University of Helsinki
and
Heart and Lung Center, Division of Pulmonary Medicine,
Helsinki University Hospital,
Helsinki, Finland

Chronic obstructive pulmonary disease (COPD)
-- from biomarkers to clinical phenotypes

Jing Gao

ACADEMIC DESSERTATION

To be presented, with the permission of
the Faculty of Medicine, University of Helsinki,
for public examination in Biomedicum Helsinki 1, Lecture hall 1
Haartmaninkatu 8, Helsinki on March 9th 2018, at 12 noon

Helsinki 2018
Supervised by
Docent Ville Pulkkinen, PhD
Principal Investigator, Group leader
Department of Medicine, Division of Pulmonary Medicine
University of Helsinki
Helsinki, Finland

Docent Witold Mazur, MD, PhD
Co-group leader, Specialist in Respiratory Medicine
Heart and Lung Center, Division of Pulmonary Medicine,
Helsinki University Hospital and University of Helsinki
Helsinki, Finland

Reviewed by
Docent, Anne Pietinalho, MD
Chief physician
Social and Health Care
Raasepori, Finland

Docent, Siri Lehtonen, PhD
Oulu University Hospital and University of Oulu
Oulu, Finland

Opponent at the Dissertation
Professor, Terttu Harju, MD, PhD
Respiratory Research Unit of Internal Medicine, Medical Research Center Oulu
Oulu University Hospital and University of Oulu
Oulu, Finland

http://ethesis.helsinki.fi

Unigrafia
Helsinki 2018
TO MY FAMILY
CONTENTS

LIST OF ORIGINAL PUBLICATIONS .................................................................................. 7
ABBREVIATIONS .............................................................................................................. 8
LIST OF TABLES AND FIGURES ..................................................................................... 11
ABSTRACT ....................................................................................................................... 12
TIIVISTELMÄ .................................................................................................................. 15
INTRODUCTION ............................................................................................................... 17
REVIEW OF THE LITERATURE ...................................................................................... 20
1. Definition of COPD and pathogenesis of COPD ......................................................... 20
   1.1. Definition of COPD ................................................................................................. 20
   1.2. Pathogenesis of COPD ............................................................................................ 20
   1.3. Systemic inflammation in COPD ............................................................................. 21
   1.4. Major inflammatory cells in COPD ......................................................................... 21
      1.4.1. Neutrophils ....................................................................................................... 21
      1.4.2. Eosinophils ...................................................................................................... 22
      1.4.3. Differences in inflammatory cells between COPD and Asthma ......................... 22
2. Clinical phenotypes in COPD ..................................................................................... 23
   2.1. Recognized COPD phenotypes on the basis of current evidence ......................... 23
   2.2. Identification of clinical COPD phenotypes with biomarkers ............................... 23
3. Diagnosis of COPD ..................................................................................................... 24
   3.1. Classification of COPD severity ............................................................................ 24
   3.2. Asthma-COPD overlap syndrome (ACOS) ............................................................ 26
4. Epidemiology and aetiology of COPD ..................................................................... 26
5. Effect of smoking in the airway in COPD ................................................................. 28
6. Treatment strategies for COPD .................................................................................. 29
   6.1. Smoking cessation ................................................................................................. 29
   6.2. Pulmonary rehabilitation therapy ........................................................................... 29
   6.3. Pharmacologic therapy for stable COPD ............................................................... 29
   6.4. Vaccination, oxygen therapy and lung-related surgery for COPD ....................... 30
7. Prognosis of COPD .................................................................................................... 30
8. Novel biomarkers in early diagnosis of COPD and in characterisation of ACOS .... 31
   8.1. Biomarkers requirement in early diagnosis of COPD and in characterisation of ACOS..31
   8.2. Potential biomarkers for early diagnosis of COPD ................................................. 32
      8.2.1. Transglutaminase 2 (TGM2) ............................................................................ 32
      8.2.2. Bactericidal/permeability-increasing protein fold containing protein B1 (BPIFB1)...32
      8.2.3. Soluble receptor for advanced glycation end-products (sRAGE) ....................... 33
   8.3. Potential biomarkers for ACOS characterisation ................................................... 33
      8.3.1. Neutrophil gelatinase-associated lipocalin (NGAL) ......................................... 33
8.3.2. Myeloperoxidase (MPO) ............................................................................. 34
8.3.3. Surfactant protein A (SP-A) ................................................................. 34
8.3.4. Chitinase-like protein YKL-40 (YKL-40) ........................................... 34
8.3.5. Interleukin 6 (IL-6) ............................................................................. 35
8.3.6. Interleukin 13 (IL-13) ........................................................................ 35
9. Proteomics study ....................................................................................... 35

AIM OF THE STUDIES ..................................................................................... 37

MATERIALS AND METHODS ........................................................................ 38
1. Projects ..................................................................................................... 38
2. Cohorts ..................................................................................................... 38
  2.1. Main cohorts for sputum and plasma samples ....................................... 38
    2.1.1. The 6-year longitudinal cohort of adult smokers for early diagnosis of COPD studies (Study I, II and III) ................................................................. 38
    2.1.2. The Finnish discovery cohort and the Japanese replication cohort for ACOS studies (Study IV and V)....................................................................................................................... 39
  2.2. Lung tissue samples ............................................................................. 39
3. Sample collection ....................................................................................... 40
  3.1. Sputum induction and processing ......................................................... 40
  3.2. Plasma collection and processing ........................................................ 40
4. Lung function measurement ...................................................................... 40
5. Laboratory methods ............................................................................... 41
  5.1. Two-dimensional difference gel electrophoresis (2D-DIGE) (Study I and II) ................ 41
  5.2. Mass spectrometry (MS) (Study I and II) ............................................. 41
  5.3. Western blot analysis (Study I and II) ................................................. 41
  5.4. Immunohistochemistry and histochemistry (Study II) ......................... 42
  5.5. Development of functional enzyme-linked immunosorbent assay (ELISA) (Study III) ................ .............................................................................................................................. 42
  5.6. Commercial ELISA (Study I-V) and Magnetic Human High Sensitivity Luminex assay (Luminex assays, Study V) .................................................................................................................. 43
6. Statistical analyses .................................................................................... 43
RESULTS .......................................................................................................... 44
1. Subject characteristics ............................................................................ 44
2. Potential biomarkers in early diagnosis of COPD .................................... 44
  2.1. TGM2 in mild-moderate COPD (Study I) ............................................ 44
    2.1.1. Proteomic changes in lung tissue ..................................................... 44
    2.1.2. Sputum and plasma TGM2 levels in smokers and COPD .............. 44
    2.1.3. Correlation between TGM2 and clinical variables ....................... 44
  2.2. Sputum BPIFB1 levels in smokers and COPD in a longitudinal study (Study II) ........ 47
    2.2.1. Presence of secreted and glycosylated BPIFB1 isoform 1 in smokers and COPD...... 47
    2.2.2. BPIFBA1 and BPIFB1 levels in smokers and COPD .................... 48
2.2.3. Correlation between BPIFB1 and clinical variables in cross-sectional and longitudinal studies…………………………………………………………………………………………………………………………………………………………………………………………………………………..49
2.2.4. BPIFB1 staining is elevated in remodelled airway epithelium in severe COPD……...50
2.3. sRAGE in airway disease in a longitudinal study (Study III) .................................50
  2.3.1. Plasma sRAGE and HMGB1 levels at baseline.............................................50
  2.3.2. Correlation between sRAGE and clinical variables in cross-sectional and longitudinal studies........................................................................................................................................................................................................................................50
3. Differences and characterisation of biomarkers for ACOS (Study IV and V) .......51
  3.1. COPD-related biomarkers (Study IV and V) .....................................................51
    3.1.1. SP-A and sRAGE as pneumocyte-derived markers in ACOS (Study IV) ........51
    3.1.2. MPO and NGAL as neutrophil-derived molecules in ACOS (Study IV and V)....51
  3.2. Asthma-related biomarkers (Study V) ...............................................................52
    3.2.1. IL-13 as a type 2 T helper cells marker in ACOS...........................................52
  3.3. Other potential biomarkers (Study V) ...............................................................53
    3.3.1. IL-6 as a biomarker of impaired lung function in ACOS.............................53
    3.3.2. YKL-40 as a biomarker of airway inflammation in ACOS............................53
DISCUSSION ...............................................................................................................54
1. Summary of background and results........................................................................54
2. Proteomics-based specific biomarkers in COPD development (Study I, II and III)......55
3. The promising biomarkers in ACOS studies with two well-defined cohorts (Study IV and V)...............................................................................................................................58
4. The difference between ACOS and ACO and the diagnostic criteria for ACOS in our studies (Study IV and V)............................................................................................................60
5. The roles of biomarkers in precision medicine of COPD........................................61
6. Strengths and limitations of the study........................................................................61
CONCLUSIONS ...........................................................................................................63
FUTURE RESEARCH ..................................................................................................65
ACKNOWLEDGEMENTS ............................................................................................66
REFERENCES .............................................................................................................68
LIST OF ORIGINAL PUBLICATIONS


<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-D DIGE</td>
<td>two-dimensional difference gel electrophoresis</td>
</tr>
<tr>
<td>AATD</td>
<td>alpha-antitrypsin deficiency</td>
</tr>
<tr>
<td>AAT</td>
<td>alpha-1-antitrypsin</td>
</tr>
<tr>
<td>AT I</td>
<td>alveolar type I</td>
</tr>
<tr>
<td>ATS</td>
<td>American Thoracic Society</td>
</tr>
<tr>
<td>ACO</td>
<td>asthma and chronic obstructive pulmonary disease overlap</td>
</tr>
<tr>
<td>ACOS</td>
<td>asthma and chronic obstructive pulmonary disease overlap syndrome</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>ANXA3</td>
<td>annexin A3</td>
</tr>
<tr>
<td>AUC</td>
<td>area under curve</td>
</tr>
<tr>
<td>B</td>
<td>regression coefficient</td>
</tr>
<tr>
<td>BALF</td>
<td>bronchoalveolar lavage fluid</td>
</tr>
<tr>
<td>BLVR</td>
<td>bronchoscopic lung volume reduction</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>BPIFA1</td>
<td>bactericidal/permeability-increasing fold-containing protein A1</td>
</tr>
<tr>
<td>BPIFB1</td>
<td>bactericidal/permeability-increasing fold-containing protein B1</td>
</tr>
<tr>
<td>BTS</td>
<td>British Thoracic Society</td>
</tr>
<tr>
<td>CAD</td>
<td>chronic airway diseases</td>
</tr>
<tr>
<td>CAT</td>
<td>COPD assessment test</td>
</tr>
<tr>
<td>CAPG</td>
<td>macrophage-capping protein</td>
</tr>
<tr>
<td>CC-16</td>
<td>clara cell secretory protein-16</td>
</tr>
<tr>
<td>CHI3L1</td>
<td>chitinase 3-like 1</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>cm</td>
<td>centimetre</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>COPD</td>
<td>chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>CRMP2</td>
<td>collapsin response mediator protein-2</td>
</tr>
<tr>
<td>CTSD</td>
<td>cathepsin D</td>
</tr>
<tr>
<td>CV</td>
<td>coefficient of variation</td>
</tr>
<tr>
<td>DLCO</td>
<td>diffusion capacity</td>
</tr>
<tr>
<td>DPYS1L2</td>
<td>dihydropyrimidinase-related protein 2</td>
</tr>
<tr>
<td>DTE</td>
<td>dithioerythritol</td>
</tr>
<tr>
<td>ECLIPSE</td>
<td>the Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>ECP</td>
<td>eosinophilic cationic protein</td>
</tr>
</tbody>
</table>
ELISA  enzyme-linked immunosorbent assay
ERS  European Respiratory Society
FABP5  fatty acid-binding protein
FeNO  fractional exhaled nitric oxide
FEV1  forced expiratory volume in one second
FTND  Fagerström test for nicotine dependence
FVC  forced vital capacity
GAPDH  glyceraldehyde-3-phosphate dehydrogenase
GINA  the Global Initiative for Asthma
GOLD  the Global Initiative for Chronic Obstructive Lung Diseases
HBB  hemoglobin subunit beta
HC  healthy control
HCgp-39  human cartilage glycoprotein 39
HMGB1  high mobility group box 1
HS  healthy smokers
HSPB8  heat shock cognate 71 kDa protein
HRQoL  health-related quality of life
HUH  Helsinki University Hospital
ICS  inhaled corticosteroid
IEF  isoelectric focusing
Ig  immunoglobulin
IL  interleukin
iNOS  inducible nitric oxide synthase
IPF  idiopathic pulmonary fibrosis
KRT7  keratin type II cytoskeletal 7
LABA  Long-acting beta2 agonist
LAMA  Long-acting muscarinic antagonist
LVRS  lung volume reduction surgery
Luminex assay  magnetic human high sensitivity luminex assay
MGG  May-Grunwald-Giemsa
mMRC  Medical Research Council Dyspnoea Scale
MMP  matrix metalloproteinase
MPO  myeloperoxidase
MS  mass spectrometry
NADPH  nicotinamide adenine dinucleotide phosphate
NIV  non-invasive ventilation
NO  nitric oxide
PARC/CCL-18  pulmonary and activation-regulated chemokine
PEF  peak expiratory flow
PBS  phosphate buffered saline
PIGR polymeric immunoglobulin receptor
PLUNC palate, lung, and nasal epithelium clone
RAGE receptor for advanced glycation end products
ROC receiver operating characteristic
ROS reactive oxygen species
RNA ribonucleic acid
SD standard deviation
SE standard error
SEM standard error of the mean
SERPINA1 α-1-antitrypsin/serpin family A member 1
SELENBP1 selenium-binding protein 1
SNP single-nucleotide polymorphism
SP surfactant protein
SPSS Statistical Package for the Social Sciences
Tc cytotoxic T cells
Th T-helper cells
Th2 type 2 T helper cells
TGM transglutaminase
TNF tumour necrosis factor
TPP1 tripeptidyl-peptidase 1
WB Western blotting
WBC white blood cells
WHO World Health Organization
YKL-40 chitinase-like protein YKL-40
LIST OF TABLES AND FIGURES

Table 1. Severity of airflow limitation and clinical severity in COPD, page 25.
Table 2. Features of asthma, COPD and ACOS, page 27.
Table 3. Subject characteristics in this work, page 45.
Table 4. Sputum cell counts and spirometry data in this work, page 46.
Table 5. Correlation of plasma and sputum TGM2 with demographics, page 46.

Figure 1. Pathway to diagnosis of COPD with key indicators, page 24.
Figure 2. Assessment ABCD groups in COPD, page 25.
Figure 3. COPD-specific changes in the human lung tissue proteome, page 47.
Figure 4. Sputum and plasma TGM2 levels, page 48.
Figure 5. Sputum BPIFB1 levels measured by a newly developed method using an ELISA-based assay, page 49.
Figure 6. Levels of plasma sRAGE, page 50.
Figure 7. Expression levels of sputum biomarkers (z-scores, median value) in patients with asthma, COPD and ACOS in the discovery and replication cohorts, page 52.
ABSTRACT

Chronic obstructive pulmonary disease (COPD) is a common, preventable and treatable chronic airway disease (CAD) and is the 4th leading cause of illness and death worldwide. CAD represents a group of diseases with different clinical phenotypes, airway inflammation, obstruction and destruction of the lung parenchyma. COPD is a heterogeneous disease, which is characterized by persistent respiratory symptoms and airflow limitation and its prevalence varies considerably across populations. Asthma-COPD overlap syndrome (ACOS) had been reported in 2009 and recognised by Global Initiative for Asthma (GINA) and Global Initiative for Chronic Obstructive Lung Disease (GOLD) in 2014. ACOS is identified by the features that it shares with both asthma and COPD. Inhaling tobacco smoke is the main risk for CAD, thus smoking cessation is not only the key step for prevention and maintenance therapy for COPD (GOLD 2017), but also it is encouraged in the patient with asthma and their families (GINA 2017). Smoking cessation, pulmonary rehabilitation, pharmacologic therapy and the other therapies for COPD may improve symptoms, health status and exercise ability and reduce frequency of exacerbations (GOLD 2017). At present, COPD is generally diagnosed by lung function measurement using spirometry, but not all individuals presenting with airflow limitation follow the paradigm. On the individual level spirometry may poorly correlate with symptoms, risk of exacerbation, prognosis and response to treatment in COPD. Due to disease heterogeneity, limitations of spirometry and lack of predictive molecular markers, the mechanisms underlying different COPD phenotypes are still unclear. Therefore, identification of the similarities and differences of COPD clinical phenotypes is important, as these phenotypes require different therapeutic approaches and have distinct clinical outcomes. We hypothesised that the potential biomarkers may play an important role and even predict the development of COPD. Accordingly, patients with asthma, COPD and ACOS would exhibit different biomarker profiles. Furthermore, distinguishing COPD phenotypes via distinct molecular and cellular pathways may facilitate the development of individualised diagnosis and precision medicine.

The first goal of this study was to identify COPD-specific proteomic changes involved in disease onset and severity. The second goal was to identify novel biomarkers for the clinical COPD phenotypes in smokers and patients with COPD, asthma and ACOS in cross-sectional and longitudinal studies. The third goal was to investigate whether the levels of these markers are associated with variables such as pack-years, lung function and sputum cell profiles. Finally, we aimed to evaluate the association of the biomarker levels with a longitudinal decline of lung function. Those studies included the 6-year longitudinal cohort of adult smokers for early diagnosis of COPD studies (Study I, II and III), and the Finnish discovery cohort and the Japanese replication cohort for ACOS studies (Study IV and V).
We selected biomarkers identified in our non-hypothesis-based proteomic approaches for early COPD diagnosis, such as transglutaminase 2 (TGM2), bactericidal/permeability-increasing (BPI) fold-containing protein B1 (BPIFB1) and soluble receptor for advanced glycation end-products (sRAGE). Based on our previous results and non-hypothesis-based proteomic approaches, we developed two panels of potential biomarkers for ACOS: COPD-related biomarkers and asthma-related biomarkers. The four COPD-related biomarkers are neutrophil gelatinase-associated lipocalin (NGAL), myeloperoxidase (MPO), surfactant protein A (SP-A) and sRAGE. The three asthma-related biomarkers are chitinase-like protein YKL-40 (YKL-40), interleukin-6 (IL-6) and interleukin-13 (IL-13).

