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

Wadah Ibrahim[1,2,5], Sushiladevi Natarajan[1,2,5], Michael Wilde [ID][3], Rebecca Cordell[3], Paul S. Monks [ID][3], Neil Greening[1,2], Christopher E. Brightling[1,2], Rachael Evans [ID][1,2] and Salman Siddiqui[1,2], on behalf of the EMBER Consortium[4]

[1]Leicester NIHR Biomedical Research Centre (Respiratory Theme), Glenfield Hospital, Leicester, UK. [2]Dept of Respiratory Sciences, College of Life Sciences, University of Leicester, Leicester, UK. [3]Dept of Chemistry, University of Leicester, Leicester, UK. [4]See Acknowledgements for contributors. [5]These authors contributed equally.

Corresponding author: Salman Siddiqui (ss338@le.ac.uk)

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 airway diseases.

Aims To 1) conduct a systematic review evaluating the diagnostic accuracy of VOCs in diagnosing airway diseases; 2) understand the relationship between reported VOCs and biomarkers of type-2 inflammation; 3) assess the standardisation of reporting according to STARD and TRIPOD criteria; 4) review current methods of breath sampling and analysis.

Methods A PRISMA-oriented systematic search was conducted (January 1997 to 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 four 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 airway 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 airway 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.

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 [1].

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 [2].

Current noninvasive 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 [3] and blood eosinophils used to stratify patients most likely to respond to inhaled corticosteroids in COPD and eosinophil-suppressing biologics therapies in severe asthma [4].

Despite these tools and the availability of consensus guidelines, there remain considerable diagnostic challenges in airway diseases [5]. Approximately 30% of patients diagnosed with asthma in primary care may have an inaccurate diagnosis [6, 7] and biomarkers that stratify high-cost therapies in severe asthma fail to identify 20–25% of patients that will fail therapy [8, 9]. Currently only FeNO is recommended for treatment titration in asthma [10].

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

Over recent years there has been noticeable interest and a growing body of evidence in using exhaled breath volatile organic compounds (VOCs) and breath signatures in the diagnosis, phenotyping and monitoring of airway diseases [11–13]. MONTUSCHI et al. [14] demonstrated that FeNO has a higher diagnostic accuracy (95%) when combined with an electronic nose (eNose) in diagnosing airway inflammation in mild asthmatics, whilst SCHLEICH et al. [15] 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 [16–18].

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 Cochrane Library, Medline and EMBASE evidence databases (online supplementary material). Further details on methodology can be found at https://www.crd.york.ac.uk/prospero/ with identifier number 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 reference 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: 1) human adult asthma and COPD studies; 2) objective evidence of asthma and COPD diagnosed by a clinician in primary or secondary care; 3) 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 rigour of the reported studies were tested using the Standards for Reporting Studies of Diagnostic Accuracy (STARD) for studies identifying diagnostic accuracy of VOCs and the 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 A).

STARD is an instrument aimed at researchers and editors with the purpose of evaluating the quality of diagnostic accuracy studies by checking their methodological rigour against a set of essential criteria [19].

TRIPOD is a 22-item checklist that explicitly covers the development and validation aspects of multivariate prediction models [20].

We also used a modified Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) to formally assess the quality of the studies included [21]. 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 the online supplementary material.

TABLE 1 Included studies and summary of findings in relation to the review objectives.

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) [26]	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) [33]	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 residual volume/total lung capacity). 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) [66]	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) [32]	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) [67]	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) [12]	58	Direct breath sampler (GC-MS)	47 compounds diagnosed asthma (86% accuracy – PPV 0.85, NPV 0.89). Thirteen compounds discriminated uncontrolled asthmatics (Asthma Control Questionnaire⩾1) with 89% accuracy (AUC 0.90) (pentadecane, heptanoic acid, O-xylene, 2-butanone, 3-methylbutanal, 2,6-diisopropylnaphthalene). Eleven compounds were able to discriminate eosinophilic asthmatics with 83% classification accuracy (AUC 0.98) (camphene, cyclohexanone, cyclohexene-4-methylene). Fourteen compounds discriminated neutrophilic asthmatics with 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) [68]	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) [27]	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.

TABLE 1 Continued				
Asthma				
Title	**Author**	**Sample size (n)**	**Breath collection and analysis**	**Summary of findings in relation to review objectives**
A mobile instrumentation platform to distinguish airway disorders	SCHIVO *et al.* (2013) [44]	31	Exhaled breath condensate collector (GC-DMS-FAIMS)	VOCs classified asthmatics 75% of the time, after executing 20 classification optimisation 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) [69]	235	Tedlar bag (GC-ToF-MS)	Sixteen unidentified VOCs diagnosed asthmatics with 100% sensitivity and 91.1% specificity. Cluster analysis based on VOCs and the clinical parameters resulted in seven 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	SCHLEICH *et al.* (2019) [15]	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 0.85). Replication=undecane and nonanal discriminated neutrophilic phenotype (AUC 0.70). Discovery=3,7-dimethylnonane, nonanal and 1-propanol 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) [34]	78	Tedlar bag (eNose)	eNose identified three 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.* (2019) [39]	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 extrinsic asthma and allergic rhinitis 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 recovery in asthma	BRINKMAN *et al.* (2017) [22]	23	Tedlar bag (GC-MS, eNose)	eNose and GC-MS VOCs (methanol, acetonitrile, bicyclo-octan-1-ol, 4-methyl-$C_9H_{16}O$) 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 analyzer	PLAZA *et al.* (2015) [70]	52	Tedlar bag (eNose)	eNose successfully diagnosed different subtypes of asthma. Neutrophilic *versus* paucigranulocytic separated with 89% cross-validation accuracy (AUC=0.88). Neutrophilic *versus* eosinophilic separated with 73% CVA (AUC=0.92). Eosinophilic *versus* paucigranulocytic separated with 74% CVA (AUC=0.79). No reported use of STARD/TRIPOD.
Electronic nose breathprints are independent of acute changes in airway caliber in asthma	LAZAR *et al.* (2010) [24]	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.

