Air pollution, metabolites and respiratory health across the life-course

Olena Gruzieva[1,2,16], Ayoung Jeong[3,4,16], Shizhen He[1], Zhebin Yu[1], Jeroen de Bont[1], Maria G.M. Pinho [5], Ikenna C. Eze [3,4], Sara Kress[6], Craig E. Wheelock [7,8,9], Annette Peters [10], Jelle Vlaanderen [11], Kees de Hoogh[3,4], Augustin Scalbert[12], Marc Chadeau-Hyam[11,13], Roel C.H. Vermeulen[11], Ulrike Gehring [11,17], Nicole Probst-Hensch [3,4,17] and Erik Melén [14,15,17]

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [2]Centre for Occupational and Environmental Medicine, Region Stockholm, Stockholm, Sweden. [3]Swiss Tropical and Public Health Institute, Basel, Switzerland. [4]University of Basel, Basel, Switzerland. [5]Dept of Epidemiology and Data Science, Amsterdam Public Health, Amsterdam UMC, location VUmc, Amsterdam, The Netherlands. [6]IUF-Leibniz Research Institute for Environmental Medicine, Düsseldorf, Germany. [7]Unit of Integrative Metabolomics, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [8]Dept of Respiratory Medicine and Allergy, Karolinska University Hospital, Stockholm, Sweden. [9]Gunma University Initiative for Advanced Research (GIAR), Gunma University, Maebashi, Japan. [10]Institute of Epidemiology, Helmholz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany. [11]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands. [12]Nutrition and Metabolism Branch, International Agency for Research on Cancer (IARC/WHO), Lyon, France. [13]Imperial College London, London, UK. [14]Dept of Clinical Science and Education Södersjukhuset, Karolinska Institutet, Stockholm, Sweden. [15]Sachs Children's Hospital, Stockholm, Sweden. [16]Both authors contributed equally to this article. [17]These authors contributed equally to this article.

Corresponding author: Olena Gruzieva (olena.gruzieva@ki.se)

Shareable abstract (@ERSpublications)
Metabolomics offers the potential to identify biomarkers linked to both environmental exposures and respiratory health. Studies with prospectively collected biosamples and harmonised exposure assessment are needed for insights into mechanisms and causality. https://bit.ly/3MRF2iq

Cite this article as: Gruzieva O, Jeong A, He S, et al. Air pollution, metabolites and respiratory health across the life-course. Eur Respir Rev 2022; 31: 220038 [DOI: 10.1183/16000617.0038-2022].

Abstract

Previous studies have explored the relationships of air pollution and metabolic profiles with lung function. However, the metabolites linking air pollution and lung function and the associated mechanisms have not been reviewed from a life-course perspective. Here, we provide a narrative review summarising recent evidence on the associations of metabolic profiles with air pollution exposure and lung function in children and adults. Twenty-six studies identified through a systematic PubMed search were included with 10 studies analysing air pollution-related metabolic profiles and 16 studies analysing lung function-related metabolic profiles. A wide range of metabolites were associated with short- and long-term exposure, partly overlapping with those linked to lung function in the general population and with respiratory diseases such as asthma and COPD. The existing studies show that metabolomics offers the potential to identify biomarkers linked to both environmental exposures and respiratory outcomes, but many studies suffer from small sample sizes, cross-sectional designs, a preponderance on adult lung function, heterogeneity in exposure assessment, lack of confounding control and omics integration. The ongoing EXposome Powered tools for healthy living in urbAN Settings (EXPANSE) project aims to address some of these shortcomings by combining biospecimens from large European cohorts and harmonised air pollution exposure and exposome data.

Introduction

The exposome concept encompasses measures of both external and internal exposures and is thought to reflect the sum of all nongenetic influences on health and disease [1–3]. Typically, the external exposome domains include factors related to ambient environmental exposures, such as air pollution; temperature; built, social and food environments; and behaviour-related and psychosocial factors, whereas the internal exposome refers to internalised exogenous exposures that comprise the fraction of dietary and nondietary environmental molecules that have entered the organism. The internal exposome also includes the gut

microbiota and the associated metabolites, arising directly from the biotransformation of environmental and endogenous metabolites and endogenous metabolites reflecting responses to exposures (*e.g.* lipid peroxidation products from oxidative stress) [3].

In recent years, it has been widely acknowledged that both environmental exposures and host characteristics influence respiratory health and disease in both children and adults [4–6]. However, little is known about the life-course perspective of exposure and disease, the critical windows of exposure, the early biomarkers of exposure and disease mechanisms, and disease progression and prognosis. Accordingly, causal inference regarding the health effects of single exposures or mixtures remains limited, in particular regarding chronic effects.

The life-course perspective is of particular relevance for lung function-related respiratory diseases. Lung function progresses over the life-course, going through growth, plateau and decline phases [7]. Timing and duration of the phases, and lung function levels at each phase, differ between individuals, shaping distinct lung function trajectories [8]. Genetic and environmental factors have been associated with these trajectories. Today, it is well accepted that lung function attained in the growth phase and maintained in the plateau phase affects respiratory (and cardiovascular) health in later life [8, 9]. COPD incidence is attributed not only to accelerated decline in lung function in later life but also to impaired attainment of lung function in earlier life [10, 11]. Early life exposures (*e.g.* maternal smoking) and diseases such as childhood asthma can alter lung development, affecting the trajectories in the growth phase. Factors in later life, such as personal smoking and adult asthma, may affect lung function decline or aggravate the effects of early life exposures. The individual effects of environmental exposures on respiratory health may differ across the life-course, potentially with distinct aetiologies at susceptible time windows [12–14], thereby complicating the understanding of the causality of effects.

Identifying metabolic profiles in different phases of life may inform us about the distinctive aetiology in each phase. Untargeted metabolomics analysis has emerged as a potentially powerful approach to investigate the biological responses of exposure to environmental pollutants. The metabolome reflects the end products of genetic and endogenous processes in response to environmental exposures and, therefore, offers the potential to investigate the aetiology of diseases arising from exposome impact, and also reflects genetic susceptibility to exposures. Metabolites identified as intermediate biomarkers between environmental exposures and respiratory health in the context of prospective cohort studies with biomaterials obtained at multiple time points can strengthen causal interpretation, conceptualised as the meet-in-the-middle (MITM) approach; namely, starting with separate investigations of the metabolites or metabolic pathways that are associated with exposures and health outcomes, respectively, followed by the identification of the overlapping pathways [15].