For the early COPD diagnosis studies, induced sputum and plasma samples were categorised into 3 groups: non-smokers, smokers without COPD and smokers with COPD (stable stage I and stage II-III). Lung tissue samples were obtained from non-smokers, smokers without COPD, smokers with COPD and patients with idiopathic pulmonary fibrosis (IPF). For the ACOS studies, induced sputum and plasma samples were categorised into 5 groups: non-smokers, healthy smokers, patients with asthma, COPD and ACOS. Sputum cell counts were evaluated from sputum cytospins. TGM2 and BPIFB1 were assessed by Western blot analysis and by cysteine-specific two-dimensional difference gel electrophoresis (2D-DIGE) coupled with mass spectrometry (MS). BPIFB1 was further detected by immunohistochemistry and functional enzyme-linked immunosorbent assay (ELISA). TGM2, sRAGE, NGAL, MPO, SP-A, YKL-40, IL-6 and IL-13 levels were measured by commercial ELISA or Magnetic Human High Sensitivity Luminex assay (Luminex assay).

Our results showed that TGM2 levels in sputum and plasma were elevated in mild-moderate COPD and associated with lung function. Sputum BPIFB1 levels were elevated in smokers with COPD, whereas plasma sRAGE levels were decreased in smokers without COPD and smokers with COPD. Sputum BPIFB1 and plasma sRAGE levels were significantly associated with reduced airflow limitation over time in smokers with COPD. In the Finnish discovery cohort of the ACOS study, sputum MPO and plasma SP-A levels were significantly elevated in ACOS, whereas plasma sRAGE levels were decreased in ACOS. Sputum IL-13 levels were increased in ACOS when compared with controls. Sputum YKL-40 and IL-6 differentiated ACOS from COPD and asthma based on receiver operating characteristic (ROC) curve analysis. However, only sputum NGAL results could be repeated from the Finnish discovery cohort to Japanese replication cohort. Importantly, NGAL differentiated ACOS from COPD and independently correlated with forced expiratory volume in one second % (FEV₁% predicted) in both cohorts.

In conclusion, our findings support the hypothesis that monitoring of these biomarkers may be useful for differential diagnosis between the clinical phenotypes of COPD. TGM2 may be a
potential diagnostic and therapeutic marker in COPD, which may relate to COPD onset and severity. BPIFB1 and sRAGE play vital roles in the pathogenesis of smoking-related COPD. High concentrations of sRAGE might be involved in the protective role of the airway. Patients with COPD and ACOS exhibit different NGAL profiles in sputum. These findings suggest that these potential biomarkers could be a novel diagnostic and therapeutic target for COPD. On the other hand, sputum biomarkers reflect both airway inflammation and tissue remodelling, and monitoring these biomarkers may be useful for differential diagnosis between asthma, COPD and ACOS.
TIIVISTELMÄ


Tämän tutkimuksen ensimmäisen osatyön tavoitteena oli löytää proteiiniprofiilien seulontamenetelmän (proteomiikka) avulla keuhkoahtaumataudin puhekeamiseen ja taudin vaikeusasteeseen liittyviä täysin uusia biomarkereita. Seuraavissa osatöissä otamme huomioon yleensä astman, keuhkoahtaumataudin varhaisempaan toteamiseen ja sen erilaisten alamuotojen taudinmuotojen tarkastamisessa käytävien tutkimusten aiheuttamaa taudin muotoa sekä astman, keuhkoahtaumataudin ja ACOS-potilaatit kerättyjä plasma- ja yskösnäytteitä. Lisäksi tavoitteena oli selvittää löydettyjen biomarkkerien ja taudin ilmiasuun liittyvien kliinisten omakseen ominaisuuksien välisiä yhteyksiä.

Proteomiikan ja kirjallisuushakujen perusteella valitsimme TGM2, BPIFB1 ja sRAGE proteiinit biomarkereiksi ja selvitimme niiden toimivuutta keuhkoahtaumataudin varhaisessa tunnistamisessa. Lisäksi selvitimme, voidaanko keuhkoahtaumataudin (NGAL, MPO, SP-A and sRAGE) ja astmaan ainetut biomarkkereita (YKL-40, IL-6 and IL-13) käyttää ACOS:n tunnistamiseen.

INTRODUCTION

By 2020, chronic obstructive pulmonary disease (COPD) is projected to be the 3rd leading cause of death in the world (Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2017). According to the World Health Organization (WHO) report, approximately 3 million people in the world die as a consequence of COPD yearly (Diaz-Guzman et al. 2014). COPD is a common, preventable and treatable disease characterized by persistent respiratory symptoms and airflow limitation (GOLD 2017). It is a heterogeneous disease with multiple phenotypes (Naz et al. 2017), thus different COPD phenotypes could be associated with different clinical features, prognostic and therapeutic ramifications. Asthma-COPD overlap syndrome (ACOS) is known as a consensus-based clinical phenotype and the patients share features common to both asthma and COPD (Gao et al. 2016, Global Initiative for Asthma (GINA) 2016, GOLD 2017). The public health efforts aimed to better understand and prevent the burden of COPD (Diaz-Guzman et al. 2014), but under-recognition (especially in early disease stages) and under-diagnosis (due to overlap with other diseases) in COPD are still very common (GOLD 2017, Capozzolo et al. 2017, Adeloye et al. 2015). Early-stage COPD is challenging to diagnose, because the patients with mild-moderate COPD may have early changes in systemic inflammation, pathology and physiology, but the patient may not have significant clinical symptoms (Hogg et al. 2004, Rodriguez-Roisin et al. 2009, Agusti et al. 2012). ACOS patients experience poorer outcomes than those with asthma or COPD alone (GINA 2016, GOLD 2017), but the pathophysiological features of ACOS are poorly understood (Iwamoto et al. 2014).

Air pollution, chronic bronchitis, infections, biomass smoke and dietary factors are important risk factors for COPD (Rosenberg et al. 2015, Diaz-Guzman et al. 2014, GOLD 2017). Cigarette smoking remains the one of the main risk factors for chronic airway diseases (CAD). Smoking is associated with COPD, increased asthma severity and with accelerated lung function decline in both current and former smokers (Vonk et al. 2017, GOLD 2017, Nakamura et al. 2009, Eisner et al. 2007). Therefore, smoking cessation has the great capacity to influence lung function trajectory and progression of COPD (GOLD 2017), and also plays the roles in asthma (GINA 2017). Nicotine replacement therapy and other medications (e.g. varenicline, bupropion and nortriptyline) are used as a component in the prevention and treatment for COPD (GOLD 2017). It has reported that long-term quitting success rates of up to 25% can be achieved if effective resources and time are dedicated (van Eerd et al. 2016). The history of severe childhood respiratory infection could be a risk factor for COPD, which is associated with declined lung function and increased respiratory symptoms in adulthood (de Marco et al. 2011). In addition, HIV infection and tuberculosis has been identified as the risk factor for COPD (Drummond et al. 2014, Byrne et al. 2015).
The patients with COPD commonly have the declined exercise tolerance on the basis of the symptoms, such as dyspnoea (Katajisto et al. 2017). Pulmonary rehabilitation is based on exercising, which has been proven to effectively influence the prevention and maintenance therapy in COPD (GOLD 2017). It improves symptoms, quality of life, and physical emotional participation in daily activities (GOLD 2017). The pharmacologic therapy is widely used for COPD, which could contribute to reducing symptoms, frequency and severity of exacerbation and improving health status and physical activity (GOLD 2017), but the role of pharmacologic therapy was limited in COPD development and progression, especially in modifying lung function decline (GOLD 2017, Burge et al. 2000, Anthonisen et al. 1994, Pauwels et al. 1999, Vestbo et al. 1999).

Spirometry provides an objective measurement of the presence of airflow limitation, and the level of forced expiratory volume in 1 second (FEV₁) has been recognised as an important value to reflect COPD staging (GOLD 2017, Pellegrino et al. 2005, Naz et al. 2017), but spirometry cannot explain a wide variety of COPD phenotypes. The diagnosis of COPD is based on a post-bronchodilator fixed FEV₁ / forced vital capacity (FVC) ratio <70 %, the presence of symptoms (Segreti et al. 2014) and the risk factors (GOLD 2017). Since the research by Lange (Lange et al. 2015), at least two trajectories showed in the development of airflow limitation, including the trajectory of rapid decline in FEV₁ from a normal pulmonary function and the trajectory of normal decline in FEV₁ from a low FEV₁ in initial. Meanwhile, the level of lung function in early adulthood may be strongly associated with the diagnosis of COPD later in life (Lange et al. 2015). Therefore, early diagnosis has a dominant role in improving prognosis or outcomes in COPD. Accordingly, considering the disease heterogeneity and lack of predictive molecular markers (Naz et al. 2017), identification of sensitive and specific biomarkers is needed to differentiate the multiple COPD phenotypes. These biomarkers could also improve both prevention and understanding of the disease and optimise patient classification for development of precision medicine in COPD. In this thesis, we investigated potential novel biomarkers for clinical COPD phenotypes in sputum, plasma and lung tissue. The studies included two projects and five studies. These projects were the early diagnosis of COPD project (Study I, II and III), which used a 6-year longitudinal cohort of adult smokers, and the characterisation of ACOS project (Study IV and V), which used a Finnish discovery cohort and a Japanese replication cohort.

In the early diagnosis of COPD project (Study I, II and III), we investigated over 80 proteins for COPD by non-hypothesis-based proteomic approaches to investigate the roles of potential biomarkers in COPD development using the 6-year longitudinal cohort of adult smokers. We identified three potential biomarkers for COPD specificity: transglutaminase 2 (TGM2, Study I), bactericidal/permeability-increasing (BPI) fold-containing protein B1 (BPIFB1, Study II) and soluble receptor for advanced glycation end-products (sRAGE, Study III). TGM2 is a
multifunctional enzyme (Lai et al. 2013) and its levels are elevated in the airways (Hallstrand et al. 2010, Maiuri et al. 2008, Witsch et al. 2014). BPIFB1 is localised in airway goblet cells (Bingle et al. 2010, Bingle et al. 2011 (1) 2005 (2)), and sputum BPIFB1 levels are elevated in smokers and in COPD (Ohlmeier et al. 2012). sRAGE is expressed in type I pneumocytes (Iwamoto et al. 2014) and is a pneumocyte-derived biomarker. Plasma sRAGE levels are significantly declined in COPD and accelerated decline is observed in exacerbation stage (Smith et al. 2011, Miniati et al. 2011, Sukkar et al. 2012). The present project investigated the biomarkers in different samples, using lung tissue (Study I and II), sputum (Study I, II and III) and plasma (Study I and III).

In the ACOS project (Study IV and V), we first investigated the four COPD-related biomarkers in sputum and plasma (Study IV), namely neutrophil gelatinase-associated lipocalin (NGAL), myeloperoxidase (MPO), surfactant protein A (SP-A) and sRAGE in the Finnish discovery cohort. In the further study (Study V), we subsequently assessed the COPD biomarkers (NGAL and MPO) and added the asthma-related biomarkers (interleukin (IL)-13) and other potential biomarkers (chitinase-like protein YKL-40 (YKL-40) and IL-6) in sputum, which were analysed in the Finnish discovery cohort and the Japanese replication cohort. NGAL and MPO are known as COPD-related biomarkers, which represent neutrophil inflammation and early airway injury (Iwamoto et al. 2014, Keatings et al. 1997). SP-A and sRAGE are pneumocyte-related biomarkers. SP-A is one of the critical markers in lung host defence related to tobacco use (Ilumets et al. 2011, Mazur et al. 2011, Ohlmeier et al. 2008). sRAGE is an alveolar type I (AT I) cell injury biomarker. IL-13 is a type 2 T helper cells (Th2) marker and is associated with asthmatic inflammation (Dente et al. 2012, Wills-Karp et al. 1998). YKL-40 is secreted by neutrophils (Volck et al. 1998) and is related to chronic airway inflammation (Tang et al. 2010, Otsuka et al. 2012, James et al. 2016). IL-6 is a pro-inflammatory cytokine (Schmidt-Arras et al. 2016) and is associated with impaired lung function (Fu et al. 2014).
1. Definition of COPD and pathogenesis of COPD

1.1. Definition of COPD

As noted in the 2017 update of GOLD strategy, the definition of COPD is a common, preventable and treatable disease which is characterised by persistent airflow limitation and with an enhanced chronic inflammatory response in the airways and the lungs to noxious particles or gases. Chronic bronchitis and emphysema are known clinical characteristics of COPD (Wright et al. 2008, Roggli et al. 2008). Chronic bronchitis is defined as a productive cough of over three months for more than two successive years in general (Barnes et al. 2003) and may be an independent disease entity correlated with development or acceleration (or both) of fixed airflow limitation (GOLD 2017).

The GOLD 2017 strategy highlights the COPD phenotypes and considers the roles of patient symptoms and exacerbation risks in individualised therapeutic decision-making and care (Tudoric et al. 2017). COPD may result in emphysema and small airway fibrosis, further leading to characteristic symptoms of COPD such as dyspnoea, cough and/or sputum production (GOLD 2017, Salvi et al. 2009). These symptoms, however, may be under-reported (GOLD 2017). The characteristic features of COPD include a mixture of small airways disease and parenchymal destruction, such as obstructive bronchiolitis and emphysema. Those features include airway inflammation, airway fibrosis, luminal plugs, increased airway resistance, loss of alveolar attachments and decrease of elastic recoil (GOLD 2017). Tobacco smoking remains the main risk factor in COPD, whereas host factors (genetic abnormalities, abnormal lung development and accelerated aging) may also play crucial roles in COPD development (GOLD 2017, Stern et al. 2007, McCloskey et al. 2001, Rennard et al. 2006, Stoller et al. 2005, Hunninghake et al. 2009, Lawlor et al. 2005). COPD exacerbations lead to acute worsening of respiratory symptoms (GOLD 2017, Hurst et al. 2007, Wedzicha et al. 2007, Seemungal et al.1998, Burge et al. 2003), which is associated with rapid disease development. Additionally, comorbidities occur frequently in patients with COPD (GOLD 2017, Chen et al. 2015, Brenner et al. 2012, Fry et al. 2012).

1.2. Pathogenesis of COPD

The pathological changes characteristic of COPD includes chronic lung inflammation and structural changes caused by repeated injury and repair located in the airways, lung parenchyma and pulmonary vasculature (Hogg et al. 2004). According to the GOLD strategy 2017, COPD pathogenesis is associated with oxidative stress (Barnes et al. 2016, Domej et al. 2014), protease-antiprotease imbalance (Stockley et al.1999, Johnson et al. 2015), inflammatory cells (Barnes et al. 2016), inflammatory mediators (Barnes et al. 2014),
peribronchiolar and interstitial fibrosis (Sze et al. 2015, Churg et al. 2006) and so on. It has been shown that smoking-induced systemic inflammation and oxidative stress are the major risk factors for COPD, and this inflammation persists after smoking cessation (Cosio et al. 2009). Inhaled cigarette smoke and other noxious particles are chronic irritants, leading to chronic inflammation in the airway. This chronic inflammation with increased presence of specific inflammatory cell types may be further amplified by oxidative stress (Rahman et al. 2005).

1.3. Systemic inflammation in COPD

COPD is mainly characterised by chronic bronchitis, emphysema, or a combination (Martinez et al. 2016) and associated with systemic inflammation. COPD involves several types of inflammatory cells and a variety of inflammatory mediators (Barnes et al. 2003). Alterations in circulating inflammatory cells (e.g. neutrophils and lymphocytes) indicate a downstream effect (Sinden et al. 2010). COPD is associated with abnormal inflammatory immune responses of the lung to inhaled substances (Brusselle et al. 2011, Sinden et al. 2010), such as cigarette smoking, passive smoking, air pollution and occupational exposures (Postma et al. 2015). These inhaled substances can enter the circulation and cause systemic inflammation (Sinden et al. 2010). Increased number of neutrophils, macrophages and T lymphocytes are found in the alveoli and small airways of COPD (Retamales et al. 2001, Forsslund et al. 2017). Neutrophilic and CD8 mediated inflammation are considered the main inflammatory players in COPD (Martinez et al. 2016). CD8+ cytotoxic T cells (Tc) are the primary T lymphocytes in COPD and can destroy lung parenchyma through cytolytic activity. On the other hand, CD4+ Th cells recruit and activate other immune cells, such as neutrophils (Forsslund et al. 2017). COPD is associated with several systemic effects (Sinden et al. 2010), such as skeletal muscle dysfunction (a statement of American Thoracic Society (ATS)/European Respiratory Society (ERS). 1999), cardiovascular disease (Sin et al. 2005), osteoporosis (Bolton et al. 2004) and diabetes (Schmidt et al. 1999). Levels of several systemic inflammation-related markers are significantly increased in COPD (Sinden et al. 2010), including C-reactive protein (CRP), IL-6, fibrinogen, activated leukocytes and tumour necrosis factor α (TNFα) (Gan et al. 2004, Sinden et al. 2010). The release of multiple inflammatory mediators, such as chemokines, cytokines and growth factors is another important feature in COPD (Barnes et al. 2003, GOLD 2017). Increased inflammatory mediators are associated with the inflammatory process and structural changes in COPD (GOLD 2017).