TABLE 1 Continued

Asthma

Title	Author	Sample size (n)	Breath collection and analysis	Summary of findings in relation to review objectives
Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label[#]	DE VRIES et al. (2018) [71]	435	Spironose (eNose)	eNose identified clusters in total population of asthma and COPD by clinical/inflammatory characteristics (ethnicity, systemic eosinophilia, systemic neutrophilia, body mass index, 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) [14]	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 et al. (2020) [72]	78	Tedlar bag (GC-ToF-MS)	Four VOCs (lysine, glycolic acid, 4-carene and octanal) were associated with traces of asthma medications in urine of severe asthmatics. Baseline AUC: 82.1 for salbutamol and 78.8 for oral corticosteroids. 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) [17]	30	Tedlar bag (eNose)	eNose diagnosed 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) [41]	132	Direct sampling (MCC/IMS)	Cyclohexanone diagnosed COPD with 60% sensitivity and 91% specificity and PPV 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) [73]	26	Full face mask – directly connected to sensors (GC-IMS and 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) [16]	79	Tedlar bag (GC-ToF-MS)	Six 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) [38]	100	Tedlar bag (eNose)	eNose diagnosed fixed airway obstruction asthma from COPD with 88% accuracy (AUC 0.95) and discriminated ex/nonsmoking asthmatics from COPD with 90% accuracy (AUC 0.96). No relation to markers of type-2 inflammation. Reported use of STARD criteria.

TABLE 1 Continued

COPD

Title	Author	Sample size	Breath collection and analysis	Summary of findings in relation to review objectives
Screening for emphysema via exhaled volatile organic compounds	CRISTESCU et al. (2011) [74]	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) [60]	43	Collection bag (eNose)	eNose is able to separate alpha-1 antitrypsin (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) [40]	130	Direct sampling (MCC/IMS)	Ten 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) [35]	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) [18]	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 ⩾1%, ⩾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-LOZANO SINUES et al. (2014) [75]	61	Direct sampling (quadruple ToF-MS)	Identified VOCs (acetone and indole) were able to discriminate COPD with 96% sensitivity and 72.7% specificity (COPD versus smoking controls), 88% sensitivity and 92% specificity (COPD versus nonsmoking controls) and 92% sensitivity and 83% specificity (GOLD I/II versus 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) [28]	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 versus control (p-value 0.005, sensitivity 70%, specificity 73%). Exacerbation of COPD versus controls (p-value 0.001, sensitivity 68%, specificity 80%). Pneumonia versus controls (p-value 0.005, sensitivity 88%, specificity 100%). Stable COPD versus pneumonia (p-value <0.001, sensitivity 88%, specificity 75%). No relation to markers of type-2 inflammation. No reported use of STARD/TRIPOD.

TABLE 1 Continued

COPD

Title	Author	Sample size	Breath collection and analysis	Summary of findings in relation to review objectives
Analysis of exhaled breath fingerprints and volatile organic compounds in COPD	Cazzolla et al. (2015) [36]	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) [11]	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 center study to compare breath volatile organic compounds from smokers and non-smokers with and without COPD	Gaida et al. (2016) [48]	190	Direct sampling (thermal desorption GC-MS)	Linear discriminant analysis correctly classified 89.4% of COPD patients in the non/ex-smoking group (CV 85.6%) of COPD patients in the actively smoking group (CV 77.9%). Ten 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) [76]	58	Direct sampling (GC-IMS, GCAPCI-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 three unidentified compounds compared to non/ex-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) [42]	157	Bio-VOC collector (GC-MS)	Hexanal discriminated COPD from healthy nonsmoking controls (p=0.015, 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=0.008, 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) [29]	54	Direct sampling via mouthpiece onto glass syringe (GC-ToF-MS)	VOCs indicative of acute exacerbations: cyclohexanone (p<0.001), n-butane (p<0.001), 4-heptanone (p=0.001), 2-pentanone (p=0.002), sensitivity 0.69, NPV 0.82, specificity 0.94, PPV 0.89, AUC 0.92. COPD specific: n-heptane (p<0.001), methyl propyl sulfide (p<0.001). Undefined association to COPD: dimethyl disulfide (p<0.001), 6-methyl-5-heptene-2-one (p<0.001). No relation to markers of type-2 inflammation. No reported use of STARD/TRIPOD.

TABLE 1 Continued				
COPD				
Title	**Author**	**Sample size**	**Breath collection and analysis**	**Summary of findings in relation to review objectives**
Real-time mass spectrometric identification of metabolites characteristic of chronic obstructive pulmonary disease in exhaled breath	BREGY *et al.* (2018) [77]	36	Direct sampling (SESI-HRMS)	SESI-HRMS can diagnose COPD with 89% accuracy, 93% sensitivity, 87% specificity, 81% PPV and 95% NPV. 11-hydroxyundecanoic acid correlates with FEV_1: forced expiratory volume in 1 s and FEV_1/FVC ratio (q=0.04 and 0.02, respectively) with positive correlation coefficients r=0.48 and 0.48, respectively). 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 *et al.* (2018) [25]	55	Direct sampling – Pneumopipe® (eNose)	Breath print patterns change in distinctive ways depending upon whether COPD patients have been prescribed long-acting muscarinic antagonists and/or long-acting beta-agonists alone or any combination of inhaled drugs including inhaled corticosteroids. 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 *et al.* (2014) [43]	181	Bio-VOC collector (GC-MS)	A substantial variation was found in the levels of VOCs 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 *et al.* (2012) [78]	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) [37]	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-acetylpyridine, 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	VAN VELZEN *et al.* (2019) [31]	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 *versus* exacerbations with 71% accuracy and sensitivity and specificity of 0.71. Exacerbation *versus* follow-up: accuracy of 75%, sensitivity of 0.92 and specificity of 0.57. Baseline *versus* 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.

TABLE 1 Continued				
COPD				
Title	**Author**	**Sample size**	**Breath collection and analysis**	**Summary of findings in relation to review objectives**
Real-time breath analysis reveals specific metabolic signatures of COPD exacerbations	GAUGG et al. (2019) [30]	52	Direct sampling (SESI-HRMS)	Metabolite levels from the ω-oxidation pathway, namely ω-hydroxy, ω-oxo and dicarboxylic acids were consistently decreased in frequent exacerbators; AUC 0.88. No relation to markers of type-2 inflammation. No reported use of STARD/TRIPOD.

