ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original article

A systematic review of the diagnostic accuracy of volatile organic compounds in airways diseases and their relation to markers of type-2 inflammation

Wadah Ibrahim, Sushiladevi Natarajan, Michael Wilde, Rebecca Cordell, Paul S Monks, Neil Greening, Christopher E Brightling, Rachael Evans, Salman Siddiqui

Please cite this article as: Ibrahim W, Natarajan S, Wilde M, *et al*. A systematic review of the diagnostic accuracy of volatile organic compounds in airways diseases and their relation to markers of type-2 inflammation. *ERJ Open Res* 2021; in press (https://doi.org/10.1183/23120541.00030-2021).

A Systematic Review of the Diagnostic Accuracy of Volatile Organic Compounds in Airways Diseases and their relation to Markers of Type-2 Inflammation

Wadah Ibrahim[1,2*], Sushiladevi Natarajan[1,2*], Michael Wilde[3], Rebecca Cordell[3], Paul S Monks[3], Neil Greening[1,2], Christopher E Brightling[1,2], Rachael Evans[1,2], Salman Siddiqui[1,2] on behalf of the EMBER[4] consortium

wadah.ibrahim@leicester.ac.uk; sushiladevi.natarajan@nhs.net; mjw77@leicester.ac.uk; rc145@leicester.ac.uk; paul.monks@leicester.ac.uk; neil.greening@leicester.ac.uk; ceb17@leicester.ac.uk; re66@leicester.ac.uk; ss338@leicester.ac.uk;

Corresponding author: Professor Salman Siddiqui,

Leicester Biomedical Research Centre (Respiratory Theme)

College of Life Sciences, Department of Respiratory Sciences, University of Leicester

Glenfield Hospital,

Groby Road

Leicester

LE3 9QP

t: +44 116 256 3841

e: ss338@ le.ac.uk

Authors' affiliations:

[1]Leicester NIHR Biomedical Research Centre (Respiratory theme), Glenfield Hospital, Groby Road, Leicester LE3 9QP

[2] College of Life Sciences, Department Respiratory Sciences, University of Leicester, University Road, Leicester LE1 7RH

[3] Department of Chemistry, University of Leicester, University Road, Leicester, LE1 7RH

[4] See acknowledgments for contributors

***Equal contribution to work**

Abstract

Background: Asthma and COPD continue to cause considerable diagnostic and treatment stratification challenges. Volatile Organic Compounds (VOCs) have been proposed as feasible diagnostic and monitoring biomarkers in airways diseases.

Aims: To conduct a systematic review evaluating (i) the diagnostic accuracy of VOCs in diagnosing airways diseases, (ii) understand the relationship between reported VOCs and biomarkers of type-2 inflammation, (iii) assess the standardisation of reporting according to STARD and TRIPOD criteria, (iv) review current methods of breath sampling and analysis

Methods: A PRISMA-oriented systematic search was conducted (January 1997-December 2020). Search terms included: '*asthma', 'volatile organic compound(s)', 'VOC', and 'COPD'.* Two independent reviewers examined the extracted titles against review objectives.

Results: 44 full-text papers were included. 40/44 studies were cross-sectional and 4 studies were interventional in design. 17/44 studies used sensor-array technologies (e.g. eNose). Cross-study comparison was not possible across identified studies due to the heterogeneity in design. The commonest airways diseases differentiating VOCs belonged to carbonyl-containing classes (i.e. Aldehydes, Esters, and Ketones) and hydrocarbons (i.e. Alkanes and Alkenes). Although individual markers that are associated with clinical biomarkers of type-2 inflammation were recognised (i.e. Ethane and 3,7–Dimethylnonane for asthma and α-Methylstyrene and Decane for COPD), these were not consistently identified across studies. Only (3/44) reported following STARD or TRIPOD criteria for diagnostic accuracy and multivariate reporting respectively.

Conclusions: Breath VOCs show promise as diagnostic biomarkers of airways diseases and for type-2 inflammation profiling. However, future studies should focus on transparent reporting of diagnostic accuracy and multivariate models and continue to focus on chemical identification of volatile metabolites.

Trial registration: **(PROSPERO - CRD42019141718).**

Introduction

Asthma and Chronic Obstructive Pulmonary Disease (COPD) are two of the most common respiratory diseases affecting millions of lives globally [1]. COPD is the fifth leading cause of death worldwide and both diseases are ranked among the top 20 conditions causing disability globally [2].

Notwithstanding their pathological differences asthma and COPD often share several clinical and immuno-pathological features e.g. varying degrees of eosinophilic and neutrophilic airway inflammation and airflow limitation, posing both clinical and diagnostic challenges for clinicians. To address these challenges, various tools have been developed to aid the diagnosis, treatment stratification and monitoring of airway diseases and identify airway inflammation that is treatable, using blood and sputum eosinophils [3, 4].

Current non-invasive biomarkers of type-2 inflammation such as Fractional Exhaled Nitric Oxide (FeNO) and blood eosinophils are gaining traction in both asthma and COPD, with FeNO being used to guide initiation of inhaled steroids in primary care [5] and blood eosinophils used to stratify patients most likely to response to inhaled corticosteroids in COPD and eosinophil suppressing biologics therapies in severe asthma [6].

Despite these tools and availability of consensus guidelines, there remains considerable diagnostic challenges in airway diseases [7]. Approximately 30% of patients diagnosed with asthma in primary care may have an inaccurate diagnosis [8, 9] and biomarkers that stratify high cost therapies in severe asthma fail to identify 20-25% of patients that will fail therapy [10, 11]. Currently only FeNO is recommended for treatment titration in asthma [12].

In view of the potential limitations of current biomarkers in airways diseases there has been a search for tools and techniques leading to the development of new novel biomarkers for enhanced disease detection, treatment stratification and monitoring.

Over recent years there has been noticeable interest and growing body of evidence in using exhaled breath Volatile Organic Compounds (VOCs) and breath signatures in the diagnosis, phenotyping and monitoring of airways diseases [13-16]. Montuschi *et al* [17] demonstrated that FeNO has a higher diagnostic accuracy (95%) when combined with Electronic nose (eNose) in diagnosing airway inflammation in mild asthmatics, whilst Schleich *et al* [18] have recently demonstrated the utility of breath analysis with Gas Chromatography and Mass Spectrometry (GC-MS) in identifying airway cellular inflammatory profiles in moderate to severe asthma. Several studies have also identified the potential diagnostic validity of breath analysis in COPD, using eNose and GC-MS approaches [19-21].

The potential challenges of widespread use of exhaled breath analysis and barriers towards future clinical adoption include: (1) The need for biomarker replication studies including internal replication with training and test datasets and external replication, (2) Standardised reporting of chemical and analytical methodologies and statistical approaches such as multivariate classification models. (3) The need for comparative studies of diagnostic accuracy of point of care sensors such as eNOSE and mass spectrometry based approaches such as GC-MS.

Aims and Objectives

1. To conduct a systematic review examining the ability of exhaled breath volatiles in aiding diagnosis and monitoring of asthma and COPD.

2. To understand the relationship between reported VOCs and existing clinical measures of type-2 airway inflammation, namely fractional exhaled nitric oxide (FeNO) and blood and sputum eosinophils.

3. To examine the standardisation of reporting of VOC diagnostic accuracy studies, and studies involving multivariate modelling approaches using STARD and TRIPOD respectively.

4. To review the current methods used for breath sampling and analysis in the studies identified in objectives 1-3.

Methods

Data sources and search criteria

A PRISMA oriented systematic search was completed from January 1997 to December 2020 using the following evidence databases; (i) Cochrane library, (ii) Medline and (iii) EMBASE **(online supplementary material).**

Further details on methodology can be found in - PROSPERO - CRD42019141718.

The keywords and mesh terms included were: '*asthma', 'volatile organic compound(s)', 'exhaled breath', 'VOC', 'VOCs', 'electronic nose', 'eNose'*, '*chronic obstructive pulmonary disease', 'airflow limitation', 'Emphysema', 'COPD', and 'chronic bronchitis'.*

Study selection and data extraction

Published peer-reviewed full-text articles concerning clinical studies of asthma and COPD diagnosis through to VOC monitoring were assessed for eligibility.

The following study types were included: observational studies: cross-sectional, case–control and cohort, and randomised controlled trials. The references lists of included studies were scrutinised to identify further relevant studies.

The following evidence sources were excluded: case reports/series, expert opinions and conference abstracts.

The following inclusion criteria were used for study selection:

a) Human adult asthma and COPD studies
b) Objective evidence of asthma and COPD diagnosed by a clinician in primary or secondary care
c) VOCs and breath profiles measured in exhaled breath using online or offline technologies

The exclusion criteria were limited to studies involving paediatric study populations as COPD does not affect this demographic and the diagnostic criteria for asthma used in paediatric populations are often confounded by pre-school wheeze and other phenotypes that are difficult to map to adult asthma.

Method of analysis

The study design, populations and methodologies including sample collection and analysis, and identified VOCs were extracted from the studies.

Studies were qualitatively assessed based on their methodology and published results.

Review process

Two independent medical doctors and researchers (SN and WI) reviewed the titles and abstracts of the identified studies. The full text articles were rigorously screened against the inclusion/exclusion criteria.

Information regarding study design, setting, population, methodologies (including sample collection, VOC analysis techniques, targeted biomarkers, environmental air and VOC assessment) and outcomes were gathered.

Any disagreements in the study inclusion or exclusion and data extraction were resolved through consultation between the two reviewers. A third independent reviewer was consulted on unresolved disagreements.

Study quality assessment

The quality and methodological rigor of the reported studies were tested using the Standards for Reporting Studies of Diagnostic Accuracy (STARD) for studies identifying diagnostic accuracy VOCs and Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) for Studies evaluating more than one breath biomarker using multivariate approaches **(Online supplementary material, Table 1).**

STARD is an instrument aimed at researchers and editors with the purpose of evaluating the quality of diagnostic accuracy studies by checking their methodological rigor against a set of essential criteria [22].

TRIPOD is a 22-item checklist that explicitly covers the development and validation aspects of multivariate prediction models [23].

We also used a modified Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) to formally assess the quality of the studies included [24]. QUADAS is a tool specifically designed to assess the risk of bias and quality of diagnostic accuracy studies and systematic reviews. It is made up of four domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed for the risk of bias and 'patient selection', 'index test' and 'reference standard' domains are assessed for risk of applicability. For the purpose of this review the index test was considered diagnostic VOC exhaled breath analysis and the reference standard was the stated diagnostic criteria for diagnosing asthma and COPD following international guidelines. The QUADAS-2 was specifically modified to enhance the applicability to breath analysis studies; these changes include: inclusion of disease state and healthy controls and whether the index test has undergone further internal or external validation. Full details of changes made are included in **(Table B)** of supplementary material.

Results

Description of studies

Our search identified 477 unique titles through database searching (EMBASE: n = 335, Medline: n=133, Cochrane Central register of Controlled Trials (CENTRAL): n = 9).

Other sources including reference lists of papers found through the database search which were also included (n=4). Of these, 427 were removed after title and abstract review, leaving 54 full-text papers for screening. Following review of the full text articles 44 papers met our inclusion criteria and were included in the final review (**Figure 1**).

Overall characteristics of the studies

The publication dates ranged from 1997 to 2020, with nearly half of the identified studies conducted in the last 5 years.

Across the 44 studies identified, there were 1793 patients with asthma and 2212 patients with COPD, with varying degrees of severity. The sample size of the included studies was relatively small, ranging from 25 to 521 (median sample size was 58), further details can be found in **(Table 1)**. There was considerable heterogeneity between the studies in terms of the characteristics of the disease cohorts and diagnostic and sampling methods. 14 studies evaluated patients with asthma and healthy volunteers only and 21 studies evaluated patients with only COPD and healthy volunteers. Three studies had a combination of patients with asthma, COPD and healthy volunteers and a further 2 studies had a combination of a patients with COPD and lung cancer and healthy volunteers.

The majority of the studies examined (90%), were cross-sectional in design, with only four interventional studies identified **(Figure 2),** these include two studies [25, 26] that focused on assessing changes in VOC profile following asthma loss of control, using both GC-MS and eNose; Lazar *et al* [27] who demonstrated no significant changes to molecular breath profile following induced changes to airway calibre and Scarlata *et al* [28] who studied the changes in VOC profile following inhaled bronchodilators and steroids in newly diagnosed COPD patients. The majority of the studies identified (37/44) were conducted in the stable state, 3 asthma studies [25, 29, 30] and 4 COPD [31-34] studies examined VOCs during acute exacerbations.

Diagnostic accuracy and monitoring studies in asthma and COPD

Asthma diagnostic accuracy studies

Several groups attempted to explore the diagnostic potential of exhaled breath VOCs in differentiating asthmatics from healthy volunteers, Dragonieri *et al* [35] reported in a cross-sectional study that exhaled breath signatures can distinguish asthmatics from healthy volunteers but were unable to differentiate between asthma severities (mild and severe asthma). In contrast, Paredi *et al* showed that Ethane levels were found to be significantly higher in severe asthma compared to mild asthma [36].

Acute asthma exacerbation and monitoring studies

Only few studies have been conducted at the exacerbation state to monitor airway responses, pioneered by Olopade *et al* [29], who showed that the levels of Pentane were increased in acute asthma but following the acute episode the levels of Pentane decrease using GC and tedlar bag. Van der Schee *et al* [26] reported that eNose can detect patients with asthma and steroid response using loss of control methodologies with higher accuracy than clinical markers of airway inflammation; blood eosinophils and FeNO. Brinkman *et al* [25] followed on by demonstrating that monitoring exhaled breath metabolites by GC-MS and eNose can discriminate stable asthmatics from those with loss of control. Another study used unbiased clustering of eNose exhaled breath profiles and identified three phenotypes of severe asthma based on blood granulocytic counts [37]. However, none of the reported studies involving acute asthma patients incorporated internal and external replication models in their study designs.

Interventional asthma studies

These were limited to assessing asthma control using Prednisolone initiation and withdrawal [26]. The study highlighted promise for breath analysis to be used for treatment stratification in asthma. In particular, it demonstrated that using VOCs can identify asthmatics and predict treatment response to oral corticosteroids.