In addition to tobacco smoking, long-term exposure to ambient air pollution is considered to be the top-ranking environmental risk factor for respiratory diseases; for example, COPD and asthma, cardiovascular diseases, premature mortality, and other noncommunicable diseases [16–18]. Following the new World Health Organization (WHO) Air Quality Guidelines 2021, an urgent call for global action has been announced by the European Respiratory Society (ERS) Environmental Health Committee [19]. A recent comprehensive overview of studies on air pollution and nontargeted metabolomics suggests that air pollution exposure is associated with metabolic pathways primarily related to oxidative stress, inflammation and steroid metabolism [20]. This is in line with findings from earlier air pollution research on candidate biomarkers [21–24] and from epigenome-wide association studies [25–27]. Yet, in recent years, the links between air pollution, metabolites as mediators, and respiratory diseases have not been reviewed from a life-course perspective. The present narrative review summarises current knowledge on the associations of metabolic profiles with air pollution exposure and respiratory health outcomes, taking different stages of life into consideration.

This work was undertaken within the framework of the EU-funded EXPANSE (EXposome Powered tools for healthy living in urbAN Settings) project on the urban exposome and cardiometabolic and pulmonary health across the life-course [28].

Methods

Air pollution-related metabolic profiles

A recently published review on air pollution and metabolomics by JIN *et al.* [20] included articles published up until 9 June 2020. As many as 23 out of 315 identified studies were included, of which 13 focused on short-term air pollution exposure, two focused on sub-chronic (*i.e.* multiple months) and eight on long-term effects (*i.e.* ⩾1 year). Two of these 23 studies investigated metabolic profiles in relation to both air pollution exposure and pulmonary function [29, 30]. As this is a rapidly evolving field, we

conducted complementary searches of articles published in PubMed (https://pubmed.ncbi.nlm.nih.gov/) up to 30 September 2021 in the English language using the search terms "metabolomics" OR "metabolic profiling" AND "air pollution" (which were also included in the search by Jɪɴ *et al.* [20]). This resulted in an additional 10 articles being included in this review (figure 1). More details on the selected studies are provided in table S1.

Lung function-related metabolic profiles

We searched for articles in PubMed published in the English language with the search term "metabolomics" OR "metabolomic profiling", "lung function" OR "pulmonary function" OR "respiratory function" restricted to human studies only (performed on 30 September 2021). We first narrowed down the articles identified by the PubMed search criteria (n=101) by screening the titles and abstracts to assess eligibility. We supplemented the results of this search with three additional studies related to the topic that did not arise from the PubMed search but were cited in the articles identified by the search (figure 1). We did not search separately for metabolic profiles associated with COPD or asthma *per se*, since these topics were recently presented in separate reviews [31–34] and our focus of interest for this review was lung function, and not clinical respiratory disease. Because cystic fibrosis (CF) is a monogenic disease with very different underlying pathophysiology, compared with other respiratory diseases such as asthma and COPD, we chose to exclude the CF papers from this review. Articles were excluded if they did not contain original research (*i.e.* reviews, n=12), did not report quantitative data on associations of metabolites with lung function parameters (n=68), or employed targeted metabolomics analysis (n=6), resulting in a total of 16 original articles included in the present review. None of these 16 studies investigated metabolites or metabolic pathways related to both air pollution exposures and lung function. More details on selected studies are provided in table S2. Figure 2 is a word cloud of the weighted list of the words extracted from the abstracts of all articles included.

Life-course perspective in the exposome research of respiratory health

Here, we relied mostly on our experience and judgement, supported by selected references. In addition, we searched for articles in PubMed with the search term "lung function trajectories" AND ("metabolomics"
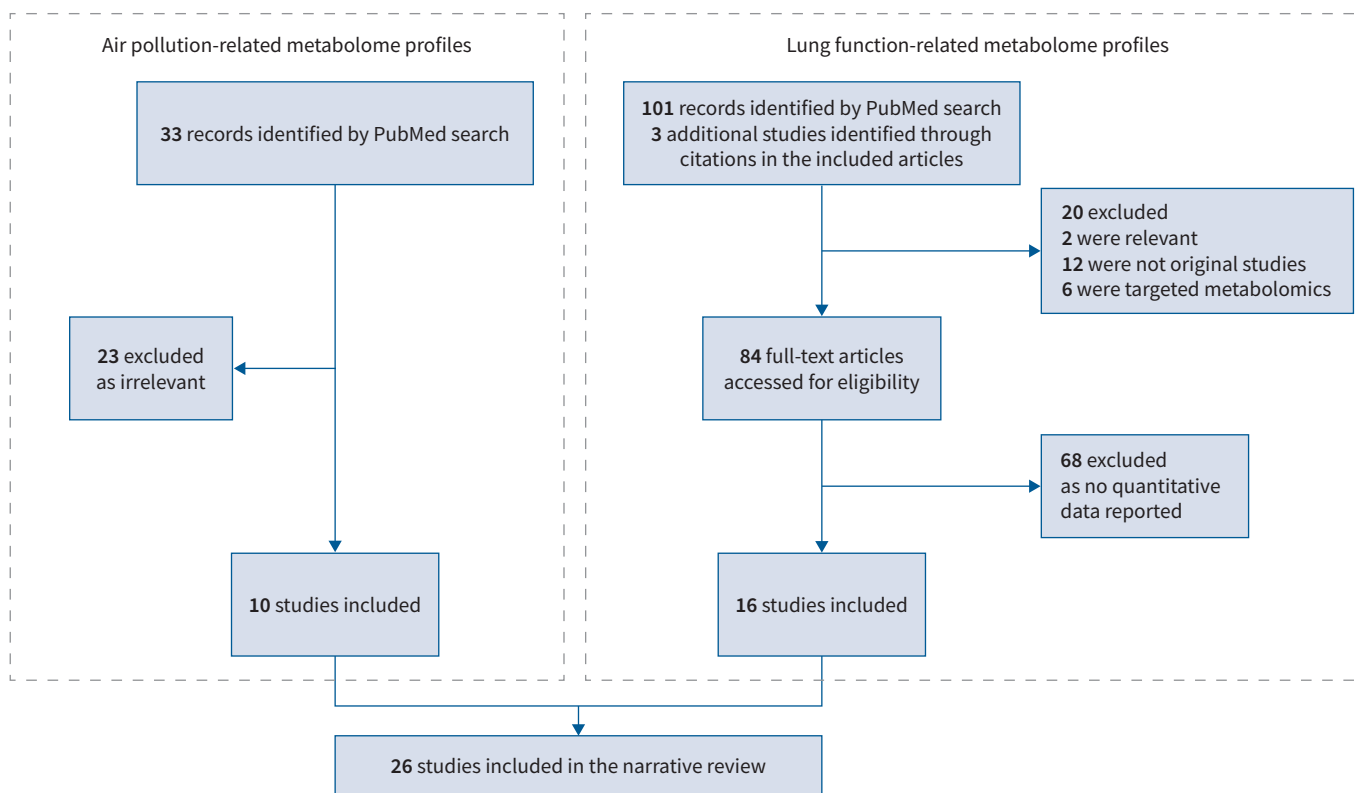