AUC: area under the curve; CV/CVA/CVV: coefficient of variation; eNose: electronic nose; FeNO: fractional exhaled nitric oxide; FEV_1 (% pred): forced expiratory volume in 1 s (% of predicted); FVC: forced vital capacity; GC-APCI-CMS: gas chromatography-atmospheric pressure chemical ionisation mass spectrometry; GC-DMS: gas chromatography-differential mass spectrometry; GC-DMS-FAIMS: gas chromatography-differential mass spectrometry-field asymmetric ion mobility spectrometry; GC-FID: gas chromatography-flame ionisation detection; GC-IMS: gas chromatography-ion mobility spectrometry; GC-MS: gas chromatography-mass spectrometry; GC-ToF-MS: gas chromatography-time of flight mass spectrometry; GCxGC-MS: comprehensive two-dimensional gas chromatography-mass spectrometry; GOLD: Global Initiative for COPD; GORD: gastro-oesophageal reflux disease; IMS: ion mobility spectrometry; MCC/IMS: multi-capillary column-ion mobility spectrometry; NPV: negative predictive value; NSCLC: nonsmall cell lung cancer; PPV: positive predictive value; PTR-MS: proton transfer reaction-mass spectrometry; SESI-HRMS: secondary electrospray ionisation-high resolution mass spectrometry; STARD: Standards for Reporting Studies of Diagnostic Accuracy; TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; VOC: volatile organic compound. #: this study included both asthma and COPD.

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 include 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 (fig. 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. Fourteen 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 two studies had a combination of 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 (fig. 2). These include two studies [22, 23] that focused on assessing changes in VOC profile following asthma loss of control, using both GC-MS and eNose; LAZAR et al. [24] who demonstrated no significant changes to molecular breath profile following induced changes to airway calibre and SCARLATA et al. [25] 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; three asthma studies [22, 26, 27] and four COPD [28–31] studies examined VOCs during acute exacerbations.

Diagnostic accuracy and monitoring studies in asthma and COPD
Asthma diagnostic accuracy studies
Several groups have attempted to explore the diagnostic potential of exhaled breath VOCs in differentiating asthmatics from healthy volunteers. DRAGONIERI et al. [32] reported in a cross-sectional study that exhaled breath signatures could distinguish asthmatics from healthy volunteers but were unable to differentiate between asthma severities (mild and severe asthma). In contrast, PAREDI et al. [33] showed that ethane levels were found to be significantly higher in severe asthma compared to mild asthma.
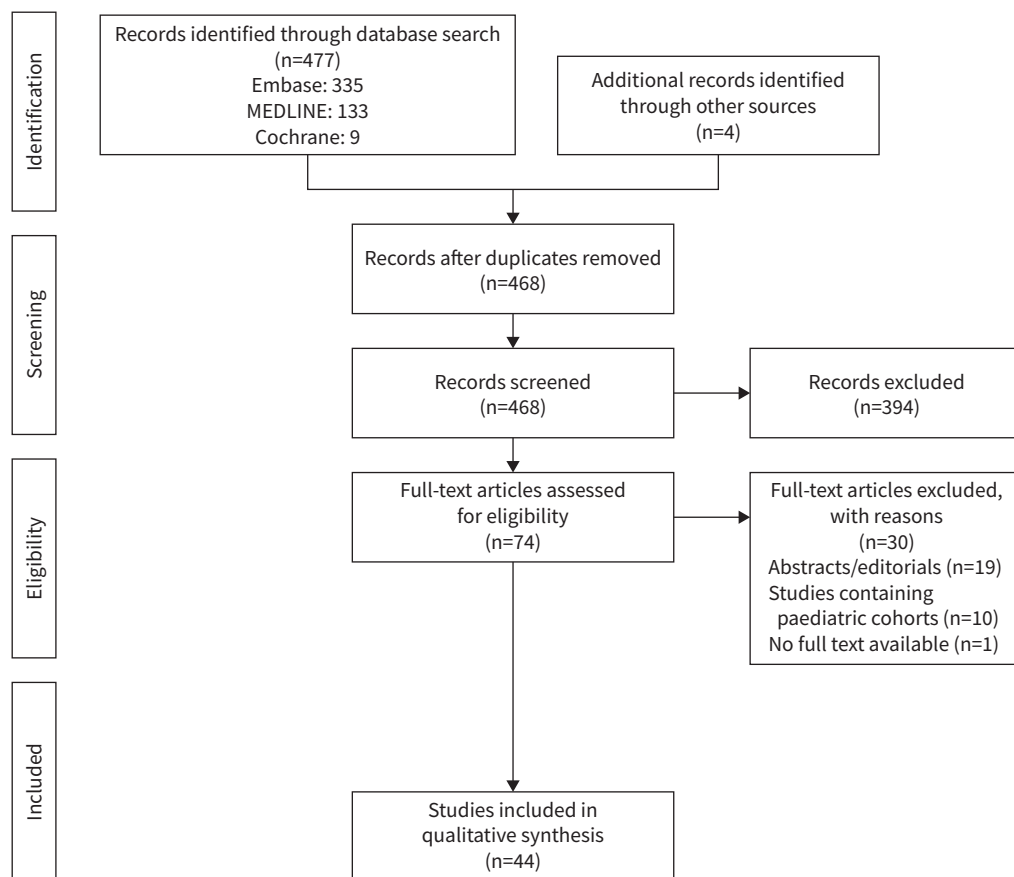
FIGURE 1 PRISMA flow chart illustrating article selection (modified from MOHER *et al.* [65]); for more information see www.prisma-statement.org.
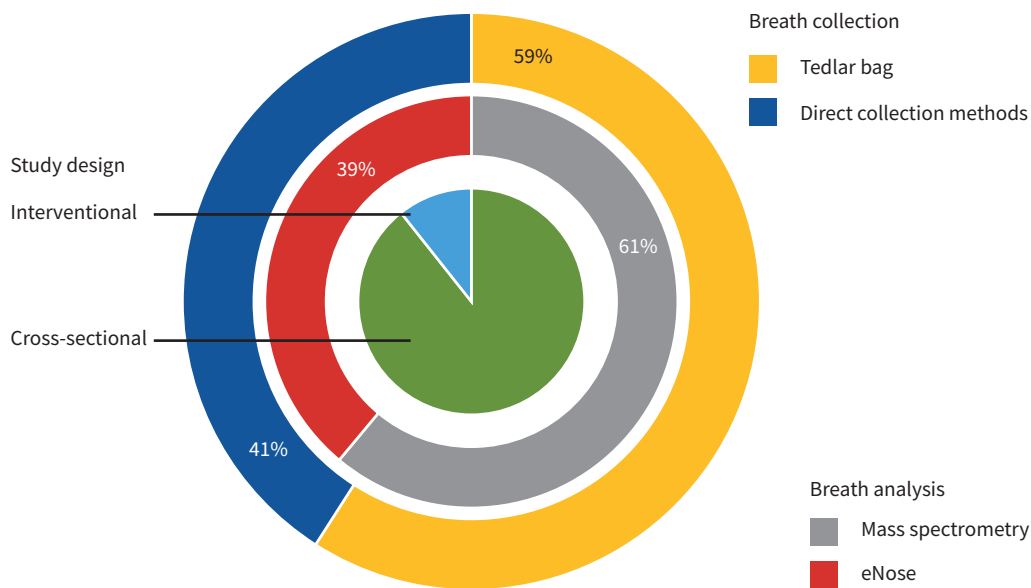
FIGURE 2 Distribution of study designs, and breath collection and analysis technologies across the identified studies.

