BMJ Open

Endotype-driven prediction of acute exacerbations in chronic obstructive pulmonary disease (EndAECOPD): protocol for a prospective cohort study

Wei Xiao [iD] , Long-yi Du, Bing Mao, Ti-wei Miao, Juan-juan Fu

Department of Integrated Traditional Chinese and Western Medicine, Sichuan University West China Hospital, Chengdu, China

Correspondence to
Juan-juan Fu;
fu.juanjuan@scu.edu.cn

ABSTRACT

Introduction Current strategies for the prevention of acute exacerbations in chronic obstructive pulmonary disease (COPD) are primarily based on clinical measurements but fail to target the pathophysiological mechanisms, namely endotypes, of the disease. Studies identifying endotypes underlying exacerbation susceptibility and discovering specific biomarkers may lead to the development of targeted therapeutics but are lacking. This study aims to assess a broad spectrum of biomarkers at multiple biological levels (genetics, airway inflammation and respiratory microbiome) for their ability in predicting acute exacerbations of COPD, thus enables high-resolution disease endotyping and may lead to precision treatment of the disease.

Methods and analysis In this prospective cohort study, participants with stable COPD (n=600) will be recruited and assessed for demographics, symptom scores, spirometry, medication use and comorbidities at baseline. Blood will be obtained for genotyping variants in a panel of nine genes. Induced sputum will be collected for the profile of microbiota using 16S rRNA gene sequencing, quantification of bacterial load, inflammatory mediators assay and sputum cytometry. Participants will be followed up for their exacerbations till 12 months and reassessed for the clinical measurements as baseline. The primary outcomes are total number of exacerbations, severe exacerbations, moderate exacerbations and time to first exacerbation. The secondary outcomes are changes in lung function and symptom scores. The effect of biomarkers representing genetic variants, airway inflammation and respiratory microbiome on predicting the frequent exacerbator phenotype and exacerbation frequency will be analysed with multivariable modelling, and time to first exacerbation with a Cox regression model.

Ethics and dissemination The study has been approved by the Clinical Trial and Biomedical Ethics Committee of West China Hospital of Sichuan University (No. 2018–298). The results of the study will be published on peer-reviewed journals.

Trial registration number ChiCTR1800019063.

INTRODUCTION

Exacerbation represents an important event in the natural history of chronic obstructive pulmonary disease (COPD), which is

---

Strengths and limitations of this study

► It is a prospective study to systematically evaluate the validity of numerous biomarkers representing multiple pathobiological mechanisms of chronic obstructive pulmonary disease in predicting exacerbations, in particular assessing the effect of respiratory microbiome and a group of inflammatory mediators on future exacerbations for the first time.

► A comprehensive picture of biomarkers covering multiple biological levels (genetics, inflammation and microbiome) will lead to high-resolution disease endotyping and provide future therapeutic targets.

► Changes of inflammation and microbiota at study entry and at endpoint will be assayed along with exacerbation frequency, so as to assess the stability of these biomarkers and the effects of exacerbations on their levels.

► This is a single-centre study that will be conducted in a tertiary hospital which somewhat limits the generalisation of the study results, although a large sample size of participants will be investigated.

► We will only evaluate a limited number of candidate biomarkers from different biological levels instead of a 'hypothesis-free' approach.

---

associated with considerable physiological deterioration, and accelerates the decline of lung function and disease progression.[1 2] The costs of treating acute exacerbations of COPD (AECOPD) account for approximately 70% of the total economic burden in COPD patients.[3] The frequency of AECOPD varies among patients, with a subset experiencing significantly more exacerbations, which is identified as the frequent exacerbator phenotype.[4] Identifying the determinants of this clinical phenotype is of great importance for early stratification of patients with high exacerbation risk so as to place targeting treatment.

Clinical factors such as exacerbation history have been found to be associated with an increased risk of future exacerbations;[4]

therefore, the guideline-based therapy for reducing AECOPD risk is initiated or adjusted accordingly. However, these clinical factors fail to adequately consider the pathobiological mechanisms of frequent exacerbation and are unfeasible to be targeted for treatment. Characterising asthma according to endotype, a term applied to 'a subtype of condition that is defined by a distinct functional or pathophysiological mechanism', has been used to account for the heterogeneity of the disease and direct therapy in clinic, for instance the well-known T helpler (Th) 2 endotype.[5] Endotyping will provide new insight in the interindividual variability of clinical presentations and treatment responses in patients with even identical phenotypes, which may lead to the development of more specific therapeutic alternatives.[6] However, apart from the quantification of blood eosinophil count in the updated Global Initiative for Chronic Obstructive Lung Disease (GOLD) document,[7] current treatments of COPD targeting exacerbation with inhaled bronchodilators or/and inhaled corticosteroids do not emphasise the measurement of endotype. The pharmacological therapy for the prevention of AECOPD calls for the identification of endotype-specific biomarkers and a new paradigm of precision medicine based on these biomarkers to better predict who is most likely to benefit from specific targeted therapies.

Susceptibility to exacerbations of COPD and the basis of the frequent exacerbator are multifactorial, where genetic predisposition,[8 9] background airway inflammation[10] and changes in lower airway bacterial colonisation[11 12] may all be involved. However, few study has assessed these important pathophysiological components of COPD for the prediction of exacerbation risk in a longitudinal study with rigorous design. Associations between a group of genetic variants and AECOPD risk have been investigated.[13] Despite the emerging of genomics-based studies, carefully conducted candidate gene study along with replication study will remain among the armamentarium of techniques used by researchers to identify reliable disease associated risk loci.[13] Therefore, in this study, we will reassess the genetic variants reported in previous studies associated with exacerbation susceptibility. Inflammatory mediators covering Th1, Th2 and Th17 immune responses and sputum cytometry at baseline will be assessed for their role in predicting exacerbations. With the development of next-generation sequencing techniques, the meaning of previously identified colonisation in the COPD airways needs to be re-evaluated,[14] and the mechanism that airway microbiota influences exacerbation frequency requires further exploration.

