (NIHES)	2022	1.4
Biostatistics I (NIHES)	2022	4.5
Good clinical practice (4x; Smelt Academy)	2020-2023	0.4
Conferences, symposia and presentations		
Virtual School on Interstitial Lung Diseases (ERS)	2020-2021	0.8
Young investigators symposium (NRS)	2020	0.2
Symposium ILD in Rheumatic diseases (ERS)	2021	0.2
Satellites conference (ERS)	2021-2022	0.9

Dutch Lung Congress (NRS; 1 poster)	2021	0.3
ERS conference – Virtual (1 poster)	2021	0.3
National lung course & Young investigators symposium (NRS)	2021	5.0
Ledendag Sarcoidose.nl (1 oral presentation)	2021	0.3
ATS conference	2022	0.8
Breath Summit (IABR; 1 oral presentation)	2022	1.0
ERS conference (1 oral presentation)	2022	1.0
Biomedical PhD Day (Graduate School Erasmus MC)	2022	0.3
European conference on neonatal and paediatric PVD (1 oral presentation)	2022	1.0
Innovation 4 health conference (1 poster)	2023	0.3
Clinical & Health sciences day (Graduate School Erasmus MC)	2023	0.3
WASOG conference (1 poster discussion)	2023	1.0
Dutch Lung Congress (NRS)	2023	0.3
Aarhus Symposium on Lung Disorders (1 invited talk)	2023	1.0
ERS conference (1 oral presentation)	2023	1.0
Wetenschapsdag Longziekten (2 oral presentations)	2022-2023	2.0
European Lung Cancer conference (1 poster presentation)	2024	0.3

Teaching activities		
Coach bachelor students (4x) including basic course	2021-2024	1
Daily supervisor master thesis medical student (2x)	2021-2022	3
Daily supervisor bachelor thesis technical medicine students (4x)	2022	1.5
Daily supervisor bachelor thesis student Erasmus University College	2023	1.5
Workshops blood gas analysis bachelor 1 (4x)	2021-2024	0.4

Other		
Peer reviewer for international journals (11x)	2021-2024	1.1
Bouwhuisprijs (Sarcoidose.nl)	2021	
Interview article (Mdedge and Healio)	2021	
Young Investigator Symposium Masterclass Award (NRS)	2021	
Sub-PI Horizon HARBOR study	2022-2023	
Member NRS webinar committee	2022-2023	
Board member AAV Erasmus MC	2021-2023	
Merit Travel Grant (ESMO)	2024	
Total ECTS		**42.65**

LIST OF PUBLICATIONS

Publications included in this thesis

van der Sar IG, Jones S, Clarke DL, Bonella F, Fourrier JM, Lewandowska K, Bermudo G, Simidchiev A, Strambu IR, Wijsenbeek MS, Parfrey H: Patient Reported Experiences and Delays During the Diagnostic Pathway for Pulmonary Fibrosis: A Multinational European Survey. Front Med (Lausanne) 2021. *doi: 10.3389/fmed.2021.711194.*

van der Sar IG, Wijbenga N, Nakshbandi G, Aerts JGJV, Manintveld OC, Wijsenbeek MS, Hellemons ME, Moor CC: The smell of lung disease: a review of the current status of electronic nose technology. Respiratory Research 2021. *doi: 10.1186/s12931-021-01835-4.*

van der Sar IG, Moor CC, Oppenheimer JC, Luijendijk ML, van Daele PLA, Maitland-van der Zee AH, Brinkman P, Wijsenbeek MS: Diagnostic Performance of Electronic Nose Technology in Sarcoidosis. Chest 2022. *doi: 10.1016/j.chest.2021.10.025.*

van der Sar IG, Wijsenbeek MS, Moor CC: Exhaled breath analysis in interstitial lung disease. Current Opinion in Pulmonary Medicine 2023. *doi: 10.1097/MCP.0000000000000978.*

van der Sar IG, van Jaarsveld N, Spiekerman IA, Toxopeus FJ, Langens QL, Wijsenbeek MS, Dauwels J, Moor CC: Evaluation of different classification methods using electronic nose data to diagnose sarcoidosis. Journal of Breath Research 2023. *doi: 10.1088/1752-7163/acf1bf.*