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. [26], who showed that the levels of pentane were increased in acute asthma but following the acute episode the levels of pentane decreased using GC and Tedlar bag. VAN DER SCHEE et al. [23] 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. [22] 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 [34]. 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 [23]. 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. [24] 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 [35]. Similarly, another study reported that using eNose, breath prints can distinguish between COPD and healthy controls [36]; this study also used GC-MS and detected two VOCs (decane, 6-ethyl-2-methyldecane) that were discriminant for patients with a diagnosis of COPD. RODRIGUEZ-AGUILAR et al. [37] demonstrated that COPD can be differentiated from healthy controls using an ultrafast gas chromatography system equipped with an electronic nose detector (FGC eNose), with modest accuracy. Seventeen VOCs, detailed in table 1 and the online supplementary material, have driven this separation.

Acute COPD exacerbation and monitoring studies

The largest exacerbation study was carried out by SHAFIEK et al. [28], 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=0.068); this was particularly evident in the presence of airway infection or pneumonia. Similarly, PIZZINI et al. [29] evaluated COPD exacerbations and, using GC-MS, identified n-butane, 2-pentanone, cyclohexanone and 4-heptanone as discriminatory biomarkers.

VAN VELZEN et al. [31] 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 versus exacerbations and 78% for exacerbation versus recovery.

Further to the aforementioned exacerbation studies, GAUGG et al. [30] 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.

All exacerbation studies show 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 VOCs with clinical measures of type-2 inflammation in asthma

Few studies have explored the relationship between exhaled breath VOCs and clinical indices of type-2 inflammation. MONTUSCHI et al. [14] 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. [33] found elevated levels of ethane in patients with asthma, which was significantly correlated with nitric oxide. IBRAHIM et al. [12] demonstrated that exhaled breath VOC profiles can identify patients with a clinically important level of sputum eosinophilia and differentiate between eosinophilic and noneosinophilic subjects using GC-MS.