We hypothesised that frequent exacerbation phenotype of COPD may be driven by a variety of cellular and molecular endotypes and related to specific biomarkers. This article describes the objectives and design of endotype-driven prediction of acute exacerbations in chronic obstructive pulmonary disease (EndAECOPD)—a 5-year prospective cohort study, aiming to systematically assess the role of host genetics, airway inflammation and

respiratory microbiome in predicting clinical outcomes of COPD, in particular exacerbation risk, which may provide potential targets in the effort to modify a patient's exacerbation phenotype.

METHODS AND ANALYSIS
Study design

This protocol adheres to the Standard Protocol Items: Recommendations for Interventional Trials reporting guidelines.[15] Participants with doctor-diagnosed COPD will be consecutively recruited in a prospective cohort study. Participants will attend a baseline clinic visit, thereafter will be followed up every 3 months for 12 months. The effect of host genetics, airway inflammation and respiratory microbiome on predicting COPD exacerbations will be examined. The primary outcomes are total number of exacerbations, severe exacerbations, moderate exacerbations and time to first exacerbation, and the secondary outcomes are changes in lung function and symptom scores. The study flow is shown in figure 1.

Patient and public involvement

Patients and the public were not and will not be involved in the design, conduct and reporting of the research.

Eligibility

Participants will be screened and recruited from a COPD patient database and a respiratory ambulatory care clinic in the West China Hospital, a 4300-bed tertiary care hospital. Inclusion criteria are an age of 40–80 years; diagnosed COPD according to a history of risk factor exposure (tobacco smoke, occupational dusts, etc.), chronic respiratory symptoms (dyspnoea, cough or sputum production) and a ratio of forced expiratory volume in 1 s (FEV_1) to forced vital capacity (FVC) of less than 0.7 after bronchodilator use[16]; GOLD stage II–IV (FEV_1 ≤80% of predicted value after bronchodilator use); and willing to provide informed consent. Particularly, participants have been stable regarding their COPD in the past 2 months which is confirmed by the absence of hospitalisations, AND emergency room visits, AND outpatient/community clinic visits, AND intravenous or oral administration of antibiotics or corticosteroids because of the deterioration of their COPD symptoms. Participants with antibiotics use in the past 2 months due to other conditions, OR other respiratory diseases, OR a history of pneumonectomy, OR solid and haematological malignancies, OR severe organic diseases, OR psychiatric disorders will be excluded. Detailed inclusion and exclusion criteria are listed in Box 1.

Clinical assessments
Baseline visit

Demographics, smoking status, symptoms assessed by the COPD assessment test[17] and the Modified Medical Research Council Dyspnoea Scale,[18] exacerbation history prior to the study entry, medical history and
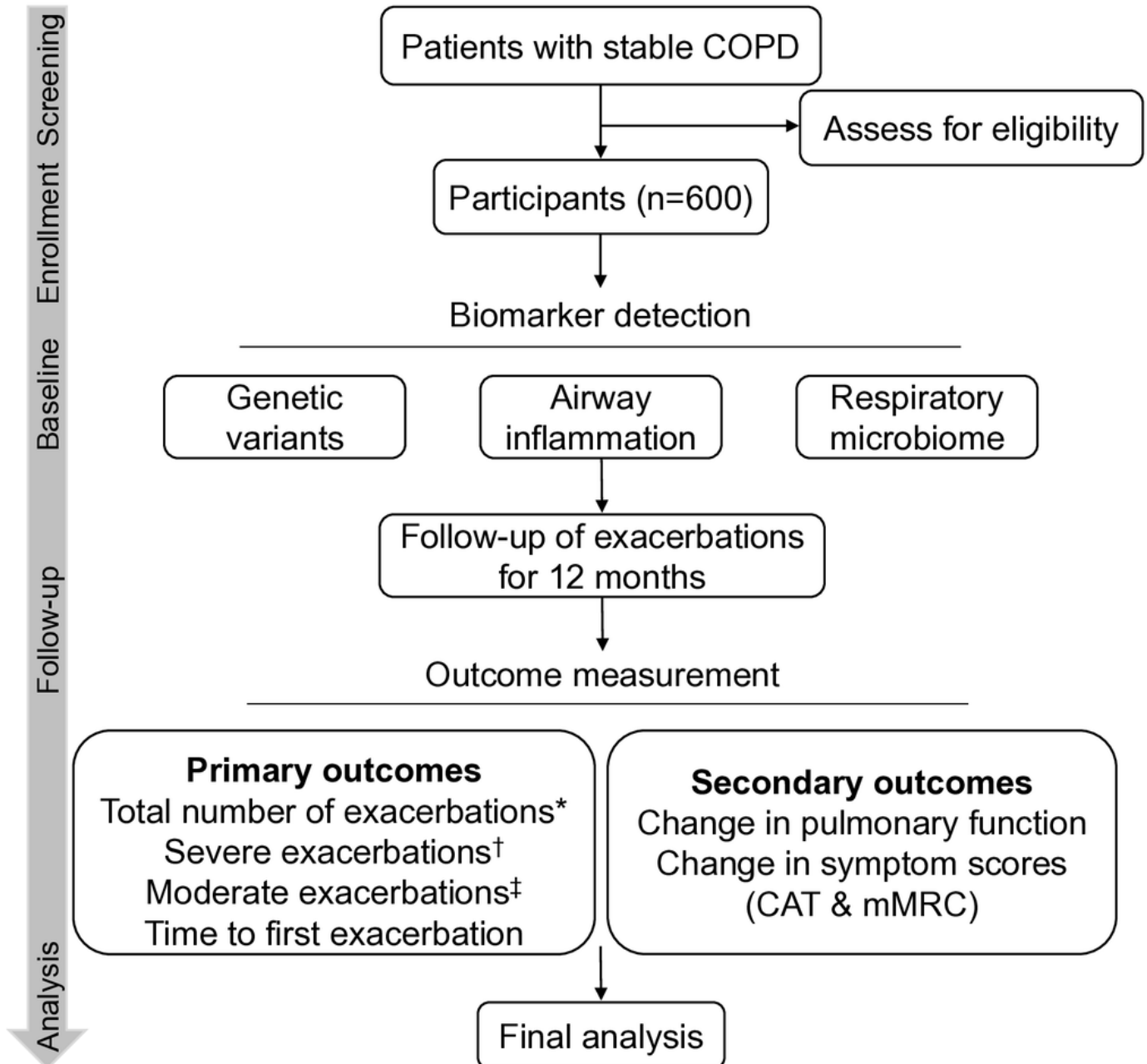
Figure 1 Study flowchart. *Total number of exacerbations refers to the sum of severe and moderate exacerbations. [†]Severe exacerbations are denoted as exacerbations requiring emergency room visits or hospitalisations due to acute worsening of respiratory symptoms. [‡]Moderate exacerbations are exacerbations requiring treatment with antibiotics or systemic steroids. CAT, COPD assessment test; COPD, chronic obstructive pulmonary disease; mMRC, the modified medical Research Council dyspnoea scale.