Furthermore, in a group of 10 stable asthmatics Lazar *et al* [27] used eNose to show that reduction in airway calibre acutely following methacholine provocation was not related to changes in breath molecular profiles in asthma. This suggests that breath biomarkers offer robust classification independently of baseline levels of airflow limitation.

COPD diagnostic accuracy studies

There are a few studies that have evaluated diagnostic accuracy in patients with COPD using breath analysis. Multivariate classification methods have been used to identify COPD patients from healthy controls using exhaled VOCs [38]. Similarly another study reported that using eNose, breath prints can distinguish between COPD and health controls [39], this study also used GC-MS and detected two VOCs (Decane, 6-ethyl-2-methyl-Decane) that were

discriminant for patients with a diagnosis of COPD. Rodriguez-Aguilar *et al* [40] demonstrated that COPD can be differentiated from healthy controls using ultrafast gas chromatography system equipped with an electronic nose detector (FGC eNose), with modest accuracy. 17 VOCs, detailed in tables (1 and 2), have driven this separation.

Acute COPD exacerbation and monitoring studies

The largest exacerbation study was carried out by Shafiek *et al* [31], who evaluated eNose-derived breath profiles of 93 hospitalised COPD patients, compared to their profiles during recovery (n=61). Exhaled breath prints were able to distinguish stable state and COPD exacerbations, with an accuracy of 70%, sensitivity of 72% and specificity of 67% (p=.068), this was particularly evident in the presence of airway infection or pneumonia. Similarly, Pizzini *et al* [32], evaluated COPD exacerbations and, using GC-MS, identified n-butane, 2-pentanone, cyclohexanone, and 4-heptanone as discriminatory biomarkers.

Van Velzen *et al* [34] recently carried out the first prospective exacerbation follow-up study using both GC-MS and eNose. Breath profiles of 68 patients were examined before, after and during a naturally occurring COPD exacerbation and breath profiles demonstrated correct classification of 71% for baseline vs. exacerbations and 78% for exacerbation vs. recovery.

Further to the aforementioned exacerbation studies, Gaugg *et al* showed that breath profiles, using a real-time breath analyser, differ between frequent and infrequent COPD exacerbators, suggesting the phenotype of frequent exacerbators is associated with a distinguishable exhaled metabolomics profile [33].

All exacerbation studies show a promising proof-of-principle data in using breath analysis to identify acute COPD exacerbations, notwithstanding the complexities associated with breath collection during an acute clinical event.

Interventional COPD studies

From the studies that we evaluated in COPD, we did not identify any studies incorporating intervention within their study design**.**

Current knowledge gaps and remaining challenges:

Exhaled breath VOCs quantified using both eNose and GC hyphenated technologies demonstrated promising diagnostic accuracy for both asthma and COPD. However, this review has identified several knowledge gaps, including the lack of specifically designed studies to assess: (1) the added value of breath biomarkers over existing diagnostic tools, (2) the usefulness of breath biomarkers in patients with high diagnostic uncertainty (3) the impact of key confounders on VOC diagnostic accuracy such as treatment intensity, ethnicity and gender as well as external factors such as diet and environment which are likely to influence metabolite concentrations in breath.

Large-scale, well designed, validation studies are needed to specifically address these issues in the context of diagnostic accuracy.

Association of exhaled breath VOCs with clinical measures of type-2 inflammation

Association of VOC with clinical measures of type-2 inflammation in asthma

Few studies explored the relationship between exhaled breath VOCs and clinical indices of type-2 inflammation, Montuschi *et al* [41] have shown that the diagnostic accuracy for asthma with eNose was higher than that of FeNO or FEV_1 alone but in combination with eNose the diagnostic performance was increased. Paredi *et al* [36] found elevated levels of ethane in patients with asthma, which was significantly correlated with Nitric oxide. Ibrahim *et al* [15] demonstrated that exhaled breath VOC profiles can identify patients with a clinically important level of sputum eosinophilia and differentiate between eosinophilic and non-eosinophilic subjects using GC-MS .

One of the largest studies in asthma to date evaluating airway inflammation was carried out by Schliech *et al* [18] where 521 patients with unselected asthma were recruited and VOC signatures identified biomarkers for eosinophilic and neutrophilic asthma quantified using induced sputum.

Details of the detected compounds can be found in (**Table C – online supplementary material**). One striking observation in the asthma studies was the consistent identification of

VOCs associated with lipid peroxidation in asthma associated with clinical biomarkers of type-2 inflammation. (**Figure 3**) is an explanatory figure that highlights potential mechanisms of metabolite generation in this context.

Association of VOC with clinical measures of type-2 inflammation in COPD

There were very limited studies which have explored the direct relationships between exhaled breath VOCs and clinical markers of type-2 inflammation in COPD. Basanta *et al* [21] evaluated sputum eosinophilia in COPD subtypes and identified VOCs that are associated with varying degrees of sputum eosinophilia in COPD **(Table 1).**

Current knowledge gaps and remaining challenges:

Exhaled breath VOCs were able to identify biomarkers of airway type-2 inflammation. However, the variability in study design across studies and lack of external replication hindered the identification of robust breath VOC multi-marker signatures in this context. To move the field forward, breath analysis needs to be conducted in well-designed clinical airway inflammatory studies with varying designs e.g. inhaled corticosteroid step-down studies, T2 biologic therapy trials, with further evaluation of its utility as a baseline stratification tool as well as a response assessment tool

Reporting of VOC diagnostic accuracy studies, and studies involving multivariate modelling approaches using STARD and TRIPOD criteria

Of the 44 included studies, only three studies reported using the STARD guidelines for testing the diagnostic accuracy of the discovered exhaled breath signatures [18, 39, 42]. Similarly, none of the studies reported their multivariate models according to TRIPOD guidance. 22/44 studies reported markers of model accuracy combined with internal or model replication and only 2 studies [42, 43] reported external model replication in independent populations.

Assessment of risk of bias and applicability using QUADAS-2, detailed by study, can be found in **(Figure 4)** and **(Table A**, **Figure A)** of the online supplementary material. The main sources of bias were (i) Index test; 19/44 of the published studies failed to validate their

results internally or externally, increasing their risk of bias. (ii) Patient selection; 12/44 studies appeared to have high or unclear concerns due to the absence of a comparable control group (iii) Reference standard; 4/44 studies failed to report diagnostic standards according to internationally recognised guidelines (i.e. GINA for asthma, GOLD for COPD). No major concerns regarding applicability were highlighted.

Current knowledge gaps and remaining challenges:

This review highlighted notable shortcomings in transparent reporting following STARD and TRIPOD criteria. It is worth noting that non-compliance with reporting guidelines does not necessarily translate to poor methodological set-up or invalidity of the published results, only that they failed to explicitly report them therefore making it difficult for readers to assess both methodological rigor and result validity. We believe it is necessary for future exhaled VOC diagnostic studies to embed STARD criteria within their reports and for clinical and metabolomics journals to mandate these to improve future compliance, enhance adjudication and replication of discovered VOCs.

Review of current methods of breath sampling and analysis

Breath sampling

There is currently no reported standardised protocol for breath sample collection. Among the methods reported, Tedlar bag was the most frequently described method (59%) **(Figure 2)**. 15 studies described direct connection of Tedlar bag to eNose sensors and 10 studies transferred the samples onto a thermal desorption (TD) tube as a medium prior to analysis.

18 (41%) of the reported studies avoided the use of breath receptacles and sampling was carried out through direct connection to analytical instruments; for example studies using multi-capillary column-ion mobility spectrometry (MCC/IMS) [44-46] advised participants to breath directly, via a mouthpiece, to the analytical instrument. Other direct collection devices included BioVOC collectors [38, 47, 48], PneumoPipe® [49] and EBC collectors (RTube ™) [50].

Breath VOC quantification

Various technologies were used to collect and analyse VOCs. This includes eNose (19 studies), GC; either combined with MS, FID or DMS (13 studies), GC-ToF-MS (6 studies), GCxGC-MS (1 study), IMS (5 studies), PTR-TOF-MS (1 study) **(Table 1).**

Current knowledge gaps and remaining challenges:

The review highlighted a wide range of breath collection and analysis technologies, with promising advances in sampling such as the development of the breath gated ReCIVA device [51, 52] and novel analytical methodologies applied to breath, such as GCxGC-MS that may offer superior sensitivity in volatile metabolite identification [18]. Standardised protocols will need to be carefully designed to accommodate the emerging methodologies.

Reported Volatile Organic Compounds

The reported VOCs from the studies evaluated are described in **(Table C – online supplementary material).** Not all studies reported the actual names of the VOCs that were identified. 21/44 studies have reported named VOCs. The most commonly identified VOCs belonged to carbonyl-containing compounds (i.e. Aldehydes, Esters, and Ketones) and hydrocarbons (i.e. Alkanes, Alkenes and monoaromatics), commonly resulting from lipid peroxidation, the free radical induced oxidative degradation of polyunsaturated fatty acids [53].

Several compounds have been reported more than once, Nonanal was first described by Basanta *et al* [54] as a potential COPD identifier and was later found to discriminate between smokers and former smokers (with and without COPD) and never smokers by Jareno-Esteban *et al* [47]. Interestingly, earlier this year Schliech *et al* [18] reported that Nonanal had the ability to classify stable asthmatics based on their sputum granulocytic cell count and was found to be higher in neutrophilic asthma. Hexanal is another aldehyde that was reported by at least three studies as a potential COPD biomarker [21, 38, 47].

Seven studies reported VOCs of alcohol and phenol classifications [15, 18, 25, 38, 39, 54, 55] with poor consistency among reported compounds. Other chemical classifications identified in small numbers included sulphides, furans, quinones and nitriles.

The majority of studies named single VOCs as potential discriminators, however some studies reported the ability of a combination of VOCs to discriminate between disease types or subgroups of disease [38].

Discussion

We have provided an overview of all published studies on breath analysis in two major airways diseases; asthma and COPD, focusing on the diagnostic accuracy of VOCs, their relationship to clinical measures of type-2 inflammation, and the reporting rigour of diagnostic accuracy and multivariate methods according to STARD and TRIPOD criteria.

The lack of standardisation and high degree of methodological heterogeneity made it impractical to conduct a meta-analysis, an issue not unique to this review [56-59]. Recently a study by Henderson *et al* [60] evaluated a standardised experiment to allow for comparisons of breath sampling and methodologies. This is a promising step forward in the development of tools and techniques to achieve standardised data capture of VOCs in breath.

Only 3 of the 28 diagnostic accuracy studies reported their results following the STARD criteria. First published in 2003, STARD guidelines were developed to tackle deficiencies in reporting research and assist in the completeness and transparency of reporting diagnostic accuracy studies. Furthermore, it provides researchers with means to critically appraise published results against potential bias [61]. None of the studies reported multivariate models using TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) criteria. TRIPOD is an essential 22 item checklist aimed at improving studies reporting development or validation of diagnostic or prognostic multivariate models [62]. Only with transparent reporting, following a set model, can the strengths and weaknesses of discovery studies be brought to light and therefore facilitating interpretation and future adoption. Large scale, real world, multi-platform breath discovery

studies, following the aforementioned construct, such as [63] are currently underway and further results are eagerly anticipated.

In this review, the QUADAS-2 tool highlighted clear bias including patient selection; 9 studies displayed high concerns with regards to their patient selection criteria, this is mainly due to the absence of comparable positive controls and inappropriate exclusion of certain sub-groups such as patients with comorbidities, recent exacerbations or patients on steroids. Whilst these exclusions are sometimes necessary, removing these groups can result in falsely positive discriminatory signals. Another source of concern was regarding the reference standard diagnosis, at least 4 studies have failed to report clear diagnostic source of airway diseases following internationally recognised guidelines e.g. global initiative for asthma (GINA) and global initiative for COPD (GOLD). The highest risk of bias however, has been that of the index test, (22/44) studies did not clearly outline an approach to validate the VOC analysis; ideally, an internal replication in an independent cohort for single centre studies and external replication for multi-centre studies to ensure the possibility of an accurate and reliable generalisation of the results [64]. The absence of external validation could be the main factor behind the lack of similarity between the published results across studies.

In our current review, the use of Tedlar bags were the commonest method of breath sampling. Although validated, this method is not without shortcomings. Beauchamp *et al* [65] described several potential limitations including; (i) decay of certain groups of compounds which significantly limits sample storage time, (ii) compound contamination, either through compound diffusion such as acetonitrile through the bag wall or contamination of the bag wall with previous substances and all have substantial effects on the limit of detection. Beauchamp *et al* went on to identify the groups of compounds that showed significant signal reduction with time. These include hexanals, acetonitrile and water vapour. The demonstrated concentration-dependent decay translates to time-consuming consideration of factors such as cleaning procedures, sampling procedures and storage procedures when using Tedlar bags. These problems have been mitigated by the use of direct breath sampling devices, these allow collection directly onto sorbent tubes with pre-concentration of VOCs. It has been previously described that samples collected onto sorbent tubes and appropriately stored, show negligible change in VOC profile concentration [66].

We examined the various technologies used for breath quantification and analysis. The discovery and development journey of any new technology is rarely a simple linear

progression but rather complicated by a series of barriers and interactive feedback loops, and VOC discovery technologies are no exception. The development pathway of any new technology is frequently described in terms of Technology readiness levels (TRL), with common reference points (1-9) to describe the level of development the technology has achieved. Various breath quantification and analysis techniques have emerged over the years and are at different stages of development with varying numbers of clinical studies **(Figure 5).**

Numerous validated studies using eNose have emerged over the last decade [67-71]. It is still debatable whether individual VOCs are crucial for establishing a diagnosis when a breath print can adequately do so. eNose breath signatures, however, are less rigorous as they can be heavily influenced by environmental factors and are unable to produce large identical reliable breath profiles [72]. VOCs that can be identified from a cellular and enzymatic metabolism pathways on the other hand, can be used to further understand complex underlying biological processes, which in turn can guide new therapeutic targets. An example for this is the practice-changing ^{14}C-Urea breath test used in diagnosing H. Pylori, which relies on the ability of H. Pylori bacteria within the gut that have functional urease activity to convert ^{14}C to CO_2 [73]. In a similar manner exogenous VOCs and labelled compounds may be used to identify specific cellular or microbial activity in airways diseases in the future, if a thorough understanding of the enzymatic pathways involved are developed.