1.4. Major inflammatory cells in COPD

1.4.1. Neutrophils

Neutrophils represent nearly 70% of peripheral blood leukocytes (Camicia et al. 2014) and is the major inflammatory cell at the site of acute inflammation. Neutrophils play a critical role in both controlling infection and in dysregulation of immune responses (Shen et al. 2017).
Neutrophils are recruited to the airways and lung parenchyma from the circulation, where the recruited neutrophils could rapidly transit through, and neutrophils in COPD airways can be further activated (Barnes et al. 2003). The levels of activated neutrophils are increased to different degrees in sputum, bronchoalveolar lavage fluid (BALF), airways and lung parenchyma (Keatings et al. 1996, Lacoste et al. 1993, Finkelstein et al. 1995). Neutrophils secrete serine proteases, which can induce tissue damage leading to alveolar destruction (Barnes et al. 2003). Neutrophils can adhere to endothelial cells in the alveolar wall and migrate into the respiratory tract (Hogg et al. 1995, Barnes et al. 2003). For example, upregulated E-selectin levels have been found in endothelial cells in COPD (Di Stefano et al. 1994). NGAL may be actively secreted by neutrophils (Eagan et al. 2010) and detected in the lungs (Cowland et al. 1997). Cigarette smoke may increase the local concentration of neutrophils in the lung (MacNee et al. 1989). In smokers, increased neutrophil levels in sputum are correlated with reduced lung function (Keatings et al. 1996) and are associated with airflow limitation severity in bronchial biopsies (Di Stefano et al. 1998).

1.4.2. Eosinophils
Eosinophils constitute nearly 1% of the leukocytes in the peripheral blood of humans. Eosinophilic inflammation is present in asthma, especially in severe asthma and Th2-high asthma (Varricchi et al. 2017). Although the role of eosinophils in COPD is not fully understood, the increased levels of eosinophil basic proteins in sputum has been observed in COPD (Barnes et al. 2003). Meanwhile, in some patients with COPD, especially with ACOS, increased level of eosinophils and Th2 were observed (GOLD 2017). Eosinophils represent repositories of some cationic proteins, cytokines, chemokines and lipid mediators (Varricchi et al. 2016), and are an important cell type in modulating the functions of a wide spectrum of immune cells (Varricchi et al. 2017). Eosinophils can migrate into allergic inflamed tissues and subsequently adhere to activated endothelial cells (Furuta et al. 2014).

1.4.3. Differences in inflammatory cells between COPD and Asthma
Asthma and COPD are different heterogeneous CAD with multiple clinical phenotypes, which are associated with differences in physiological effects, symptoms, and response to therapy (Fabbri et al. 2003). While asthma and COPD exhibit differences in inflammatory cells and mediators, ACOS patients may combine asthma and COPD inflammatory patterns. Eosinophils, neutrophils or both in the lung are an important characteristic in asthma patients, whereas COPD patients may exhibit systemic inflammation. Severe eosinophilic inflammation-predominant asthma and persistent airflow limitation may be an important cluster in adult-onset asthma (Amelink et al. 2013). Neutrophilic inflammation with increased eosinophil counts are observed in some COPD (Barnes et al. 2016). ACOS as the intersection between asthma and COPD, may have the features of eosinophilic- and neutrophilic-like inflammation.
2. Clinical phenotypes in COPD

2.1. Recognized COPD phenotypes on the basis of current evidence

A phenotype is an observable characteristic obtained from different scales, including the interactions between genes and environment (Barker et al. 2013). The clinical COPD phenotypes are associated with clinical characteristics, natural history or response to therapy (or both) (Burgel et al. 2014). COPD phenotypes can be described at the gene (genomic, epigenomic), cell (cellular arrays, proteomics, metabolomics), tissue (histopathology), organ (lung function and imaging) and individual (symptoms and signs) (Barker et al. 2013) level. Several studies have reported that COPD covers multiple clinical phenotypes with clearly different characteristics (Vanfleteren et al. 2016), for example, chronic bronchitis, emphysematous and ACOS (Cheng et al. 2017, de Oca et al. 2012, Oh et al. 2014, Gelb et al. 2016, Izquierdo-Alonso et al. 2013). Those phenotypes of COPD could be with frequent exacerbations and with infrequent exacerbations (Cheng et al. 2017). It has been reported that approximately 16% the patients with frequent exacerbations have emphysematous (Blasi et al. 2017). Furthermore, ACOS has been recognized as a predictor for frequent exacerbations in COPD (Wan et al. 2011). While chronic bronchitis is an important clinical diagnostic factor in COPD and related to airflow limitation development, chronic bronchitis also exists in the normal lung function population (GOLD 2017). ACOS is not represent single discrete disease entity (Woodruff et al. 2017), defined by shared clinical features from both asthma and COPD, particularly in smokers and older adults (GINA 2016, GOLD 2017). ACOS outcomes are often poorer than those of asthma or COPD alone. The underlying mechanisms behind different ACOS are still not identified (Iwamoto et al. 2014). Recently, ACO has been reconsidered, which is without syndrome (Woodruff et al. 2017). It mainly used to describe two types patients, such as the asthma patients who have features of COPD and the COPD patients with features of asthma (Woodruff et al. 2017). Hence, distinguishing COPD phenotypes can improve the understanding of distinct subgroups, prognostic information and clinical outcome (Barker et al. 2013, Han et al. 2010).

2.2. Identification of clinical COPD phenotypes with biomarkers

The Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE) and Burgel studies confirmed that COPD patients with comparable airflow limitation (e.g. FEV1 % predicted) showed marked heterogeneity in clinical characteristics (Burgel et al. 2014, Agusti et al. 2010). These subjects belonged to different phenotypes and marked differences in clinical characteristics, symptoms, exacerbation rates, exercise capacity, comorbidities, predicted mortality and health status (Burgel et al. 2010, 2014). Identifying the clinical phenotypes of COPD can contribute to understanding of disease pathophysiology, clarifying specific biomarkers and improving the precision, effectiveness and safety of COPD patient treatment (Burgel et al. 2014, Agusti et al. 2017). The newest GOLD strategy redefined
the classification of COPD into ABCD groups, based on assessment of symptoms and risk of exacerbations. Biomarkers may facilitate defining phenotypes of patients with chronic airway diseases (Burgel et al. 2014), which might reflect the pathophysiological mechanism of COPD. Apart from biomarkers, other characteristics can also be used to define COPD phenotypes, including clinical manifestations and outcomes, imaging, pulmonary function and exercise tests (Burgel et al. 2014). It is a challenging to identify COPD-relevant phenotypes and their specific biomarkers on the basis of integrating knowledge of bioinformatics, clinical science, systems biology and network medicine.

3. Diagnosis of COPD

3.1. Classification of COPD severity

In COPD diagnosis, symptom status and risk factors must be evaluated before spirometry. Further spirometry is required to establish a diagnosis of COPD (Figure 1). The assessment of COPD patient symptoms, severity of spirometric abnormality, exacerbation risk and presence of comorbidities may contribute to the classification of COPD severity (GOLD 2017). The COPD Assessment Test (CAT) and modified Medical Research Council Dyspnoea Scale (mMRC) can evaluate the symptoms and quality of life status in COPD (Jones et al. 2009, Kankaanranta et al. 2015, GOLD 2017). The obstruction severity of COPD is assessed based on the FEV\(_1\) % predicted and FEV\(_1\)/FVC (Table 1). In addition, exercise tolerance, level of symptoms, exacerbations and co-morbidities are also considered in evaluation of clinical COPD severity (Table 1). According to the GOLD 2017 recommendations, COPD can be classified into ABCD groups, defined on the basis of symptoms and history of exacerbations (Figure 2), this is in contrast with previous GOLD strategy (Tudoric et al. 2017). The limitations of the FEV\(_1\) affects therapeutic decisions in COPD (GOLD 2017).

Figure 1: Pathway to diagnosis of COPD with key indicators

![Figure 1](image_url)

Figure 1 is modified from GOLD strategy 2017.
Table 1: Severity of airflow limitation and clinical severity in COPD

<table>
<thead>
<tr>
<th>Severity of Airflow Limitation in COPD*</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOLD</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

Clinical severity of COPD*

<table>
<thead>
<tr>
<th>Type</th>
<th>Severity</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mild</td>
<td>CAT &lt;10, no frequent exacerbations and FEV₁ &gt;50% predicted</td>
</tr>
</tbody>
</table>
| 2 | Moderate-Severe** | FEV₁ <50% predicted  
At least 2 exacerbations/year OR 1 hospitalisation because of COPD  
CAT ≥10 OR causes poor quality of life OR impaired exercise tolerance |
| 3 | Very Severe** | FEV₁ <30% predicted  
Chronic respiratory failure  
Frequent exacerbations OR hospitalisations regardless of treatment to COPD  
CAT ≥20 OR causes very poor quality of life OR exercise tolerance |

Table 1 is modified from GOLD strategy 2017 and Kankaanranta et al. Basic Clin Pharmacol Toxicol. 2015. *: all patients with FEV₁/FVC <0.70; **: at least one condition is fulfilled.

Figure 2: Assessment ABCD groups in COPD

Figure 2 is modified from GOLD strategy 2017.
3.2. Asthma-COPD overlap syndrome (ACOS)
According to GINA strategy (GINA 2016) and the Finnish Guidelines (Kankaanranta et al. 2015), ACOS is a phenotype of COPD and is defined by the features that it shares with both asthma and COPD (Table 2). ACOS is characterised by persistent (but not fully reversible) airflow limitation, often with current or historical variability (GINA 2016). ACOS outcomes are often poorer than for asthma or COPD alone, such as lower health-related quality of life (HRQoL), faster disease progression, higher morbidity and mortality, increased health care utilisation and frequency of exacerbation (de Marco et al. 2013, Kauppi et al. 2011, Louie et al. 2013, Miravitlles et al. 2013, Shaya et al. 2008, Soriano et al.2005). ACOS patients often have asthma and a history of noxious substance exposure. The age of ACOS onset is usually > 40 years and is more prevalent in the elderly (Zeki et al. 2011, GINA 2016). The inflammatory cells of ACOS are sputum eosinophils, neutrophils or both, suggesting asthma and inflammation characteristic of COPD (GINA 2016). As ACOS is a group of heterogeneous diseases with multiple phenotypes, defining ACOS remains challenging. Several studies have reported the partial characterisation of ACOS based on clinical features and pathophysiological and genetic identifiers (Hardin et al. 2011, Carolan et al. 2013, Wardlaw et al. 2005). However, further research is needed to understand the mechanisms of ACOS in broad population. Diagnostic confusion between COPD, asthma and ACOS is common based on spirometry, clinical data and symptoms (Rogliani et al. 2016, Silva et al. 2004, Zeki et al. 2011), particularly in distinguishing between asthma with fixed airflow obstruction, COPD and ACOS (GINA 2016). Overall, ACOS is a group of consensus-based clinical diseases with different characteristics, consisting of several phenotypes (Rogliani et al. 2016). Although guidelines for ACOS are continuously updated, knowledge of the disease is still limited, particularly in smokers and older adults. As a heterogencous disease with poorly defined clinical features, it is a challenging to diagnosis ACOS (Tommola et al. 2017). Novel biomarkers may facilitate identification of the similarities and differences in asthma, COPD and ACOS. Such biomarkers may also be useful in clarifying the clinical phenotype (Iwamoto et al. 2014, Gao et al. 2016).

4. Epidemiology and aetiology of COPD
Globally, COPD is the one of the main diseases causing disability, illness and death (Machin et al. 2017, GOLD 2017). The number of COPD cases increased to approximately 390 million in 2010 globally (Adeloye et al. 2015). The prevalence and burden of COPD are high and growing both globally and regionally. However, it is challenging to accurately assess COPD prevalence due to differences in research methods and diagnostic criteria (Halbert et al. 2006). COPD prevalence ranged from 8.4 -15.0% in 2010 (Adeloye et al. 2015, Machin et al.2017). In general, the prevalence of COPD is less than 6% in most countries (GOLD 2017). COPD prevalence is nearly 50% in heavy smokers (Rennard et al. 2006). Age and cigarette smoking
Table 2: Features of asthma, COPD and ACOS

<table>
<thead>
<tr>
<th>Feature</th>
<th>Asthma</th>
<th>COPD</th>
<th>ACOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of onset (usually)</td>
<td>Childhood</td>
<td>&gt;40 years</td>
<td>&gt;40 years</td>
</tr>
<tr>
<td>Pattern of respiratory symptoms</td>
<td>Vary over time</td>
<td>Chronic</td>
<td>Persistent exertional dyspnoea</td>
</tr>
<tr>
<td></td>
<td>Often limiting activity</td>
<td>Continuous symptoms</td>
<td>Variability may be prominent</td>
</tr>
<tr>
<td></td>
<td>Often triggered by exercise, emotions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Airflow limitation</td>
<td>Variable</td>
<td>Persistent</td>
<td>Not fully reversible</td>
</tr>
<tr>
<td>History</td>
<td>Frequently a history of allergies in childhood</td>
<td>Exposure to noxious particles and gases</td>
<td>Frequently a history of doctor-diagnosed asthma</td>
</tr>
<tr>
<td></td>
<td>Frequently a personal history of asthma in childhood</td>
<td></td>
<td>Frequently a history of allergies</td>
</tr>
<tr>
<td></td>
<td>May have family history of asthma</td>
<td></td>
<td>Frequently a family history of asthma</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>May have a history of noxious exposures</td>
</tr>
<tr>
<td>Time course</td>
<td>Often improves spontaneously</td>
<td>Slowly progressive</td>
<td>Reduced symptoms by treatment</td>
</tr>
<tr>
<td></td>
<td>May result in fixed airflow limitation</td>
<td></td>
<td>Usual progression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High-level treatment</td>
</tr>
<tr>
<td>Chest X-ray</td>
<td>Usually normal</td>
<td>Severe hyperinflation</td>
<td>Similar to COPD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other changes of COPD</td>
<td></td>
</tr>
<tr>
<td>Exacerbations</td>
<td>Considerably reduced by treatment</td>
<td>Considerably reduced by treatment</td>
<td>May be more common than in COPD</td>
</tr>
<tr>
<td></td>
<td>Risk can be considerably reduced by treatment</td>
<td></td>
<td>Reduced by treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Airway inflammation</td>
<td>Eosinophils, neutrophils or both</td>
<td>Neutrophils ± eosinophils in sputum</td>
<td>Eosinophils, neutrophils or both in sputum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lymphocytes in airways</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May have systemic inflammation</td>
<td></td>
</tr>
<tr>
<td>Spirometry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-BD* FEV1/FVC &lt;0.7</td>
<td>May improve spontaneously or on treatment</td>
<td>Required for diagnosis</td>
<td>Usually present</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1 ≥80% predicted</td>
<td>Compatible with diagnosis</td>
<td>Compatible with GOLD mild stage</td>
<td>Compatible with mild ACOS</td>
</tr>
<tr>
<td>FEV1 &lt;80% predicted</td>
<td>Compatible with diagnosis</td>
<td>Severity of airflow limitation</td>
<td>Severity of airflow limitation</td>
</tr>
<tr>
<td></td>
<td>Risk factor for exacerbations</td>
<td>Risk of future events</td>
<td>Risk of future events</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-BD* increase in FEV1 &gt;12% and 200 ml from baseline</td>
<td>Usual</td>
<td>Usual</td>
<td>Usual</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-BD* increase in FEV1 &gt;12% and 400 ml from baseline</td>
<td>High probability of asthma</td>
<td>Unusual in COPD</td>
<td>Compatible</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consider ACOS</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 is modified from GINA strategy 2016. * post-BD: post-bronchodilator spirometry.
are well-known major risk factors in COPD (Jyrki-Tapani et al. 2005, Lokke et al. 2006, Mannino et al. 2007, Lindberg et al. 2006, Shahab et al. 2006) and have a synergistic effect (Lindberg et al. 2006). Moreover, advanced age and smoking history are also considered as clinical characteristics in ACOS (GINA 2016). An early smoking starting age is a strong risk factor for COPD (Jyrki-Tapani et al. 2005). Recently, Lopez-Campos and co-workers reported that the standardised mortality of COPD has declined in Western countries, based on the development of smoking cessation tools and other anti-smoking efforts (Lopez-Campos et al. 2016). Nevertheless, COPD prevalence is still high due to the impact of aging and the increasing of air pollution and regional smoking. Tuberculosis history and exposure to biomass are other known risk factors for COPD (Amaral et al. 2015, Hooper et al. 2012, Rivera et al. 2008). In addition, several studies reported that never smokers may account for approximately 30% of patients with COPD (Celli et al. 2005, Behrendt et al. 2005, Lamprecht et al. 2008, Bridevaux et al. 2010). Altogether, COPD still is a challenge for clinicians and for social healthcare worldwide.