medication use will be recorded (table 1). Spirometry will be performed according to the American Thoracic Society standard.[19] Participants will withhold bronchodilators for their duration of action before testing. Three reproducible measurements of FEV in 1 s and FVC will be obtained (Medgraphics MGC Diagnostic Corporation, USA) before and after inhalation of 400 μg salbutamol via a metered dose inhaler with valved holding chamber (Taian Character Polymer Co, China). Predicted values of lung function will be calculated according to the Intermountain Thoracic Society for Asian populations.[20]

Follow-up and exacerbation capture

An exacerbation episode is identified as acute worsening of respiratory symptoms (cough, sputum, wheezing, dyspnoea or chest tightness) with a duration of at least 3 days requiring treatment with short acting bronchodilators only (mild exacerbation) OR antibiotics or systemic steroids (moderate exacerbation) OR emergency room visits or hospitalisations (severe exacerbation).[7 21] Only moderate and severe COPD exacerbations will be analysed in this study. Frequent exacerbator is defined as a

---

Box 1   Study inclusion and exclusion criteria

Inclusion criteria*
► Aged 40–80 years.
► Written informed consent obtained from the participant.
► Confirmed diagnosis of COPD by 2017 GOLD criteria:
  – COPD symptoms (dyspnoea, chronic cough or sputum production).
  – Risk factors (tobacco smoke, occupational dusts, etc.).
  – FEV1/FVC<0.70 after bronchodilator use.
► GOLD stage II–IV (FEV$_1$ ≤80% of predicted value after bronchodilator use).
► Having been stable in the previous 2 months confirmed by the absence of the deterioration of COPD symptoms leading to:
  – Hospitalisations/emergency room visits.
  – Outpatient/community clinic visits.
  – Self-administration of systemic antibiotics or corticosteroids.

Exclusion criteria†
► Confirmed diagnosis of respiratory diseases other than COPD (asthma, cystic fibrosis, non-cystic fibrosis bronchiectasis, lung cancer, tuberculosis, etc.).
► History of pneumonectomy.
► Solid or haematological malignancies.
► Severe organic diseases (cerebral infarction, liver failure, chronic renal failure, etc.).
► Psychiatric disorders.
► Antibiotic use in the previous 2 months due to other conditions, for example, intestinal infection.

*Participants must satisfy all of the inclusion criteria at study entry.
†Participants satisfying any item of the exclusion criteria will be excluded.
COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

---

participant who has two or more moderate or severe exacerbations during the 12 months of follow-up.[4]

Participants will be asked to contact the researchers and referred to respiratory physicians in our study centre to evaluate whether the patient is undergoing an episode of exacerbation and the severity of exacerbation when experiencing deteriorations of their daily symptoms according to above criteria. Respiratory physicians will decide whether to perform chest radiography to exclude acute pneumonia in case patients develop two or more of the following symptoms and signs: fever, new or increasing cough or sputum production, dyspnoea, chest pain or new focal signs on chest examination. Community-acquired pneumonia is defined as the presence of two or more of these signs and symptoms and a consolidation in the chest radiograph consistent with acute infection.[22] Patients will be re-evaluated after the treatment of exacerbation to further differentiate the nature of the episode from pneumonia. Telephone reviews will also be routinely conducted with participants to assess exacerbations, medication use and symptom scores every 3 months for 12 months after study entry (table 1). Respiratory hospitalisations, emergency room visits, unscheduled out-patient clinic visits, and medication use including antibiotics and systemic corticosteroids will be recorded at each assessment. To minimise recall bias, an exacerbation follow-up calendar was designed and participants will be asked to record on the calendar for the worsening of their symptoms. A clinical visit will be conducted at the end of study including the same measurements as baseline visit.

Sputum induction, processing and quality control

Prior to sputum induction, patients will be requested to brush their teeth for 3 min and gargle with sterile saline for three times. Sputum induction using hypertonic saline will be conducted as previously reported.[23] Briefly, subjects will breathe in 4.5% saline from an ultrasonic nebuliser for doubling time periods (from 30 s to 4 min), and a fixed cumulative sputum induction time of 15 min will be applied.

Gathered sputum will be transferred to the laboratory on ice within 30 min. First, 100 µL of sputum plug will be isolated for 16S ribosomal RNA (rRNA) gene analysis and bacterial DNA quantification. The rest sputum plug (≥100 µL) will be processed to prepare supernatant for inflammatory mediators assay and cytocentrifuge slides for inflammatory cell counting as described previously.[23]

Sputum samples with a squamous cell percentage greater than 50% or cell viability less than 40% will be excluded. Saliva samples after oral cleaning were collected from seven eligible participants and compared with sputum plugs to assess whether sputum microbiota would be contaminated by oral microbiota.

Biological assessments

Blood and induced sputum will be collected for single nucleotide polymorphism (SNP) genotyping, inflammatory cell counting, mediators assay and microbiota analysis (table 1). Biomarkers assayed in this study are shown in figure 2.