One of the largest studies in asthma to date evaluating airway inflammation was carried out by SCHLIECH et al. [15], 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 of the 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 VOCs 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. [18] 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 reported using the STARD guidelines to test the diagnostic accuracy of the discovered exhaled breath signatures [15, 36, 38]. Similarly, none of the studies reported their multivariate models according to TRIPOD guidance. Twenty-two out of 44 studies reported markers of model accuracy combined with internal or model replication and only two studies [38, 39] reported external model replication in independent populations.
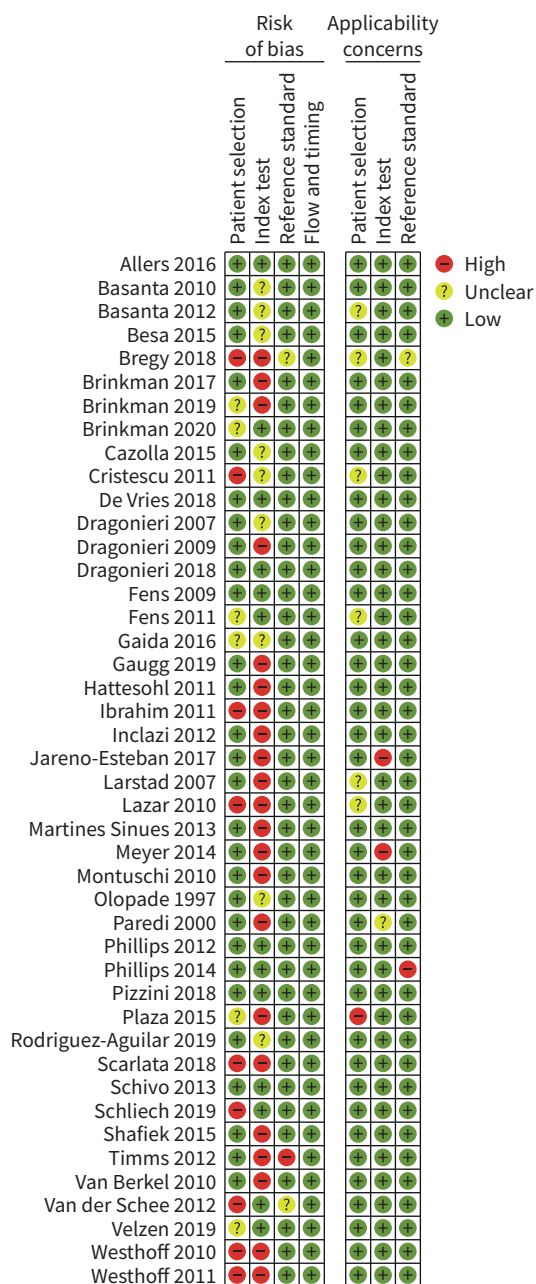
Assessment of risk of bias and applicability using QUADAS-2, detailed by study, can be found in fig. 4 and table A, figure A of the online supplementary material. The main sources of bias were: 1) index test; 19/44 of the published studies failed to validate their results internally or externally, increasing their risk of bias; 2) patient selection; 12/44 studies appeared to have high or unclear concerns due to the absence of a comparable control group; 3) 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.
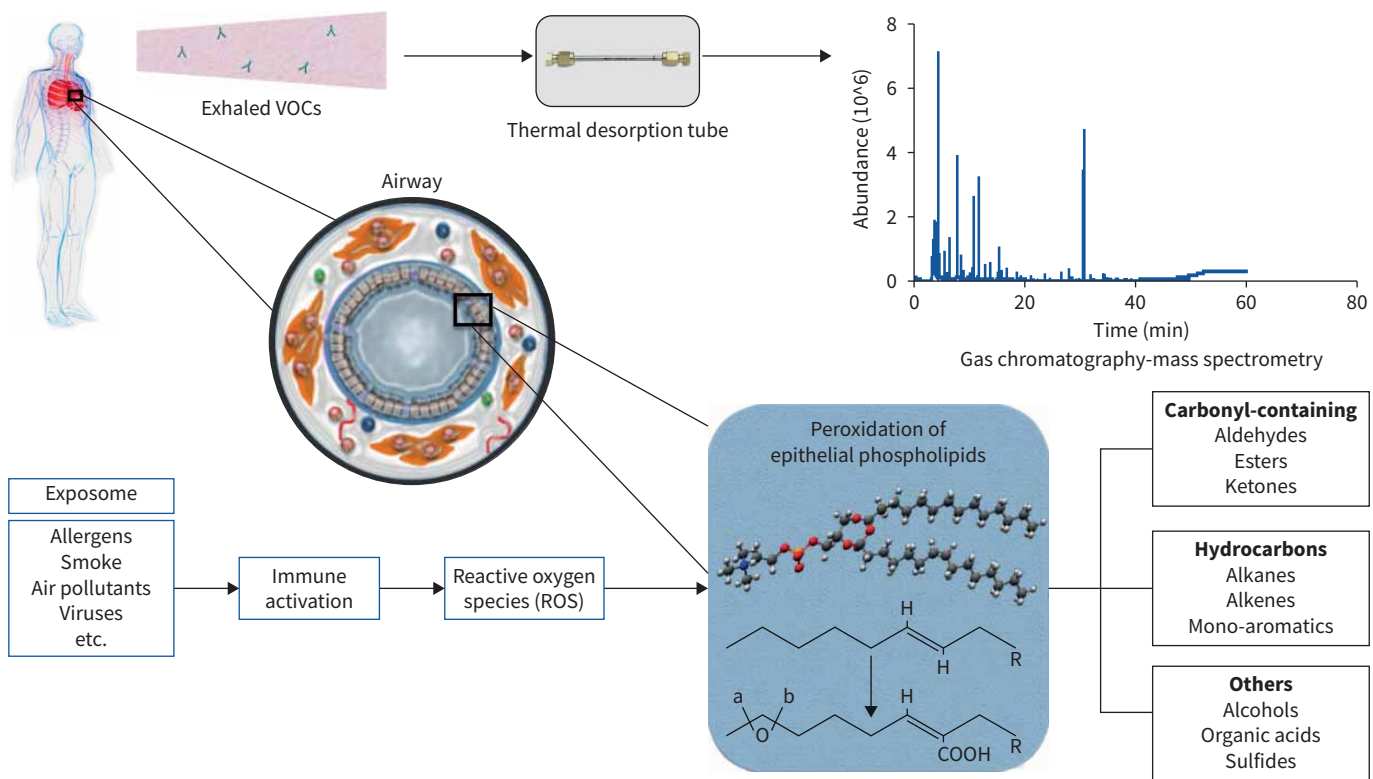
FIGURE 3 Top panel: The precise cellular metabolic processes that underpin the majority of reported VOCs described in breath using GC-MS have yet to be determined. Bottom panel: Simplified illustration of the different causes of lipid peroxidation and the resultant formation of aldehydes and alkanes as well as other VOC classes, which have been consistently described across several breath volatile association studies in asthma and COPD (table 1 and the online supplementary material).

Current knowledge gaps and remaining challenges

This review has highlighted notable shortcomings in transparent reporting when following STARD and TRIPOD criteria. It is worth noting that noncompliance 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 rigour 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, and enhance adjudication and replication of discovered VOCs.

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

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%) (fig. 2). Fifteen studies described the direct connection of a Tedlar bag to eNose sensors and 10 studies transferred the samples onto a thermal desorption (TD) tube as a medium prior to analysis.

Eighteen (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) [12, 40, 41] advised participants to breath directly, *via* a mouthpiece, to the analytical instrument. Other direct collection devices included Bio-VOC collectors [35, 42, 43], PneumoPipe® [28] and exhaled breath condensate collectors (RTube™) [44].

Breath VOC quantification

Various technologies were used to collect and analyse VOCs, includes eNose (19 studies); GC combined with MS, flame ionisation detection (FID) or differential mass spectrometry (DMS) (13 studies); gas chromatography-time of flight mass spectrometry (GC-ToF-MS) (six studies); comprehensive two-dimensional gas chromatography and mass spectrometry (GCxGS-MS) (one study); IMS (five studies); PTR-ToF-MS (one study) (table 1).

Current knowledge gaps and remaining challenges

The review has 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 [45, 46] and novel analytical methodologies applied to breath, such as GCxGC-MS, that may offer superior sensitivity in volatile metabolite identification [15]. Standardised protocols will need to be carefully designed to accommodate the emerging methodologies.

Reported VOCs

The reported VOCs from the studies evaluated are described in table C of the online supplementary material. Not all studies reported the actual names of the VOCs that were identified. Twenty-one out of 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 [47].

Several compounds have been reported more than once. Nonanal was first described by BASANTA *et al.* [18] 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.* [42]. Interestingly, in 2019 SCHLIECH *et al.* [15] 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 [18, 35, 42].

Seven studies reported VOCs of alcohol and phenol classifications [12, 15, 21, 22, 35, 36, 48] with poor consistency among reported compounds. Other chemical classifications identified in small numbers included sulfides, 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 [35].

Discussion

We have provided an overview of all published studies on breath analysis in two major airway 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 [49–52]. Recently a study by HENDERSON *et al.* [53] 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 three 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, they provide researchers with the means to critically appraise published results against potential bias [54]. None of the studies reported multivariate models using TRIPOD criteria. TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) is an essential 22-item checklist aimed at improving studies reporting development or validation of diagnostic or prognostic multivariate models [55]. Only with transparent reporting, following a set model, can the strengths and weaknesses of discovery studies be brought to light, therefore facilitating interpretation and future adoption. Large-scale, real-world, multi-platform breath discovery studies, following the aforementioned construct, such as the one by IBRAHIM et al. [56], are currently under way and further results are eagerly anticipated.