5. Effect of smoking in the airway in COPD

COPD is recognised as a smoking-related lung disorder, but not all smokers develop COPD. Susceptible smokers may easily develop irreversible lung obstruction (Fletcher et al. 1977, Faner et al. 2014), which is associated with their genetic and epigenetic background (Vestbo et al. 2013, Lomas et al. 2001, Silverman et al. 2000, DeMeo et al. 2004). Small airway (<2 mm in diameter) remodelling is the major factor leading to airflow obstruction (McDonough et al. 2011). COPD with progressive irreversible airway obstruction is associated with smoking-induced changes in the small airways, and smoking-induced pathological changes play an important role in COPD development (Hogg et al. 2004). Hogg study reported that exposure to tobacco smoke leads to thickened small airway epithelium and increased occlusion of small airways by mucus, which may lead to airway remodelling and further airway obstruction (Hogg et al. 2004). Smoking behaviour is associated with multiple factors and their interactions, including genetic, social and psychological factors. Tobacco-related deaths account for a large proportion of preventable deaths in the world, such as COPD (Rigotti 2013). It is recommended that clinicians encourage smoking cessation in smokers (Wu et al. 2006). Importantly, the GOLD 2016 guideline emphasised that smoking cessation is the key intervention in COPD, based on strong (type A) evidence. Meanwhile, Fletcher study reported that smoking cessation could revert the average further rates of FEV\textsubscript{1} loss to normal in susceptible smokers (Fletcher et al. 1977).
6. Treatment strategies for COPD

6.1. Smoking cessation
Smoking cessation is in central role to avoid COPD. While relapse is common in patients that stop smoking, the ‘5 A’ program (Ask, Advise, Assess, Assist and Arrange) provides a strategic framework for smoking cessation. Nicotine replacement products are an important part of this program (GOLD 2017, Glynn et al.1990), which have different types of in market, including nicotine gum, inhaler, nasal spray, transdermal patch, sublinual tablet, and lozenge (GOLD 2017). Nicotine replacement therapy has been proved to reliably increases long-term smoking abstinence rates (GOLD 2017, the Tobacco Use and Dependence Clinical Practice Guideline Panel. 2000, van der Meer et al. 2003). E-cigarettes are popularly used, but the efficacy still is controversial (GOLD 2017, McNeill et al. 2015, McRobbie et al. 2014). There are several other pharmacological products for smoking cessation, that would be used as a component of a supportive intervention program, such as varenicline, bupropion and nortriptyline (GOLD 2017, Cahill et al. 2013, Tashkin et al. 2011(1) (2)).

6.2. Pulmonary rehabilitation therapy
Pulmonary rehabilitation therapy is an important method in the integrated care of the patient with chronic respiratory disease. As the note of the 2013 updated official ATS/ERS statement (Spruit et al. 2013), pulmonary rehabilitation has been defined as “a comprehensive intervention based on a thorough patient assessment followed by patient-tailored therapies, which include, but are not limited to, exercise training, education, and behavior change, designed to improve the physical and psychological condition of people with chronic respiratory disease and to promote the long-term adherence of health-enhancing behaviors.” Pulmonary rehabilitation may be used for increasing exercise capacity (Lacasse et al. 2002, GOLD 2017), which is efficient when measured by saved hospital days in severe COPD (Katajisto et al. 2017). Meanwhile, new pulmonary rehabilitation methods are currently being investigated.

6.3. Pharmacologic therapy for stable COPD
To date, the pharmacologic therapy for stable COPD mainly include bronchodilators, antimuscarinic drugs, methylxanthines, combination bronchodilator therapy, anti-inflammatory agents, inhaled corticosteroids (ICS), triple inhaled therapy, oral glucocorticoids, phosphodiesterase-4 inhibitors, antibiotics, mucolytic and antioxidant agent, and other pharmacologic treatments (GOLD 2017). Pharmacologic therapy for COPD is to reduce symptoms, reduce the frequency and severity of exacerbations, and to improve exercise tolerance and health status (GOLD 2017). Comparing combination therapy with monotherapy in COPD, combination bronchodilator therapy may increase the degree of bronchodilation with a similar safety profile (GOLD 2017, Vogelmeier et al. 2008). The combination

6.4. Vaccination, oxygen therapy and lung-related surgery for COPD
Prevention of respiratory infections plays an important role in COPD treatment. Accordingly, influenza and pneumococcal vaccination are offered to COPD patients (Wongsurakiat et al. 2003, 2004, Nichol et al. 1994, Edwards et al. 1994, Jackson et al. 2003, Centers for Disease Control and Prevention 2010, Alfageme et al. 2006). Oxygen therapy and lung-related surgery assist COPD treatment. Oxygen therapy has been widely used in chronic respiratory failure patients, and non-invasive ventilation (NIV) is used in severe COPD patients. In addition, invasive COPD treatments, such as lung volume reduction surgery (LVRS), bronchoscopic lung volume reduction (BLVR) and lung transplantation may improve patient quality of life (Christie et al. 2010, Trulock et al.1997, Naunheim et al. 2006, Sciruba et al. 2010). The further research is needed to evaluate the outcomes of these interventions. However, while the existing medications for COPD may improve symptoms, health status and exercise ability and reduce frequency of exacerbations, they do not modify the long-term decline in lung function (GOLD 2017, Burge et al. 2000, Anthonisen et al. 1994, Pauwels et al. 1999, Vestbo et al.1999).

7. Prognosis of COPD

COPD is a common, preventable and treatable chronic airway disease. The mortality of COPD is close to 5% of all deaths globally in 2015 (www.who.int). According to the GOLD strategy 2017, the different GOLD grades of COPD (GOLD 1-4) show varying characteristics. This classification system is defined in 4 grades (from mild to very severe), and can be a predictor for future adverse events. COPD may lead to progressive breathlessness, exacerbations, other serious related illnesses and low quality of life (www.who.int, Donaldson et al. 2005, Flattet et al. 2017, Miravitlles et al. 2000). Frequent exacerbations requiring hospitalisation are related to poor prognosis (Soler-Cataluña et al. 2005). Meanwhile, higher exacerbation rates and high CAT scores (scores ≥ 10) are factors related to reduced quality of life in COPD (Spencer et al. 2004, Jones et al. 2011). Baseline lung function and demographic factors are strongly associated with COPD outcome (Flattet et al. 2017, www.who.int). Tobacco use, biomass fuel exposure and air pollution may contribute to COPD development and may affect COPD prognosis. ACOS outcomes are poorer than asthma or COPD alone, such as poor HRQoL, rapid disease progression, higher morbidity and mortality, greater exacerbation frequency and

8. Novel biomarkers in early diagnosis of COPD and in characterisation of ACOS

8.1. Biomarkers requirement in early diagnosis of COPD and in characterisation of ACOS

COPD is not a single disease and includes multiple phenotypes. Identification of the similarities and differences of COPD clinical phenotypes is increasingly important, as these phenotypes have distinct clinical outcomes and different treatment guidelines. Accurate assessment of COPD phenotype contributes to improved diagnosis, treatment, and reduced disease progression. Currently, COPD is generally diagnosed by symptoms, spirometry and other tests (GOLD 2017). Distinguishing the phenotypes of COPD remains challenging, particularly in the early and overlapping stage. Some clinical phenotypes exhibit poor or absent correlation with lung function decline. In COPD, symptom aggravation may be earlier than lung function decline (Seemungal et al. 2000). Meanwhile, the biological features of ACOS are still poorly understood (Gao et al. 2016, Bonten et al. 2016). Based on lung function tests, current imaging and physiological testing techniques, it remains challenging to unambiguously distinguish ACOS from asthma and COPD (Kauppi et al. 2011, Iwamoto et al. 2014, Hardin et al. 2011, Gao et al. 2016). Accordingly, novel biomarker solutions are needed for COPD clinical assessment, which could be used to accelerate therapeutic discoveries and to improve the care and outcomes (Sin et al. 2015). Cross-sectional and longitudinal research on COPD biomarkers is on-going globally, as it is likely that sensitive and specific biomarkers would improve COPD clinical assessment (Gao et al. 2016, Sin et al. 2015, Hollander et al. 2017). It is crucial to evaluate multipanel biomarkers for COPD, which could not only be used to understand the pathophysiological features and phenotypes for COPD, but could also contribute to improved early COPD diagnosis and clarified the characterisation of ACOS. These findings may ultimately be used for developing precision medicine in COPD.

Several potential biomarkers in sputum and blood have been investigated in COPD in the ECLIPSE study, referring to novel cellular, proteomic, genetic, transcriptomic and metabolomic biomarkers (Faner et al. 2014). Some of these biomarkers have shown promising results in COPD, such as neutrophils (Singh et al. 2010), circulating white blood cells (WBC) (Agusti et al. 2012, Hurst et al. 2010, Celli et al. 2012), fibrinogen (Celli et al. 2012, Duvoix et al. 2013, Casaburi et al. 2013), clara cell secretory protein-16 (CC16) (Lomas et al. 2008, Vestbo et al. 2011 Hanania et al. 2011), surfactant protein D (SP-D) (Hurst et al. 2010, Lomas et al. 2009), Pulmonary and activation-regulated chemokine (PARC/CCL-18) (Sin et al. 2011), soluble receptor for advanced glycation end-products (sRAGE) (Cheng et al. 2013), inflammome (Agusti et al. 2012, Miller et al. 2013), adipokines (Breyer et al. 2012) and Vitamin D (Berg et al. 2013). Although the ECLIPSE study identified a panel of biomarkers
that were associated with risks of poor clinical outcome (Faner et al. 2014), the utility of these biomarkers were still limited in COPD. Sputum neutrophils may not reflect the pathophysiological abnormalities in COPD and its specificity is limited (Faner et al. 2014). Circulating WBC counts associated with COPD, but its specificity is limited in COPD (Faner et al. 2014, Agusti et al. 2012, Hurst et al. 2010, Celli et al. 2012). Plasma fibrinogen has been considered as a robust biomarker in COPD (Faner et al. 2014), which is associated with COPD prognosis (Casaburi et al. 2013). However, the clinical utility of fibrinogen is still limited, particularly with regards to outcomes and COPD risk factors (Faner et al. 2014). Accordingly, further investigations are needed to evaluate new biomarkers and to validate previous results.

8.2. Potential biomarkers for early diagnosis of COPD

8.2.1. Transglutaminase 2 (TGM2)
Transglutaminases (TGM) is associated with lung extracellular matrix proteins (ECM) via protein modifications and crosslinking. TGM may block lung structure modifications and prevent remodelling processes in alveolarisation (Witsch et al. 2014). TGM2 is a multifunctional enzyme localised to the developing alveolar septa and is associated with cancer and autoimmune, inflammatory and neurodegenerative diseases (Ohlmeier et al 2016, Lai et al. 2013, Witsch et al. 2014). To the best of our knowledge, there are no studies of TGM2 in relation to COPD. Some studies have implicated TGM2 levels in asthma (Hallstrand et al. 2010), pulmonary fibrosis (Oh et al. 2011, Olsen et al. 2011) and bronchopulmonary dysplasia (Witsch et al. 2014). These results suggest that TGM2 may play an important role in chronic inflammatory diseases (Maiuri et al. 2008), and could contribute to the development of new treatments for airway inflammatory diseases.

8.2.2. Bactericidal/permeability-increasing protein fold containing protein B1 (BPIFB1)
The bactericidal/permeability-increasing fold-containing (BPIF) / palate, lung, and nasal epithelium clone (PLUNC) genes are located on a single locus on chromosome 20q11.2 (Bingle et al. 2002). Eight of the 11 genes generate functional proteins (Gao et al. 2015). BPIF protein B1 (BPIFB1, previously termed long-PLUNC1) is localised in a population of goblet cells in the nasal passages and in the airway epithelium. BPIFB1 is also present within the serous cells of airway submucosal glands (Bingle et al. 2010, Bingle et al. 2005, 2011). BPIFB1 is strongly expressed in the upper respiratory tract (Bingle et al. 2002, Bingle et al. 2007) and may have a function in innate defence of the airways (Gao et al. 2015). Elevated BPIFB1 levels have been reported in several lung diseases, such as advanced cystic fibrosis (Bingle et al. 2012) and usual interstitial pneumonia (Bingle et al. 2013). In addition, our previous proteomics study revealed that BPIFB1 levels in sputum were increased in smokers and in COPD (Ohlmeier et al. 2012). Therefore, we hypothesised that BPIFB1 is involved in COPD (Gao et al. 2015) and developed a functional enzyme-linked immunosorbent assay (ELISA) kit for BPIFB1.
8.2.3. Soluble receptor for advanced glycation end-products (sRAGE)

The receptor for advanced glycation end-products (RAGE) is a pattern-recognition receptor for an unspecific ligand (Fritz et al. 2011). RAGE belongs to the immunoglobulin superfamily of cell-surface receptors (Fritz et al. 2011, Neeper et al. 1992, Schmidt et al. 1992) and participates in host response to injury, infection and inflammation (Sukkar et al. 2012). RAGE is associated with the inflammatory response and is involved in development of chronic pathologies (Fritz et al. 2011), such as chronic inflammation (Hofmann et al. 1999) and cancer (Taguchi et al. 2000). Interesting, RAGE levels differ between those in the lung and in most normal tissues (Iwamoto et al. 2014). RAGE levels are lowest in patients with severe COPD in lung tissue (Ohlmeier et al. 2010). RAGE is expressed at a high level in AT I cells under normal physiological conditions and is associated with lung homeostatic function (Buckley et al. 2010, Demling et al. 2006). Soluble RAGE (sRAGE) is known as an AT I cell injury biomarker (Buckley et al. 2010) and has anti-inflammatory properties (Park et al. 1998, Englert et al. 2008). Several studies showed that sRAGE is associated with emphysema, reduced diffusion capacity and neutrophilic COPD (Sukkar et al. 2012, Miniati et al. 2011, Iwamoto et al. 2014, Cheng et al. 2013).

8.3. Potential biomarkers for characterisation of ACOS

8.3.1. Neutrophil gelatinase-associated lipocalin (NGAL)

Neutrophil gelatinase-associated lipocalin (NGAL) is an antimicrobial 25-kDa protein that belongs to the lipocalin superfamily (Flower et al. 2000). NGAL covalently binds to gelatinase from human neutrophils (Eagan et al. 2010, Mishra et al. 2003). NGAL is actively secreted by neutrophils, epithelial cells and tubular cells in the kidneys (Kjeldsen et al. 1994, Eagan et al. 2010, Nielsen et al. 1996, Bu et al. 2006, Mishra et al. 2003) and may be detected in the lungs, stomach, colon and kidney (Cowland et al. 1997). The study by Bolignano reported that NGAL levels reflect tubular cell injury and correlate with renal impairment severity (Bolignano et al. 2008). Increased NGAL levels in COPD patients have been detected in sputum, bronchial lavage and plasma (Keatings et al. 1997, Eagan et al. 2010, Ekberg-Jansson et al. 2001). NGAL inhibits of bacterial growth and enhances matrix degradation, which may be associated with early airway injury in smoking-related diseases and involved in COPD pathogenesis (Keatings et al. 1997, Kjeldsen et al. 1994, Eagan et al. 2010). NGAL has a high affinity for siderophores, which helps reduce available iron and inhibit bacterial infection by binding to siderophores secreted by bacteria (Fischbach et al. 2006, Flo et al. 2004). NGAL inhibits matrix metalloprotease 9 (MMP-9) inactivation, which contributes to prolonged effects on collagen degradation (Gupta et al. 2007, Yan et al. 2001). These effects suggest that NGAL may be a potential mediator in acute and chronic respiratory infections and in protease-antiprotease imbalance in COPD (Kjeldsen et al. 1994, Eagan et al. 2010, Fischbach et al. 2006, Flo et al. 2004).
8.3.2. Myeloperoxidase (MPO)
Myeloperoxidase (MPO) is a heme-containing peroxidase and a mediator of neutrophil activity (Zhu et al. 2014). MPO is released from secondary granules following neutrophil activation (van der Veen et al. 2009, Zhu et al. 2014, Klebanoff. 2005). MPO not only affects with the oxidative properties, but also correlates with the processes involved in cell signaling and cell-cell interactions (Van der Veen et al. 2009). Excessive generation of oxidants by MPO is not beneficial in the immune response and may result in host tissue damage associated with acute or chronic inflammatory diseases. (Zhu et al. 2014). Elevated MPO levels have been reported in asbestos-induced lung impairment (Haegens et al. 2005) and in COPD (Keatings et al. 1997, Yamamoto et al. 1997).

8.3.3. Surfactant protein A (SP-A)
Alveolar epithelium repair is involved in the maintenance of lung homeostasis (Nathan et al. 2016). Pulmonary surfactant proteins are significant molecular components in host defence and in inflammatory regulation in the lung (Nathan et al. 2016, Whitsett. et al. 2005). In particular, surfactant protein A (SP-A) plays a critical role in pulmonary defence against inflammation and oxidative stress (Whitsett et al. 2005). SP-A is a highly ordered, collagen-like glycoprotein and a member of the collectin family. Collectins include a calcium-dependent phospholipid-binding protein and the lectins (Thiel et al. 1989, Sastry et al. 1993, McCormack. 2001). SP-A is secreted primarily by alveolar type II and clara cells, and binds to specific cell surface receptors in the lung (Pison et al. 1992, Wright et al. 1989). SP-A affects surfactant structure and function and regulates various microbial species (McCormack. 1998 and 2001, LeVine. et al. 2000). In our previous proteomics studies, we revealed that SP-A is associated with smoking-related lung injury (Ilumets et al. 2011, Mazur et al. 2011, Ohlmeier et al. 2008), and that SP-A levels are elevated in plasma and sputum samples from smokers (Mazur et al. 2011). Increased SP-A levels have been detected in serum (Behera et al. 2005, Kobayashi et al. 2008, Robin et al. 2002) and lung tissue samples (Ohlmeier et al. 2008) from COPD patients.

8.3.4. Chitinase-like protein YKL-40 (YKL-40)
Chitinase-like protein YKL-40, YKL-40 is for short, which also called human cartilage glycoprotein 39 (HCgp-39) and chitinase 3-like 1, encoded by the chitinase 3–like 1 gene CHI3L1 (Chupp et al. 2007, Ober et al. 2008). The chitinase family consists of YKL-40, chitotriosidase and mammalian chitinase; the levels of these proteins are associated with airway inflammation (Lee et al. 2011). YKL-40 binds expressed chitin without the enzymatically active, which lacks of chitinase activity (Chupp et al. 2007, Lee et al. 2011, James et al. 2016, Ober et al. 2008). YKL-40 has several important roles, such as inhibiting oxidant-induced lung injury, augmenting adaptive Th2 immunity, regulating apoptosis, stimulating alternative macrophage activation and improving fibrosis and wound healing (Lee et al. 2011). Elevated YKL-40 levels were observed in serum of patients with asthma and COPD (Tang et al. 2010,
Otsuka et al. 2012, James et al. 2016, Chupp et al. 2007). Serum YKL-40 levels in COPD are higher than those in asthma (Chupp et al. 2007). YKL-40 levels are associated with asthma severity, airway obstruction and airway remodelling (Bargagli et al. 2010, Chupp et al. 2007, Ober et al. 2008). The -131 C/G polymorphism in the CHI3L1 gene (rs4950928) is associated with increased serum YKL-40 levels (Ober et al. 2008). Some studies have shown that YKL-40 levels are associated with subepithelial basal membrane thickness in serum and BALF (Chupp et al. 2007) and linked to profibrotic factors and inflammatory cells after allergen challenge (Gavala et al. 2013).