Blood collection and SNP genotyping

Venepuncture will be performed to collect 2 mL of peripheral venous blood into Vacutainer EDTA tubes (BD, USA). Genomic DNA will be extracted using Dzup Genomic DNA Isolation Reagent (Sangon Biotech, China) according to the manufacturer's instructions.

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) method allows highly accurate allelic discrimination of greater than 99.9% by direct measurement of molecular mass and enables high-throughput SNP genotyping of the blood group,[24] which will be used for SNP genotyping of *FUT2* (rs1047781), *ADRB2* (rs1042713), *F2R* (rs2227744), *GC* (rs4588), *MBL2* (rs11003125, rs7096206 and rs1800450), *SIGLEC9* (rs2075803), *HHIP* (rs13118928), *SFTPB* (rs3024791, rs2118177, rs2304566 and rs1130866) and *CCL1* (rs2282691). Briefly, the SNP-containing regions will be amplified by multiplex PCR and undergo subsequent extension reactions with specific single base extension primers. After purification with a cationic exchange resin, the final nucleotide extension products will be

Table 1 Overview of study assessments performed at baseline and at scheduled follow-up visits

Clinical assessments	Clinical visit 1 Baseline	Phone interview 1 3 months	Phone interview 2 6 months	Phone interview 3 9 months	Clinical visit 2 12 months
Demographics*	X				X
Smoking status	X				X
Medication use†	X				
Current medication†	X	X	X	X	X
Chest CT scan‡	X				
Comorbidities (Charlson Index)	X				X
Exacerbations	X	X	X	X	X
Symptom score (mMRC, CAT)	X	X	X	X	X
Spirometry§	X				X
Biological assessments					
Blood sampling					
SNP genotyping	X				
Sputum sampling					
Inflammatory cell counting	X				X
Inflammatory mediators detection	X				X
16S rRNA gene sequencing	X				X
Bacterial DNA quantification	X				X

*Including age, sex, height and weight.

†COPD related medication use including short/long acting beta$_2$-agonists, short/long acting muscarinic antagonists, inhaled/oral corticosteroids, theophylline as well as leukotriene receptor antagonists.

‡Required for the exclusion of other respiratory diseases: bronchiectasis, lung cancer, pulmonary fibrosis, etc.

§Including post-bronchodilator FEV_1, rate of decline in FEV_1, FVC and FEV_1/FVC.

CAT, COPD assessment test; mMRC, the Modified Medical Research Council Dyspnoea Scale; SNP, single nucleotide polymorphism.

transferred to the 384-format SpectroCHIP and analysed on MassARRAY MALDI-TOF System (Sequenom). Genotype determination will be performed using TYPER software (Sequenom).

Microbiota analysis
Bacterial DNA isolation
To isolate bacterial DNA, 100 µL of sputum plug will be lysed and homogenised using the FastDNA SPIN Kit and the FastPrep Instrument (MP Biomedicals, Santa Ana, CA). Specimens will be randomly processed in order to minimise the risk of false pattern formation due to reagent contamination.[25]

Bacterial DNA quantification
Bacterial DNA will be quantified using an ABI7500 qPCR system (Applied Biosystems, USA) using the PCR primers 5′- ACTCCTACGGGAGGCAGCAG-3′ (Eub338) and 5′- ATTACCGCGGCTGCTGG-3′ (Eub518). PCR run will be initiated at 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 58°C for 30 s and 72°C for 40 s.

pMD18 T vector (Takara, Japan) carrying 16S gene will be serially diluted to generate a standard curve for accurate quantification. Three replicates will be performed for each sample.