FIGURE 1 Flowchart of the selection of eligible metabolomics studies on air pollution and lung function or COPD related metabolic profiles. The PubMed literature search was performed following the search strategy by Jɪɴ *et al.* [20], extending the search period up to 30 September 2021.
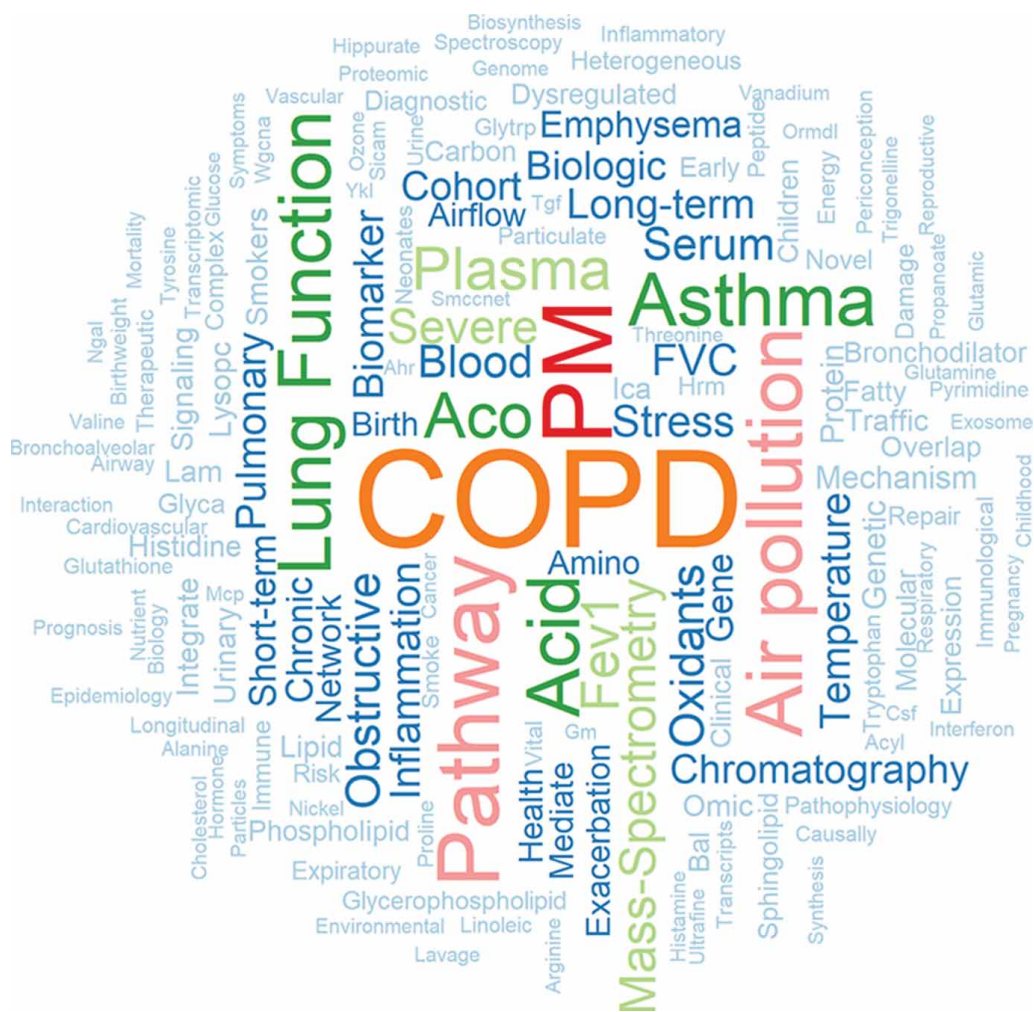
FIGURE 2 The word cloud of the weighted list of the words extracted from the abstracts of included articles. This word cloud mainly highlights the clinical outcomes "asthma", "COPD" and "lung function" and general exposure terms such as "air pollution". Apart from the broad term "pathway", few specific metabolomic pathways are highlighted, which is likely to be because of the issues with the studies included, such as small sample size, cross-sectional design, heterogeneity in exposure assessment, lack of confounding control and omics integration.

OR "metabolomics profiling"), restricted to the English language and human studies (performed on 30 September 2021).

Results

Air pollution exposure and metabolic profiles

Previous review

A recent literature review by JIN et al. [20] reported a wide range of metabolites linked to air pollution exposure, the majority of which were endogenous (i.e. lipids, amino acids, steroids, nucleotides, carbohydrates and vitamins). Several of the included studies also found xenobiotics, such as polycyclic aromatic hydrocarbon metabolites; for example, catechol, 3-(2-hydroxyphenyl) propanoate, naphthylamine, nicotine metabolites and benzoate. Pathway analyses identified that most detected metabolites were related to oxidative stress or inflammatory responses. Metabolites with pro-inflammatory effects (e.g. leukotrienes) generally tended to be upregulated, while anti-inflammatory metabolites (e.g. histidine, linolenic acid) appeared to be downregulated following air pollution exposure. Furthermore, differential perturbations of air pollution-related metabolites and pathways by asthma status [35], age [36] and sex [37, 38] have been detected, providing evidence on mechanisms for susceptible subpopulations. Of note, the vast majority of studies included in the review by JIN et al. [20] were based on adults aged 18 or older, while only two studies focused on children.

New studies

Studies on short-term air pollution exposure

In a series of newly published studies employing plasma untargeted metabolomics in elderly males from the US Normative Aging Study (NAS), the authors identified 19 metabolites that were significantly associated with short-term exposure to nitrogen dioxide (NO_2), mainly lipids (*e.g.* dihomolinoleoylcarnitine (C20:2), palmitoloelycholine, oleoylcholine, 8-hydroxyoctanoate, linolenate, phosphoethanolamine and choline phosphate), and hypotaurine, maltotriose and 3-phosphoglycerate [39]. In the pathway analysis, short-term exposure to NO_2 was associated with sphingolipid and butanoate metabolism. In a subsequent study looking at the relative contribution of the different $PM_{2.5}$ species (particulate matter with a diameter of 2.5 µm or smaller), the authors revealed 12 metabolic pathways that were significantly associated with short-term exposure to ultrafine particles and $PM_{2.5}$ elemental composition (nickel, vanadium, potassium, silicon and lead), including glycerophospholipid, sphingolipid, glutathione, β-alanine, pyrimidine, propanoate, purine, arginine and butanoate metabolism [40]. Another US study of 180 adults from an Emory University-based employee cohort identified several thousands of metabolic features significantly associated with traffic-related air pollutants. Further, the authors reported 21 biological pathways enriched by metabolic features linked to short-term air pollution exposure, including nucleic acids damage and repair (pyrimidine and purine metabolism), nutrient metabolism (*e.g.* fatty acid β-oxidation, tryptophan and vitamin A metabolism) and acute inflammation (*e.g.* histidine, tyrosine, alanine and aspartate metabolism) [41]. Higher levels of exposure to $PM_{2.5}$, black carbon (BC) and SO_2 in healthy adults living in Beijing, China, were also related to reductions in plasma levels of alanine, threonine and glutamic acid [42]. The authors also identified several metabolic pathways affected by high levels of air pollution, most of them involved in amino acid metabolism. In the US prospective cohort of women undergoing assisted reproduction, the top identified metabolic pathways associated with short-term exposure to NO_2, ozone, $PM_{2.5}$ and BC, comprised butanoate, β-alanine, tryptophan, linolenic acid, urea cycle/amino group and vitamin B3 metabolism [43]. Several studies based on personal exposure measurements reported a number of metabolites found to be significantly associated with traffic-related air pollution (TRAP) exposure, mainly involved in amino acid, glucose and fatty acid metabolism [44], and metabolites involved in haem, energy, oxidative stress, phospholipid and tryptophan metabolism related to $PM_{2.5}$ exposure levels [45]. In the latter study, sex-specific analysis revealed males to be more susceptible to $PM_{2.5}$ exposure than females. Du *et al.* [44] also reported novel pathways activated in response to air pollution, including growth hormone signalling, adrenomedullin signalling and arachidonic acid metabolism. Importantly, there is limited existing evidence on the short-term effects in children and, therefore, on life-course perspective.