8.3.5. Interleukin 6 (IL-6)
Interleukin-6 (IL-6) is a proinflammatory cytokine with multiple functions and is mainly synthesised by monocytes, macrophages, T cells, fibroblasts and endothelial cells (Schmidt-Arras et al. 2016, Attaran et al. 2010). IL-6 is involved in systemic inflammation and may play a considerable role in impaired lung function in COPD (He et al. 2009, Karadag et al. 2008, Donaldson et al. 2005) and in mild-moderate asthma (Neveu et al. 2010). Several studies have reported elevated serum IL-6 levels in COPD (Attaran et al. 2010) and ACOS (Fu et al. 2014) patients. Increased IL-6 levels are associated with airflow limitation (Attaran et al. 2010). Only a few studies have investigated the association of IL-6 with respiratory function in asthma (Grubek-Jaworska et al. 2012) and ACOS.

8.3.6. Interleukin 13 (IL-13)
Interleukin-13 (IL-13) is a classical Th2 pneumocyte-derived biomarker and a central mediator in allergic asthma (Grubek-Jaworska et al. 2012). IL-13 is associated with eosinophilic inflammatory disease (Bel et al. 2017). IL-13 is involved in multiple roles in the airway, such as eosinophil recruitment and retention and regulation of airway hyperresponsiveness (Bel et al. 2017). A panel of anti-IL-13 could improve airway inflammation mediated by eosinophils (Bel et al. 2017). Thus far, two monoclonal antibodies against IL-13 (lebrikizumab and tralokinumab) have been developed. Based on phase III trial results, lebrikizumab may be effective in reducing the rate of exacerbations in uncontrolled asthmatics with moderate to severe disease (Hanania et al. 2015, Bel et al. 2017)

9. Proteomics study
Proteomics is a non-biased large-scale screening technique, which is an important method for investigating COPD pathology (Ohlmeier et al. 2016). We have investigated the pathological mechanisms of COPD using proteomics for many years (Ohlmeier et al. 2008, 2010). Based on our proteomic smoking-related and COPD-related studies, a variety of promising molecular markers have been reported. These include SP-A (Ohlmeier et al. 2008), RAGE (Ohlmeier et al. 2008, 2010), polymeric immunoglobulin receptor (PIGR) (Ohlmeier et al. 2012), α-1-
antitrypsin (SERPINA1) and BPIFB1 (Ohlmeier et al. 2012). In this thesis, we selected SP-A, sRAGE and BPIFB on the basis of our proteomic study. However, further COPD research is needed via other methods, as proteomics do not verify the specificity of the identified changes (Ohlmeier et al. 2008). Therefore, our selected biomarkers were further measured by quantitative ELISA after our proteomic studies.
AIM OF THE STUDIES

In general, the aim of the studies was to identify COPD-related proteomic involved COPD development. Another goal was to identify novel biomarkers for the clinical phenotypes of COPD. Finally, we aimed to investigate whether the levels of these markers are associated with clinical variables.

The specific aims of my thesis in each study were:

**Study I:** to identify COPD-specific proteomic changes involved in disease onset and severity in non-smokers, smokers, smokers with mild to moderate (stage I-II) COPD, severe to very severe COPD (stage III-IV) and patients with α-1-antitrypsin deficiency (AATD) and idiopathic pulmonary fibrosis (IPF).

**Study II:** to evaluate sputum BPIFB1 levels in smokers and in COPD in a longitudinal study, and further to investigate the association of sputum BPIFB1 levels with smoking and longitudinal changes in lung function.

**Study III:** to evaluate the association of plasma sRAGE levels with a longitudinal decline of lung function.

**Study IV:** to clarify the similarities and differences between ACOS and COPD or asthma in SP-A, sRAGE, MPO and NGAL levels in sputum and plasma.

**Study V:** to clarify the inflammatory mediators typical for ACOS, namely five sputum biomarkers (IL-13, MPO, NGAL YKL-40 and IL-6) in a Finnish discovery cohort and to validate these results in a Japanese replication cohort.
MATERIALS AND METHODS

1. Projects

This thesis work contained the two projects, including the project of early diagnosis of COPD in cross-sectional and longitudinal studies (Study I, II and III) and the project of ACOS in cross-sectional studies (Study IV and V). The studies were approved by the Ethics Committees Helsinki University Hospital (HUH) and Lapland Central Hospital and conducted in accordance with the ethical standards established in the Helsinki Declaration of 1975 and Dnro 125/E0/2004. All participants provided written informed consent.

2. Cohorts

2.1. Main cohorts for sputum and plasma samples

2.1.1. The 6-year longitudinal cohort of adult smokers for early diagnosis of COPD studies (Study I, II and III)

Induced sputum and plasma samples were collected from the 6-year longitudinal cohort of adult smokers (Rovaniemi cohort, Finland) (Toljamo et al. 2015). These participants were recruited through newspaper announcements (Toljamo et al. 2015). This longitudinal study included three visits (visit 1, visit 2 after two or three years and visit 3 after six years). All visits consisted of a personal interview, and all the assessments were conducted by the same experienced nurse during the three visits (Toljamo et al. 2015). Based on a detailed self-reported questionnaire, all participants felt healthy without any chronic diseases and were symptom-free at visit 1. Sputum samples, plasma samples and spirometric data were collected at visits 2 and 3 (Toljamo et al. 2010, 2015). Therefore, we used visit 2 as baseline and visit 3 as follow-up in our studies (Iwamoto et al. 2014, Mazur et al. 2011). From 2007 to 2008 (visit 2, baseline), spirometric measurements and samples were obtained from 345 participants (Iwamoto et al. 2014). From 2011 to 2012 (visit 3, 4-year follow-up), there were 295 participants with a baseline blood sample as well as baseline spirometry and follow-up spirometry (Iwamoto et al. 2014). Sputum samples were required to fulfil selection criteria (less than 70% squamous epithelial cells), 252 participants fulfilled the criteria and enrolled in our studies. The participants had no history of chronic pulmonary diseases, allergies, medications, the risk factors for other pulmonary diseases or lung infections during the 2 months prior to entering the study. COPD patients fulfilled the GOLD criteria for obstruction (FEV1/FVC < 0.7); severity was distinguished by using the bronchodilator FEV1 value. In addition, the lung tissue samples were retrieved from the patients at HUH (Ishikawa et al. 2010, Ohlmeier et al. 2012). COPD, IPF and AATD tissues were obtained from lung transplantations, and control tissues were obtained from lung surgery from hamartomas or from the surgery for local tumors (Gao et al. 2015, Ohlmeier et al. 2016). The control plasma and sputum samples...
were obtained from the non-smokers if they were healthy, middle-aged (>40 years), not taking any medications, and had normal lung function.

2.1.2. The Finnish discovery cohort and the Japanese replication cohort for ACOS studies (Study IV and V)
Induced sputum and plasma samples were collected from the Finnish discovery cohort, (part of the longitudinally followed cohort of Finnish asthma and COPD patients, or FinnCAD Study) (Kauppi et al. 2011, Laitinen et al. 2009) and the replication cross-sectional cohort (Hiroshima cohort, Japan) (Ishikawa et al. 2015) for ACOS studies. Based on medical history and self-reported questionnaire data, the study subjects were categorised into five groups in the discovery and replication cohort: healthy non-smokers (NS, n=26 and n=22), asymptomatic healthy smokers with normal lung function (HS, n=23 and n=40) and patients with asthma (Asthma, n=32 and n=21), COPD (COPD, n=39 and n=35) and ACOS (ACOS, n=14 and n=17).

In the discovery cohort, asthma diagnosis was based on the two criteria according to the British Guidelines on Asthma Management (British Thoracic Society Scottish Intercollegiate Guidelines Network. 2008). The first one indicates that the subjects should fulfil one or more of the following findings, including a post-bronchodilator increase in FEV1 of ≥12%, a bronchodilator response of ≥15% or diurnal variation of ≥20% in peak expiratory flow (PEF) recording, moderate-to-severe bronchial hyper-reactivity or a decrease in FEV1 of ≥15% in the exercise test. The second one is a clinical history compatible with asthma. COPD diagnosis in the discovery cohort was based on ATS/ERS Task Force recommendations (Laitinen et al. 2009). Specifically, COPD was diagnosed if FEV1/FVC <0.7 or FEV1/FVC <88% predicted was in post-bronchodilatation spirometry related to long-term smoking (Viljanen et al. 1982, Rabe et al. 2007). In the replication cohort, specialists in respiratory medicine diagnosed asthma according to GINA strategy (Reddel et al. 2015). COPD patients had airway obstruction with a pre-bronchodilator FEV1/FVC < 0.7 (Vestbo et al. 2013, Ishikawa et al. 2015, Viljanen et al. 1982) and over 10 pack-years long-term smoking. ACOS was diagnosed if patients fulfilled both diagnostic criteria for asthma and COPD. Controls were recruited at HUH through hospital and local media announcements, and were never-smokers or current smokers with normal lung function and no history of lung disease.

2.2. Lung tissue samples
Lung tissue specimens were collected from patients of the HUH (Ohlmeier et al. 2016), that was used for the proteomic study. Samples from non-smokers and smokers without or with mild to severe COPD (stage I–III) were obtained from tumor-free peripheral lung in operations for local tumors. Lung specimens with very severe COPD (stage IV), IPF, and AATD were obtained from lung transplantations (Ohlmeier et al. 2016). Malignancy had been excluded
The history of smoking is ≥10 years in smokers with and without COPD (Ohlmeier et al. 2016).

3. Sample collection

3.1. Sputum induction and processing
According to the ERS Task Force and described in detail previously (Djukanovic et al. 2002, Kelly et al. 2002), sputum was induced by inhalation of 4.5% hypertonic saline at 5-minute intervals for a maximum of 20 minutes, after the participants cleaned their mouths and noses. Control and healthy smoker subjects did not use any premedication before induction, whereas patients with COPD used 200 μg salbutamol (GSK) as premedication. After induction, sputum without salivary contamination was treated with dithioerythritol (DTE, Sigma, Munich, Germany) then mixed for 30 minutes. The homogenised sputum was further centrifuged at 400g for 10 minutes to separate the supernatant from the cell pellet. Cell viability was verified with trypan blue in a Burker chamber using the resuspended pellet with phosphate-buffered saline (PBS) (Louhelainen et al. 2010). Cytospins (Thermo Scientific, Wilmington, DE, USA) were performed on the resuspended pellet after centrifugation at 400g for 6 minutes. Coded cytospins were stained by May-Grunwald-Giemsa staining (MGG) (Merck, Darmstadt, Germany). The acceptable criteria for each cytospin were based on squamous epithelial cells below 70% per 400 cells. The acceptable cytospins were frozen at -20°C (Louhelainen et al. 2010). Our group technician (Tiina Marjomaa) was responsible for collecting the sputum in Helsinki and teaching sputum processing in Rovaniemi.

3.2. Plasma collection and processing
Peripheral whole venous blood and sputum samples were collected at the same visit day for the individual. Whole blood in EDTA tubes was centrifuged at 400g for 10-15 minutes to prepare plasma. The plasma samples were frozen at -80°C until analysed.

4. Lung function measurement
Standard spirometry and the diffusing capacity of the lung for carbon monoxide (DLCO) were performed according to ATS/ERS recommendations (Enright et al. 2008). Spirometry values were assessed by standard spirometry (Medikro M 903/4, Medikro Oy, Kuopio, Finland). In addition, the Finnish cohort used data from published Finnish research as a reference value (Viljanen et al. 1982). All participants were tested.
5. Laboratory methods

5.1. Two-dimensional difference gel electrophoresis (2D-DIGE) (Study I-II)
Lung tissues obtained from non-smokers, smokers, COPD, AATD or IPF were purified by acetone precipitation. Supernatant protein was quantified according to the manufacturer's protocol using a Bradford-based Roti-Nanoquant (Roth, Karlsruhe, Germany). Protein extracts were labelled using the CyDye DIGE Fluor labelling kit (saturation DIGE, GE Healthcare, Piscataway, NJ, USA). Isoelectric focusing (IEF) after anodic sample cup loading was performed with the Multiphor II system (GE Healthcare). SDS-PAGE was performed overnight with the Ettan DALT II system (GE Healthcare). 2-D gels were scanned with a Typhoon 9400 imager (GE Healthcare) to detect fluorescence signals and were analysed with the 2-DE image analysis software Delta2D 4.3 (Decodon, Greifswald, Germany). According to the known protein sequence, theoretical spot positions were calculated with the Compute pI/Mw tool (http://web.expasy.org/compute_pi/). The spot positions were further analysed by the 2-DE image analysis software Melanie 3.09 (GeneBio, Geneva, Switzerland)

5.2. Mass spectrometry (MS) (Study I and II)
Additional 2-D gels were run with the unlabelled protein (400-600 μg) combined with Cy3-labeled internal standard (5 μg, Ohlmeier et al. 2016) for protein identification. The proteins were analysed and matched with the 2-DE image analysis software Melanie 3.09 (GeneBio, Geneva, Switzerland). The matched spots were further excised and digested with trypsin (recombinant, Roche, Switzerland) and prepared as previously described (Hallstrand et al. 2010). Peptide masses were analysed using MALDI-TOF/TOF (UltrafleXtreme, Bruker, Billerica, USA). Theoretical spot positions were calculated according to their spot-specific peptide mass fingerprint or peptide sequence (or both) with the bioinformatic tool BioTools Version 3.2 (Bruker). The following search parameters were used: MS tolerance, 30 ppm; MSMS tolerance, 0.7 Da; enzyme, Trypsin; engine, Mascot version 2.4.0.; database, NCBInr; modifications, carbamidomethyl and optional oxidation of Met, up to one missed cleavage.

5.3. Western blot analysis (Study I and II)
Lung tissue homogenates (Ishikawa et al. 2010) (n=42, Study I and II) and sputum supernatants (n=42, Study II) were analysed by Western blotting. Ponceau S staining (Sigma Aldrich, USA) was used to standardize the loading of the sputum and lung homogenates (Ohlmeier et al. 2010) to avoid problems with the conventional loading markers in COPD (Casado et al. 2007, Glare et al. 2002, Rabe et al. 2007, Gao et al. 2015). The monoclonal antibodies mouse cathepsin D (CTSD) (BD Transduction Laboratories, USA), mouse collapsin response mediator protein-2 (CRMP2)/dihydropyrimidinase-related protein 2 (DPYSL2) (Immuno-Biological Laboratories, USA), rabbit TGM2 and rabbit tripeptidyl-peptidase 1 (TPP1) (Cell Signaling Technology, Danvers, USA) were used in Study I. BPIFA1 (MAB1897) monoclonal antibodies (RD
Systems, Minneapolis, USA), polyclonal BPIFB1-1 and BPIFB1-2 antibodies (Bingle et al. 2010) were used in Study II. Immunodetection was performed by the chemiluminescent HPR-substrate immune-detection kit (Millipore, USA). Membranes were exposed to X-ray film (Fujifilm, Tokyo, Japan). ImageJ software was used in further analysis and calculation.

5.4. Immunohistochemistry and histochemistry (Study II)
Lung tissue was fixed by formalin and then embedded in paraffin. The prepared tissues were further cut into serial sections (4 μm) and stained (Bingle et al. 2010, Bingle et al. 2007). Detection was performed with anti-rabbit biotin-labelled secondary antibody from the Vectastain Elite ABC kit (Vector Laboratories, USA). Rabbit IgG (DAKO, USA) was used as a negative control. Positive cells were stained red though peroxidase enzymatic development using Vector NovaRed substrate kit. Sections were counterstained with haematoxylin, dehydrated with xylene and mounted in DPX. Alcian blue staining of acidic mucins was performed by a standard histological method (Gao et al. 2016).