16S rRNA gene amplification and sequencing
The V3–V4 hypervariable region of the 16S rRNA gene will be amplified using barcoded 338 F-806R primer pairs (338F: 5′-barcode-ACTCCTACGGGAGGCAGCAG-3′; 806R: 5′-barcode-GGACTACHVGGGTWTCTAAT-3′), where the barcode is sample specific. The cycling protocol will be 1 cycle at 95°C for 3 min, 27 cycles at 95°C for 30 s, 55°C for 30 s and 72°C for 45 s, and 1 cycle at 72°C for 10 min. The PCR product will then be visualised on a 2% agarose gel, from which the bands of the specific size (approximately 750 bp) will be cut and purified with AxyPrep DNA gel extraction kit (Axygen, USA). Recovered DNA will be quantified using the QuantiFluor-ST (Promega, USA) and amplicons will be pooled in equimolar amounts. Paired-end sequencing will be performed
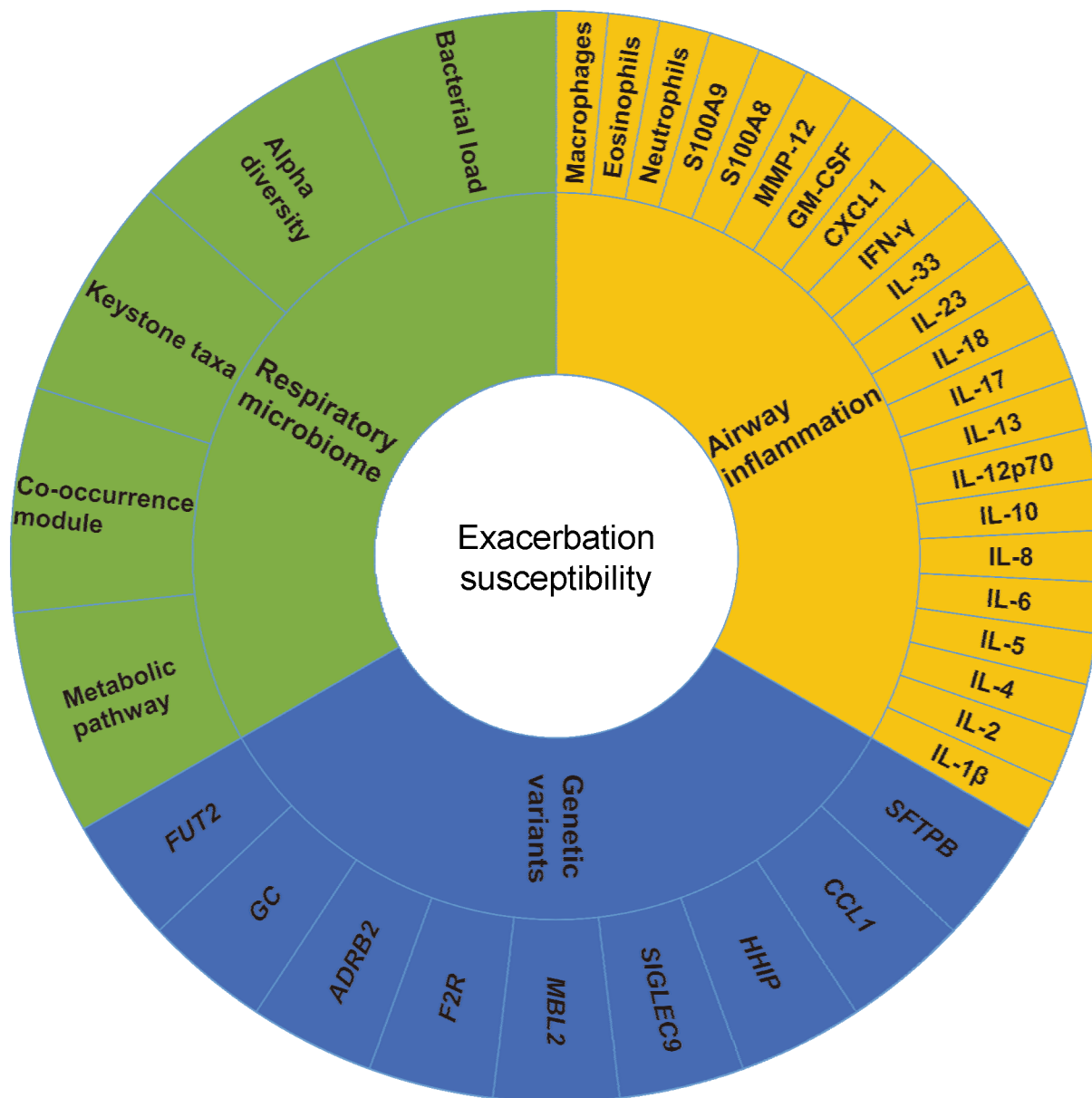
Figure 2 Overview of multiple biomarkers analysed in this study.

on an Illumina MiSeq PE300 platform according to published protocols.[26]

Sputum inflammatory mediators detection

Inflammatory mediators including interleukin (IL)–1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, IL-13, IL-17, IL-18, IL-23, IL-33, interferon-γ, C-X-C motif chemokine ligand 1, granulocyte–macrophage colony-stimulating factor, S100A8, S100A9 and matrix metalloproteinase-12 in the sputum supernatant will be assayed in duplicates using a Human Magnetic Luminex Assay kit (R&D Systems, USA) on the Luminex LX-200 platform (Luminex, USA) according to the manufacturer's instructions (R&D Systems, USA).

Sample size calculation

Sample size was calculated based on the primary outcome of exacerbation frequency. An optimal sample size that

enables accurate estimation of exacerbation frequency is the premise to assess the effects of biomarkers on predicting exacerbation frequency. In China, each patient with COPD experiences an average of two exacerbations per year.[27] If 180 participants are followed for 1 year, around 360 exacerbation events would be captured. Assuming the distribution of exacerbation events per participant follows a Poisson distribution with no overdispersion, an overdispersion factor of 1.5 or an overdispersion factor of 2, the approximate values of the lower and upper bounds of the 95% CI around the point estimate of two events per participant per year would be 1.8–2.2, 1.7–2.3 and 1.7–2.3, respectively, which should ensure sufficient precision in estimating exacerbation frequency of COPD. Given that COPD is a highly heterogenous disease with diverse clinical phenotypes and multiple biomarkers will be investigated in this study, sample size

will be expanded to a total of 600 participants to address potential endotypes underlying exacerbations of COPD. Since frequent exacerbators account for approximately 1/3 of patients with COPD across all GOLD stages and exacerbations in preceding year is the strongest predictor of future exacerbation risk,[4] at least 200 patients with high exacerbation risk based on a history of exacerbations will be recruited to ensure sufficient patients with high exacerbation risk during study follow-up.

Sequence data processing and statistical analysis

Raw paired reads will be demultiplexed, quality filtered and merged using Fast Length Adjustment of Short Reads V1.2.11 software. After chimaeras are removed, sequences will be rarefied to a fixed depth corresponding to the sample containing the fewest sequences. The classification of operational taxonomic units (OTUs) clustered at an identity threshold of 97% will be performed by implementation of the Ribosomal Database Project Classifier against the SILVA 16S rRNA database (release 132) with a 70% confidence cut-off using Quantitative Insights Into Microbial Ecology (QIIME) V2 according to published protocols.[28]

Microbial ecology analysis will be performed in QIIME and R software V3.5.1. Alpha-diversity (taxa abundance and evenness within sample) and beta-diversity (differences between samples) will be calculated based on rarefied OTU tables. Principal co-ordinates analysis will be performed for visualisation of differences in microbiota between groups and permutational multivariate analysis of variance with 999 permutations will be used to determine the significance. Co-occurrence network analysis at genus level will be performed by using Sparse Correlations for Compositional data (SparCC)[29] and co-occurrence modules and keystone species will be detected by using R package *igraph*. Metabolic pathway prediction from 16S data will be performed by Phylogenetic Investigation of Communities by Reconstruction of Unobserved States.[30]

Univariate analysis will be performed to evaluate the association of each biomarker with exacerbation frequency, lung function and symptom scores. Multivariable logistic and Poisson regression analyses will be performed to further examine these biomarkers as predictors of frequent exacerbator phenotype and exacerbation frequency. The receiver operating characteristic curve analysis will be used to assess the ability of continuous biomarkers in identifying frequent exacerbator phenotype. Kaplan-Meier survival analysis will be used to investigate the prediction of biomarkers on time-to-first exacerbation and confounders will be adjusted by a Cox proportional hazard model.