Studies on long-term air pollution exposure

A study based on plasma untargeted metabolomic profiling of the NAS cohort of men revealed several metabolites perturbed by long-term exposure to $PM_{2.5}$, belonging to glycerophospholipids, sphingolipids, glutathione, β-alanine, propanoate, purine, taurine and hypotaurine, and unsaturated fatty acids [46]. In addition, 18 pathways related to the elemental composition of long-term exposure to $PM_{2.5}$ (BC, nickel, vanadium, zinc, iron, copper, selenium), including glycerophospholipid, sphingolipid, purine and glutathione metabolism, were identified [40]. These pathways are involved in inflammation, oxidative stress, immunity, and nucleic acid damage and repair. In a Chinese prospective cohort study of college students, long-term exposure to $PM_{2.5}$, but not particulate matter with a diameter of less than 10 µm (PM_{10}), was associated with 25 plasma metabolic markers, most of which were phospholipids [47]. Further, prenatal exposure to $PM_{2.5}$, particularly during the third trimester of pregnancy, has been shown to influence the newborn metabolome, contributing to alterations of fatty acid, glycerophospholipid, methionine and cysteine metabolism [48]. Thus, for long-term exposure, the present studies cover a broad age range that suggest effects across the whole life-course.

Metabolomics and respiratory health

In the present review, we identified 16 studies of metabolic profiles in relation to lung function measures. Only two of the included studies were performed in children, while the remaining ones were predominantly conducted in adults with existing lung diseases such as COPD. Most of the studies performed metabolic profiling of blood samples (n=11), followed by urine (n=2), bronchoalveolar lavage fluid (BALF) (n=2) and exhaled breath (n=1). The studies are presented in the following sections stratified by paediatric, adult and population-based studies. Studies are not further stratified by short- and long-term associations, given the fact that many studies were cross-sectional case–control studies and given that acute impacts on physiology are in part thought to be at the heart of subsequent chronic changes (*i.e.* disentangling short- and long-term effects is challenging).

Studies on paediatric patient cohorts (lung function in the presence of asthma)

In a cohort of Costa-Rican children with asthma (6–14 years, n=380), the researchers identified plasma metabolic profiles that distinguish children with and without asthma by their phenotypic aspects of lung function, *i.e.* the ratio of forced expiratory volume in 1 s (FEV_1) to forced vital capacity (FVC) (FEV_1/FVC) and airway hyperresponsiveness (AHR) [49]. Pre- and post-bronchodilator FEV_1/FVC were associated with 102 and 155 metabolites, respectively, with a large overlap (97 common metabolites). The majority of these were amines, a metabolite class consistently linked with asthma [50]. The metabolic profile for AHR (measured as a 20% decrement in FEV_1 after methacholine administration) included metabolites different from the FEV_1/FVC-related ones, largely polar and nonpolar lipids. Further downstream, metabolite-set enrichment analysis revealed several significantly enriched pathways, namely glycerophospholipid, linoleic acid and pyrimidine metabolism. Two other pathways, sphingolipid metabolism and D-glutamine/glutamate metabolism, were exclusive for AHR, while pantothenate and CoA biosynthesis was enriched only among the FEV_1/FVC post-bronchodilator (post-BD) metabolites. In subsequent integrated omics analyses, this research group interrogated blood transcriptomic and metabolomic data, and discovered associations of *ORMDL3* and dysregulated sphingolipid metabolism with impaired lung function [51]. *ORMDL3* is known to be one of the key genes associated with childhood asthma along with other genes at the chromosome *17q12–21* locus [52, 53].

Of note, in all the abovementioned studies, biospecimens for metabolomics profiling were collected after asthma occurrence, thus limiting the opportunities to gain insights into disease aetiology from a clearly time-resolved longitudinal perspective.

Studies on adult patient cohorts (lung function in the presence of COPD and/or asthma)

In a US cohort of individuals with and without COPD, proton nuclear magnetic resonance (^{1}H NMR) spectroscopy-based metabolomics revealed three particular urinary metabolites that correlated with lung function (trigonelline, hippurate and formate), with trigonelline evidencing the strongest correlation with baseline FEV_1 measurements [54]. Hippurate and formate originate from production by the gut microbiome and dietary sources [55]. Therefore, the differential urine levels of these compounds might be indicative of the differences in microbiome composition or functionalities between individuals with different lung function. In the same study, no significant associations with lung function or lung function decline were observed for plasma metabolites.

By integrating plasma proteomic and metabolomic data from 1008 former and current smokers of the US COPD Gene study, MASTEJ *et al.* [56] found a protein–metabolite network consisting of 13 proteins and seven metabolites that negatively correlated with FEV_1 % predicted. Troponin T, phosphocholine and ergothioneine metabolites demonstrated high connectivity (*i.e.* number of edges linked to a metabolite node) and strong edges in the FEV_1 network, where edges represent associations between metabolite–protein pairs relative to FEV_1. Further, enrichment analysis found metabolites in the diacylglycerol and branched-chain amino acids (leucine, isoleucine and valine) sub-pathways to be enriched for associations with FEV_1/FVC [57]. For FEV_1 % pred, 79 metabolites, including lipid phosphocholine, ergothioneine and carbohydrate N6-carboxymethyllysine, were most significantly associated. By combining blood metabolomics and transcriptomics data within the same cohort, the authors identified glycerophospholipids, an important component of lung surfactant, along with the lipid classes lysophosphatidylethanolamine (LysoPE) and phosphatidylglycerol (PG) as highly significant in relation to FEV_1 % pred, and sphingolipids to be associated with FEV_1/FVC [58].