In summary, we have demonstrated that exhaled VOCs show promise as potential non-invasive biomarkers for airways diseases, and correlate with clinical biomarkers of type-2 inflammation. Many of the biomarkers reviewed in our study appear to be associated with lipid peroxidation in the airway suggesting a role for VOC as diagnostic biomarkers and markers of oxidant induced airway inflammation.

These observations highlight a conceivable role for VOCs as diagnostic and monitoring tools for airways diseases. Advancing the field has been hindered by numerous challenges, as highlighted by this review, including the lack of standardisation and transparent reporting. Well-designed clinical studies are strongly warranted to further advance breath volatile biomarkers toward clinical adoption.

References

1. *Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015.* Lancet Respir Med, 2017. **5**(9): p. 691-706.
2. Soriano, J.B., et al., *Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015.* The Lancet Respiratory Medicine, 2017. **5**(9): p. 691-706.
3. Lange, P., C.S. Ulrik, and O. Hilberg, *[Fraction of exhaled nitric oxide as biomarker of asthmatic airway inflammation].* Ugeskr Laeger, 2019. **181**(19).
4. Wells, C.D. and M.J. Joo, *COPD and asthma: Diagnostic accuracy requires spirometry.* J Fam Pract, 2019. **68**(2): p. 76-81.
5. Price, D., et al., *Using fractional exhaled nitric oxide (FeNO) to diagnose steroid-responsive disease and guide asthma management in routine care.* Clinical and translational allergy, 2013. **3**(1): p. 37-37.
6. Kostikas, K., C. Brindicci, and F. Patalano, *Blood Eosinophils as Biomarkers to Drive Treatment Choices in Asthma and COPD.* Current drug targets, 2018. **19**(16): p. 1882-1896.
7. Tinkelman, D.G., et al., *Misdiagnosis of COPD and asthma in primary care patients 40 years of age and over.* J Asthma, 2006. **43**(1): p. 75-80.
8. Magnoni, M.S., et al., *Asthma under/misdiagnosis in primary care setting: an observational community-based study in Italy.* Clinical and molecular allergy : CMA, 2015. **13**: p. 26-26.
9. Kavanagh, J., D.J. Jackson, and B.D. Kent, *Over- and under-diagnosis in asthma.* Breathe, 2019. **15**(1): p. e20.
10. Kim, H., et al., *Asthma biomarkers in the age of biologics.* Allergy, Asthma & Clinical Immunology, 2017. **13**(1): p. 48.
11. Pavord, I.D., et al., *The current and future role of biomarkers in type 2 cytokine-mediated asthma management.* Clin Exp Allergy, 2017. **47**(2): p. 148-160.
12. Excellence, N.I.f.H.a.C., *Asthma: diagnosis, monitoring and chronic asthma management (NICE guidleines NG80)*. 2020: Available at https://www.nice.org.uk/guidance/ng80.
13. Besa, V., et al., *Volatile organic compounds (VOCs) in COPD patients with exacerbation.* European Respiratory Journal, 2011. **38**.
14. Esteban, J.J., et al., *Volatile organic compounds (VOC) in exhaled breath in patients with lung cancer, using the analytical technique thermal desorber-gase chromatography-spectrometer mases.* European Respiratory Journal, 2012. **40**.
15. Ibrahim, B., et al., *Non-invasive phenotyping using exhaled volatile organic compounds in asthma.* Thorax, 2011. **66**(9): p. 804-9.
16. Samara, M.A., et al., *Single breath exhaled gas analysis by selected ion-flow tube mass-spectrometry (SIFT-MS) in patients with acute decompensated heart failure.* Journal of Cardiac Failure, 2011. **17**(8).
17. Montuschi, P., et al., *Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma.* Chest, 2010. **137**(4): p. 790-796.
18. Schleich, F.N., et al., *Exhaled Volatile Organic Compounds are Able to Discriminate between Neutrophilic and Eosinophilic Asthma.* Am J Respir Crit Care Med, 2019.
19. Van Berkel, J.J.B.N., et al., *A profile of volatile organic compounds in breath discriminates COPD patients from controls.* Respiratory Medicine, 2010. **104**(4): p. 557-563.
20. Dragonieri, S., et al., *An electronic nose in the discrimination of patients with non-small cell lung cancer and COPD.* Lung Cancer, 2009. **64**(2): p. 166-170.
21. Basanta, M., et al., *Exhaled volatile organic compounds for phenotyping chronic obstructive pulmonary disease: a cross-sectional study.* Respir Res, 2012. **13**: p. 72.

22. Cohen, J.F., et al., *STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration.* BMJ Open, 2016. **6**(11): p. e012799.
23. Moons, K.G.M., et al., *Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and ElaborationThe TRIPOD Statement: Explanation and Elaboration.* Annals of Internal Medicine, 2015. **162**(1): p. W1-W73.
24. Whiting, P.F., et al., *QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.* Ann Intern Med, 2011. **155**(8): p. 529-36.
25. Brinkman, P., et al., *Exhaled breath profiles in the monitoring of loss of control and clinical recovery in asthma.* Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology, 2017. **47**(9): p. 1159-1169.
26. Van Der Schee, M.P., et al., *Predicting steroid responsiveness in patients with asthma using the electronic nose.* American Journal of Respiratory and Critical Care Medicine, 2012. **185**.
27. Lazar, Z., et al., *Electronic nose breathprints are independent of acute changes in airway caliber in asthma.* Sensors (Basel, Switzerland), 2010. **10**(10): p. 9127-9138.
28. Scarlata, S., et al., *Cluster analysis on breath print of newly diagnosed COPD patients: effects of therapy.* Journal of breath research, 2018. **12**(3): p. 036022.
29. Olopade, C.O., et al., *Exhaled pentane levels in acute asthma.* Chest, 1997. **111**(4): p. 862-5.
30. van der Schee, M.P., et al., *Predicting steroid responsiveness in patients with asthma using exhaled breath profiling.* Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology, 2013. **43**(11): p. 1217-1225.
31. Shafiek, H., et al., *Using the Electronic Nose to Identify Airway Infection during COPD Exacerbations.* PLoS One, 2015. **10**(9): p. e0135199.
32. Pizzini, A., et al., *Analysis of volatile organic compounds in the breath of patients with stable or acute exacerbation of chronic obstructive pulmonary disease.* Journal of breath research, 2018. **12**(3): p. 036002.
33. Gaugg, M.T., et al., *Real-Time Breath Analysis Reveals Specific Metabolic Signatures of COPD Exacerbations.* Chest, 2019. **156**(2): p. 269-276.
34. van Velzen, P., et al., *Exhaled Breath Profiles Before, During and After Exacerbation of COPD: A Prospective Follow-Up Study.* COPD, 2019. **16**(5-6): p. 330-337.
35. Dragonieri, S., et al., *An electronic nose in the discrimination of patients with asthma and controls.* J Allergy Clin Immunol, 2007. **120**(4): p. 856-62.
36. Paredi, P., S.A. Kharitonov, and P.J. Barnes, *Elevation of exhaled ethane concentration in asthma.* Am J Respir Crit Care Med, 2000. **162**(4 Pt 1): p. 1450-4.
37. Brinkman, P., 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.
38. Phillips, C.O., et al., *Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls.* J Breath Res, 2012. **6**(3): p. 036003.
39. Cazzola, M., et al., *Analysis of exhaled breath fingerprints and volatile organic compounds in COPD.* COPD Research and Practice, 2015. **1**(1).
40. Rodríguez-Aguilar, M., et al., *Ultrafast gas chromatography coupled to electronic nose to identify volatile biomarkers in exhaled breath from chronic obstructive pulmonary disease patients: A pilot study.* Biomedical chromatography : BMC, 2019. **33**(12): p. e4684.
41. Montuschi, P., et al., *Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma.* Chest, 2010. **137**(4): p. 790-6.
42. Fens, N., 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 and experimental allergy : journal of the British Society for Allergy and Clinical Immunology, 2011. **41**(10): p. 1371-1378.

43. Dragonieri, S., et al., *Exhaled breath profiling by electronic nose enabled discrimination of allergic rhinitis and extrinsic asthma.* Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals, 2018: p. 1-24.
44. Besa, V., et al., *Exhaled volatile organic compounds discriminate patients with chronic obstructive pulmonary disease from healthy subjects.* International Journal of Chronic Obstructive Pulmonary Disease, 2015. **10**: p. 399-406.
45. Westhoff, M., et al., *Statistical and bioinformatical methods to differentiate chronic obstructive pulmonary disease (COPD) including lung cancer from healthy control by breath analysis using ion mobility spectrometry.* International Journal for Ion Mobility Spectrometry, 2011. **14**(4): p. 139-149.
46. Westhoff, M., et al., *Differentiation of chronic obstructive pulmonary disease (COPD) including lung cancer from healthy control group by breath analysis using ion mobility spectrometry.* International Journal for Ion Mobility Spectrometry, 2010. **13**(3-4): p. 131-139.
47. Jareno-Esteban, J.J., et al., *Study of 5 Volatile Organic Compounds in Exhaled Breath in Chronic Obstructive Pulmonary Disease.* Archivos De Bronconeumologia, 2017. **53**(5): p. 251-256.
48. Phillips, C., et al., *Short-Term Intra-Subject Variation in Exhaled Volatile Organic Compounds (VOCs) in COPD Patients and Healthy Controls and Its Effect on Disease Classification.* Metabolites, 2014. **4**(2): p. 300-318.
49. Scarlata, S., et al., *Cluster analysis on breath print of newly diagnosed COPD patients: effects of therapy.* J Breath Res, 2018. **12**(3): p. 036022.
50. Schivo, M., et al., *Mobile instrumentation platform to distinguish airway disorders.* Journal of breath research, 2013. **7**(1): p. 017113-017113.
51. Kitchen, S., et al., *LATE-BREAKING ABSTRACT: Breathe free: Open source development of a breath sampler by a consortium of breath researchers.* European Respiratory Journal, 2015. **46**(suppl 59).
52. Holden, K.A., et al., *Use of the ReCIVA device in breath sampling of patients with acute breathlessness: a feasibility study.* ERJ Open Research, 2020. **6**(4): p. 00119-2020.
53. Gardner, H.W., *Oxygen radical chemistry of polyunsaturated fatty acids.* Free Radic Biol Med, 1989. **7**(1): p. 65-86.
54. Basanta, M., et al., *Exhaled volatile organic compounds for phenotyping chronic obstructive pulmonary disease: a cross-sectional study.* Respiratory research, 2012. **13**: p. 72.
55. Gaida, A., et al., *A dual center study to compare breath volatile organic compounds from smokers and non-smokers with and without COPD.* Journal of breath research, 2016. **10**(2): p. 026006.
56. Azim, A., et al., *Exhaled volatile organic compounds in adult asthma: a systematic review.* Eur Respir J, 2019. **54**(3).
57. Catino, A., et al., *Breath Analysis: A Systematic Review of Volatile Organic Compounds (VOCs) in Diagnostic and Therapeutic Management of Pleural Mesothelioma.* Cancers (Basel), 2019. **11**(6).
58. Christiansen, A., et al., *A systematic review of breath analysis and detection of volatile organic compounds in COPD.* J Breath Res, 2016. **10**(3): p. 034002.
59. Töreyin, Z.N., et al., *Exhaled Breath Analysis in Diagnosis of Malignant Pleural Mesothelioma: Systematic Review.* Int J Environ Res Public Health, 2020. **17**(3).
60. Henderson, B., et al., *A benchmarking protocol for breath analysis: the peppermint experiment.* Journal of Breath Research, 2020. **14**(4): p. 046008.
61. Bossuyt, P.M., et al., *STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies.* BMJ, 2015. **351**: p. h5527.
62. Collins, G.S., et al., *Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement.* Br J Surg, 2015. **102**(3): p. 148-58.

63. Ibrahim, W., et al., *Assessment of breath volatile organic compounds in acute cardiorespiratory breathlessness: a protocol describing a prospective real-world observational study.* BMJ open, 2019. **9**(3): p. e025486.
64. Fijten, R.R.R., et al., *The necessity of external validation in exhaled breath research: a case study of sarcoidosis.* Journal of Breath Research, 2017. **12**(1): p. 016004.
65. Beauchamp, J., et al., *On the use of Tedlar(R) bags for breath-gas sampling and analysis.* J Breath Res, 2008. **2**(4): p. 046001.
66. Harshman, S.W., et al., *Storage stability of exhaled breath on Tenax TA.* J Breath Res, 2016. **10**(4): p. 046008.
67. Fens, N., et al., *Exhaled molecular patterns change after experimental rhinovirus 16 infection in asthma.* European Respiratory Journal, 2011. **38**.
68. Hattesohl, A.D.M., et al., *Discrimination between COPD patients with and without alpha 1-antitrypsin deficiency using an electronic nose.* Respirology (Carlton, Vic.), 2011. **16**(8): p. 1258-1264.
69. Scarlata, S., et al., *Effects of topical therapy on the breath pattern of patients with newly diagnosed COPD: An e-nose based study.* European Respiratory Journal, 2015. **46**.
70. Shafiek, H., et al., *Using the Electronic Nose to Identify Airway Infection during COPD Exacerbations.* PloS one, 2015. **10**(9): p. e0135199.
71. Finamore, P., et al., *Validation of exhaled volatile organic compounds analysis using electronic nose as index of COPD severity.* International journal of chronic obstructive pulmonary disease, 2018. **13**: p. 1441-1448.
72. Fens, N., et al., *Exhaled breath analysis by electronic nose in airways disease. Established issues and key questions.* Clin Exp Allergy, 2013. **43**(7): p. 705-15.
73. Hawtin, P.R., *Serology and urea breath test in the diagnosis of H. pylori infection.* Mol Biotechnol, 1999. **11**(1): p. 85-92.
74. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.
75. Larstad, M.A., et al., *Determination of ethane, pentane and isoprene in exhaled air--effects of breath-holding, flow rate and purified air.* Acta Physiol (Oxf), 2007. **189**(1): p. 87-98.
76. Dragonieri, S., et al., *An electronic nose in the discrimination of patients with asthma and controls.* The Journal of allergy and clinical immunology, 2007. **120**(4): p. 856-862.
77. Fens, N., 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): p. 1076-1082.
78. Timms, C., P.S. Thomas, and D.H. Yates, *Detection of gastro-oesophageal reflux disease (GORD) in patients with obstructive lung disease using exhaled breath profiling.* J Breath Res, 2012. **6**(1): p. 016003.
79. Schivo, M., et al., *A mobile instrumentation platform to distinguish airway disorders.* Journal of breath research, 2013. **7**(1): p. 017113.
80. Meyer, N., et al., *Defining adult asthma endotypes by clinical features and patterns of volatile organic compounds in exhaled air.* Respiratory research, 2014. **15**: p. 136.
81. Plaza, V., et al., *Inflammatory Asthma Phenotype Discrimination Using an Electronic Nose Breath Analyzer.* Journal of investigational allergology & clinical immunology, 2015. **25**(6): p. 431-437.
82. de Vries, R., et al., *Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label.* Eur Respir J, 2018. **51**(1).
83. Brinkman, P., et al., *Exhaled Volatile Organic Compounds as Markers for Medication Use in Asthma.* European Respiratory Journal, 2019: p. 1900544.
84. Basanta, M., et al., *Non-invasive metabolomic analysis of breath using differential mobility spectrometry in patients with chronic obstructive pulmonary disease and healthy smokers.* Analyst, 2010. **135**(2): p. 315-20.