PILOT RESULTS

To examine the possibility of oral microbiome contamination on sputum samples, we quantified and compared bacterial 16S rRNA gene copies in sputum plug and saliva
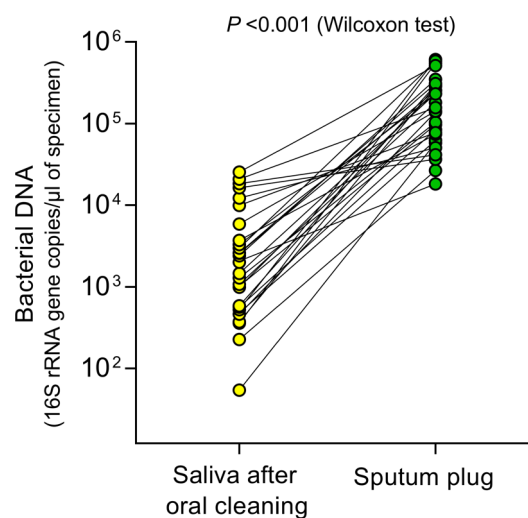

Figure 3 Comparison of bacterial load in paired sputum and saliva specimens. Sputum plug contained significantly more bacterial DNA than saliva after oral cleaning (p<0.001).

after oral cleaning (teeth brushing plus oral wash) from 30 stable COPD patients. Remarkably, bacterial DNA copy number per µL in sputum plug was 124.8 (20.1–351.0) times greater than that in saliva after oral cleaning (Wilcoxon test, p<0.001) (figure 3).

DISCUSSION

Although reducing exacerbation risk has been recognised as one of the therapeutic goals of COPD, strategies for the prevention of AECOPD have not been altered significantly for many years. The major difficulty in the prevention of exacerbations in COPD is the heterogeneity of the disease that is manifested by diverse phenotypes and endotypes. Current treatment strategies based on clinical measurements of symptoms and exacerbation history fail to target the pathophysiology of the frequent exacerbator phenotype, therefore certain patients with COPD will continue to suffer frequent exacerbations despite aggressive medical maintenance therapy.[31]

Advance has been made towards an endotype-driven prediction and treatment of exacerbations represented by the use of ICS based on blood eosinophil counts.[7] Although a number of studies have been conducted to identify AECOPD-predicting biomarkers that enables disease endotyping, most of which are retrospective or cross-sectional design,[32–35] or only examine a single candidate gene or biomarker.[36–38] This is the first study to assess the associations of biological mechanisms relating to COPD pathophysiology (genetics, inflammation and microbiome) with exacerbation susceptibility using a wide range of biomarkers. The current study will provide a comprehensive picture of an endotype-driven prediction of AECOPD and potential targets for treatment and prevention.

SNPs in a group of genes have been found to be associated with COPD.[39 40] Remarkably, the protein products encoded by several of these genes, for example, *CHRM3,*

have become pharmacological targets for the treatment of COPD and asthma.[39] We conducted a thorough search on the genetic predispositions of AECOPD, and identified a panel of eight candidate genes (*ADRB2*,[41] *F2R*,[42] *GC*,[43] *MBL2*,[9] *SIGLEC9*,[44] *SFTPB*,[45] *HHIP*[46] and *CCL1*[8]) that are significantly associated with exacerbation susceptibility. Besides, recent studies on asthma[47] and bronchiectasis[48] have identified a new genetic variant, that is, *FUT2* nonsense mutation, associated with respiratory exacerbations, which will also be investigated in this study for its association with AECOPD. Non-replication or conflicting reports of association between genetic variants and AECOPD risk are common in genetic studies, even when identical variants are interrogated,[13] therefore validating these candidate genes in a different cohort is necessary. It is known that most human diseases are not due to a single genetic defect but involve the interaction of many different genes. We will incorporate these risk loci in one study to validate their role in exacerbation prediction.

Persistent airway inflammation is a cardinal feature of COPD, which is exaggerated at exacerbations. We have previously shown that sputum IL-1β was a significant predictor of AECOPD.[49] A hypothetic association between airway inflammation and exacerbation susceptibility was established; however, exacerbation frequencies were collected retrospectively in vast majority of studies,[50] which may be due to the fact that a history of frequent exacerbation aggravates airway inflammation. COPD lungs are infiltrated by various inflammatory cells and enriched with mediators released by these cells, covering Th1, Th2 and Th17 immune responses.[51] Therefore, it is necessary to acquire a comprehensive airway inflammatory profile for their role in predicting future exacerbations, while adjusting for previous exacerbations. Identifying specific inflammatory mediators responsible for frequent exacerbation will lead to the development of specific targeted anti-inflammation therapy for the prevention of AECOPD.

Airway bacterial colonisation increases rate of exacerbations and elevated levels of both airway and systemic inflammation.[52] Use of macrolides has been recommended for patients treated with triple therapies but still experiencing frequent exacerbations,[7 21] while a recent study found the effect of macrolides on shifting microbial composition in the airways.[53] Investigation into respiratory microbiome using next generation sequencing technologies provides an opportunity to understand COPD heterogeneity and exacerbations. Recent evidences have shown that airway microbiota is associated with historical exacerbations, exacerbation severity and 1-year mortality following COPD hospitalisations.[11 12 54 55] However, whether baseline airway microbiota could predict future exacerbations remains unknown. A longitudinal study is required to answer this question and untangle the causality between microbiota and AECOPD.

16S rRNA gene sequencing is a well-established and sensitive technology for identification of microbes. However, regarding sputum specimens, oral flora especially under circumstance of poor dental hygiene can be the potential sources of sampling contamination. We developed the method of stringent oral cleaning including teeth brushing and oral wash with sterile saline before sputum induction. By comparing paired oral and sputum samples in the pilot study, we showed that bacterial burden in sputum plug was much higher than that in saliva after oral cleaning. This sputum sampling approach minimises the microbial contamination from oral cavity into sputum specimens, which guarantees the accuracy of the study results to the utmost.