Similarly, examination of the serum metabolome in a large UK cohort of women with lymphangioleiomyomatosis revealed a link between sphingolipid, fatty acid and phospholipid metabolites and FEV_1 [59]. HALPER-STROMBERG *et al.* [60] showed that the concentrations of lipid metabolites were negatively correlated with FEV_1/FVC in BALF from patients with COPD.

In a US study, metabolic profiling of BALF in patients with COPD demonstrated a significant increase in peptides compared with healthy controls that was strongly associated with lung function [61]. The detected associations were strongest for lung function tests associated with airflow obstruction (FEV_1 and FEV_1/FVC).

Several studies have examined metabolomics of lung function in the context of asthma-COPD overlap. They report significant positive correlations between lung function parameters (*i.e.* FEV_1 and FEV_1/FVC) and urinary L-histidine [62], serum valine, citrate and glutamate metabolites [63], serine, threonine, glucose, cholesterol, D-mannose and succinic acid [64], and fatty acid, propionate, isopropanol, lactate, acetone, valine, methanol and formate in exhaled breath condensate [65]. Of note, in the abovementioned

studies, biosamples for metabolomics analysis have mostly been collected at the time of lung function measurements, whereas studies with prospectively assessed metabolic profiles are scarce.

Population-based studies on metabolic profiles and lung function

Metabolic profiling of fasting blood from 6055 individuals from the population-based UK twins study identified C-glycosyl tryptophan (C-glyTrp) to be strongly correlated with FEV_1 [66]. This finding was further replicated in the independent German KORA cohort (also cross-sectional). Further, by comparing metabolite levels of C-glyTrp with genome-wide DNA methylation profiles, the authors found three differentially methylated CpG sites, annotated to the *WDR85, EDN2* and *GLB1L3* genes, respectively, that have previously been implicated in human early development and age-related phenotypes, such as retinal degeneration, renal inflammation and hypertension.

By combining data from over 4700 individuals across two population-based studies, YU *et al.* [67] identified 95 serum metabolites associated with FEV_1 and 100 with FVC (73 overlapping), including inverse associations with branched-chain amino acids and positive associations with glutamine. More metabolites were found to be associated with FEV_1 and FVC than with FEV_1/FVC, suggesting better abilities to detect lung and airway size as opposed to obstructive airflow associations with serum compounds. Subsequent pathway analysis revealed enriched pathways of amino acid metabolism. Several of the lung function-related metabolic pathways overlapped with those reported in studies on air pollution-related metabolic profiles, including the metabolism of glycerophospholipid, pyrimidine, sphingolipid, purine, glutamate, histidine and tryptophan. More detailed descriptions of study-specific findings, including identified metabolites and metabolic pathways, are provided in tables S1 and S2. However, since none of the included studies investigated both sets of metabolites (associated with air pollution exposure and with lung function) measured in the same population and with the same analytical platform, inference regarding overlapping findings is limited.

So far, only a limited number of studies investigated correlated metabolites or pathways for the association between air pollution exposure and respiratory conditions. Most of these studies applied the MITM concept, starting with separate investigations of the metabolites or pathways that are associated with air pollution and health outcomes, followed by identification of the overlapping pathways [68]. JIN *et al.* [20] found, in their recent review, eight studies examining air pollution-related metabolites that are also perturbed by health outcomes, such as COPD and lung function, ischaemic heart disease, as well as biomarkers of oxidative stress, or inflammation, all but one of which collected biological samples after or at the same time of the assessment of health outcomes. Only two out of these eight studies applied metabolomic analyses to investigate molecular and biochemical pathways that link environmental exposures to pulmonary function. In a Dutch experimental study based on a panel of 31 volunteers exposed to ambient air pollution, a causal mediation analysis of the effects of air pollutants on lung function parameters through a change in blood metabolic features demonstrated metabolic perturbations within eight pathways, including tyrosine metabolism, urea cycle/amino group metabolism and N-glycan degradation [29]. A cross-sectional study conducted within the TwinsUK cohort found eight metabolites, *i.e.* asparagine, glycine, N-acetylglycine, serine, glycerate, threonate, α-tocopherol and benzoate, associated with both long-term particulate matter air pollution exposure and lung function [30]. Some of these (*i.e.* tyrosine, guanosine, glycine, α-tocopherol, benzoate, and urea cycle/amino group metabolism) were also reported in the studies on air pollution-related metabolic profiles included in the present review [41–43, 46]. Also, the studies on lung function-related metabolic profiles included in the present review showed associations with tyrosine [58, 67], glycine [58, 67], N-acetylglycine [66, 69], serine [58, 66], glycerate [67] and urea cycle/amino group metabolism [57, 66].

These results highlight the promising role of metabolomics for identification of biomarkers both linked to environmental exposures and to intermediate respiratory health end-points and of the MITM concept.

Conclusions and recommendations for future research

Metabolomics offers the potential to fill the gap in the exposome research of respiratory health, from exposure assessment, and the improved causal understanding and identification of individuals at risk. There is accumulating evidence for distinct variations in circulating metabolites related to both air pollution exposure and lung function, including compounds of amino acid and lipid metabolism. Identifying subtle alterations in metabolic profiles related to lung function in the general population holds the potential to identify biomarkers of the early stages of asthma and COPD pathogenesis that precede diagnostic reductions in lung function. However, our review of the literature identified several shortcomings that need to be addressed in future studies (table 1), as listed below.

1) Most reports were based on a small number of subjects recruited from clinical settings for COPD or other respiratory pathology. Studies of metabolic profiles and quantitative lung function traits in the general population (*e.g.* trajectories) are still limited.

2) Very few studies have investigated the overlap between markers of exposure and predictive markers of disease outcomes within the same study. Implementation of the MITM and/or other causal mediation analytical approaches may facilitate identification of intermediate biomarkers involved in health effects related to air pollution.

3) It is difficult to fully benefit from increasingly available metabolomics findings because of heterogeneities in the analyses, *i.e.* the choice of biological matrix analysed, the type of analytical platform employed and the technical variabilities in how data are processed and annotated. There is a need for large-scale data acquisition, including replication, from the same analytical platform and with uniform metabolite annotation protocols (and exposure assessment).

4) Many of the reviewed studies measured metabolic profiles at the same time as health markers were evaluated (*i.e.* cross-sectional studies), thus making it challenging to draw conclusions on whether these overlapping metabolites mediate the impact of air pollution on lung function, or whether impaired lung function is exacerbated through these shared metabolic features. This temporal disconnect between sampling and exposure and/or health marker renders it difficult to identify causality. Therefore, future studies should prospectively assess air pollution exposures, metabolic profiles and health outcomes to effectively identify such intermediate metabolites in the causal pathway linking exposure to the outcome.