85. Cristescu, S.M., et al., *Screening for emphysema via exhaled volatile organic compounds.* Journal of breath research, 2011. **5**(4): p. 046009.
86. Martinez-Lozano Sinues, P., et al., *Breath analysis in real time by mass spectrometry in chronic obstructive pulmonary disease.* Respiration, 2014. **87**(4): p. 301-10.
87. Allers, M., et al., *Measurement of exhaled volatile organic compounds from patients with chronic obstructive pulmonary disease (COPD) using closed gas loop GC-IMS and GC-APCI-MS.* Journal of breath research, 2016. **10**(2): p. 026004.
88. Bregy, L., et al., *Real-time mass spectrometric identification of metabolites characteristic of chronic obstructive pulmonary disease in exhaled breath.* Clinical Mass Spectrometry, 2018. **7**: p. 29-35.
89. Incalzi, R.A., et al., *Reproducibility and Respiratory Function Correlates of Exhaled Breath Fingerprint in Chronic Obstructive Pulmonary Disease.* PLOS ONE, 2012. **7**(10): p. e45396.

Tables

Table 1: Table outlining included studies and summary of findings in relation to the review objectives.

Legend: ¥: Studies including both asthma and COPD. Abbreviations: COPD = chronic obstructive pulmonary disease, GC-MS = gas chromatography & mass spectrometry, GC-FID = gas chromatography flame ionisation detector, eNose = electronic nose, GORD = gastro-oesophageal reflux disease, EBC = exhaled breath condensate, GC-DMS = Differential Mobility Spectrometry, GC-TOF-MS = Gas Chromatography – Time-of-Flight Mass Spectrometry, GCxGC-MS = comprehensive two dimensional gas chromatography & mass spectrometry, FeNO = fractional exhaled nitric oxide, MCC/IMS = multi-capillary column-ion mobility spectrometer, PTR-MS = proton-transfer-reaction mass spectrometry, APCI-CMS = atmospheric-pressure chemical ionisation mass spectrometry, SESI – HRMS = Secondary Electrospray Ionisation High-resolution Mass Spectrometry.

Figure legends

Figure 1. PRISMA flow chart illustrating article selection. (Modified from Moher *et al)[74]*

Figure 2: Distribution of study designs, breath collection and analysis technologies across the identified studies.

Figure 3. Top Panel: The precise cellular metabolic processes that underpin the majority of reported VOC described in breath using GC-MS have yet to be determined. **Bottom Panel:** Simplified illustration of the different causes of lipid peroxidation and the resultant formation of aldehydes and alkanes as well as other VOC classes, which have been consistently described across several breath volatile association studies in asthma and COPD (**Tables 1 and 2).**

Figure 4: Risk of bias and applicability concerns using QUADAS-2 tool. Green = low risk, yellow = unclear, red = high concern.

Figure 5: Qualitative assessment of breath sampling and analytical technologies level on a nine-point technology readiness level (TRL) scale (*adapted from TRL guidance published by the Nuclear Decommissioning Authority NDA*), comparative to exiting biomarkers used in clinical practice for airway disease stratification.

Acknowledgments

This research was funded by the Medical Research Council (MRC), Engineering and Physical Sciences Research Council (EPSRC) Stratified Medicine Grant for Molecular Pathology Nodes (Grant No. MR/N005880/1), Midlands Asthma and Allergy Research

Association (MAARA) and British Lung Foundation (Grant No. BLFPHD17-1). The work was carried out at the University Hospitals of Leicester NHS Trust, University of Leicester and Loughborough University, supported by the NIHR Leicester Biomedical Research Centre and the NIHR Leicester Clinical Research Facility. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The authors would like to acknowledge the invaluable efforts of Dr. Pip Divall and the wider East Midlands Breathomics Pathology Node consortium (members list can be found at: https://ember.le.ac.uk/web).

The authors declare no conflicts of interest in preparing this article

Asthma

Title	Author	Sample size (n)	Breath Collection and analysis	Summary of findings in relation to review objectives
Exhaled Pentane Levels in Acute Asthma	Olopade *et al* (1997) [29]	**40**	Tedlar bag (GC-FID)	• Pentane used as a diagnostic marker in asthma. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Elevation of Exhaled Ethane Concentration in Asthma	Paredi *et al* (2000) [36]	**40**	Tedlar bag (GC-FID)	• Ethane showed potential as a diagnostic marker in asthma Levels significantly correlated with NO and air trapping (as measured by RV/TLC). Ethane and NO were reduced in steroid-treated patients. • No reported use of STARD/TRIPOD
Determination of ethane, pentane and isoprene in exhaled air – effects of breath-holding, flow rate and purified air	Larstad *et al* (2007) [75]	**27**	Tedlar bag (GC-FID)	• Ethane concentrations were slightly flow dependent in subjects with asthma. Isoprene levels were significantly lower in asthmatics with marked increase after breath-holding. Not used in diagnostic accuracy context • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
An electronic nose in the discrimination of patients with asthma and controls	Dragonieri *et al* (2007) [76]	**40**	Tedlar bag (eNose)	• eNose successfully diagnosed mild asthmatics (cross-validation of 100% correct with M-distance of 5.32) with comparable differentiation in severe asthmatics. Unable to discriminate mild from severe asthma. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Exhaled Breath Profiling Enables Discrimination of Chronic Obstructive Pulmonary Disease and Asthma	Fens *et al* (2009) [77]	**90**	Tedlar bag (eNose)	• eNose successfully discriminated asthmatics from COPD patients (cross validated accuracy of 96%, $p < 0.0001$). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Non-invasive phenotyping using exhaled volatile organic compounds in asthma	Ibrahim *et al* (2011) [15]	**58**	Direct breath sampler (GC-MS)	• 47 compounds diagnosed asthma (86% accuracy - PPV 0.85, NPV 0.89). 13 compounds discriminated uncontrolled asthmatics (ACQ ≥1) with 89% accuracy (AUC 0.90) (**Pentadecane, Heptanoic acid, O-xylene, 2-Butanone, 3-methyl/butanal, 2,6-diisopropylnaphtalene**) • 11 compounds were able to discriminate eosinophilic asthmatics with 83% classification accuracy (AUC 0.98) (**Camphene, Cyclohexanone, Cyclohexene-4-methylene**). 14

				compounds discriminated neutrophilic asthmatics wit 72% accuracy (AUC 0.90) (**Cyclopentene, Naphthalene, Cyclohexanol, Tetradecane, Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene**). • No reported use of STARD/TRIPOD
Detection of gastro-oesophageal reflux disease (GORD) in patients with obstructive lung disease using exhaled breath profiling	Timms *et al* (2012) [78]	**44**	Tedlar bag (eNose)	• eNose diagnosed asthma patients with GORD (p = 0.015, accuracy 85%, interclass m distance > 2.8). Weak statistically significant difference between COPD and COPD with GORD (p < 0.05, accuracy 64.7%). Significant difference distinguishing controls from COPD (interclass M distance 3.601, p < 0.01) and COPD with GORD (interclass M distance 2.974, p < 0.01). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Predicting steroid responsiveness in patients with asthma using exhaled breath profiling	Van der Schee *et al* (2012) [57]	**45**	Tedlar bag (eNose)	• eNose successfully diagnosed asthma patients (AUC: 0.766; p = 0.002), with maintained discrimination after prednisolone administration (AUC = 0.842 p < 0.001). eNose also predicted responsiveness to subsequent treatment with oral prednisone (AUC=0.883; p=0.008). • Unlike FeNO, both eNose and sputum eosinophils were able to distinguish loss of control (AUC 0.81; p= 0.008 and 0.868; p < 0.002 respectively). Additionally, eNose showed a strong correlation with the percentage of sputum eosinophils (r = 0.601, P = 0.002).FeNO was not associated with the VOC- profile (r = 0.141, p = 0.502). • No reported use of STARD/TRIPOD
A mobile instrumentation platform to distinguish airway disorders	Schivo *et al* (2013) [79]	**31**	EBC collector (GC-DMS - FAIMS)	• VOCs classified asthmatics **75%** of the time, after executing 20 classification optimization loops and discriminated subjects taking omalizumab from subjects not taking this medication **70%** of the time after executing 40 loops. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Defining adult asthma endotypes by clinical features and patterns of volatile organic compounds in exhaled air	Meyer *et al* (2014) [80]	**235**	Tedlar bag (GC-ToF-MS)	• 16 unidentified VOCs diagnosed asthmatics with 100% sensitivity and 91.1% specificity. Cluster analysis based on VOCs and the clinical parameters resulted in 7 different asthma endotype clusters.

				• No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Exhaled volatile organic compounds are able to discriminate between neutrophilic and eosinophilic asthma	Schliech *et al* (2019) [18]	**521**	Tedlar bag (GC-MS, GCxGC-MS)	• VOCs discriminated asthma inflammatory phenotypes. Discovery study = **3- tetradecene** and **pentadecene** distinguished between neutrophilic and paucigranulocytic phenotypes (AUC of 0.85). Replication = **undecane** and **nonanal** discriminated neutrophilic phenotype (AUC of 0.70) • Discovery = **3, 7 –Dimethylnonane**, **nonanal** and **1-propano**l discriminated neutrophilic phenotype compared to eosinophilic (AUC 0.92). Replication = **Hexane** and **nonanal** discriminated neutrophilic from eosinophilic phenotypes (AUC 0.71) • Combining VOCs (**hexane** and **2-hexanone**), blood eosinophil and FeNO demonstrated the highest specificity for predicting a sputum eosinophilia (AUC 0.87). **Nonanal, 1-propanol**, and **hexane** identified neutrophilic inflammation compared to other phenotypes (AUC 0.73) • Reported use of STARD criteria
Identification and prospective stability of electronic nose (eNose)-derived inflammatory phenotypes in patients with severe asthma	Brinkman *et al* (2019) [37]	**78**	Tedlar bag (eNose)	• eNose identified 3 phenotypes of severe asthma based on their blood granulocytic count • No reported use of STARD/TRIPOD
Exhaled breath profiling by electronic nose enabled discrimination of allergic rhinitis and extrinsic asthma.	Dragonieri *et al* (2018) [43]	**42**	Tedlar bag (eNose)	• eNose diagnosed allergic rhinitis with and without extrinsic asthma, CVA = of 85.7% ($p < 0.01$), AUC: 0.93. Breathprints of AAR differed from those of controls with CVA% of 75.0 ($p < 0.05$) with an AUC of 0.87. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Exhaled breath profiles in the monitoring of loss of control and clinical recover y in asthma	Brinkman *et al* (2017) [25]	**23**	Tedlar bag (GC-MS, eNose)	• eNose and GC-MS VOCs (**Methanol, Acetonitrile, Bycyclo-octan-1-ol, 4 methyl-C9H16O**) diagnosed loss of asthma control from clinically stable patients. The accuracies of distinguishing baseline, loss of control and recovery were (68%-77%) for GC-MS and (86%-95%) for eNose. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Inflammatory Asthma Phenotype Discrimination Using an Electronic Nose Breath Analyser	Plaza *et al* (2015) [81]	**52**	Tedlar bag (eNose)	• eNose successfully diagnosed different subtypes of asthma • Neutrophilic vs. paucigranulocytic separated with 89% Cross-validation accuracy (AUC = 0.88). Neutrophilic vs. eosinophilic separated with 73% CVA (AUC = 0.92). Eosinophilic vs. paucigranulocytic separated with 74 CVA (AUC = 0.79).