5.5. Development of a functional ELISA (Study III)
Since no commercial ELISA kit was available for BPIFB1, we developed a functional ELISA for detecting BPIFB1 concentration in sputum based on a previous study (Bingle et al. 2010). The detection range of the assay was 0.5-20 μg/ml and the lowest standard was 312 ng/ml. BPIFB1-1 and BPIFB1-2 antibodies were produced in rabbits by Genscript (Piscataway, USA) against the peptides C+AQATIRMDTOASGPTRLVLS and C+KDALVLTAPASLWKSPVQS (corresponding to BPIFB1-1 residues 137-156 and BPIFB1-2 residues 465-484) (Bingle et al. 2010, Gao et al. 2015). The antibodies were purified by affinity chromatography and the rabbit anti-BPIFB1-2 antibody was biotinylated with Lightning Link kit (704-0010, Innova Biosciences, UK) (Gao et al. 2015). The anti-BPIFB1-1 antibody was captured on microtiter plates (MaxiSorp NUNC immuno module, Thermo Scientific, USA) overnight at 4°C. Nonspecific binding sites were blocked with 1% BSA-PBS for 2 hours at room temperature. The general diluted concentration was 1:20 for samples and standards (recombinant human BPIFA1, 13275-H08H, Sino Biological). A 1:200 dilution was used if the concentration exceeded the detection range. The prepared samples and standards were incubated as duplicates for 1 hour with gentle shaking. Bound BPIFB1 was measured using biotinylated rabbit anti-BPIFB1-2 and Eu-labelled streptavidin (1244-360, Perkin Elmer, Finland). Delfia enhancement solution (118-100, Perkin Elmer, Finland) was added. Fluorescence of europium was detected with an EnVision Xcite fluorometer (Perkin Elmer, Finland).
5.6. Commercial ELISA (Study I-V) and Magnetic Human High Sensitivity Luminex assay (Luminex assays, Study V)

For the measurement of TGM2, sRAGE, high mobility group box 1 (HMGB1), SP-A, MPO, NGAL, YKL-40 and IL-13 (Study I-V) commercially available ELISA kits (LifeSpan BioSciences, Seattle, USA; RD Systems, Minneapolis, USA; Uscn Life Science, Wuhan, China; Kokusai-F kit Sysmex, Kobe, Japan; Abnova Inc. Walnut, USA; Uscn Life Science, Wuhan, China; Uscn Life Science, Wuhan, China; RD Systems, Minneapolis, USA) were used according to the manufacturers’ instructions. IL-6 levels were measured by Magnetic Human High Sensitivity Luminex assay (RD Systems, Inc. Minneapolis, USA) according to the manufacturers’ instructions. The detection limits were: 0.156 ng/ml for TGM2, 78 pg/mL for sRAGE, 0.238 ng/mL for HMGB1, 1 ng/mL for SP-A, 0.78 ng/mL for MPO, 39 pg/mL (Study IV) and 17 pg/mL for NGAL (Study IV), 3.55 pg/mL for YKL-40, 5.7 pg/mL for IL-13 and 0.14 pg/mL for IL-6, respectively.

6. Statistical analyses

The distribution of basic subject characteristics is shown as mean ± standard deviation (SD, Study I-IV), mean ± standard error (SEM, Study V) or as percentages, if not stated otherwise. Biomarker concentrations are shown as mean ± SEM or median ± interquartile range. Comparisons between groups were evaluated by analysis of variance (ANOVA) (Study I, III and V), χ²-Test (Study I) and t-test (Study III). The data for groups was also analysed by the Kruskal-Wallis test followed by the Mann-Whitney U test (Study I, II and V). Linear regression analysis was used to assess the association between TGM2 levels and COPD severity or lung function in Study I. Intra-assay coefficient of variation (CV) and inter-assay CV% were determined for the functional BPIFB1 ELISA in study II. The intra-assay CV% was calculated in one plate, while the inter-assay CV% was analysed between different plates. Multivariate stepwise regression analysis was performed in Studies III, IV and V to obtain the variables to model ACOS. The biomarkers were further analysed by area under the curve (AUC) statistics to assess their predictive ability to distinguish ACOS from COPD and ACOS and COPD from asthma in Study IV. AUC was also used to distinguish ACOS from asthma, and ACOS from asthma and COPD in Study V. Raw data was used in the comparisons within each cohort and standardised data (Z-scores) in the comparisons between the two cohorts (Study V). Spearman's rank correlation was used to evaluate the associations between biomarker concentrations with other variables (Study I-V). All statistical analyses were performed with the SPSS software program (version 16.0, 20.0 and 21.0, SPSS Inc., Chicago, IL, USA). GraphPad Prism (version 6.0, San Diego, USA) was used for statistical analyses in Studies I and II. A p-value of <0.05 was considered statistically significant.
RESULTS

1. Subject characteristics

The main demographic characteristics of the subjects are shown in Table 3 (Study I-V). The sputum cell counts and spirometry data are shown in Table 4. Smokers with COPD had lower FEV$_1$ than the other two groups, and smokers had lower post-bronchodilator FEV$_1$/FVC than non-smokers (Study I, II and III). The post-bronchodilator FEV$_1$/FVC was over 0.70 in IPF patients (Study II). In the Finnish discovery cohort, patients with COPD and ACOS had more pack-years and significantly lower FEV$_1$ than the asthma groups (Study IV and V). The levels of pack-years and neutrophil and eosinophil percentages in ACOS in the discovery cohort were as similar as those in the replication cohort (Study V).

2. Potential biomarkers in early diagnosis of COPD

2.1. TGM2 in mild-moderate COPD (Study I)

2.1.1. Proteomic changes in lung tissue

Lung tissue samples were analysed using 2D-DIGE in non-smokers, smokers, smokers with mild to moderate COPD (stages I-II) and with very severe COPD (stages III-IV), AATD and IPF (Figure 3). TGM2 and CTSD were observed in COPD-specific expression profiles. TGM2 levels were highest in COPD stages III-IV, whereas TGM2 levels were not elevated in smokers. CTSD levels were highest in COPD stages I-II, and the levels were also increased in smokers compared with controls. Macrophage-capping protein (CAPG), DPYSL2 and TPP1 represented putative changes. CAPG and TPP1 levels were increased in COPD stages I-II and smokers when compared with non-smokers.

2.1.2. Sputum and plasma TGM2 levels in smokers and COPD

TGM2 in sputum and plasma was further investigated in a larger cohort (n=120, Figure 4). TGM2 levels in sputum and plasma were significantly increased in smokers with COPD (stages II-III) than in smokers (p=0.004 and p=0.007, respectively) and in non-smokers (p=0.005 and p=0.033, respectively). There were no significant differences in sputum TGM2 levels between smokers with COPD stage I and smokers (p>0.05), whereas plasma TGM2 levels were increased in COPD stage I compared with smokers (p=0.008).

2.1.3. Correlation between TGM2 and clinical variables

Sputum and plasma TGM2 levels both were correlated with FEV$_1$% predicted and FEV$_1$/FVC in all subjects. Sputum TGM2 levels also were correlated with FVC, FEV$_1$, age and pack-years in univariate analysis (Table 5). In linear regression analyses of smokers, age was an
independent factor for sputum TGM2 (p=0.038), whereas gender and pack-years were not a confounding factor (p>0.05 for both).

Table 3: Subject characteristics in this work

<table>
<thead>
<tr>
<th>Study</th>
<th>Subject n</th>
<th>Gender F/M</th>
<th>Age</th>
<th>Pack-years</th>
<th>Smoking status Ex-/current-</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung tissue study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>9</td>
<td>7/2</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Smokers</td>
<td>9</td>
<td>5/4</td>
<td>66</td>
<td>43</td>
<td>3/6</td>
</tr>
<tr>
<td>COPD I-II</td>
<td>8</td>
<td>1/7</td>
<td>62</td>
<td>43</td>
<td>2/6</td>
</tr>
<tr>
<td>COPD III+</td>
<td>8</td>
<td>4/4</td>
<td>59</td>
<td>31</td>
<td>3/5</td>
</tr>
<tr>
<td>AATD</td>
<td>8</td>
<td>4/4</td>
<td>57</td>
<td>28</td>
<td>3/NA</td>
</tr>
<tr>
<td>IPF</td>
<td>9</td>
<td>2/7</td>
<td>55</td>
<td>12</td>
<td>3/NA</td>
</tr>
<tr>
<td><strong>Sputum and plasma studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>25</td>
<td>18/7</td>
<td>56</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Smokers</td>
<td>46</td>
<td>26/20</td>
<td>49</td>
<td>26</td>
<td>10/36</td>
</tr>
<tr>
<td>COPD I-II</td>
<td>24</td>
<td>5/19</td>
<td>55</td>
<td>34</td>
<td>2/22</td>
</tr>
<tr>
<td>COPD II-III</td>
<td>25</td>
<td>3/22</td>
<td>62</td>
<td>43</td>
<td>8/17</td>
</tr>
<tr>
<td><strong>Study II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sputum study (baseline; follow-up)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>31; 31</td>
<td>21/10; 21/10</td>
<td>56; 58</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Smokers without COPD</td>
<td>169; 145</td>
<td>81/88; 68/77</td>
<td>52; 55</td>
<td>29; 32</td>
<td>39/130; 33/112</td>
</tr>
<tr>
<td>Smokers with COPD</td>
<td>52; 70</td>
<td>7/45; 19/51</td>
<td>59; 62</td>
<td>39; 40</td>
<td>11/41; 20/50</td>
</tr>
<tr>
<td>Lung tissue study (set 1; set 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>14; 4</td>
<td>7/7; 4/0</td>
<td>64; 70</td>
<td>0</td>
<td>3/0; NA</td>
</tr>
<tr>
<td>Smokers without COPD; IPF</td>
<td>14; 4</td>
<td>5/9; 0/4</td>
<td>61; 56</td>
<td>30; NA</td>
<td>1/13; NA</td>
</tr>
<tr>
<td>Smokers with COPD</td>
<td>14; 5</td>
<td>7/7; 2/3</td>
<td>61; 60</td>
<td>40; 27</td>
<td>NA/11; 3/1</td>
</tr>
<tr>
<td><strong>Study III (baseline)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>32</td>
<td>22/10</td>
<td>56</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Smokers without COPD</td>
<td>212</td>
<td>103/109</td>
<td>52</td>
<td>28</td>
<td>44/168</td>
</tr>
<tr>
<td>Smokers with COPD</td>
<td>51</td>
<td>9/42</td>
<td>59</td>
<td>38</td>
<td>12/39</td>
</tr>
<tr>
<td><strong>Study IV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>26</td>
<td>10/16</td>
<td>47</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Health smokers</td>
<td>23</td>
<td>11/12</td>
<td>48</td>
<td>26</td>
<td>0/23</td>
</tr>
<tr>
<td>Asthma</td>
<td>32</td>
<td>16/16</td>
<td>58</td>
<td>16</td>
<td>17/4</td>
</tr>
<tr>
<td>COPD</td>
<td>39</td>
<td>13/26</td>
<td>62</td>
<td>48</td>
<td>13/26</td>
</tr>
<tr>
<td>ACOS</td>
<td>14</td>
<td>5/9</td>
<td>61</td>
<td>37</td>
<td>8/6</td>
</tr>
<tr>
<td><strong>Study V (discovery; replication)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>14; 22</td>
<td>5/9; 13/9</td>
<td>16; 58</td>
<td>0; 0</td>
<td>0/0; NA</td>
</tr>
<tr>
<td>Health smokers</td>
<td>14; 40</td>
<td>5/9; 5/35</td>
<td>48; 63</td>
<td>27; 37</td>
<td>0/14; NA</td>
</tr>
<tr>
<td>Asthma</td>
<td>24; 21</td>
<td>15/9; 17/4</td>
<td>58; 42</td>
<td>13; 4</td>
<td>9/6; NA</td>
</tr>
<tr>
<td>COPD</td>
<td>20; 35</td>
<td>8/12; 1/34</td>
<td>59; 72</td>
<td>37; 48</td>
<td>7/12; NA</td>
</tr>
<tr>
<td>ACOS</td>
<td>18; 17</td>
<td>6/12; 3/14</td>
<td>62; 66</td>
<td>43; 41</td>
<td>11/7; NA</td>
</tr>
</tbody>
</table>
### Table 4: Sputum cell counts and spirometry data in this work

<table>
<thead>
<tr>
<th>Study I</th>
<th>Neutrophils %</th>
<th>Eosinophils %</th>
<th>Macrophages %</th>
<th>FEV$_1$/FVC %</th>
<th>FEV$_1$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sputum and plasma studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>24</td>
<td>NA</td>
<td>27</td>
<td>85</td>
<td>106</td>
</tr>
<tr>
<td>Smokers</td>
<td>35</td>
<td>NA</td>
<td>22</td>
<td>82</td>
<td>97</td>
</tr>
<tr>
<td>COPD I-II</td>
<td>40</td>
<td>NA</td>
<td>24</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>COPD II-III</td>
<td>45</td>
<td>NA</td>
<td>28</td>
<td>60</td>
<td>63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study II</th>
<th>Neutrophils</th>
<th>Eosinophils</th>
<th>Macrophages</th>
<th>FEV$_1$/FVC</th>
<th>FEV$_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sputum study (baseline; follow-up)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>22; NA</td>
<td>NA</td>
<td>23; NA</td>
<td>86; 82</td>
<td>108; 103</td>
</tr>
<tr>
<td>Smokers without COPD</td>
<td>32; NA</td>
<td>NA</td>
<td>24; NA</td>
<td>82; 76</td>
<td>96; 93</td>
</tr>
<tr>
<td>Smokers with COPD</td>
<td>43; NA</td>
<td>NA</td>
<td>27; NA</td>
<td>62; 58</td>
<td>74; 71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study III</th>
<th>Neutrophils %</th>
<th>Eosinophils %</th>
<th>Macrophages %</th>
<th>FEV$_1$/FVC %</th>
<th>FEV$_1$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma study (baseline)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>84</td>
<td>111</td>
</tr>
<tr>
<td>Smokers without COPD</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>82</td>
<td>97</td>
</tr>
<tr>
<td>Smokers with COPD</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>63</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study IV</th>
<th>Neutrophils %</th>
<th>Eosinophils %</th>
<th>Macrophages %</th>
<th>FEV$_1$/FVC %</th>
<th>FEV$_1$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-smokers</td>
<td>35</td>
<td>2</td>
<td>59</td>
<td>84</td>
<td>104</td>
</tr>
<tr>
<td>Smokers</td>
<td>42</td>
<td>1</td>
<td>54</td>
<td>80</td>
<td>96</td>
</tr>
<tr>
<td>Asthma</td>
<td>53</td>
<td>6</td>
<td>41</td>
<td>81</td>
<td>77</td>
</tr>
<tr>
<td>COPD</td>
<td>63</td>
<td>1</td>
<td>33</td>
<td>63</td>
<td>58</td>
</tr>
<tr>
<td>ACOS</td>
<td>68</td>
<td>8</td>
<td>22</td>
<td>70</td>
<td>55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study V (discovery; replication)</th>
<th>Neutrophils %</th>
<th>Eosinophils %</th>
<th>Macrophages %</th>
<th>FEV$_1$/FVC %</th>
<th>FEV$_1$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-smokers</td>
<td>37; 64</td>
<td>2; 3</td>
<td>NA</td>
<td>84; 83</td>
<td>106; 96</td>
</tr>
<tr>
<td>Smokers</td>
<td>43; 70</td>
<td>0; 2</td>
<td>NA</td>
<td>81; 77</td>
<td>99; 94</td>
</tr>
<tr>
<td>Asthma</td>
<td>54; 46</td>
<td>5; 12</td>
<td>NA</td>
<td>77; 81</td>
<td>79; 95</td>
</tr>
<tr>
<td>COPD</td>
<td>63; 72</td>
<td>2; 3</td>
<td>NA</td>
<td>56; 60</td>
<td>58; 70</td>
</tr>
<tr>
<td>ACOS</td>
<td>65; 66</td>
<td>7; 9</td>
<td>NA</td>
<td>56; 55</td>
<td>52; 64</td>
</tr>
</tbody>
</table>

### Table 5: Correlation of plasma and sputum TGM2 with demographics

<table>
<thead>
<tr>
<th></th>
<th>Plasma TGM2</th>
<th>Sputum TGM2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>P-value</td>
</tr>
<tr>
<td>Age</td>
<td>0.027</td>
<td>0.772</td>
</tr>
<tr>
<td>Pack-years V2 (n=95)</td>
<td>0.158</td>
<td>0.475</td>
</tr>
<tr>
<td>Post-bronchodilator</td>
<td>-0.098</td>
<td>0.285</td>
</tr>
<tr>
<td>FVC %</td>
<td>-0.119</td>
<td>0.194</td>
</tr>
<tr>
<td>FEV$_1$ L</td>
<td>-0.233</td>
<td>0.010</td>
</tr>
<tr>
<td>FEV$_1$ % predicted</td>
<td>-0.237</td>
<td>0.009</td>
</tr>
<tr>
<td>FEV$_1$/FVC %</td>
<td>-0.237</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Table 5 is modified from table in Ohlmeier et al. Am J Physiol Lung Cell Mol Physiol. 2016.
Figure 3. COPD-specific changes in the human lung tissue proteome

<table>
<thead>
<tr>
<th>Spot</th>
<th>Protein</th>
<th>Ratio NS vs. COPD</th>
<th>CI-II</th>
<th>CI-IV</th>
<th>AATD</th>
<th>IPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>-91</td>
<td>CAPG</td>
<td>2.04</td>
<td>4.43</td>
<td>2.94</td>
<td>(1.56)</td>
<td>(2.02)</td>
</tr>
<tr>
<td>138</td>
<td>CTSD</td>
<td>1.69</td>
<td>3.36</td>
<td>2.03</td>
<td>(1.47)</td>
<td>(-1.17)</td>
</tr>
<tr>
<td>-139</td>
<td>CTSD</td>
<td>2.37</td>
<td>5.67</td>
<td>(3.04)</td>
<td>(1.22)</td>
<td>(1.14)</td>
</tr>
<tr>
<td>(93⁺)</td>
<td>DPYSL2/CRMP2</td>
<td>2.32</td>
<td>2.01</td>
<td>2.44</td>
<td>(1.49)</td>
<td>(1.50)</td>
</tr>
<tr>
<td>-35</td>
<td>TGM2</td>
<td>(-1.05)</td>
<td>1.45</td>
<td>1.93</td>
<td>(1.16)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>36</td>
<td>TGM2</td>
<td>(1.03)</td>
<td>1.58</td>
<td>2.36</td>
<td>(1.41)</td>
<td>1.51</td>
</tr>
<tr>
<td>37</td>
<td>TGM2</td>
<td>(-1.06)</td>
<td>2.00</td>
<td>2.68</td>
<td>1.98</td>
<td>(2.27)</td>
</tr>
<tr>
<td>-87</td>
<td>TPP1</td>
<td>2.18</td>
<td>5.09</td>
<td>3.65</td>
<td>(2.47)</td>
<td>1.81</td>
</tr>
</tbody>
</table>

Figure 3: Gel parts represent the exact spot positions and expression profiles of the corresponding protein in non-smokers (NS), smokers (S), smokers with COPD stage I–II (CI–II) and stage III–IV (CI–IV), α-1-antitrypsin deficiency (AATD), and idiopathic pulmonary fibrosis (IPF). The expression profile of the most prominent spot (underlined) is shown. Protein fragments are indicated with asterisks and putative changes with brackets. In the expression profiles statistically significant changes are indicated, whereas study groups without statistical significance are marked with asterisks. Spot numbers according to Fig. 3. Spots with similar molecular weight but different charge belonging to the same protein (“spot trains”) are clustered. Ratio corresponds to the change in the mean normalized spot volumes when NS, and S, and COPD I–II or COPD III–IV, and AATD and IPF were compared. The figure is modified from figure and table in Ohlmeier et al. Am J Physiol Lung Cell Mol Physiol. 2016.