In this review, the QUADAS-2 tool highlighted clear bias, including patient selection; nine studies displayed high concerns with regards to their patient selection criteria, mainly due to the absence of comparable positive controls and inappropriate exclusion of certain subgroups 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 four studies failed to report a 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 multicentre studies to ensure the possibility of an accurate and reliable generalisation of the results [57]. 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 was the most common method of breath sampling. Although validated, this method is not without its shortcomings. BEAUCHAMP et al. [58] described several potential limitations including: 1) decay of certain groups of compounds, which significantly limits sample storage time; 2) 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. [58] 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, sampling and storage procedures when using Tedlar bags. These problems have been mitigated by the use of direct breath sampling devices, which 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 [59].

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 (fig. 5).

Numerous validated studies using eNose have emerged over the last decade [31, 60–62]. 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 [63]. VOCs that can be identified from 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 [64]. In a similar manner exogenous VOCs and labelled compounds may be used to identify specific cellular or microbial activity in airway 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 noninvasive biomarkers for airway diseases, and correlate with clinical biomarkers of type-2 inflammation. Many of the biomarkers
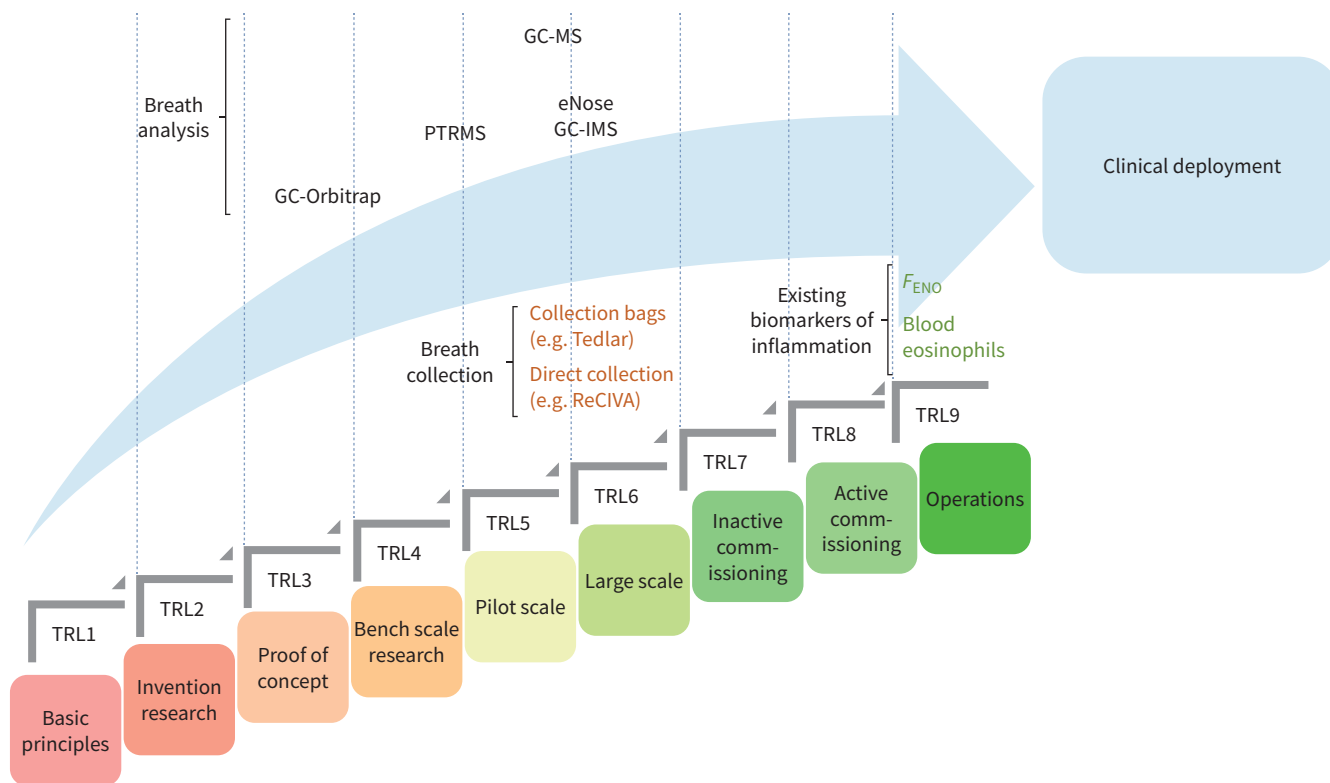
FIGURE 5 Qualitative assessment of breath sampling and analytical technologies level on a nine-point technology readiness level (TRL) scale (adapted from TRL guidance published by the Nuclear Decommissioning Authority, NDA), comparative to existing biomarkers used in clinical practice for airway disease stratification. GC: gas chromatography; MS: mass spectrometry; IMS: ion mobility spectrometry; F_{ENO}: exhaled nitric oxide fraction.

reviewed in our study appear to be associated with lipid peroxidation in the airway, suggesting a role for VOCs as diagnostic biomarkers and markers of oxidant-induced airway inflammation.

These observations highlight a conceivable role for VOCs as diagnostic and monitoring tools for airway 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   GBD 2015 Chronic Respiratory Disease Collaborators. 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: 691–706.

2   Wells CD, Joo MJ. COPD and asthma: diagnostic accuracy requires spirometry. *J Fam Pract* 2019; 68: 76–81.

3   Price D, Ryan D, Burden A, *et al.* Using fractional exhaled nitric oxide (FeNO) to diagnose steroid-responsive disease and guide asthma management in routine care. *Clin Transl Allergy* 2013; 3: 37.

4   Kostikas K, Brindicci C, Patalano F. Blood eosinophils as biomarkers to drive treatment choices in asthma and COPD. *Curr Drug Targets* 2018; 19: 1882–1896.

5   Tinkelman DG, Price DB, Nordyke RJ, *et al.* Misdiagnosis of COPD and asthma in primary care patients 40 years of age and over. *J Asthma* 2006; 43: 75–80.

6   Magnoni MS, Caminati M, Senna G, *et al.* Asthma under/misdiagnosis in primary care setting: an observational community-based study in Italy. *Clin Mol Allergy* 2015; 13: 26.

7   Kavanagh J, Jackson DJ, Kent BD. Over- and under-diagnosis in asthma. *Breathe* 2019; 15: e20–e27.

8   Kim H, Ellis AK, Fischer D, *et al.* Asthma biomarkers in the age of biologics. *Allergy, Asthma Clin Immunol* 2017; 13: 48.