van der Sar IG, Wijsenbeek MS, Braunstahl GJ, Loekabino JO, Dingemans AC, In 't Veen JCCM, Moor CC: Differentiating interstitial lung diseases from other respiratory diseases using electronic nose technology. Respiratory Research 2023. *doi: 10.1186/s12931-023-02575-3.*

van der Sar IG, Wijsenbeek MS, Dumoulin DW, Jager A, van der Veldt AAM, Rossius MJP, Dingemans AC, Moor CC: Detection of Drug-induced Interstitial Lung Disease Caused by Cancer Treatment Using Electronic Nose Exhaled Breath Analysis. Annals of the American Thoracic Society 2024. *doi: 10.1513/AnnalsATS.202401-112RL.*

van der Sar IG, Moor CC, Wijsenbeek MS: Classifying Interstitial Lung Disease: Omics are in the Air. American Journal of Respiratory and Critical Care Medicine 2024. *doi: 10.1164/rccm.202404-0748LE.*

Marges ER, van der Sar IG, Vries-Bouwstra JK, Huizinga TWJ, van Daele PLA, Wijsenbeek MS, Moor CC, Geelhoed JJM: Detection of Systemic Sclerosis-Associated Interstitial Lung Disease by Exhaled Breath Analysis Using Electronic Nose Technology. American Journal of Respiratory and Critical Care Medicine 2024. *doi: 10.1164/rccm.202402-0272LE.*

Published conference abstracts

van der Sar IG, Moor CC, Luijendijk ML, Brinkman P, Maitland-Van Der Zee A-H, Aerts JGJV, Wijsenbeek MS: Unsupervised analysis of electronic nose data identifies patient clusters in fibrosing interstitial lung disease. European Respiratory Journal 2021. *doi: 10.1183/13993003.congress-2021.PA475.*

van der Sar IG, Moor CC, Vellekoop BP, Wijsenbeek MS: Predicting treatment response in patients with interstitial lung disease using electronic nose technology. European Respiratory Journal 2022. *doi: 10.1183/13993003.congress-2022.345.*

Moor CC, van der Sar IG, Wijsenbeek MS: Electronic nose technology detects connective tissue disease-associated interstitial lung disease. European Respiratory Journal 2023. *doi: 10.1183/13993003.congress-2023.OA852.*

van der Sar IG, Wijsenbeek MS, Braunstahl G-J, Loekabino JO, Dingemans AC, In ‘t Veen JCCM, Moor CC: Electronic nose technology differentiates interstitial lung diseases from other chronic respiratory diseases. European Respiratory Journal 2023. *doi: 10.1183/13993003.congress-2023.OA1427.*

van der Sar IG, Wijsenbeek MS, Dumoulin DW, Jager A, van der Veldt AAM, Rossius MJP, Dingemans AC, Moor CC: 246P Detection of drug-induced interstitial lung disease caused by cancer treatment using electronic nose exhaled breath analysis. ESMO open 2024. *doi: 10.1016/j.esmoop.2024.102708.*

ABOUT THE AUTHOR

Iris Gerdina van der Sar was born at November 1st 1993 in Delft, the Netherlands. She was raised in ‘s-Gravenzande. There, she attended primary (Prins Willem Alexanderschool) and secondary school (Interconfessionele Scholengroep Westland) graduating in 2011 at atheneum level. She started medical school in September 2011 at the Erasmus University Rotterdam. During her studies, she was chair of the full-time board of the medical student association Rotterdam (MFVR), gained medical research experience abroad at the Cardiology department of the Royal Prince Alfred Hospital (Sydney, Australia) for her master’s thesis, and got involved in educational committees at the faculty of Medicine and Liberal Arts & Sciences.

After obtaining her master’s degree in 2019, she started working as resident (not in training) at the internal medicine department at Reinier de Graaf hospital in Delft. She continued this position at the respiratory medicine department at Leiden University Medical Center in 2020, seeing her increasing interests for patients with pulmonary diseases. In October that year, she started as a PhD candidate supervised by promotor prof.dr. M.S. Wijsenbeek (prof.dr. J.G.J.V. Aerts until October 2022) and co-promotor dr. C.C. Moor. Conducted research projects since then have resulted in the presented thesis.

From September 2024 onwards, Iris will continue her medical career as General Practitioner in training.