There are several highlights of this study. We will conduct a prospective study to systematically evaluate the validity of numerous biomarkers representing multiple pathobiological mechanisms of COPD in predicting exacerbations. Second, a comprehensive picture of biomarkers covering multiple biological levels (genetics, inflammation and microbiome) will lead to high-resolution disease endotyping and provide future therapeutic targets. Third, changes of inflammation and microbiota at study entry and at endpoint will be assayed along with exacerbation frequency, so as to assess the stability of these biomarkers and the effects of exacerbations on their levels. Fourth, the method of sputum sampling by stringent oral cleaning prevents oral contamination on sputum microbiota, which will ultimately represent authentic airway microbiota.

Certain limitations of the study have to be considered. This is s single-centre study that will be conducted in a tertiary hospital which somewhat limits the generalisation of the study results, although a large sample size of participants will be investigated. We will only evaluate a limited number of candidate biomarkers from different biological levels instead of a 'hypothesis-free' approach. We are unable to determine the types of exacerbations, for example, bacterial or eosinophilic exacerbations, and relate them to different biomarkers. However, this study will assess a large number of biomarkers at baseline for the prediction of future exacerbations including total number of exacerbations, severe exacerbations, moderate exacerbations and time to first exacerbation in a large cohort of patients, thus has great potential in identification of different endotypes. Significant physiological biomarkers can be further investigated for their role in specific exacerbation phenotypes in future studies.

CONCLUSION

A precision medicine strategy for the management of COPD is not yet ready to be implemented in clinical practice. Endotypes of the disease need to be better understood and their specific biomarkers need to be validated. We present here the protocol of the EndAECOPD study aiming at identifying pathophysiological biomarkers from multiple levels for the prediction of AECOPD. Identifying endotypes underlying exacerbation susceptibility and specific biomarkers may help to move current management strategy towards precision treatment of the disease.

Trial status

The trial was registered on 24 October 2018 and is currently in the early stage of recruitment.

Ethics and dissemination

The Declaration of Helsinki will always be complied with. All participants will volunteer to attend the study, sign an informed consent and have the right to withdraw from the study at any time. The study results will be published on peer-reviewed journals.

ORCID iD
Wei Xiao http://orcid.org/0000-0003-3170-8629

REFERENCES

1. Wedzicha JA, Seemungal TAR. Copd exacerbations: defining their cause and prevention. *Lancet* 2007;370:786–96.
2. Seemungal TAR, Donaldson GC, Paul EA, et al. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;157:1418–22.
3. Strassels SA, Smith DH, Sullivan SD, et al. The costs of treating COPD in the United States. *Chest* 2001;119:344–52.
4. Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363:1128–38.
5. Lötvall J, Akdis CA, Bacharier LB, et al. Asthma endotypes: a new approach to classification of disease entities within the asthma syndrome. *Journal of Allergy and Clinical Immunology* 2011;127:355–60.
6. De Greve G, Hellings PW, Fokkens WJ, et al. Endotype-driven treatment in chronic upper airway diseases. *Clin Transl Allergy* 2017;7:22.
7. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, 2019. Available: http://wwwgoldcopdcom
8. Takabatake N, Shibata Y, Abe S, et al. A single nucleotide polymorphism in the CCL1 gene predicts acute exacerbations in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006;174:875–85.
9. Dicker AJ, Crichton ML, Cassidy AJ, et al. Genetic mannose binding lectin deficiency is associated with airway microbiota diversity and reduced exacerbation frequency in COPD. *Thorax* 2018;73:510–8.
10. Bhowmik A*et al*. Relation of sputum inflammatory markers to symptoms and lung function changes in COPD exacerbations. *Thorax* 2000;55:114–20.
11. Wang Z, Singh R, Miller BE, et al. Sputum microbiome temporal variability and dysbiosis in chronic obstructive pulmonary disease exacerbations: an analysis of the COPDMAP study. *Thorax* 2018;73:331–8.
12. Wang Z, Bafadhel M, Haldar K, et al. Lung microbiome dynamics in COPD exacerbations. *Eur Respir J* 2016;47:1082–92.
13. Wan ES. Examining genetic susceptibility in acute exacerbations of COPD. *Thorax* 2018;73:507–9.
14. Wedzicha JA, Brill SE, Allinson JP, et al. Mechanisms and impact of the frequent exacerbator phenotype in chronic obstructive pulmonary disease. *BMC Med* 2013;11:181.
15. Chan A-W, Tetzlaff JM, Altman DG, et al. Spirit 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200–7.
16. Global initiative for chronic obstructive lung disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, 2017. Available: http://wwwgoldcopdcom
17. Jones PW, Harding G, Berry P, et al. Development and first validation of the COPD assessment test. *Eur Respir J* 2009;34:648–54.
18. Bestall JC, Paul EA, Garrod R, et al. Usefulness of the medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54:581–6.
19. Anon. Standardization of spirometry, 1994 update. American thoracic Society. *Am J Respir Crit Care Med* 1995;152:1107–36.
20. Morris AH, Kanner RE, Crapo RO. *Clinical pulmonary function testing: a manual of uniform laboratory procedures*. Salt Lake City, UT: Intermountain Thoracic Society, 1984.
21. Albert RK, Connett J, Bailey WC, et al. Azithromycin for prevention of exacerbations of COPD. *N Engl J Med* 2011;365:689–98.
22. Pifarre R, Falguera M, Vicente-de-Vera C, et al. Characteristics of community-acquired pneumonia in patients with chronic obstructive pulmonary disease. *Respir Med* 2007;101:2139–44.
23. Gibson PG, Wlodarczyk JW, Hensley MJ, et al. Epidemiological association of airway inflammation with asthma symptoms and airway hyperresponsiveness in childhood. *Am J Respir Crit Care Med* 1998;158:36–41.
24. Ross PL, Lee K, Belgrader P. Discrimination of single-nucleotide polymorphisms in human DNA using peptide nucleic acid probes detected by MALDI-TOF mass spectrometry. *Anal Chem* 1997;69:4197–202.
25. Salter SJ, Cox MJ, Turek EM, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;12:87.
26. Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme J* 2012;6:1621–4.
27. Cai B-qiang, Cai S-xi, Chen R-chang, et al. Expert consensus on acute exacerbation of chronic obstructive pulmonary disease in the people's Republic of China. *Int J Chron Obstruct Pulmon Dis* 2014;9:381–95.
28. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–6.
29. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 2012;8:e1002687.
30. Langille MGI, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 2013;31:814–21.
31. Wan ES, DeMeo DL, Hersh CP, et al. Clinical predictors of frequent exacerbations in subjects with severe chronic obstructive pulmonary disease (COPD). *Respir Med* 2011;105:588–94.
32. Kim TH, Oh DK, Oh Y-M, et al. Fibrinogen as a potential biomarker for clinical phenotype in patients with chronic obstructive pulmonary disease. *J Thorac Dis* 2018;10:5260–8.
33. Oh JY, Lee YS, Min KH, et al. Difference in systemic inflammation and predictors of acute exacerbation between smoking-associated COPD and tuberculosis-associated COPD. *Int J Chron Obstruct Pulmon Dis* 2018;13:3381–7.
34. Osterburg AR, Lach L, Panos RJ, et al. Peripheral blood natural killer cell phenotypes associated with exacerbation risk in patients with chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2018;15:S288–9.
35. Gaugg MT, Nussbaumer-Ochsner Y, Bregy L, et al. Real-Time breath analysis reveals specific metabolic signatures of COPD exacerbations. *Chest* 2019;156:269–76.
36. Miłkowska-Dymanowska J, Białas AJ, Szewczyk K, et al. The usefulness of soluble receptor for advanced glycation end-products in the identification of COPD frequent exacerbator phenotype. *Int J Chron Obstruct Pulmon Dis* 2018;13:3879–84.