				• No reported use of STARD/TRIPOD
Electronic nose breathprints are independent of acute changes in airway calibre in asthma.	Lazar *et al* (2010) [27]	**10**	Tedlar bag (eNose)	• Acute reduction in airway calibre was not associated with an altered breath molecular profile • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label ¥	De Vries *et al* (2018) [82]	**435**	Spironose (eNose)	• eNose identified clusters in total population of asthma and COPD by clinical/inflammatory characteristics (ethnicity, systemic eosinophilia, systemic neutrophilia, BMI, atopy, recent exacerbation and FeNO) • No reported use of STARD/TRIPOD
Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma	Montuschi *et al* (2010) [17]	**51**	Tedlar bag (eNose)	• Diagnostic accuracy for asthma is highest (95%) when eNose and FeNo are combined. • No reported use of STARD/TRIPOD
Exhaled Volatile Organic Compounds as Markers for Medication Use in Asthma	Brinkman P *et al (2019) [83]*	**78**	Tedlar bag (GC-ToF-MS)	• 4 VOCs (**Lysine**, **Glycolic acid**, **4-Carene** and **Octanal.)** weer associated with traces of asthma medications in urine of severe asthmatics. Baseline AUC: 82.1 for salbutamol and 78.8 for OCS. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD

COPD				
Title	**Author**	**Sample size**	**Breath Collection and analysis**	**Summary of findings in relation to review objectives**
An electronic nose in the discrimination of patients with non-small cell lung cancer and COPD	Dragonieri *et al* (2009) [20]	**30**	Tedlar bag (eNose)	• eNose diagnosed non-small cell lung cancer (NSCLC) from COPD subjects (CVV: 85%; M-distance: 3.73) and NSCLC patients from healthy controls in duplicate measurements (CVV: 90% and 80%, respectively; M-distance: 2.96 and 2.26). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Differentiation of chronic obstructive pulmonary disease (COPD) including lung cancer from healthy control group by breath analysis using ion mobility spectrometry	Westhoff *et al* (2010) [46]	**132**	Direct sampling (MCC/IMS)	• **Cyclohexanone** diagnosed COPD with 60% sensitivity and 91% specificity and positive predictive value of 95%. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Non-invasive metabolomic analysis of breath using differential mobility spectrometry in patients with chronic obstructive pulmonary disease and healthy smokers	Basanta *et al* (2010) [84]	**26**	Full face mask - directly connected to sensors (IMS & GC-DMS)	• No specific diagnostic peaks related to COPD reported. AUC 0.79 for one feature, further work needed to identify chemical identity. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
A profile of volatile organic compounds in breath discriminates COPD patients from controls	Van berkel *et al* (2010) [19]	**79**	Tedlar bag (GC-TOF-MS)	• 6 VOCs (mainly long chain hydrocarbons) named as potential discriminators of COPD **(Isoprene, C_{16} hydrocarbon, 4,7-Dimethyl-undecane, 2,6-Dimethyl-heptane, 4-Methyl-octane, Hexadecane**) with a 91% diagnostic accuracy, specificity of 81% and a sensitivity of 100%. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD

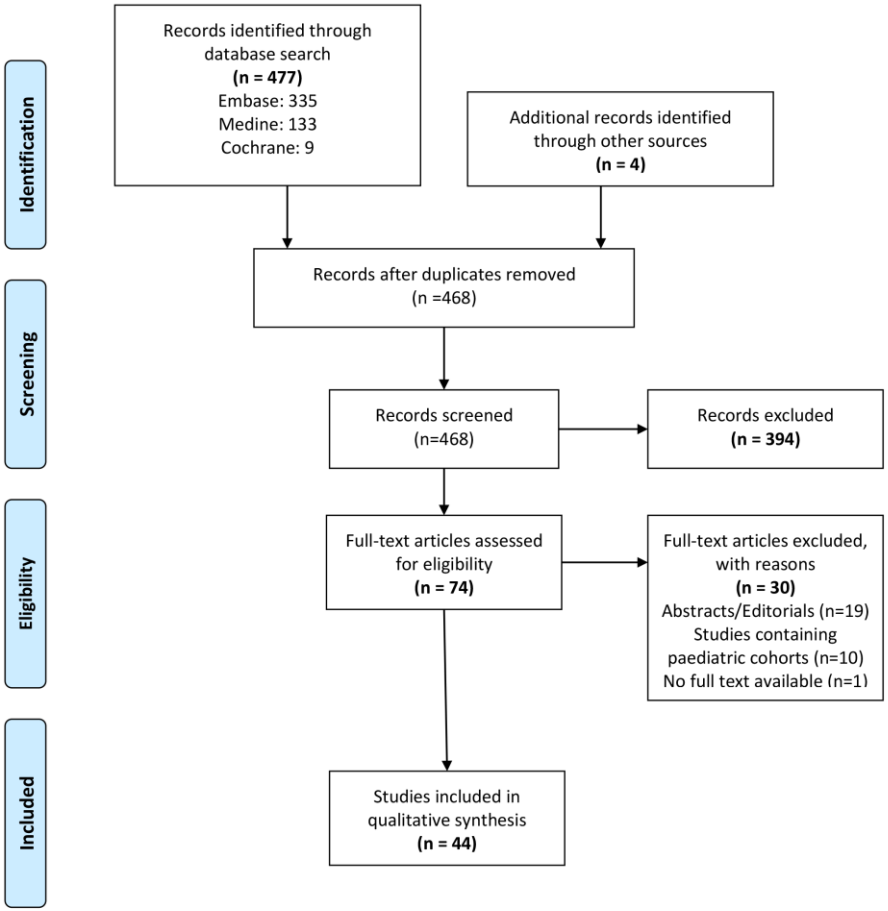
External validation of exhaled breath profiling using an electronic nose in the discrimination of asthma with fixed airways obstruction and chronic obstructive pulmonary disease	Fens *et al* (2011) [42]	**100**	Tedlar bag (eNose)	• eNose diagnosed fixed airway obstruction asthma from COPD with 88% accuracy (AUC 0.95) and discriminated ex/non-smoking asthmatics from COPD with 90% accuracy (AUC 0.96). • No relation to markers of type-2 inflammation • Reported use of STARD criteria
Screening for emphysema via exhaled volatile organic compounds	Cristescu *et al* (2011) [85]	**204**	Tedlar bag (PTR-MS)	• Exhaled VOCs did not provide valuable diagnostic information. AUC (0.58). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Discrimination between COPD patients with and without alpha 1-antitrypsin deficiency using an electronic nose	Hattesohl *et al* (2011) [68]	**43**	Collection bag (eNose)	• eNose is able to separate AAT from non AAT COPD (p<0.0001, sensitivity of 1.00, specificity of 1.00), and COPD from healthy controls (p<0.0001, sensitivity of 1.00, specificity of 1.00), and AAT from healthy controls (p<0.0001, sensitivity of 1.00, specificity of 1.00). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Statistical and bioinformatical methods to differentiate chronic obstructive pulmonary disease (COPD) including lung cancer from healthy control by breath analysis using ion mobility spectrometry	Westhoff *et al* (2011) [45]	**130**	Direct sampling (MCC-IMS)	• 10 peaks with high likelihood of discriminating COPD and controls. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls	Phillips *et al* (2012) [38]	**182**	Bio-VOC collector (Markes) (GC-MS)	• Identified VOCs **(Isoprene, Acetic acid, Benzaldehyde, Benzene, Butane, Carbon dioxide, hexanal, nonadecane, Phenol, Phthalic anhydride, Sulphur dioxide and Toluene)** as potential discriminators of COPD. AUC 0.82. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Exhaled volatile organic compounds for phenotyping chronic obstructive pulmonary disease: a cross-sectional study	Basanta *et al* (2012) [21]	**71**	Direct sampling onto sorbent tubes (GC-TOF-MS)	• Identified VOCs diagnosed COPD with 85% sensitivity, 50% specificity, 67% precision, 69% accuracy and AUC 0.74. • VOCs separated clinically relevant COPD subgroups (eosinophil count $\geq$ 1%, $\geq$ 2%) with logistic regression models accuracy of 79% and 92%, AUC of 0.90 and 0.94. • No reported use of STARD/TRIPOD
Breath analysis in real time by mass spectrometry in chronic obstructive pulmonary disease	Martines Sinues *et al* (2013) [86]	**61**	Direct sampling (Quadruple TOF-MS)	• Identified VOCs (Acetone and Indole) were able to discriminate COPD with 96% sensitivity and 72.7% specificity (COPD vs. smoking controls), 88% sensitivity and 92% specificity (COPD vs. non-smoking controls) and 92 % sensitivity and 83% specificity (GOLD I/II vs. GOLD III/IV).

				• No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Using the Electronic Nose to Identify Airway Infection during COPD Exacerbations	Shafiek *et al* (2015) [31]	**192**	Tedlar bag (eNose)	• Significant difference between breath profiles from infective COPD exacerbation with and without confirmed bacterial infection. In presence of microorganisms: Stable COPD vs. control (*p* value 0.005, sensitivity 70%, and specificity 73%). ECOPD vs. controls (*p* value 0.001, sensitivity 68%, specificity 80%). Pneumonia vs. controls (*p* value 0.005, sensitivity 88%, specificity 100%). Stable COPD vs. Pneumonia (*p* value <0.001, sensitivity 88%, specificity 75%). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Analysis of exhaled breath fingerprints and volatile organic compounds in COPD	Cazzolla *et al* (2015) [39]	**34**	Tedlar bag (GC-MS and eNose)	• eNose discriminated COPD from controls, p value ≤ 0.001. The cross validated model provided the correct classification of 26 of 27 COPD patients and 5 of 7 control subjects, 96 % sensitivity, 71 % specificity, 83 % NPV, 93 % PPV and 91 % diagnostic accuracy. • GC-MS detected two VOCs that were positively correlated to COPD (**Decane, 6-ethyl-2-methyl-Decane**) (Pearson Correlation: 0.35 ± 0.01; $p < 0.05$), seven VOCs negatively correlated with COPD (**Benzene, 1,3,5-tri-tert-butyl-, Butylated hydroxytoluene, Hexane, 3-ethyl-4-methyl-, Hexyl ethylphosphonofluoridate, Limonene, 1-Pentene, 2,4,4-trimethyl-, 2-Propanol**) - Pearson Correlation: -0.43 ± 0.01; $p < 0.01$. • Reported use of STARD criteria
Exhaled volatile organic compounds discriminate patients with chronic obstructive pulmonary disease from healthy subjects	Besa *et al* (2015) [44]	**96**	Direct sampling – (IMS)	• Six unidentified VOCs were able to distinguish COPD patients from healthy subjects (with an accuracy of 71%, 70%, 70%, 71%, 70%, and 67%, respectively). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
A dual centre study to compare breath volatile organic compounds from smokers and non-smokers with and without COPD	Gaida *et al* (2016) [55]	**190**	Direct sampling (TD-GC-MS)	• Linear discriminant analysis correctly classified 89.4% of COPD patients in the non/ex-smoking group, CV: 85.6%, and 82.6% of COPD patients in the actively smoking group (CV 77.9%). • 10 novel breath VOCs appear to be related to COPD (**m/p-Xylene, 1,6-Dimethyl-1, 3,5-heptatriene, o-xylene, 1-ethyl-3-methyl benzene, phenole, m/p-Cresol, Linalyl acetate, tridecane, indole**). • No relation to markers of type-2 inflammation

				• No reported use of STARD/TRIPOD
Measurement of exhaled volatile organic compounds from patients with chronic obstructive pulmonary disease (COPD) using closed gas loop GC-IMS and GC-APCI-MS	Allers *et al* (2016) [87]	**58**	Direct sampling (GC-IMS, APCI – CMS)	• **2-pentanone**, a potential differentiator between COPD and healthy and 16 unidentified VOCs significantly different ($p < 0.01$) between smokers and non/ex-smokers • In active smokers, there were higher levels of **ethanol, acetonitrile, 2-butanone** and 3 unidentified compounds compared to ex- and non-smokers ($p < 0.01$). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Study of 5 Volatile Organic Compounds in Exhaled Breath in Chronic Obstructive Pulmonary Disease	Jareno-estaban *et al* (2017) [47]	**157**	Bio-VOC collector (GC-MS)	• **Hexanal** discriminated COPD from healthy non-smoking controls (p=.015, 95% CI: 1.13–8.40) (sensitivity 31.58%, specificity 89.74%, PPV 81.82% and NPV 47.30%). **Nonanal** discriminated smokers and former smokers (with and without COPD) and never smokers (p=.008, 95% CI: 1.17–5.08). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Analysis of volatile organic compounds in the breath of patients with stable or acute exacerbation of chronic obstructive pulmonary disease	Pizzini *et al* (2018) [32]	**54**	Direct sampling via mouth piece onto glass syringe (GC-ToF-MS)	• **VOCs indicative of acute exacerbations: cyclohexanone** (p value <0.001), **n-butane** (p value <0.001), **4-heptanone** (p value 0.001), **2-pentanone** (p value 0.002), sensitivity 0.69, NPV 0.82, specificity 0.94, PPV 0.89, AUC 0.92**. COPD specific: N-heptane** (p value <0.001), **methyl propyl sulphide** (p value <0.001). • **Undefined association to COPD: Dimethyl disulphide** (p value <0.001), **6-methyl-5-heptene-2-one** (p value <0.001) • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Real-time mass spectrometric identification of metabolites characteristic of chronic obstructive pulmonary disease in exhaled breath	Bregy *et al* (2018) [88]	**36**	Direct sampling (SESI-HRMS)	• SESI-MS can diagnose COPD with 89% accuracy, 93% sensitivity, 87% specificity, 81% PPV, and 95% NPV. • **11-hydroxyundecanoic acid** correlates with FEV1% predicted and FEV_1/FVC ratio (q = 0.04 respectively q = 0.02) with positive correlation coefficients r = 0.48 resp. 0.48). • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Cluster analysis on breath print of newly diagnosed COPD patients: effects of therapy	Scarlata S *et al* (2018) [49]	**55**	Direct sampling - Pneumopipe ® (eNose)	• Breath prints patterns change in distinctive ways depending upon whether COPD patients have been prescribed LAMA and/or LABA alone or any combination of inhaled drugs including ICS.