5) A lack of long-term longitudinal data with information on respiratory health and air pollution exposure, and the heterogeneity of air pollution metrics and sources across studies poses major challenges in the interrogation of latency. Also, to which extent the metabolome may aid in characterising exposure history (including the stability of biomarkers) needs to be explored. The latency may vary widely from days to decades depending on which respiratory outcome is involved. While some studies of the short-term effects of air pollution incorporated *a priori* determined lag periods into their analysis, very few studies conducted a systematic search for latency of air pollution effects on respiratory health.

6) Many potential confounders and effect modifiers remain unaccounted for; for example, noise, heat, diet, physical activity, socioeconomic factors and detailed smoking histories. For a complete understanding of how exposures influence health, a full exposome perspective will be needed.

7) Susceptibility factors are often ignored, such as genetics and other host factors. Given that gene–air pollution interactions are well described for respiratory outcomes [70], genetics are likely to also influence associations between exposure, metabolic profiles and respiratory disease. In this context, the use of polygenic risk scores offers an attractive analytical approach, as exemplified in recent COPD studies [71].

8) The integration of metabolomics with other types of omics data, such as epigenetics and transcriptomics [72], will improve the interrogation and understanding of pathophysiological pathways mediating air pollution effects on lung function across the life-course [9].

TABLE 1 Limitations of current air pollution, metabolomics and respiratory health studies and directions for future research	
Limitations to date	**Future directions**
Most studies are small and based on patient datasets (*e.g.* COPD)	Large-scale, population-based studies (including replication) needed
Limited number of studies on intermediate biomarkers linking exposure and disease	Implementation of the meet-in-the-middle (MITM) and/or other causal mediation approaches
Heterogeneities in metabolomics assessment	Uniform analytical platform and metabolite annotation protocols
Cross-sectional data	Exposure and metabolomic profiles assessed before disease occurrence (*e.g.* in longitudinal studies)
Difficult to evaluate latency and long-term effects	Longitudinal studies with long-term follow-up
Lack of confounding control	A complete exposome concept taking all relevant exposures into account
Susceptibility factors ignored	Inclusion of genetics (*e.g.* polygenic risk scores) and other host factors in the analyses
Lack of integration with other omics data	Apply multi-omics models (including *e.g.* epigenetics, transcriptomics, proteomics)

In addition to the above shortcomings, the challenges of the life-course perspective in exposome research for respiratory health need to be acknowledged, discussed and considered to optimise the use of the metabolome as a mediator.

Metabolic profiling at different time points can shed light on the biological mechanisms through which environmental exposures affect respiratory health in different growth phases. Causal mediation analysis [73] is a widely used epidemiological method to assess the role of intermediate variables (*i.e.* metabolites) that lie along causal pathways from exposure (*i.e.* air pollution) to outcome (*i.e.* lung function trajectories). This method has several advantages, including allowing for nonlinearity, interactions, as well as multiple mediators with path-specific effects [74–76]. Ideally, this approach would require long-term follow-up of the same subjects, not only with information on their respiratory health, but also with prospectively collected biological samples and air pollution exposure estimates. Such long-term longitudinal data are unfortunately scarce and to our best knowledge there has been no study that has investigated metabolic profiles as mediators across a broad age spectrum. However, many studies of different age groups have applied the MITM concept (as described in the Population-based studies on metabolic profiles and lung function section). However, combining findings across studies is challenging because of the heterogeneity in the biological matrices and in the analytical technologies used. More detailed methodological issues of adopting the metabolome as the mediator between air pollution exposure and adverse health outcomes have been reviewed elsewhere [77]. Furthermore, by comparing metabolic profiles related to air pollution and lung function between different age groups, we ignore the fact that older people have different childhood air pollution-related metabolic profiles than the current childhood/adolescent group. In the absence of biosamples collected prospectively decades ago and in the absence of exposures back-extrapolated many decades back in long-term studies with repeat lung function measures from childhood to late adulthood, we cannot learn about the long-term health effects of specific metabolic profiles arising from childhood exposure to air pollution. Yet, if we do find that air pollution affects overlapping metabolic profiles in all age groups and that these are associated with lung function irrespective of age, this strengthens our causal and biological understanding of life-long respiratory effects of air pollution.

In the EXPANSE project, we aim to address several of the abovementioned challenges [28]. We will: 1) bring together large datasets including individual adult and matured birth cohorts with prospectively obtained biosamples, longitudinal high-quality lung function measurements and incident respiratory diagnoses covering different age groups towards a life-course perspective; 2) model back-extrapolated air pollution exposure as well as exposure to potential confounders (*e.g.* noise, temperature) and the external exposome in a harmonised manner to participating cohorts; 3) systematically analyse metabolic profiles using prospectively collected biosamples from several cohorts on a same analytical platform for both untargeted liquid- and gas-chromatography coupled with high-resolution mass spectrometry; and 4) link both exposures and metabolomics to respiratory health outcomes using a life-course perspective and gain insights into mediating biological mechanisms and causality by applying an MITM approach.

References

1 Vermeulen R, Schymanski EL, Barabási AL, *et al.* The exposome and health: where chemistry meets biology. *Science* 2020; 367: 392–396.

2 Peters A, Nawrot TS, Baccarelli AA. Hallmarks of environmental insults. *Cell* 2021; 184: 1455–1468.

3 Zhang P, Carlsten C, Chaleckis R, *et al.* Defining the scope of exposome studies and research needs from a multidisciplinary perspective. *Environ Sci Technol Lett* 2021; 8: 839–852.

4 Murrison LB, Brandt EB, Myers JB, *et al.* Environmental exposures and mechanisms in allergy and asthma development. *J Clin Invest* 2019; 129: 1504–1515.

5 Wheelock CE, Rappaport SM. The role of gene-environment interactions in lung disease: the urgent need for the exposome. *Eur Respir J* 2020; 55: 1902064.

6 Afshar-Mohajer N, Wu TD, Shade R, *et al.* Obesity, tidal volume, and pulmonary deposition of fine particulate matter in children with asthma. *Eur Respir J* 2022; 59: 2100209.

7 Jobe AH, Whitsett JA, Abman SH. Fetal and Neonatal Lung Development: clinical correlates and technologies for the future. New York, Cambridge University Press, 2016.

8 Agusti A, Faner R. Lung function trajectories in health and disease. *Lancet Respir Med* 2019; 7: 358–364.

9 Agusti A, Melen E, DeMeo DL, *et al.* Pathogenesis of chronic obstructive pulmonary disease: understanding the contributions of gene-environment interactions across the lifespan. *Lancet Respir Med* 2022; 10: 512–524.