2.2. Sputum BPIFB1 levels in smokers and COPD in a longitudinal study (Study II)

2.2.1. The presence of secreted and glycosylated BPIFB1 isoform 1 in smokers and COPD

Identification of BPIFB1 by 2D-DIGE and MS using proteins extracted from purified sputum from non-smokers, smokers and COPD (Ohlmeier et al. 2012). The elevated BPIFB1 levels in sputum had been observed in smokers and COPD (Ohlmeier et al. 2012). In 2D gels, the localisation of the two observed BPIFB1 spots (54 kDa) were similar to theoretical localisation (52 kDa). The two spots also had similar expression profiles, suggesting the presence of full-
length BPIFB1 isoform 1. Glycosylated BPIFB1 in sputum was identified using the shift of the 55-kDa band to a lower molecular weight after treatment. Furthermore, we clarified that the secreted BPIFB1 isoform 1 is the upregulated BPIFB1 in smokers and COPD.

**Figure 4.** Sputum and plasma TGM2 levels

![Figure 4](image)

Figure 4: Sputum and plasma TGM2 levels in sputum (A) and plasma (B) from non-smokers (NS, n=25), smokers (S, n=46), smokers with COPD stage I (CI, n=24) and COPD stage II-III (CII-III, n=25) were measured with ELISA. The solid line shows the median value. * p<0.05 and ** p < 0.01. The figure is modified from figure in Gao et al. Am J Physiol Lung Cell Mol Physiol. 2015.

**2.2.2. BPIFBA1 and BPIFB1 levels in smokers and COPD**

Immunoblotting was used to investigate the expression of BPIFA1 and BPIFB1 in sputum and lung tissue samples (Bingle et al. 2010). There were no significant differences in BPIFBA1 levels among non-smokers, smokers and COPD in sputum and lung tissue. Sputum BPIFB1 levels were increased in COPD compared with non-smokers (p<0.05) and smokers (p=0.006) by image analysis, whereas the BPIFB1 levels in lung tissue were higher in COPD compared with non-smokers (p=0.006) by Western blot analysis. There was no significant difference between COPD and IPF when compared to the BPIFB1 levels in lung tissue (p=0.064). Furthermore, sputum BPIFB1 levels at baseline were investigated in a longitudinal cohort (n=252) using functional ELISA (Figure 5). Sputum BPIFB1 levels were increased in smokers with COPD when compared with smokers without COPD (p=0.007) and non-smokers (p<0.001). Based on smoking status at baseline, the smokers and COPD groups were further divided into ex-smokers vs. current smokers. Interestingly, sputum BPIFB1 levels were significantly elevated in current smokers with COPD when compared with current smokers without COPD (p=0.004).
Figure 5: Sputum BPIFB1 levels measured by a newly developed method using an ELISA-based assay

2.2.3. Correlation between BPIFB1 and clinical variables in cross-sectional and longitudinal studies

The study subjects in the non-smokers, smokers without COPD and smokers with COPD groups were pooled to analyse the correlation between BPIFB1 concentrations and clinical variables in univariate and multivariate analysis. The correlations were mainly analysed in two studies: 1) with cross-section variables: BPIFB1 vs. characteristics at baseline. 2) with longitudinal variables: BPIFB1 vs. characteristics after a 4-year follow-up and BPIFB1 vs. changes in lung function over a 4-year follow-up period. e.g. the value of longitudinal changes (Δ). The cross-section study showed that baseline sputum BPIFB1 levels were correlated with BMI, pack-years, FEV1% predicted and FEV1/FVC% in all subjects and were correlated with pack-years in smokers with COPD at baseline. In the longitudinal study, baseline sputum BPIFB1 levels were correlated with pack-years, FEV1% predicted and FEV1/FVC% in all subjects and were associated with FVC% and FEV1% predicted after the 4-year follow-up in smokers with COPD. Meanwhile, elevated BPIFB1 levels were correlated with ΔFEV1% predicted (p=0.001) and ΔFEV1/FVC% (p=0.012) in smokers with COPD, especially those levels were associated with ΔFEV1% predicted (0.007) and ΔFEV1/FVC% (p=0.001) in
current smokers with COPD. There was no significant correlation between BPIFB1 levels and ΔFEV₁ % predicted in all participants.

2.2.4. BPIFB1 staining is elevated in remodelled airway epithelium in severe COPD
Lung section samples were further analysed using immunohistochemistry for BPIFB1 and MUC5B in non-smokers and in severe COPD. MUC5B is a goblet cell marker. BPIFB1 and MUC5B exhibited strong staining in the airway from severe COPD, no staining was observed in non-smokers. In addition, no staining was observed in peripheral lung tissue or within the sub-epithelial and fibrotic regions.

2.3. sRAGE in airway disease in a longitudinal study (Study III)
2.3.1. Plasma sRAGE and HMGB1 levels at baseline
Plasma sRAGE levels were significantly lower in smokers without COPD and smokers with COPD than in non-smokers (p<0.05 and p<0.05, respectively) (Figure 6). There were no significant differences in plasma sRAGE levels between smokers with COPD and smokers without COPD. In addition, plasma HMGB1 levels were similar among non-smokers, smokers without COPD and smokers with COPD.

Figure 6: Levels of plasma sRAGE

![Figure 6: Plasma levels of soluble receptor for advanced glycation end-products (sRAGE) from non-smokers (NS, n=32), smokers (S, n=212) and smokers with COPD (COPD, n=51). *: p < 0.05. The figure is modified from the figure in Iwamoto et al. BMC Pulm Med. 2014.](image)

2.3.2. Correlation between sRAGE and clinical variables in cross-sectional and longitudinal studies
In cross-sectional study, plasma sRAGE levels were correlated with BMI, pack-years, FVC, FVC% predicted, FEV₁ and FEV₁% predicted in all subjects at baseline. Subgroup analysis showed that plasma sRAGE concentrations were associated with BMI, FVC, FVC% predicted and FEV₁% predicted at baseline in smokers without COPD. In addition, plasma HMGB1
levels did not correlate with baseline cross-sectional parameters in all subjects and subgroups. In the longitudinal study, reduced sRAGE levels were correlated with longitudinal changes ($\Delta$) of FEV$_1$/FVC ($\Delta$FEV$_1$/FVC, $p=0.010$) in all subjects. Subgroup analysis of smokers with COPD showed that plasma sRAGE levels were correlated with $\Delta$FEV$_1$/FVC ($p=0.019$), and that sRAGE levels had a correlation trend with $\Delta$FEV$_1$/% predicted ($p=0.056$). Multivariate analysis showed that baseline sRAGE levels are an independent predictor for $\Delta$FEV$_1$/FVC ($p=0.019$).

3. Differences and characterisation of biomarkers for ACOS (Study IV and V)

3.1. COPD-related biomarkers (Study IV and V)

3.1.1. SP-A and sRAGE as pneumocyte-derived markers in ACOS (Study IV)

Plasma SP-A levels were increased in the COPD and ACOS groups when compared with the asthma ($p<0.05$ and $p<0.01$, respectively) and smokers ($p<0.05$ and $p<0.05$, respectively) groups. SP-A levels in COPD were not significantly different when compared with those in ACOS ($p=0.124$). Based on Spearman's correlation, plasma SP-A levels were correlated with age, pack-years, FEV$_1$%, FEV$_1$/FVC%, DLCO, sputum neutrophils and sputum eosinophils levels. However, FEV$_1$/FVC was an independent predictor for plasma SP-A in multivariate analysis.

Plasma sRAGE levels were decreased in the COPD and ACOS groups when compared with asthma groups ($p=0.025$ and $p=0.013$, respectively). There were no significant differences in sRAGE levels between subgroups (ACOS vs. smokers, $p>0.05$; ACOS vs. COPD, $p>0.05$; and COPD vs. smokers, $p>0.05$, respectively). In univariate analysis, plasma sRAGE levels were correlated with pack-years, FEV$_1$%, DLCO and sputum neutrophil levels.

3.1.2. MPO and NGAL as neutrophil-derived molecules in ACOS (Study IV and V)

MPO levels did not distinguish COPD from ACOS ($p=0.129$, Study IV and $p=0.128$, Study V, respectively). In Study V, sputum MPO levels were elevated in ACOS compared with non-smokers ($p=0.012$), smokers ($p=0.008$) and asthma ($p=0.004$) in the Finnish discovery cohort, whereas there were no differences in sputum MPO levels among asthma, COPD and ACOS in the Japanese replication cohort (Figure 7). In relationship analysis in Study IV, sputum MPO levels were correlated with age, pack-years, FEV$_1$%, DLCO and sputum neutrophil levels, whereas sputum neutrophil levels were an independent predictor for sputum MPO. Interestingly, sputum neutrophil levels showed pack-years and age to be independent predictors for sputum MPO in the Finnish discovery cohort in multivariate analysis in Study V.

Sputum NGAL showed highly repeatable results in the Finnish discovery cohort and the Japanese replication cohort (Figure 7). In the discovery and replication cohorts, sputum NGAL
levels were significantly higher in ACOS than in asthma (p<0.001 and p<0.001, respectively) and COPD (p=0.026 and p=0.002, respectively) (Study V). Sputum NGAL levels in COPD were increased compared with those in asthma (p=0.025) in the discovery cohort (Study V). Importantly, area under the curve (AUC) statistics showed that sputum NGAL distinguished ACOS from COPD and asthma in discovery cohort and replication cohort, such as ACOS vs. COPD (AUC=0.711, p=0.026 and AUC=0.794, p=0.002, respectively) and ACOS vs. asthma (AUC=0.833, p<0.001 and AUC=0.840, p<0.001, respectively). In multivariate analysis, sputum NGAL levels were independently associated with post-FEV1% predicted (p<0.001) in Study IV. In Study V, asthma, COPD, and ACOS groups in both cohorts were pooled to analyse the correlations in the univariate and multivariate analyses. Importantly, NGAL levels were independently associated with pre-FEV1% predicted in the discovery cohort (p=0.001) and replication cohort (p=0.002).

**Figure 7:** Expression levels of sputum biomarkers (z-scores, median value) in patients with asthma, COPD and ACOS in the discovery and replication cohorts

![Figure 7](image_url)

3.2. Asthma-related biomarkers (Study V)

3.2.1. IL-13 as a type 2 T helper cells marker

Sputum IL-13 levels were significantly increased in ACOS when compared to smokers (p<0.001) in the discovery cohort, whereas these results could not be repeated in replication cohort. Meanwhile, increased IL-13 levels in COPD when compared with asthma (p=0.025) in
the replication cohort (Figure 7), whereas these results were not replicated in discovery cohort also. Sputum IL-13 levels did not distinguish ACOS from asthma and COPD in both cohorts. In multivariate correlation analysis, sputum IL-13 were correlated with pack-years in the replication cohort, but not in the discovery cohort.

3.3. Other potential biomarkers (Study V)

3.3.1. IL-6 as a biomarker of impaired lung function in ACOS

In the discovery cohort, promising sputum IL-6 levels were observed in ACOS (Figure 7), including ACOS vs. asthma (p<0.001) and ACOS vs. COPD (p=0.006). Sputum IL-6 levels were also significantly increased in COPD when compared with asthma (p=0.034). Furthermore, sputum IL-6 differentiated ACOS from COPD and asthma (AUC >0.7). However, these results could not be repeated in the replication cohort. In addition, sputum IL-6 were correlated with pre-FEV$_1$% predicted and pack-years in the discovery cohort. There was no significant correlation between IL-6 and clinical variables in the replication cohort.

3.3.2. YKL-40 as a biomarker of airway inflammation in ACOS

In the discovery cohort, sputum YKL-40 levels were significantly increased in ACOS compared with non-smokers (p<0.001), smokers (p<0.001), asthma (p=0.001) and COPD (p=0.002). ROC analysis showed sputum YKL-40 could distinguish ACOS from COPD and asthma (AUC >0.7). In univariate and multivariate analyses, YKL-40 strongly was correlated with age and pre-FEV$_1$%. However, there was no significant difference in sputum YKL-40 levels among groups in the replication cohort (Figure 7).
DISCUSSION

1. Summary of background and results

The diagnosis of COPD is often too late. With COPD development, late diagnosis could lead to poorer outcomes and increased burdens on society. Currently, diagnosis and staging of COPD depends on the airflow limitation level (e.g. FEV₁) in clinical practice, whereas FEV₁ based COPD diagnosis is a crude measure in individuals (Vestbo et al. 2008). FEV₁ maybe insensitive to changes in small airway calibre, and a similar FEV₁ may also be of a different functional status (Vestbo et al. 2008). Although several risk factors for COPD have been identified (GOLD 2017), tobacco smoking remains the major environmental risk factor for COPD (Brusselle et al. 2011). Smoking time may be associated with the mortality rate of COPD (Bai et al. 2017). Smoking cessation is the critical method for prevention and maintenance in COPD, which has the greatest capacity to influence the natural history of COPD, including lung function trajectory and progression of COPD (GOLD 2017, Bai et al. 2017). Thus smoking, along with numerous features related to smoking susceptibility and smoking status, should be considered when potential COPD biomarkers are evaluated, which may assist in identifying the COPD phenotype. Pulmonary rehabilitation is an effective therapy in the integrated care of the patient with chronic respiratory disease (Spruit et al. 2013, Evans et al. 2017). Exercise-based pulmonary rehabilitation has proven to reduce exacerbations, subjective dyspnoea, and frequency for re-hospitalization (Katajisto et al. 2017, Casaburi et al. 2009, Garcia-Rio et al. 2009). To date, the existing medications are effective therapy for COPD, but it could not modify the long-term reduced lung function (GOLD 2017, Burge et al. 2000, Anthonisen et al. 1994, Pauwels et al. 1999, Vestbo et al.1999). ACOS has been confirmed as one of the clinical phenotypes of COPD and is associated with poorer outcomes than those in patients with asthma and COPD alone, especially in smokers and the elderly.

Therefore, the first project of this thesis intended to investigate novel stable biomarkers for early diagnosis of COPD using the longitudinal cohort in sputum, plasma or both. TGM2, sRAGE and BPIFB1 were selected primarily on the basis of our proteomic results using lung tissue. Our second project was focused on biomarkers for ACOS using the two cohorts, and the aim was to distinguish ACOS from asthma and COPD and to validate our results. We investigated biomarkers relevant for airway inflammation from sputum or plasma (or both) in asthma (IL-13), COPD (MPO, NGAL, sRAGE and SP-A), or in both asthma and COPD (YKL-40 and IL-6). The first project confirmed that three potential biomarkers (TGM2, BPIFB1 and sRAGE) for COPD development reflect different aspects in disease pathology. These biomarkers had also different correlation with smoking and decline of lung function. In the second project, we determined that one of our selected sputum biomarkers might be a characteristic feature of ACOS, reflecting both airway inflammation and tissue remodelling.
2. Proteomics-based specific biomarkers in COPD development (Study I, II and III)

Study I

Proteomics is a non-biased large-scale screening technique that is an important method for investigating COPD pathology (Ohlmeier et al. 2016). We investigated the pathological mechanisms of COPD using proteomics (Ohmeier et al. 2008, 2010, 2012, 2016). In study I, we detected specific and unspecific changes using lung tissue samples. Although we identified 82 proteins, only 18 proteins showed disease specificity (Ohlmeier et al. 2016). We verified the expression profiles specific to COPD (CTSD, TPP1, CAPG, DPYSL2 and TGM2), AATD (e.g. SERPINA1, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), hemoglobin subunit beta (HBB), and heat shock cognate 71 kDa protein (HSPA8)) and IPF (e.g. annexin A3 (ANXA3), fatty acid-binding protein (FABP5), HBB, selenium-binding protein 1 (SELENBP1) and keratin type II cytoskeletal 7 (KRT7)). The other proteins could reflect the unspecific changes, overlapping changes or both (e.g. COPD+AATD, COPD+IPF and AATD+IPF) in lung tissue (Ohlmeier et al. 2016). CTSD is present in azurophil granules (Almansa et al. 2012) and may be expressed by pulmonary macrophages in response to cigarette smoke (Bracke et al. 2012). The expression of CTSD in BALF is elevated in patients with COPD (Tu et al. 2014), and gene expression of CTSD and CAPG in blood are both increased in COPD patients with critical respiratory illness (Almansa et al. 2012). Consistent with study I, CAPG and CTSD are linked to COPD. Our results showed that CAPG, CTSD, DPYSL2 and TPP1 are associated with smoking, whereas previous studies revealed that levels of CAPG, CTSD and DPYSL2 are increased in lung cancer (Pastor et al. 2013, Pernemalm et al. 2013). These results can be partially explained by the observations that smoking is associated with reactive oxygen species (ROS) (Kelsen et al. 2008) and is also a known risk factor for COPD and lung cancer.