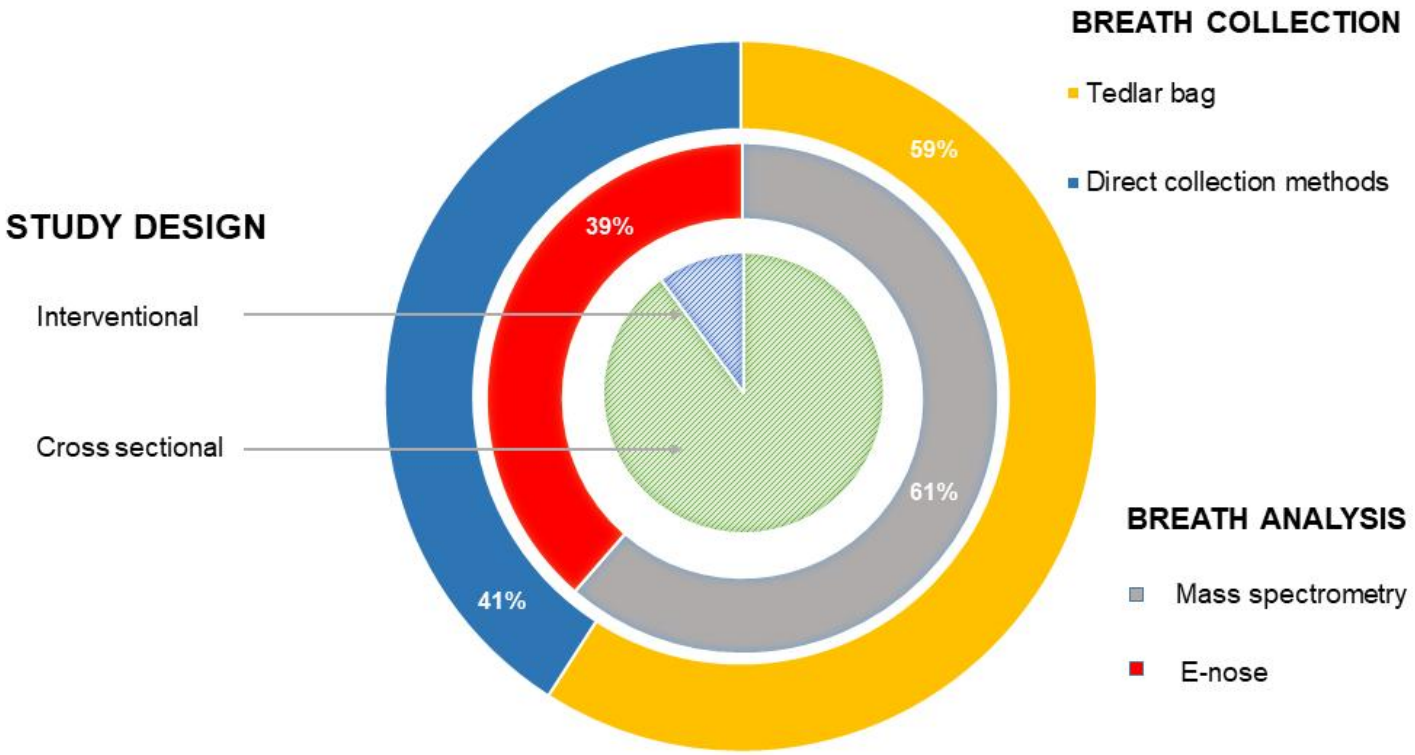
				• No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Short-Term Intra-Subject Variation in Exhaled Volatile Organic Compounds (VOCs) in COPD Patients and Healthy Controls and Its Effect on Disease Classification	Phillips C *et al* (2014) [48]	**181**	Bio-VOC collector (GC-MS)	• A substantial variation was found in the levels of VOC in breath during three repeat samples over a short time period. The extent of the variation in VOC levels differed between COPD and healthy subjects and the patterns of variation differed for isoprene *versus* the majority of other VOCs. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Reproducibility and respiratory function correlates of exhaled breath fingerprint in chronic obstructive pulmonary disease	Incalzi R *et al* (2012) [89]	**25**	Tedlar bag (eNose)	• eNose has the potential to assess COPD severity and study phenotypic variability. Suboptimal reproducibility within GOLD 1–3 patients. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Ultrafast gas chromatography coupled to electronic nose to identify volatile biomarkers in exhaled breath from chronic obstructive pulmonary disease patients: A pilot study	Rodriguez-Aguilar *et al* (2019) [40]	**56**	Tedlar bag (Ultrafast gas chromatography equipped with eNose detector (FGC eNose))	• 17 VOCs distinguished COPD from controls (Sensitivity 96%, specificity 91%). There was an increase in the AUC of **alpha-pyrene, acetaldehyde, 2-butyloctanol, octane, methylisobutyrate, butanal, 2-propranolol, 3-hexanone, cyclopentanone, and 3-methyl-propanal** and a decrease in the AUC of **delta-dodecalactone, 2-methyl butanoic acid, 2-acetylpyridie, tetradecane, [E]-cinnamaldehyde** and **vinylpyrazine** in the exhaled breath of COPD patients compared to controls. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Exhaled breath profiles before, during and after exacerbation of COPD: a prospective follow-up study	Velzen *et al (2019) [34]*	**68**	Tedlar bag (GC-MS and eNose)	• GC-MS identified 10 compounds of interest: **acetone, 1,2-pentadiene, toluene, butyrolactone, ethylbenzene, 2-decenal, limonene, 4,7-dimethyl-undecane, eicosane,** and **1-undecanol.** • Breath profiles by eNose and GC-MS showed correct classification of: **Baseline Vs exacerbations** with 71% accuracy and sensitivity and specificity of 0.71. **Exacerbation Vs follow up**: accuracy of 75%, sensitivity of 0.92 and specificity of 0.57. **Baseline Vs follow up** accuracy of 57%, sensitivity of 0.64 and specificity of 0.5. • No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD
Real-time breath analysis reveals specific metabolic signatures of COPD exacerbations	Gaugg *et al* (2019) [33]	**52**	Direct sampling (Secondary electrospray	• Metabolite levels from the ω-oxidation pathway, namely ω-hydroxy, ω-oxo, and dicarboxyclic acids were consistently decreased in frequent exacerbators.AUC 0.88.

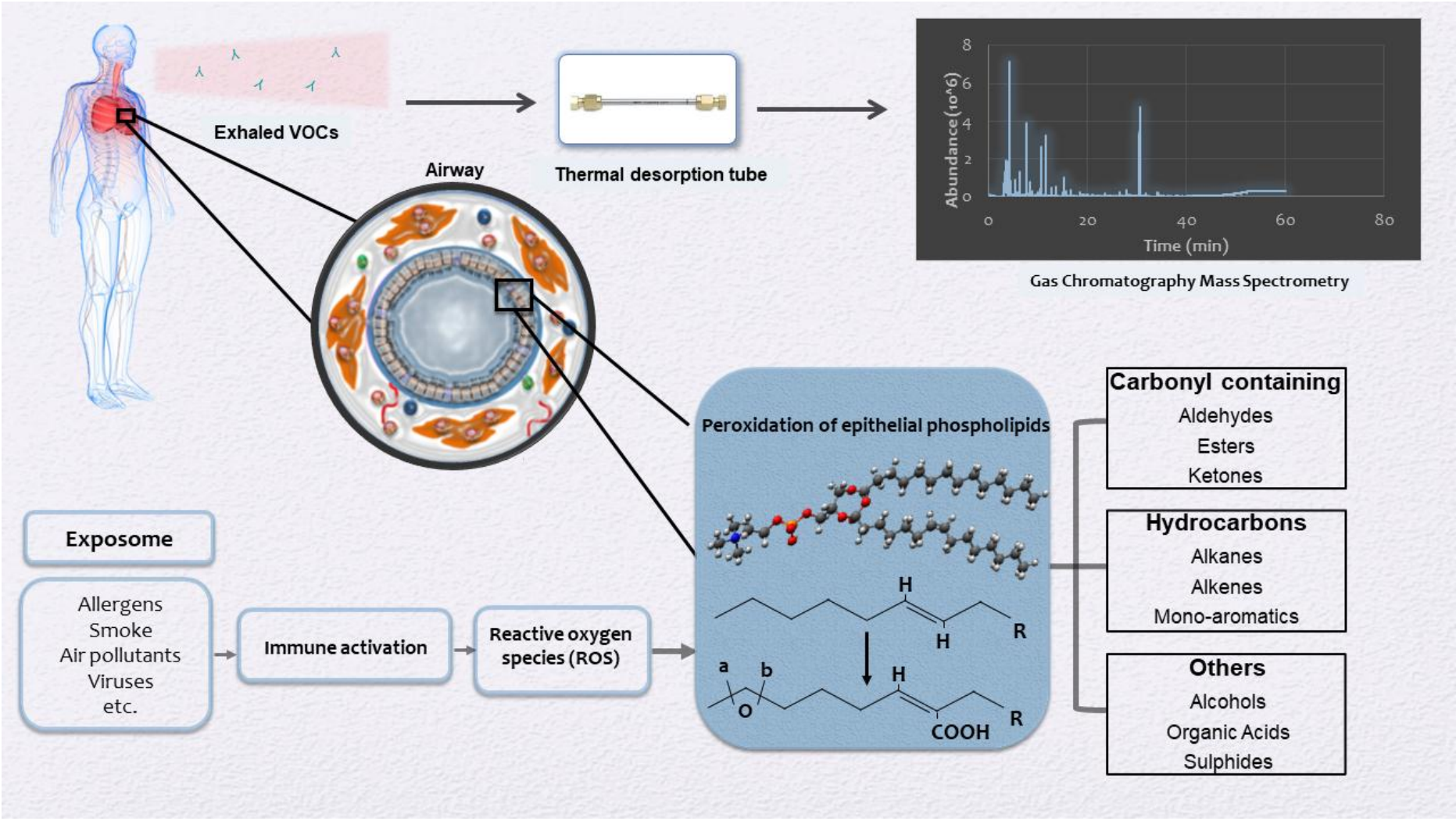
			ionisation high resolution mass spectrometry)	• No relation to markers of type-2 inflammation • No reported use of STARD/TRIPOD

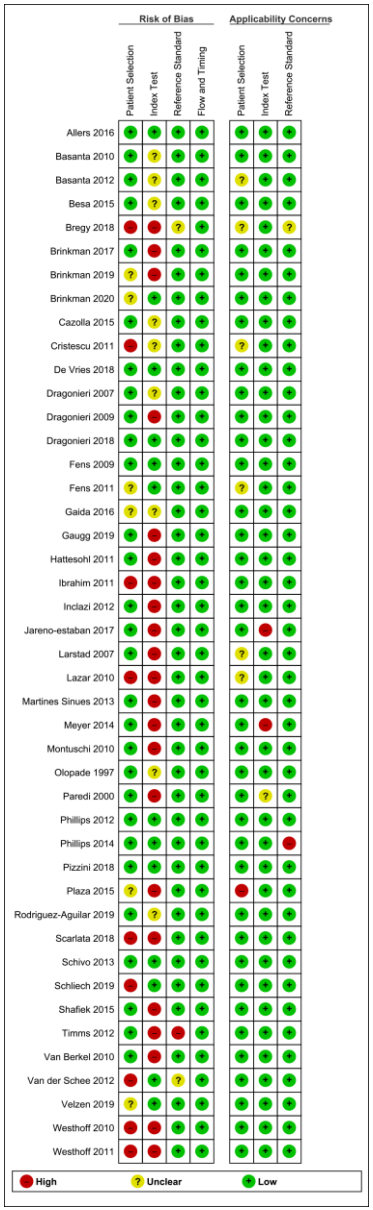


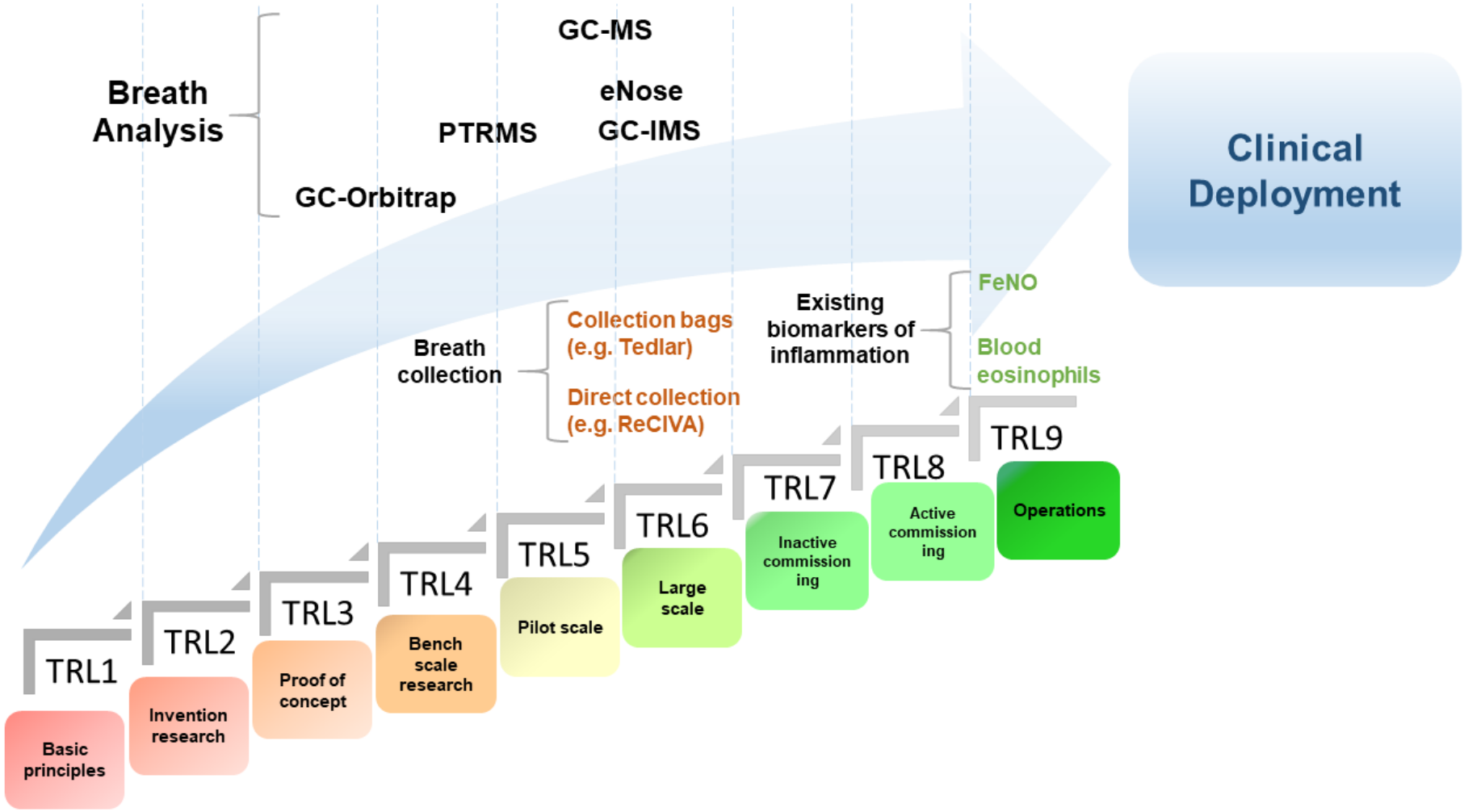
From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.