9   Pavord ID, Afzalnia S, Menzies-Gow A, *et al.* The current and future role of biomarkers in type 2 cytokine-mediated asthma management. *Clin Exp Allergy* 2017; 47: 148–160.

10   National Institute for Health and Care Excellence. Asthma: diagnosis, monitoring and chronic asthma management (NICE guideline [NG80]). 2020. www.nice.org.uk/guidance/ng80.

11   Besa V, Teschler H, Kurth I, *et al.* Exhaled volatile organic compounds discriminate patients with chronic obstructive pulmonary disease from healthy subjects. *Int J Chron Obstruct Pulmon Dis* 2015; 10: 399–406.

12   Ibrahim B, Basanta M, Cadden P, *et al.* Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 2011; 66: 804–809.

13   Samara MA, Paschke KM, Tang WHW, *et al.*, Single breath exhaled gas analysis by selected ion-flow tube mass-spectrometry (SIFT-MS) in patients with acute decompensated heart failure. *J Cardiac Fail* 2011; 17: S4.

14   Montuschi P, Santonico M, Mondino C, *et al.* Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma. *Chest* 2010; 137: 790–796.

15   Schleich FN, Zanella D, Stefanuto P-H, *et al.* Exhaled volatile organic compounds are able to discriminate between neutrophilic and eosinophilic asthma. *Am J Respir Crit Care Med* 2019; 201: A2373.

16   Van Berkel JJBN, Dallinga JW, Möller GM, *et al.* A profile of volatile organic compounds in breath discriminates COPD patients from controls. *Respir Med* 2010; 104: 557–563.

17   Dragonieri S, Annema JT, Schot R, *et al.* An electronic nose in the discrimination of patients with non-small cell lung cancer and COPD. *Lung Cancer* 2009; 64: 166–170.

18   Basanta M, Ibrahim B, Dockry R, *et al.* Exhaled volatile organic compounds for phenotyping chronic obstructive pulmonary disease: a cross-sectional study. *Respir Res* 2012; 13: 72.

19   Cohen JF, Korevaar DA, Altman DG, *et al.* STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; 6: e012799.

20   Moons KGM, Altman DG, Reitsma JB, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; 162: W1–W73.

21   Whiting PF, Rutjes AW, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.

22   Brinkman P, van de Pol MA, Gerritsen MG, *et al.* Exhaled breath profiles in the monitoring of loss of control and clinical recovery in asthma. *Clin Exp Allergy* 2017; 47: 1159–1169.

23   van Der Schee MP, Liley J, Palmay R, *et al.* Predicting steroid responsiveness in patients with asthma using the electronic nose. *Am J Respir Crit Care Med* 2012; 185: A3946.

24   Lazar Z, Fens N, van der Maten J, *et al.* Electronic nose breathprints are independent of acute changes in airway caliber in asthma. *Sensors (Basel)* 2010; 10: 9127–9138.

25   Scarlata S, Finamore P, Santangelo S, *et al.* Cluster analysis on breath print of newly diagnosed COPD patients: effects of therapy. *J Breath Res* 2018; 12: 036022.

26   Olopade CO, Zakkar M, Swedler WI, *et al.* Exhaled pentane levels in acute asthma. *Chest* 1997; 111: 862–865.

27   van der Schee MP, Palmay R, Cowan JO, *et al.* Predicting steroid responsiveness in patients with asthma using exhaled breath profiling. *Clin Exp Allergy* 2013; 43: 1217–1225.

28   Shafiek H, Fiorentino F, Merino JL, *et al.* Using the electronic nose to identify airway infection during COPD exacerbations. *PLoS ONE* 2015; 10: e0135199.

29   Pizzini A, Filipiak W, Wille J, *et al.* Analysis of volatile organic compounds in the breath of patients with stable or acute exacerbation of chronic obstructive pulmonary disease. *J Breath Res* 2018; 12: 036002.

30   Gaugg MT, Nussbaumer-Ochsner Y, Bregy L, *et al.* Real-time breath analysis reveals specific metabolic signatures of COPD exacerbations. *Chest* 2019; 156: 269–276.

31   van Velzen P, Brinkman P, Knobel HH, *et al.* Exhaled breath profiles before, during and after exacerbation of COPD: a prospective follow-up study. *COPD* 2019; 16: 330–337.

32   Dragonieri S, Schot R, Mertens BJ, *et al.* An electronic nose in the discrimination of patients with asthma and controls. *J Allergy Clin Immunol* 2007; 120: 856–862.

33 Paredi P, Kharitonov SA, Barnes PJ. Elevation of exhaled ethane concentration in asthma. *Am J Respir Crit Care Med* 2000; 162: 4 Pt 1, 1450–1454.

34 Brinkman P, Wagener AH, Hekking P-P, *et al.* Identification and prospective stability of electronic nose (eNose)-derived inflammatory phenotypes in patients with severe asthma. *J Allergy Clin Immunol* 2019: 143: 1811–1820.e7.

35 Phillips CO, Syed Y, Parthaláin NM, *et al.* Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls. *J Breath Res* 2012; 6: 036003.

36 Cazzola M, Segreti A, Capuano R, *et al.* Analysis of exhaled breath fingerprints and volatile organic compounds in COPD. *COPD Res Pract* 2015; 1: 7.

37 Rodríguez-Aguilar M, Ramirez-Garcia S, Ilizaliturri-Hernández C, *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. *Biomed Chromatogr* 2019; 33: e4684.

38 Fens N, Roldaan AC, van der Schee MP, *et al.* External validation of exhaled breath profiling using an electronic nose in the discrimination of asthma with fixed airways obstruction and chronic obstructive pulmonary disease. *Clin Exp Allergy* 2011; 41: 1371–1378.

39 Dragonieri S, Quaranta VN, Carratu P, *et al.* Exhaled breath profiling by electronic nose enabled discrimination of allergic rhinitis and extrinsic asthma. *Biomarkers* 2019; 24: 70–75.

40 Westhoff M, Litterst P, Maddula S, *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. *Int J Ion Mobil Spectrom* 2011; 14: 139–149.

41 Westhoff M, Litterst P, Maddula S, *et al.* Differentiation of chronic obstructive pulmonary disease (COPD) including lung cancer from healthy control group by breath analysis using ion mobility spectrometry. *Int J Ion Mobil Spectrom* 2010; 13: 131–139.