10 Lange P, Celli B, Agustí A, *et al.* Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med* 2015; 373: 111–122.

11 Melén E, Guerra S, Hallberg J, *et al.* Linking COPD epidemiology with pediatric asthma care: implications for the patient and the physician. *Pediatric Allergy Immunol* 2019; 30: 589–597.

12 Schultz ES, Litonjua AA, Melen E. Effects of long-term exposure to traffic-related air pollution on lung function in children. *Curr Allergy Asthma Rep* 2017; 17: 41.

13 Wang G, Hallberg J, Um Bergstrom P, *et al.* Assessment of chronic bronchitis and risk factors in young adults: results from BAMSE. *Eur Respir J* 2021; 57: 2002120.

14 Milanzi EB, Koppelman GH, Smit HA, *et al.* Air pollution exposure and lung function until age 16 years: the PIAMA birth cohort study. *Eur Respir J* 2018; 52: 1800218.

15 Vineis P, Demetriou CA, Probst-Hensch N. Long-term effects of air pollution: an exposome meet-in-the-middle approach. *Int J Public Health* 2020; 65: 125–127.

16 Cohen AJ, Brauer M, Burnett R, *et al.* Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* 2017; 389: 1907–1918.

17 Thurston GD, Balmes JR, Garcia E, *et al.* Outdoor air pollution and new-onset airway disease. An official American Thoracic Society Workshop report. *Ann Am Thorac Soc* 2020; 17: 387–398.

18 Liu S, Jorgensen JT, Ljungman P, *et al.* Long-term exposure to low-level air pollution and incidence of asthma: the ELAPSE project. *Eur Respir J* 2021; 57: 2003099.

19 Andersen ZJ, Gehring U, De Matteis S, *et al.* Clean air for healthy lungs - an urgent call to action: European Respiratory Society position on the launch of the WHO 2021 Air Quality Guidelines. *Eur Respir J* 2021; 58: 2102447.

20 Jin L, Godri Pollitt KJ, Liew Z, *et al.* Use of untargeted metabolomics to explore the air pollution-related disease continuum. *Curr Environ Health Rep* 2021; 8: 7–22.

21 Gruzieva O, Merid SK, Gref A, *et al.* Exposure to traffic-related air pollution and serum inflammatory cytokines in children. *Environ Health Perspect* 2017; 125: 067007.

22 Hew KM, Walker AI, Kohli A, *et al.* Childhood exposure to ambient polycyclic aromatic hydrocarbons is linked to epigenetic modifications and impaired systemic immunity in T cells. *Clin Exp Allergy* 2015; 45: 238–248.

23 Mostafavi N, Jeong A, Vlaanderen J, *et al.* The mediating effect of immune markers on the association between ambient air pollution and adult-onset asthma. *Sci Rep* 2019; 9: 8818.

24 Fiorito G, Vlaanderen J, Polidoro S, *et al.* Oxidative stress and inflammation mediate the effect of air pollution on cardio- and cerebrovascular disease: a prospective study in nonsmokers. *Environ Mol Mutagen* 2018; 59: 234–246.

25 Gruzieva O, Xu CJ, Breton CV, *et al.* Epigenome-wide meta-analysis of methylation in children related to prenatal NO_2 air pollution exposure. *Environ Health Perspect* 2017; 125: 104–110.

26 Gruzieva O, Xu CJ, Yousefi P, *et al.* Prenatal particulate air pollution and DNA methylation in newborns: an epigenome-wide meta-analysis. *Environ Health Perspect* 2019; 127: 57012.

27 Eze IC, Jeong A, Schaffner E, *et al.* Genome-wide DNA methylation in peripheral blood and long-term exposure to source-specific transportation noise and air pollution: the SAPALDIA study. *Environ Health Perspect* 2020; 128: 67003.

28 Vlaanderen J, de Hoogh K, Hoek G, *et al.* Developing the building blocks to elucidate the impact of the urban exposome on cardiometabolic-pulmonary disease: The EU EXPANSE project. *Environ Epidemiol* 2021; 5: e162.

29 Vlaanderen JJ, Janssen NA, Hoek G, *et al.* The impact of ambient air pollution on the human blood metabolome. *Environ Res* 2017; 156: 341–348.

30   Menni C, Metrustry SJ, Mohney RP, *et al.* Circulating levels of antioxidant vitamins correlate with better lung function and reduced exposure to ambient pollution. *Am J Respir Crit Care Med* 2015; 191: 1203–1207.

31   Sim S, Choi Y, Park HS. Potential metabolic biomarkers in adult asthmatics. *Metabolites* 2021; 11: 430.

32   Papamichael MM, Katsardis C, Sarandi E, *et al.* Application of metabolomics in pediatric asthma: prediction, diagnosis and personalized treatment. *Metabolites* 2021; 11: 251.

33   Nambiar S, Bong How S, Gummer J, *et al.* Metabolomics in chronic lung diseases. *Respirology* 2020; 25: 139–148.

34   Reinke SN, Naz S, Chaleckis R, *et al.* Urinary metabotype of severe asthma evidences decreased carnitine metabolism independent of oral corticosteroid treatment in the U-BIOPRED study. *Eur Respir J* 2021: 2101733.

35   Liang D, Ladva CN, Golan R, *et al.* Perturbations of the arginine metabolome following exposures to traffic-related air pollution in a panel of commuters with and without asthma. *Environ Int* 2019; 127: 503–513.

36   Chen CS, Yuan TH, Shie RH, *et al.* Linking sources to early effects by profiling urine metabolome of residents living near oil refineries and coal-fired power plants. *Environ Int* 2017; 102: 87–96.

37   Zhang Y, Chu M, Zhang J, *et al.* Urine metabolites associated with cardiovascular effects from exposure of size-fractioned particulate matter in a subway environment: a randomized crossover study. *Environ Int* 2019; 130: 104920.

38   Li H, Cai J, Chen R, *et al.* Particulate matter exposure and stress hormone levels: a randomized, double-blind, crossover trial of air purification. *Circulation* 2017; 136: 618–627.

39   Nassan FL, Kelly RS, Koutrakis P, *et al.* Metabolomic signatures of the short-term exposure to air pollution and temperature. *Environ Res* 2021; 201: 111553.

40   Nassan FL, Wang C, Kelly RS, *et al.* Ambient $PM_{2.5}$ species and ultrafine particle exposure and their differential metabolomic signatures. *Environ Int* 2021; 151: 106447.

41   Li Z, Liang D, Ye D, *et al.* Application of high-resolution metabolomics to identify biological pathways perturbed by traffic-related air pollution. *Environ Res* 2021; 193: 110506.

42   Feng B, Liu C, Yi T, *et al.* Perturbation of amino acid metabolism mediates air pollution associated vascular dysfunction in healthy adults. *Environ Res* 2021; 201: 111512.

43   Gaskins AJ, Tang Z, Hood RB, *et al.* Periconception air pollution, metabolomic biomarkers, and fertility among women undergoing assisted reproduction. *Environ Int* 2021; 155: 106666.