BREATH COLLECTION
Tedlar bag
Direct collection methods
59%
39%
STUDY DESIGN
Interventional
Cross sectional
61%
41%
BREATH ANALYSIS
Mass spectrometry
E-nose


Exhaled VOCs
Airway
Thermal desorption tube
Abundance(10^6)
Time (min)
Gas Chromatography Mass Spectrometry
Peroxidation of epithelial phospholipids
Exposome
Allergens
Smoke
Air pollutants
Viruses
etc.
Immune activation
Reactive oxygen species (ROS)
Carbonyl containing
Aldehydes
Esters
Ketones
Hydrocarbons
Alkanes
Alkenes
Mono-aromatics
Others
Alcohols
Organic Acids
Sulphides


Risk of Bias
Applicability Concerns
Patient Selection
Index Test
Reference Standard
Flow and Timing
Patient Selection
Index Test
Reference Standard
Allers 2016
Basanta 2010
Basanta 2012
Besa 2015
Bregy 2018
Brinkman 2017
Brinkman 2019
Brinkman 2020
Cazolla 2015
Cristescu 2011
De Vries 2018
Dragonieri 2007
Dragonieri 2009
Dragonieri 2018
Fens 2009
Fens 2011
Gaida 2016
Gaugg 2019
Hattesohl 2011
Ibrahim 2011
Inclazi 2012
Jareno-estaban 2017
Larstad 2007
Lazar 2010
Martines Sinues 2013
Meyer 2014
Montuschi 2010
Olopade 1997
Paredi 2000
Phillips 2012
Phillips 2014
Pizzini 2018
Plaza 2015
Rodriguez-Aguilar 2019
Scarlata 2018
Schivo 2013
Schliech 2019
Shafiek 2015
Timms 2012
Van Berkel 2010
Van der Schee 2012
Velzen 2019
Westhoff 2010
Westhoff 2011
High
Unclear
Low


Breath Analysis
GC-MS
eNose
PTRMS
GC-IMS
GC-Orbitrap
Clinical Deployment
Breath collection
Collection bags (e.g. Tedlar)
Direct collection (e.g. ReCIVA)
Existing biomarkers of inflammation
FeNO
Blood eosinophils
TRL1
Basic principles
TRL2
Invention research
TRL3
Proof of concept
TRL4
Bench scale research
TRL5
Pilot scale
TRL6
Large scale
TRL7
Inactive commissioning
TRL8
Active commissioning
TRL9
Operations

A Systematic Review of the Diagnostic Accuracy of Volatile Organic Compounds in Airways Diseases and their relation to Markers of Type-2 Inflammation

Wadah Ibrahim[1,2*], Sushiladevi Natarajan[1,2*], Michael Wilde[3], Rebecca Cordell[3], Paul S Monks[3], Neil Greening[1,2], Christopher E Brightling[1,2], Rachael Evans[1,2], Salman Siddiqui[1,2]

wadah.ibrahim@leicester.ac.uk; sushiladevi.natarajan@nhs.net; mjw77@leicester.ac.uk; rc145@leicester.ac.uk; paul.monks@leicester.ac.uk; neil.greening@leicester.ac.uk; ceb17@leicester.ac.uk; re66@leicester.ac.uk; ss338@leicester.ac.uk;

Corresponding author: Professor Salman Siddiqui,

Leicester Biomedical Research Centre (Respiratory Theme)

College of Life Sciences, Department of Respiratory Sciences, University of Leicester

Glenfield Hospital,

Groby Road

Leicester

LE3 9QP

t: +44 116 256 3841

e: ss338@ le.ac.uk

Authors' affiliations:

[1]Leicester NIHR Biomedical Research Centre (Respiratory theme), Glenfield Hospital, Groby Road, Leicester LE3 9QP

[2] College of Life Sciences, Department Respiratory Sciences, University of Leicester, University Road, Leicester LE1 7RH

[3] Department of Chemistry, University of Leicester, University Road, Leicester, LE1 7RH

***Equal contribution to work**

Supplementary material

1. Search strategy

A PRISMA oriented systematic search was completed from **January 1997 to December 2020** using the following sources of evidence, (i) Cochrane library, (ii) Medline and (iii) EMBASE.

Further details on methodology can be found in (PROSPERO - CRD42019141718).

Strategy 577083

#	Database	Search term	Results
1	Medline	("volatile organic compound*").ti,ab	7726
2	Medline	"VOLATILE ORGANIC COMPOUNDS"/	6691
3	Medline	(VOC OR vocs).ti,ab	7212
4	Medline	(exhal*6 OR breath*).ti,ab	113917
5	Medline	EXHALATION/	3377
6	Medline	(1 OR 2 OR 3)	14894
7	Medline	(4 OR 5)	117153
8	Medline	(6 AND 7)	1107
9	Medline	exp DYSPNEA/	19644
10	Medline	(breathless* OR dyspn*4).ti,ab	46465
11	Medline	(9 OR 10)	55084
12	Medline	(8 AND 11)	5
13	Medline	(asthma*).ti,ab	146117
14	Medline	exp ASTHMA/	121260

15	Medline	exp "PULMONARY DISEASE, CHRONIC OBSTRUCTIVE"/	50255
16	Medline	(copd).ti,ab	37549
17	Medline	(((chronic AND obstructive) AND pulmonary) AND disease).ti,ab	40946
18	Medline	exp PULMONARY DISEASE, CHRONIC OBSTRUCTIVE/	47893
19	Medline	(((chronic AND obstructive) AND airway) AND (disease OR coad)).ti,ab	5928
20	Medline	(((chronic AND obstructive) AND lung) AND disease).ti,ab	15699
21	Medline	((chronic AND airflow) AND limitation).ti,ab	1694
22	Medline	(((chronic AND obstructive) AND respiratory) AND disease).ti,ab	11050
23	Medline	(emphysema).ti,ab	21250
24	Medline	exp EMPHYSEMA/	13966
25	Medline	(chronic AND bronchitis).ti,ab	11694
26	Medline	exp BRONCHITIS, CHRONIC/	1665
27	Medline	exp PNEUMONIA/	86684
28	Medline	(pneumonia* OR lung inflammation*or respiratory tract infection* OR respiratory infection*).ti,ab	214438
29	Medline	exp "HEART FAILURE"/	110645
30	Medline	(heart failure).ti,ab	165767
31	Medline	(13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30)	678339
32	Medline	(8 AND 31)	163

33	EMBASE	("volatile organic compound*").ti,ab	10048
34	EMBASE	"VOLATILE ORGANIC COMPOUND"/	14083
35	EMBASE	"VOLATILE ORGANIC COMPOUNDS"/	10777
36	EMBASE	(voc OR vocs).ti,ab	9092
37	EMBASE	(33 OR 34 OR 35 OR 36)	20436
38	EMBASE	(exhal*6 OR breath*).ti,ab	165810
39	EMBASE	EXHALATION/	4059
40	EMBASE	(38 OR 39)	177809
41	EMBASE	(37 AND 40)	1675
42	EMBASE	(copd).ti,ab	70312
43	EMBASE	(((chronic AND obstructive) AND pulmonary) AND disease).ti,ab	59947
44	EMBASE	"CHRONIC OBSTRUCTIVE LUNG DISEASE"/ OR "CHRONIC OBSTRUCTIVE LUNG DISORDER"/ OR "CHRONIC OBSTRUCTIVE PSEUDOEMPHYSEMA"/ OR "CHRONIC OBSTRUCTIVE PULMONARY DISEASE"/ OR "CHRONIC OBSTRUCTIVE PULMONARY DISORDER"/ OR "CHRONIC OBSTRUCTIVE RESPIRATORY DISEASE"/	112227
45	EMBASE	(((chronic AND obstructive) AND airway) AND (disease OR coad)).ti,ab	9382
46	EMBASE	(((chronic AND obstructive) AND lung) AND disease).ti,ab	24650
47	EMBASE	((chronic AND airflow) AND limitation).ti,ab	2646

48	EMBASE	(((chronic AND obstructive) AND respiratory) AND disease).ti,ab	17649
49	EMBASE	(emphysema).ti,ab	28141
50	EMBASE	exp EMPHYSEMA/	41697
51	EMBASE	(chronic AND bronchitis).ti,ab	15855
52	EMBASE	"CHRONIC BRONCHITIS"/	11611
53	EMBASE	exp ASTHMA/	237742
54	EMBASE	(asthma*).ti,ab	205990
55	EMBASE	exp PNEUMONIA/	265756
56	EMBASE	(pneumonia* OR lung inflammation*or respiratory tract infection* OR respiratory infection*).ti,ab	225025
57	EMBASE	exp "HEART FAILURE"/	443663
58	EMBASE	(heart failure).ti,ab	242128
59	EMBASE	(42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49 OR 50 OR 51 OR 52 OR 53 OR 54 OR 55 OR 56 OR 57 OR 58)	1204355
60	EMBASE	(41 AND 59)	375

61 Cochrane search:

ID	Search
#1	"volatile organic compound*"
#2	MeSH descriptor: [Volatile Organic Compounds] this term only
#3	VOC OR vocs
#4	MeSH descriptor: [Exhalation] this term only
#5	exhal*6 OR breath*
#6	#1 OR #2 OR #3

#7	#4 OR #5
#8	#6 AND #7
#9	asthma*
#10	MeSH descriptor: [Asthma] this term only
#11	MeSH descriptor: [Pulmonary Disease, Chronic Obstructive] this term only
#12	copd
#13	(((chronic AND obstructive) AND pulmonary) AND disease)
#14	(((chronic AND obstructive) AND airway) AND (disease OR coad))
#15	(((chronic AND obstructive) AND lung) AND disease)
#16	((chronic AND airflow) AND limitation)
#17	(((chronic AND obstructive) AND respiratory) AND disease)
#18	(emphysema)
#19	MeSH descriptor: [Emphysema] this term only
#20	(chronic AND bronchitis)
#21	MeSH descriptor: [Bronchitis, Chronic] this term only
#22	{OR #9-#21}
#23	{AND #8-#22}
#24	#8 and #22

2. Risk of Bias table (QUADAS2)

Table A: Breakdown of included studies' risk of bias and applicability concerns

Asthma studies	Risk of Bias					Applicability concerns		
Author	**Patient selection**	**Index test**	**Reference standard**	**Flow and timing**		**Patient selection**	**Index test**	**Reference standard**
Olopade *et al* (1997)	Low	Unclear	Low	Low		Low	Low	Low
Paredi *et al* (2000)	Low	High	Low	Low		Low	Unclear	Low
Larstad *et al* (2007)	Low	High	Low	Low		Unclear	Low	Low
Dragonieri *et al* (2007)	Low	Unclear	Low	Low		Low	Low	Low
Fens *et al* (2009)	Low	Low	Low	Low		Low	Low	Low
Ibrahim *et al* (2011)	High	High	Low	Low		Low	Low	Low
Timms *et al* (2012)	Low	High	High	Low		Low	Low	Low
Van der Schee *et al* (2012)	High	Low	Unclear	Low		Low	Low	Low
Schivo *et al* (2013)	Low	Low	Low	Low		Low	Low	Low
Meyer *et al* (2014)	Low	High	Low	Low		Low	High	Low
Schliech *et al* (2019)	High	Low	Low	Low		Low	Low	Low
Brinkman *et al* (2019)	Unclear	Low	Low	Low		Low	Low	Low
Dragonieri *et al* (2018)	Low	Low	Low	Low		Low	Low	Low
Brinkman *et al*	Low	High	Low	Low		Low	Low	Low

(2017)								
Plaza *et al* (2015)	Unclear	High	Low	Low		Low	Low	Low
Lazar *et al* (2010)	High	High	Low	Low		Unclear	Low	Low
De Vries *et al* (2018)	Low	Low	Low	Low		Low	Low	Low
Montuschi *et al* (2010)	Low	High	Low	Low		Low	Low	Low
Brinkman et al (2020)	Unclear	Low	Low	Low		Low	Low	Low
COPD studies								
Dragonieri *et al* (2009)	Low	High	Low	Low		Low	Low	Low
Westhoff *et al* (2010)	Low	Low	High	High		Low	Low	Low
Basanta *et al* (2010)	Low	Unclear	Low	Low		Low	Low	Low
Van berkel *et al* (2010)	Low	High	Low	Low		Low	Low	Low
Fens *et al* (2011)	Unclear	Low	Low	Low		Unclear	Low	Low
Cristescu *et al* (2011)	High	Unclear	Low	Low		Unclear	Low	Low
Hattesohl *et al* (2011)	Low	High	Low	Low		Low	Low	Low
Westhoff *et al* (2011)	High	High	Low	Low		Low	Low	Low
Phillips *et al* (2012)	Low	Low	Low	Low		Low	Low	Low
Basanta *et al* (2012)	Low	Unclear	Low	Low		Unclear	Low	Low

Martines Sinues *et al* (2013)	Low	High	Low	Low		Low	Low	Low
Shafiek *et al* (2015)	Low	High	Low	Low		Low	Low	Low
Cazzolla *et al* (2015)	Low	Unclear	Low	Low		Low	Low	Low
Besa *et al* (2015)	Low	Unclear	Low	Low		Low	Low	Low
Gaida *et al* (2016)	Unclear	Unclear	Low	Low		Low	Low	Low
Allers *et al* (2016)	Low	Low	Low	Low		Low	Low	Low
Jerno-estaban *et al* (2017)	Low	High	Low	Low		Low	High	Low
Pizzini *et al* (2018)	Low	Low	Low	Low		Low	Low	Low
Bregy *et al* (2018)	High	High	Unclear	Low		Unclear	Low	Unclear
Scarlata *et al* (2018)	High	High	Low	Low		Low	Low	Low
Phillips *et al* (2014)	Low	Low	Low	Low		Low	Low	High
Incalzi *et al* (2012)	Low	High	Low	Low		Low	Low	Low
Rodriguez-Aguilar *et al* (2019)	Low	Unclear	Low	Low		Low	Low	Low
Velzen *et al* (2019)	Unclear	Low	Low	Low		Low	Low	Low
Gaugg *et al* (2019)	High	High	Low	Low		Low	Low	Low