42 Jareno-Esteban JJ, Munoz-Lucas MA, Gomez-Martin O, *et al.* Study of 5 volatile organic compounds in exhaled breath in chronic obstructive pulmonary disease. *Arch Bronconeumol* 2017; 53: 251–256.

43 Phillips C, Mac Parthaláin N, Syed Y, *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: 300–318.

44 Schivo M, Seichter F, Aksenov AA, *et al.* A mobile instrumentation platform to distinguish airway disorders. *J Breath Res* 2013; 7: 017113.

45 Kitchen S, Edge A, Smith R, *et al.* Breathe free: open source development of a breath sampler by a consortium of breath researchers. *Eur Respir J* 2015; 46: Suppl. 59, PA3987.

46 Holden KA, Ibrahim W, Salman D, *et al.* Use of the ReCIVA device in breath sampling of patients with acute breathlessness: a feasibility study. *ERJ Open Res* 2020; 6: 00119-2020.

47 Gardner HW. Oxygen radical chemistry of polyunsaturated fatty acids. *Free Radic Biol Med* 1989; 7: 65–86.

48 Gaida A, Holz O, Nell C, *et al.* A dual center study to compare breath volatile organic compounds from smokers and non-smokers with and without COPD. *J Breath Res* 2016; 10: 026006.

49 Azim A, Barber C, Dennison P, *et al.* Exhaled volatile organic compounds in adult asthma: a systematic review. *Eur Respir J* 2019; 54: 1900056.

50 Catino A, de Gennaro G, Di Gilio 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: 831.

51 Christiansen A, Davidsen JR, Titlestad I, *et al.* A systematic review of breath analysis and detection of volatile organic compounds in COPD. *J Breath Res* 2016; 10: 034002.

52 Töreyin ZN, Ghosh M, Göksel Ö, *et al.* Exhaled breath analysis in diagnosis of malignant pleural mesothelioma: systematic review. *Int J Environ Res Public Health* 2020; 17: 1110.

53 Henderson B, Ruszkiewicz DM, Wilkinson M, *et al.* A benchmarking protocol for breath analysis: the peppermint experiment. *J Breath Res* 2020; 14: 046008.

54 Bossuyt PM, Reitsma JB, Bruns DE, *et al.* STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015; 351: h5527.

55 Collins GS, Reitsma JB, Altman DG, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. *Br J Surg* 2015; 102: 148–158.

56 Ibrahim W, Wilde M, Cordell R, *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: e025486.

57 Fijten RRR, Smolinska A, Drent M, *et al.* The necessity of external validation in exhaled breath research: a case study of sarcoidosis. *J Breath Res* 2017; 12: 016004.

58 Beauchamp J, Herbig J, Gutmann R, *et al.* On the use of Tedlar(R) bags for breath-gas sampling and analysis. *J Breath Res* 2008; 2: 046001.

59 Harshman SW, Mani N, Geier BA, *et al.* Storage stability of exhaled breath on Tenax TA. *J Breath Res* 2016; 10: 046008.

60 Hattesohl ADM, Jörres RA, Dressel H, et al. Discrimination between COPD patients with and without alpha 1-antitrypsin deficiency using an electronic nose. Respirology 2011; 16: 1258–1264.

61 Scarlata S, Santangelo S, Pennazza G, et al. Effects of topical therapy on the breath pattern of patients with newly diagnosed COPD: an e-nose based study. Eur Respir J 2015; 46: PA2099.

62 Finamore P, Pedone C, Scarlata S, et al. Validation of exhaled volatile organic compounds analysis using electronic nose as index of COPD severity. Int J Chron Obstruct Pulmon Dis 2018; 13: 1441–1448.

63 Fens N, van der Schee MP, Brinkman P, et al. Exhaled breath analysis by electronic nose in airways disease. Established issues and key questions. Clin Exp Allergy 2013; 43: 705–715.

64 Hawtin PR. Serology and urea breath test in the diagnosis of H. pylori infection. Mol Biotechnol 1999; 11: 85–92.

65 Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic reviews and Meta-Analyses: the PRISMA statement. PLoS Med 2009; 6: e1000097.

66 Larstad MA, Torén K, Bake B, 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: 87–98.

67 Fens N, Zwinderman AH, van der Schee MP, et al. Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma. Am J Respir Crit Care Med 2009; 180: 1076–1082.

68 Timms C, Thomas PS, Yates DH. Detection of gastro-oesophageal reflux disease (GORD) in patients with obstructive lung disease using exhaled breath profiling. J Breath Res 2012; 6: 016003.

69 Meyer N, Dallinga JW, Nuss SJ, et al. Defining adult asthma endotypes by clinical features and patterns of volatile organic compounds in exhaled air. Respir Res 2014; 15: 136.

70 Plaza V, Crespo A, Giner J, et al. Inflammatory asthma phenotype discrimination using an electronic nose breath analyzer. J Investig Allergol Clin Immunol 2015; 25: 431–437.

71 de Vries R, Dagelet YWF, Spoor P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. Eur Respir J 2018; 51: 1701817.

72 Brinkman P, Ahmed WM, Gómez C, et al. Exhaled volatile organic compounds as markers for medication use in asthma. Eur Respir J 2020; 55: 1900544.

73 Basanta M, Jarvis RM, Xu Y, 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: 315–320.

74 Cristescu SM, Gietema HA, Blanchet L, et al. Screening for emphysema via exhaled volatile organic compounds. J Breath Res 2011; 5: 046009.

75 Martinez-Lozano Sinues P, Meier L, Berchtold C, et al. Breath analysis in real time by mass spectrometry in chronic obstructive pulmonary disease. Respiration 2014; 87: 301–310.

76 Allers M, Langejuergen J, Gaida A, 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. J Breath Res 2016; 10: 026004.

77 Bregy L, Nussbaumer-Ochsner Y, Sinues PM-L, et al. Real-time mass spectrometric identification of metabolites characteristic of chronic obstructive pulmonary disease in exhaled breath. Clin Mass Spectrom 2018; 7: 29–35.

78 Incalzi RA, Pennazza G, Scarlata S, et al. Reproducibility and respiratory function correlates of exhaled breath fingerprint in chronic obstructive pulmonary disease. PLoS ONE 2012; 7: e45396.