37 Wells JM, Parker MM, Oster RA, *et al*. Elevated circulating MMP-9 is linked to increased COPD exacerbation risk in SPIROMICS and COPDGene. *JCI Insight* 2018;3.

38 Gulati S, Wells JM, Urdaneta GP, *et al*. Fibroblast growth factor 23 is associated with a frequent Exacerbator phenotype in COPD: a cross-sectional pilot study. *Int J Mol Sci* 2019;20:2292.

39 Wain LV, Shrine N, Artigas MS, *et al*. Genome-Wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet* 2017;49:416–25.

40 Busch R, Hobbs BD, Zhou J, *et al*. Genetic association and risk scores in a chronic obstructive pulmonary disease meta-analysis of 16,707 subjects. *Am J Respir Cell Mol Biol* 2017;57:35–46.

41 Rabe KF, Fabbri LM, Israel E, *et al*. Effect of ADRB2 polymorphisms on the efficacy of salmeterol and tiotropium in preventing COPD exacerbations: a prespecified substudy of the POET-COPD trial. *Lancet Respir Med* 2014;2:44–53.

42 Platé M, Lawson PJ, Hill MR, *et al*. Impact of a functional polymorphism in the PAR-1 gene promoter in COPD and COPD exacerbations. *Am J Physiol Lung Cell Mol Physiol* 2014;307:L311–6.

43 Ishii T, Motegi T, Kamio K, *et al*. Association of group component genetic variations in COPD and COPD exacerbation in a Japanese population. *Respirology* 2014;19:590–5.

44 Ishii T, Angata T, Wan ES, *et al*. Influence of *SIGLEC9* polymorphisms on COPD phenotypes including exacerbation frequency. *Respirology* 2017;22:684–90.

45 Foreman MG, DeMeo DL, Hersh CP, *et al*. Polymorphic variation in surfactant protein B is associated with COPD exacerbations. *Eur Respir J* 2008;32:938–44.

46 Pillai SG, Kong X, Edwards LD, *et al*. Loci identified by genome-wide association studies influence different disease-related phenotypes in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010;182:1498–505.

47 Innes AL, McGrath KW, Dougherty RH, *et al*. The H antigen at epithelial surfaces is associated with susceptibility to asthma exacerbation. *Am J Respir Crit Care Med* 2011;183:189–94.

48 Taylor SL, Woodman RJ, Chen ACH, *et al*. *FUT2* genotype influences lung function, exacerbation frequency and airway microbiota in non-CF bronchiectasis. *Thorax* 2017;72:304–10.

49 Fu J-J, McDonald VM, Baines KJ, *et al*. Airway IL-1β and systemic inflammation as predictors of future exacerbation risk in asthma and COPD. *Chest* 2015;148:618–29.

50 Tumkaya M, Atis S, Ozge C, *et al*. Relationship between airway colonization, inflammation and exacerbation frequency in COPD. *Respir Med* 2007;101:729–37.

51 Brusselle GG, Joos GF, Bracke KR. New insights into the immunology of chronic obstructive pulmonary disease. *The Lancet* 2011;378:1015–26.

52 Fuschillo S, Martucci M, Donner CF, *et al*. Airway bacterial colonization: the missing link between COPD and cardiovascular events? *Respir Med* 2012;106:915–23.

53 Taylor SL, Leong LEX, Mobegi FM, *et al*. Long-Term Azithromycin Reduces *Haemophilus influenzae* and Increases Antibiotic Resistance in Severe Asthma. *Am J Respir Crit Care Med* 2019;200:309–17.

54 Mayhew D, Devos N, Lambert C, *et al*. Longitudinal profiling of the lung microbiome in the aeris study demonstrates repeatability of bacterial and eosinophilic COPD exacerbations. *Thorax* 2018;73:422–30.

55 Leitao Filho FS, Alotaibi NM, Ngan D, *et al*. Sputum microbiome is associated with 1-year mortality following COPD hospitalizations. *Am J Respir Crit Care Med* 2018;199:1205–13.