Figure A: Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies

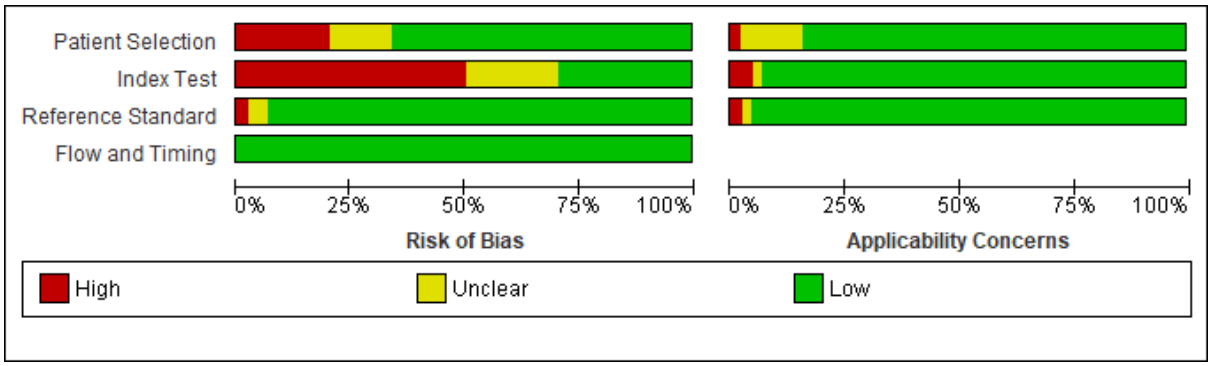


Table B: Description of modification of QUADAS-2:

Risk of Bias	QUADAS-2	Modified QUADAS-2
Patient selection	Was a consecutive or random sample of patients enrolled?	Was patient selection representative of the intended population
	Was a case-control design avoided?	Did the study include disease state and healthy controls
	Did the study avoid inappropriate exclusions?	Did the study avoid inappropriate exclusions?
Index test	Were the index test results interpreted without knowledge of the results of the reference standard?	Was the index test and data interpretation completed in a standardised and reproducible way?
	If a threshold was used, was it pre-specified?	Was any biomarker validation performed (internal or external)?
Reference standard	Is the reference standard likely to correctly classify the target condition?	Is the reference standards likely to correctly classify the target condition?
	Were the reference standard results interpreted without knowledge of the results of the index test?	Removed as not applicable
Flow and timing	Was there an appropriate interval between index test and reference standard?	Was there an appropriate interval between index test and reference standard?
	Did all patients receive the same reference standard?	Did all patients receive the same reference standard?
	Were all patients included in the analysis?	Were all patients included in the analysis?

Applicability	**QUADAS-2**	**Modified QUADAS-2**
Patient selection	Are there concerns that the included patients and setting do not match the review question?	Are there concerns that the included patients and setting do not match the review question?
Index test	Are there concerns that the index test, its conduct, or interpretation differs from the review question?	Could the conduct or interpretation of the index test have introduced bias?
Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Are there concerns that the target condition as defined by the reference standard does not match the question?

Index test: Exhaled breath analysis

Reference standard: Internationally accepted standard for diagnosing asthma and COPD (i.e. following GINA and GOLD guidelines)

Target condition: Asthma and COPD

Intended use of the index test: diagnostic

3. PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	3
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7-8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7-8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementary material
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7-8

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7-8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7-8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	N/A
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8 and Supplementary material
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	9 and Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8,9,10,11 and Tables 1-2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8 and supplementary material
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9-11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Supplementary material
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A

DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	14-15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Declarations

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Compound and chemical classification	Author	Disease	Comments
1. Carbonyl containing			
1.1. Aldehydes			
2-oxoglutaric acid semi-aldehyde	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - high levels in COPD patients
aspartic acid semi-aldehyde	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - high levels in COPD patients
Benzaldehyde	Phillips *et al* [38]	COPD vs. controls	
Butanal	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Positively correlated to COPD
Decanal	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
Dodecanal	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
Hexanal	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
Hexanal	Jareno-estaban *et al* [46]	COPD vs. controls	Discriminates between COPD and healthy controls
Hexanal	Phillips *et al* [38]	COPD vs. controls	
Nonanal	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
Nonanal	Jareno-estaban *et al* [46]	COPD vs. controls	Discriminates smokers and former smokers (with and without COPD) and never smokers
Nonanal	Schliech *et al* [18]	Asthma	Discriminates paucigranulocytic and neutrophilic asthma (Higher in neutrophilic asthma)
Nonanal	Schliech *et al* [18]	Asthma	Discriminates eosinophilic and neutrophilic asthma (Higher in neutrophilic asthma)
Octanal	Brinkman *et al [75]*	Asthma	Association between exhaled breath VOCs and urinary levels of salbutamol and OCS

Pentadecanal	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
Pentadecanal	Ibrahim *et al* [15]	Asthma vs. controls	Negatively correlated to asthma
Undecanal	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
3-methyl-propanal	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Positively correlated to COPD
1.2. Esters			
Ethyl 2,2-dimethylacetoacetate	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
Linalylacetate	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
1.3. Ketones			
2-butanone	Allers *et al* [82]	COPD vs. controls	IMS - smoking related compounds
2-butanone	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
2-hexanone	Schliech *et al* [18]	Asthma	Discriminates pauci-granulocytic and eosinophilic asthma (lower in eosinophilic asthma)
2-pentanone	Allers *et al* [82]	COPD vs. controls	Detected by GC-APCI-MS discriminates COPD from healthy volunteers
2-pentanone	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Indicative of acute exacerbation of COPD (positive correlation)
4-heptanone	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Indicative of acute exacerbation of COPD (positive correlation)
6-methyl-5-hepten-2-one	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Non-specific. Significant difference between COPD and healthy volunteers (higher in healthy)
Acetone	Martines *et al* [80]	COPD vs. controls	Discriminate COPD from healthy volunteers using IMS
Cyclohexanone	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Indicative of acute exacerbation of COPD (positive correlation)

Cyclohexanone (CAS 108-94-1)	Westhoff *et al* [76]	COPD vs. controls	IMS identified - raised in COPD patients
Cyclopentanone	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Positively correlated to COPD
1.4. Organic acids			
11-hydroxyundecanoic acid	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - low levels in COPD patients
2-hydroxyisobutyric acid	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - compound predictive that breath is from a COPD patient
2-methyl butanoic acid	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Negatively correlated to COPD
Acetic acid	Phillips *et al* [38]	COPD vs. controls	
Butanoic acid	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS) - negative correlation
Dodecanedioic acid	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - low levels in COPD patients
Oxoheptadecanoic acid	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - low levels in COPD patients
Pentanoic acid	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS) - negative correlation
Lysine	Brinkman *et al [75]*	Asthma	Association between exhaled breath VOCs and urinary levels of salbutamol and OCS
Glycolic acid	Brinkman *et al* [75]	Asthma	Association between exhaled breath VOCs and urinary levels of salbutamol and OCS
ω-oxo-alkenoic acids	Gaugg *et al* [33]	COPD	Levels significantly reduced in frequent COPD exacerbators
ω-hydroxy acids	Gaugg *et al [33]*	COPD	Levels significantly reduced in frequent COPD exacerbators
2. Hydrocarbons			
2.1. Alkanes			

3,7-dimethylnonane	Schliech *et al* [18]	Asthma	Discriminates eosinophilic and neutrophilic asthma (Higher in neutrophilic asthma) and lower eosinophilic asthma
Hexane	Schliech *et al* [18]	Asthma	Discriminates paucigranulocytic and eosinophilic asthma (lower in eosinophilic asthma)
Undecane	Schliech *et al* [18]	Asthma	Discriminates paucigranulocytic and neutrophilic asthma (Higher in paucigranulocytic asthma)
2,6-Dimethyl-heptane	Van Berkel *et al* [19]	COPD vs. controls	Classification model differentiated COPD from healthy volunteers
4,7-Dimethyl-undecane	Van Berkel *et al* [19]	COPD vs. controls	Classification model differentiated COPD from healthy volunteers
4-Methyl-octane	Van Berkel *et al* [19]	COPD vs. controls	Classification model differentiated COPD from healthy volunteers
Hexadecane	Van Berkel *et al* [19]	COPD vs. controls	Classification model differentiated COPD from healthy volunteers
6-ethyl-2-methyl-Decane	Cazzola *et al* [39]	COPD vs. controls	Positively correlated to COPD
Decane	Cazzola *et al* [39]	COPD vs. controls	Positively correlated to COPD
Hexane, 3-ethyl-4-methyl-	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
Tridecane	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
Tetradecane	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Negatively correlated to COPD
2 methyl-decane	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
2,6,10-trimethyl-dodecane	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
2,6,11-trimethyl-dodecane	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
5,5-Dibutylnonane	Ibrahim *et al* [15]	Asthma vs. controls	Negatively correlated to asthma
Pentane	Olopade *et al* [29]	Acute and stable asthma vs. controls	Raised in asthma

Ethane	Paredi *et al* [36]	Asthma vs. controls	Raised in asthma
Butane	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
2,4-dimethylheptane	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Lower in COPD exacerbations compared to healthy volunteers
2,6-dimethyloctane	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Lower in COPD exacerbations compared to healthy volunteers
2-methylhexane	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Higher in COPD exacerbations compared to healthy volunteers
cyclohexane	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Higher in COPD exacerbations compared to healthy volunteers
n-butane	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Indicative of acute exacerbation of COPD (negative correlation)
n-Heptane	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Indicative of stable COPD (positive correlation)
2.2. Alkenes			
3- tetradecene	Schliech *et al* [18]	Asthma	Discriminates paucigranulocytic and neutrophilic asthma (Higher in neutrophilic asthma)
Pentadecene	Schliech *et al* [18]	Asthma	Discriminates paucigranulocytic and neutrophilic asthma (Higher in neutrophilic asthma)
Isoprene	Van Berkel *et al* [19]	COPD vs. controls	Classification model differentiated COPD from healthy volunteers
1-Pentene, 2,4,4-trimethyl-	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
1,6-Dimethyl-1,3,5-heptatriene	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
3,5-heptatriene	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
Isoprene	Phillips *et al* [38]	COPD vs. controls	Differentiating stable COPD patients

2.3. Hydrocarbons			
C16 hydrocarbon	Van Berkel *et al* [19]	COPD vs. controls	Classification model differentiated COPD from healthy volunteers
4-ethyl-o-xylene	Ibrahim *et al* [15]	Asthma vs. controls	Negatively correlated to asthma
Isoprene	Lastard *et al* [67]	Asthma vs. controls	Low in asthmatics
Nonadecane	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
Octane	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Positively correlated to COPD
2.4. Monoaromatics			
Benzene, 1,3,5-tri-tert-butyl-	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
1-Ethyl-3-methyl benzene	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
m/p-Xylene	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
O-xylene	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
Benzene	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
Toluene	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
2.5. Terpenes			
Limonene	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
Terpinolene	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
4-Carene	Brinkman *et al [75]*	Asthma	Association between exhaled breath VOCs and urinary levels of salbutamol and OCS
3. Alcohol and Phenols			

3.1. Alcohols			
Cyclohexanol	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
Bicyclo[2.2.2]octan-1-ol, 4-methyl - C9H16O	Brinkman *et al* [25]	Acute and stable asthma vs. controls	Correlated with sputum eosinophils during loss of asthma control and with FENO during loss of asthma control
Methanol CH3OH	Brinkman *et al* [25]	Acute and stable asthma vs. controls	Correlated with FEV1% predicted during loss of asthma control
2-Propanol	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
2-Propanol	Rodriguez-Aguilar *et al [40]*	COPD vs. controls	Positively correlated to COPD
Phenole	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
2-butylcyclohexanol	Ibrahim *et al* [15]	Asthma vs. controls	Negatively correlated to asthma
2-butyloctanol	Rodriguez-Aguilar *et al* [40]	COPD vs. controls	Positively correlated to COPD
Benzyl alcohol	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
Phenol	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
1-propanol	Schliech *et al* [18]	Asthma	Discriminates eosinophilic and neutrophilic asthma (Higher in neutrophilic asthma) and lower in eosinophilic asthma
3.2. Phenol derivatives			
Butylated hydroxytoluene	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
m/p-Cresol,	Gaida *et al* [47]	COPD vs. controls	VOCs seem to be related to COPD
4. Others			

4.1.Sulphides			
Phthalic anhydride	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
Sulphur dioxide	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
dimethyl disulfide	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Non-specific. Significant difference between COPD and healthy volunteers (higher in healthy)
methyl propyl sulfide	Pizzini *et al* [32]	Acute and stable COPD vs. controls	Indicative of stable COPD (positive correlation)
4.2. Permanent gases			
Ethyl 4-nitrobenzoate	Ibrahim *et al* [15]	Asthma vs. controls	Negatively correlated to asthma
Indole	Martines *et al* [80]	COPD vs. controls	Discriminate COPD from healthy volunteers using IMS
Indole	Gaida *et al* [47]	COPD vs. controls	VOCs seems to be related to COPD
4.3. Heterocycles			
Oxirane-dodecyl	Basanta *et al* [21]	COPD vs. controls	VOCs seem to be related to COPD
γ-hydroxy-L-homoarginine	Bregy *et al* [83]	COPD vs. controls	Analysis by SESI-MS - compound predictive that breath is from a COPD patient
4.4. Nitriles			
Ace-tonitrile - C2H3N	Brinkman *et al* [25]	Acute and stable asthma vs. controls	Correlated with sputum eosinophils and FEV1% during loss of asthma control
Hexyl ethylphosphonofluoridate	Cazzola *et al* [39]	COPD vs. controls	Negatively correlated to COPD
4.5. Anhydrides			
Acetonitrile	Allers *et al* [82]	COPD vs. controls	IMS - smoking related compound
4.6. Furans			

2-pentylfuran	Basanta *et al* [21]	COPD vs. controls	Identifying COPD (GC-TOF-MS)
4.7. Quinones			
Carbon dioxide	Phillips *et al* [38]	COPD vs. controls	Undetermined correlation
4.8.Others			
2,6-Di-tert-butylquinone	Ibrahim *et al* [15]	Asthma vs. controls	Negatively correlated to asthma
3,4-Dihydroxybenzonitrile	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma
Allyl methyl sulphide	Ibrahim *et al* [15]	Asthma vs. controls	Positively correlated to asthma

Table C: Table outlining reported VOC biomarkers in Asthma and COPD.