44   Du X, Zhang Q, Jiang Y, *et al.* Dynamic molecular choreography induced by traffic exposure: a randomized, crossover trial using multi-omics profiling. *J Hazard Mater* 2021; 424: 127359.

45   Huan S, Jin S, Liu H, *et al.* Fine particulate matter exposure and perturbation of serum metabolome: a longitudinal study in Baoding, China. *Chemosphere* 2021; 276: 130102.

46   Nassan FL, Kelly RS, Kosheleva A, *et al.* Metabolomic signatures of the long-term exposure to air pollution and temperature. *Environ Health* 2021; 20: 3.

47   Chu H, Huang FQ, Yuan Q, *et al.* Metabolomics identifying biomarkers of PM2.5 exposure for vulnerable population: based on a prospective cohort study. *Environ Sci Pollut Res Int* 2021; 28: 14586–14596.

48   Ritz B, Yan Q, He D, *et al.* Child serum metabolome and traffic-related air pollution exposure in pregnancy. *Environ Res* 2021; 203: 111907.

49   Kelly RS, Virkud Y, Giorgio R, *et al.* Metabolomic profiling of lung function in Costa-Rican children with asthma. *Biochim Biophys Acta Mol Basis Dis* 2017; 1863: 1590–1595.

50   Chang C, Guo ZG, He B, *et al.* Metabolic alterations in the sera of Chinese patients with mild persistent asthma: a GC-MS-based metabolomics analysis. *Acta Pharmacol Sin* 2015; 36: 1356–1366.

51   Kelly RS, Chawes BL, Blighe K, *et al.* An integrative transcriptomic and metabolomic study of lung function in children with asthma. *Chest* 2018; 154: 335–348.

52   Melén E. Asthma genetics revisited: understanding disease mechanisms by studying ethnically diverse groups. *Lancet Respir Med* 2020; 8: 427–429.

53   Rago D, Pedersen CT, Huang M, *et al.* Characteristics and mechanisms of a sphingolipid-associated childhood asthma endotype. *Am J Respir Crit Care Med* 2021; 203: 853–863.

54   McClay JL, Adkins DE, Isern NG, *et al.* (1)H nuclear magnetic resonance metabolomics analysis identifies novel urinary biomarkers for lung function. *J Proteome Res* 2010; 9: 3083–3090.

55   Pallister T, Jackson MA, Martin TC, *et al.* Hippurate as a metabolomic marker of gut microbiome diversity: modulation by diet and relationship to metabolic syndrome. *Sci Rep* 2017; 7: 13670.

56   Mastej E, Gillenwater L, Zhuang Y, *et al.* Identifying protein-metabolite networks associated with COPD phenotypes. *Metabolites* 2020; 10: 124.

57   Gillenwater LA, Pratte KA, Hobbs BD, *et al.* Plasma metabolomic signatures of chronic obstructive pulmonary disease and the impact of genetic variants on phenotype-driven modules. *Netw Syst Med* 2020; 3: 159–181.

58   Cruickshank-Quinn CI, Jacobson S, Hughes G, *et al.* Metabolomics and transcriptomics pathway approach reveals outcome-specific perturbations in COPD. *Sci Rep* 2018; 8: 17132.

59   Bottolo L, Miller S, Johnson SR. Sphingolipid, fatty acid and phospholipid metabolites are associated with disease severity and mTOR inhibition in lymphangioleiomyomatosis. *Thorax* 2020; 75: 679–688.

60   Halper-Stromberg E, Gillenwater L, Cruickshank-Quinn C, *et al.* Bronchoalveolar lavage fluid from COPD patients reveals more compounds associated with disease than matched plasma. *Metabolites* 2019; 9: 157.

61   Wendt CH, Nelsestuen G, Harvey S, *et al.* Peptides in bronchoalveolar lavage in chronic obstructive pulmonary disease. *PloS one* 2016; 11: e0155724.

62   Oh JY, Lee YS, Min KH, *et al.* Increased urinary L-histidine in patients with asthma COPD overlap: a pilot study. *Int J Chron Obstruct Pulmon Dis* 2018; 13: 1809–1818.

63   Ghosh N, Choudhury P, Subramani E, *et al.* Metabolomic signatures of asthma-COPD overlap (ACO) are different from asthma and COPD. *Metabolomics* 2019; 15: 87.

64   Ghosh N, Choudhury P, Kaushik SR, *et al.* Metabolomic fingerprinting and systemic inflammatory profiling of asthma COPD overlap (ACO). *Respir Res* 2020; 21: 126.

65   Ghosh N, Choudhury P, Joshi M, *et al.* Global metabolome profiling of exhaled breath condensates in male smokers with asthma COPD overlap and prediction of the disease. *Sci Rep* 2021; 11: 16664.

66   Menni C, Kastenmuller G, Petersen AK, *et al.* Metabolomic markers reveal novel pathways of ageing and early development in human populations. *Int J Epidemiol* 2013; 42: 1111–1119.

67   Yu B, Flexeder C, McGarrah RW, 3rd, *et al.* Metabolomics identifies novel blood biomarkers of pulmonary function and COPD in the general population. *Metabolites* 2019; 9: 61.

68   Chadeau-Hyam M, Athersuch TJ, Keun HC, *et al.* Meeting-in-the-middle using metabolic profiling – a strategy for the identification of intermediate biomarkers in cohort studies. *Biomarkers* 2011; 16: 83–88.

69   Prokic I, Lahousse L, de Vries M, *et al.* A cross-omics integrative study of metabolic signatures of chronic obstructive pulmonary disease. *BMC Pulm Med* 2020; 20: 193.

70   Gref A, Merid SK, Gruzieva O, *et al.* Genome-wide interaction analysis of air pollution exposure and childhood asthma with functional follow-up. *Am J Respir Crit Care Med* 2017; 195: 1373–1383.

71   Moll M, Sakornsakolpat P, Shrine N, *et al.* Chronic obstructive pulmonary disease and related phenotypes: polygenic risk scores in population-based and case-control cohorts. *Lancet Respir Med* 2020; 8: 696–708.

72   Merid SK, Bustamante M, Standl M, *et al.* Integration of gene expression and DNA methylation identifies epigenetically controlled modules related to PM2.5 exposure. *Environ Int* 2021; 146: 106248.

73   Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986; 51: 1173–1182.

74   Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992; 3: 143–155.

75   Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science* 2010; 25: 51–71.

76   Daniel RM, De Stavola BL, Cousens S, *et al.* Causal mediation analysis with multiple mediators. *Biometrics* 2015; 71: 1–14.

77   Inoue K, Yan Q, Arah OA, *et al.* Air pollution and adverse pregnancy and birth outcomes: mediation analysis using metabolomic profiles. *Curr Environ Health Rep* 2020; 7: 231–242.