The most striking finding of study I is related to TGM2, which may be associated to COPD onset and severity. TGM2 is a multifunctional enzyme (Ohlmeier et al 2016) and it expressed in asthma (Hallstrand et al. 2010), cystic fibrosis (Maiuri et al. 2008), bronchopulmonary dysplasia (Witsch et al. 2014) and non-small cell lung cancer (Choi et al. 2011). Although TGM2 levels have been analysed in multiple diseases, TGM2 levels have not been investigated in COPD (Ohlmeier et al. 2016). In study I, increased levels of TGM2 were detected in the moderate-to-severe COPD group compared with those in smokers and in non-smokers with different approaches and sample sets (Ohlmeier et al. 2016). In addition, increased levels of TGM2 in mild COPD were shown in lung tissue and plasma samples, whereas differences in sputum TGM2 levels were not statistically significant between mild COPD and non-smokers, although there was an increasing trend in COPD (p=0.075) (Ohlmeier et al. 2016). These results suggest that TGM2 may be a promising biomarker for COPD on the basis of COPD specificity and concordance with previous results (Ohlmeier et al. 2016). In correlation analysis,
sputum TGM2 levels were correlated with age, pack years and lung function, especially lung function showed the strongest correlation. As increased age is associated with declined lung function in general, we further adjusted the smoking factor in our correlation model. These results showed that sputum TGM2 significantly correlates with lung function. Thus based on the relationship between TGM2 and lung function, and taking into account its cellular functions in transglutaminase crosslinking, G protein signalling and kinase activities, we suggest that TGM2 may be related to the disease process itself and is involved in lung tissue remodelling (Ohlmeier et al. 2016). Accordingly, TGM2 inhibitors or anti-TGM2 may have therapeutic potential in COPD (Ohlmeier et al. 2016).

Study II
Nashida et co-workers reported that BPIFB1 mRNA was expressed in the palate, lung, and nasal epithelium clone protein family (Nashida et al. 2012). Previous proteomic studies have shown elevated BPIFB1 levels in cystic fibrosis (Bling et al. 2012), severe uncontrolled asthma (Lee et al. 2013) and asthma patients after segmental allergen challenge (Wu et al. 2005). While our previous proteomic study investigated BPIFB1 in a smoker cohort (Ohlmeier et al. 2012), BPIFB1 levels in COPD have not been previously analysed. In study II, we first detected sputum BPIFB1 levels using quantitative analysis in non-smokers, smokers without COPD and smokers with COPD (Gao et al. 2015). Furthermore, we examined the relationship between BPIFB1 and airflow limitation using a 4-year follow-up cohort (Gao et al. 2015). When compared with levels in non-smokers and smokers without COPD, we observed increased sputum BPIFB1 levels in smokers with COPD (Gao et al. 2015). BPIFB1 was observed in airway goblet cells (Bling et al. 2010), and ciliated epithelial cells are replaced by goblet cells when airway irritation increased (Gao et al. 2015). Our immunohistochemistry results confirmed that BPIFB1 was located in airway goblet cells, and we further revealed that BPIFB1 staining is restricted to the airway epithelium (Gao et al. 2015). Together with the quantitative results and immunohistochemistry analysis for BPIFB1, it is possible that sputum BPIFB1 levels may be upregulated in COPD on the basis of enhanced physiological response and BPIFB1 protein expression may be maintained in severe disease (Gao et al. 2015). In subgroup relationship analysis, increased sputum BPIFB1 levels were correlated with longitudinal declined lung function in COPD. In the smoking-related relationship analysis, sputum BPIFB1 levels were correlated with longitudinal declined lung function in current smokers with COPD during the 4-year follow-up period. This correlation was not observed in ex-smokers with COPD. Furthermore, sputum BPIFB1 levels showed a strong correlation with pack-years only in smokers with COPD. Thus BPIFB1 may be associated with smoking-related pathological changes in the airway wall (Gao et al. 2015). Smoking, chronic infection, and enhanced immune responses may result in upregulating the BPIFB1 expression in COPD (Gao et al. 2015).
We also detected BPIFA1 in sputum and lung tissue in this study. BPIFA1 levels were not significantly different between COPD and other groups, but there was an increasing trend in COPD compared with those in non-smokers (Gao et al. 2015). Several studies analysed BPIFA1 levels (Steiling et al. 2009, Casado et al. 2007) with contradictory results. Steiling and co-workers detected BPIFA1 in airway epithelial cells of current smokers (Steiling et al. 2009), whereas Casado and co-workers identified BPIFA1 in sputum of patients with COPD, but not in smokers (Casado et al. 2007). Together with our BPIFA1 negative results, the more systematic studies in larger cohorts are required to clarify the importance of BPIFA1 and BPIFB1 in COPD.

Study III

RAGE is a multiligand signal transduction receptor (Miniati et al. 2011). In soluble forms, sRAGE provides protection against inflammation (Miniati et al. 2011) and sRAGE deficiency is associated with enhanced inflammation in multiple chronic conditions (Sukkar et al. 2012). Several studies have shown that sRAGE expression is reduced in COPD (Smith et al. 2011, Sukkar et al. 2012, Miniati et al. 2011). Our previous studies also showed reduced expression of sRAGE in patients with COPD, and that full length-RAGE and the C-terminal processed variant RAGE were associated with COPD and IPF (Ohlmeier et al. 2010, Iwamoto et al. 2014). As the longitudinal study for RAGE was limited, thus our study III aimed to evaluate plasma sRAGE and HMGB1 levels in a 4-year longitudinal cohort with 3 groups (non-smokers, smokers without COPD and smokers with COPD). To the best of our knowledge, our study was the first to compare circulatory HMGB1 levels in smokers and in early COPD without comorbidities (Iwamoto et al. 2014). In study III, reduced plasma sRAGE levels were observed in smokers with COPD when compared with the corresponding levels in non-smokers (Iwamoto et al. 2014). Importantly, this was the first study to reveal the relationship between reduced plasma sRAGE levels and progression of airflow limitation, even after controlling for demographics and baseline lung function (Iwamoto et al. 2014). sRAGE levels are associated with emphysema (Cheng et al. 2013, Cockayne et al. 2012, Coxson et al. 2013, Miniati et al. 2011), chronic cor pulmonale, neutrophilic airway inflammation in COPD and correlates with impaired diffusion capacity (Miniati et al. 2011, Sukkar et al. 2012). The ECLIPSE studies confirmed the association between low sRAGE levels and increased emphysema. Circulatory sRAGE levels are significantly reduced in COPD patients with advanced GOLD stage (Cheng et al. 2013). Another longitudinal study demonstrated that circulatory sRAGE levels correlate with decline of lung density in moderate-to-severe COPD using CT scans (Coxson et al. 2013). Study III revealed lower plasma sRAGE levels in COPD using CT scans (Coxson et al. 2013). Study III revealed lower plasma sRAGE levels in COPD using CT scans (Coxson et al. 2013). Study III revealed lower plasma sRAGE levels in COPD using CT scans (Coxson et al. 2013). Study III revealed lower plasma sRAGE levels in COPD using CT scans (Coxson et al. 2013). Study III revealed lower plasma sRAGE levels in COPD using CT scans (Coxson et al. 2013). As one of the common
RAGE ligands, HGMB1 binds Toll-like receptors and activates a pro-inflammatory cascade and is associated with inflammation and human airway epithelial cell repair (Ferhani et al. 2010, Li et al. 2015, Ojo et al. 2015, Ulloa et al. 2006). Diener and Zhang reported that plasma HMGB1 levels are decreased in COPD (Diener et al. 2013, Zhang et al. 2014), whereas another study using sputum and plasma demonstrated increased HMGB1 levels in asthma and COPD (Hou et al. 2011). In light of these inconsistent results, future large-scale studies are needed to further identify the levels and specific roles of HMGB1 in COPD.

3. Promising biomarkers in ACOS studies with two well-defined cohorts (Studies IV and V)

To characterise ACOS biomarkers, we investigated seven potential candidates for ACOS in the two cohorts in non-smokers, smokers, the patients with asthma, COPD and ACOS, including pneumocyte-derived markers (SP-A and sRAGE), neutrophil-derived molecules (MPO and NGAL), and Th2 (IL-13), impaired lung function (IL-6) and airway inflammation (YKL-40) biomarkers.

**Study IV**

To evaluate the utility of COPD-related biomarkers in ACOS, we evaluated the levels of plasma SP-A, plasma sRAGE, sputum MPO and sputum NGAL in the Finnish discovery cohort (Iwamoto et al. 2014). The levels of SP-A, sRAGE and MPO in ACOS and COPD showed a similar pattern when compared with the corresponding levels in asthma. However, these biomarkers did not differentiate ACOS from COPD. Overall, AUC statistics revealed that all four biomarkers could differentiate ACOS from asthma. However, only sputum NGAL levels differentiated ACOS from COPD and asthma. These results suggested that increased sputum NGAL levels may be a characteristic feature of ACOS. SP-A is synthesised in alveolar type II cells, which are associated with leakage of the lung epithelial barrier (Robin et al. 2002, Iwamoto et al. 2014). Our previous study revealed increased plasma SP-A levels in smokers and in COPD (Mazur et al. 2011, Ilumet et al. 2011). In study IV, we confirmed increased plasma SP-A levels in COPD and the relationship between SP-A and smoking. Further, our new findings showed that plasma SP-A levels increased in ACOS compared with those in asthma. Moreover, plasma SP-A levels were also associated with lung function and peripheral tissue damage. Interestingly, the degree of airway obstruction could be an independent factor for SP-A expression (Iwamoto et al. 2014). These results suggested that smoking-induced lung injury was strongly associated with both COPD and ACOS.

RAGE can be expressed on AT I pneumocytes and associated with cell adherence and spreading (Dahlin et al. 2014, Demling et al. 2006). We investigated plasma sRAGE levels in smokers and COPD (Iwamoto et al. 2014). Our current results confirmed our previous results.
and also revealed that plasma sRAGE levels are decreased in ACOS. Previous evidence suggested that reduced plasma sRAGE may be a biomarker for emphysema (Cheng et al. 2013, Yonchuk et al. 2015). Consistent with those results, our study demonstrated that plasma sRAGE levels are independently associated with DLCO (Iwamoto et al. 2014). In addition, plasma sRAGE levels were also associated with smoking, airway limitation and neutrophils. Thus our results suggest that peripheral lung destruction may be associated with reduced plasma sRAGE in ACOS and COPD (Iwamoto et al. 2014). MPO is a neutrophil marker and is released from the primary granules of neutrophils (Iwamoto et al. 2014, Keatings et al. 1997). Our results showed that the increased sputum MPO levels were in COPD when compared with asthma. The ratio of sputum MPO levels and neutrophils are associated with exacerbations of COPD (Liesker et al. 2011). Increased sputum MPO levels have been observed in acute COPD exacerbations, which may be evidence of neutrophil activation (Aaron et al. 2001). Sputum neutrophil levels were an independent factor for sputum MPO in our study, which was in a good agreement with this.

The most important findings in this study were associated with sputum NGAL. NGAL is a neutrophil-related biomarker and is produced by epithelial cells in the respiratory tract (Cowland et al. 2003, Bolignano et al. 2008). NGAL inhibits bacterial growth and enhances matrix degradation in COPD (Yan et al. 2001). Previously, elevated NGAL levels were also shown in bronchial lavage and in plasma from patients with COPD (Eagan et al. 2010, Ekberg-Jansson et al. 2001). To the best of our knowledge, our study was the first to show increased sputum NGAL levels in COPD and ACOS. This increased sputum NGAL may be produced by activated neutrophils and respiratory epithelial cells after inflammatory stimuli (Cowland et al. 2003, Karlsen et al. 2010). Moreover, sputum NGAL was significantly elevated in ACOS compared with COPD and was independently associated with airflow obstruction. Thus ACOS can be identified by increased levels of sputum NGAL, and NGAL might reflect the enhanced neutrophilic airway inflammation or airway epithelial injury (or both) in ACOS (Iwamoto et al. 2014).

**Study V**

In study V, we investigated five biomarkers to identify inflammatory profiles typical for ACOS in the discovery cohort and the replication cohort (Gao et al. 2016). Two biomarkers (NGAL and MPO) were selected on the basis of our previous studies (Iwamoto et al. 2014 (1) (2)) and the other biomarkers (YKL-40, IL-6 and IL-13) were based on the existing literature. Expression of sputum MPO and YKL-40 were not statistically significant different between the study groups in the replication cohort, whereas these two biomarkers showed some promising results in the discovery cohort. There were no differences in sputum IL-13 levels among ACOS, asthma and COPD in both cohorts. While YKL-40, IL-6 and MPO levels in ACOS were promising, we were unable to replicate these results in the replication cohort.
Serum YKL-40 levels were increased in two clinical phenotypes of asthma (asthma with irreversible airway obstruction and asthma with severe exacerbations) (Gomez et al. 2017) and in acute exacerbations COPD (Lai et al. 2016). Acute exacerbations in COPD are associated with IL-6 (Dickens et al. 2011), and Fu and co-workers observed that there were no significant differences in IL-6 levels between ACOS and COPD in serum (Fu et al. 2014). Considering IL-6 is a biomarker of systematic inflammation, thus the specificity of IL-6 in airway disease may be limited. As a bactericidal compound, MPO diminishes dendritic cell activation and is associated with local activation of neutrophils. Accordingly, elevated MPO activity is associated with a number of inflammatory diseases (Strzepa et al. 2017, Iwamoto et al. 2014, Keatings et al. 1997, Aaron et al. 2001). Our patients with asthma, COPD and ACOS were all in stable phase, and the difference of aetiological variability and airway limitation were shown in the two cohorts, which could explain the differing results in our study.

The key findings of the study were the following: sputum NGAL showed highly repeatable results in the two cohorts, distinguished ACOS from COPD and asthma and independently correlated with airflow limitation. NGAL is associated with inflammatory stimuli in respiratory epithelial cells (Iwamoto et al. 2014, Cowland et al. 2003) and with Toll-like receptor activation during bacterial infections (Flo et al. 2004, Chan et al. 2009). NGAL may also be induced by reactive oxygen species in the lung (Roudkenar et al. 2007, Sunil et al. 1009). Microbial colonisation and augmented oxidative stress may be involved in ACOS, resulting in airway inflammation, airway reactivity, injury and remodelling (Iwamoto et al. 2016). Consistent with our study, our validated results confirmed that sputum NGAL showed potential in differentiating between asthma, COPD and ACOS and may be associated with acute viral infections and exacerbations in ACOS (Iwamoto et al. 2016). Together, our identified sputum biomarkers may be candidate molecules for identifying airway inflammation and airway remodelling (Iwamoto et al. 2016).

4. The difference between ACOS and ACO and the diagnostic criteria for ACOS in our studies (Study IV and V)

ACOS is an important clinical phenotype in chronic airway disease. However, this phenotype is currently not clearly defined. The clinical, pathophysiological and genetic features of ACOS are poorly understood (Carolan et al. 2013, Hardin et al. 2011, Wardlaw et al. 2005). ACOS has been re-defined as asthma-COPD overlap (ACO) based on the newest GINA strategy (GINA 2017), which highlighted that ACO is not a single discrete disease (GINA 2017, Bonten et al. 2016). The collection of our ACOS project began in 2005, when there was no confirmed name and diagnosis criteria for ACOS available on the basis of GINA strategy. Therefore, in study IV, we selected ACOS patients based on asthma diagnosis. This decision was based on the British Guidelines on Asthma Management (British Thoracic Society Scottish
Intercollegiate Guidelines Network. 2008) and the ATS/ERS Task Force recommendations (Laitinen et al. 2009, Miller et al. 2005). COPD defined by irreversible chronic airway obstruction using FEV1/ FVC <0.70 or FEV1/FVC% <88% predicted in post-bronchodilator spirometry (Iwamoto et al. 2014). However, when we began study V, the GINA strategy formally named "ACOS" and described the diagnosis criteria. Therefore we removed the term of FEV1/FVC% <88% predicted, selected the age of onset (>40 years) and included all patients with a history of smoking. Taken together with the consumption of sputum samples, there were some group size differences between study IV and study V in the Finnish cohort.

5. The roles of biomarkers in precision medicine of COPD

COPD and asthma are common CAD, and both are well-known complex heterogeneous diseases. The overlapping syndromes have been recognised as one of the phenotypes in COPD or asthma. Current guidelines are useful for COPD management, but a ‘one size fits all’ approach is insufficient (Bel et al. 2017). Assessment of ‘treatable traits’ could be a vital step in precision medicine of CAD. This approach could deconstruct the diseases into individual traits that can be measured and modified (Agusti et al. 2016, Bel et al. 2017) and that could be associated with clinical phenotypes of diseases. Biomarker development may be an effective approach in precision medicine, which could improve the treatable traits. Investigating the biological mechanisms via different molecular pathways or endotypes for COPD may facilitate selection of promising biomarkers. Endotypes, biomarkers and clinical characters may contribute to distinguishing the clinical phenotype. On the basis of our study, BPIFB1 may be a promising biomarker for COPD. We developed a BPIFB1 ELISA kit, which may assist other researchers in investigating this biomarker. So far, several targeted drugs have been developed and are used for improving the treatable trait related to eosinophils. This includes new therapeutics based on Th2 biomarkers, such as IL-5 receptor alpha (mepolizumab, reslizumab, benralizumab) and IL-13 (lebrikizumab, tralokinumab (Mukherjee et al. 2014, Bel et al. 2017). Therapy targeting the IL-5 receptor alpha may be effective in long-term management of hypereosinophilic syndrome (Busse et al. 2010, Mukherjee et al. 2014). This approach is also effective in COPD related to eosinophilic inflammation (Bel et al. 2017). Our study showed that NGAL might reflect ACOS characteristics, and could be a potential biomarker for improving treatable traits related to neutrophils in CAD.

6. Strengths and limitations of the study

We investigated human samples from multiple well-defined cohorts, including the 6-year longitudinal cohort of adult smokers for early diagnosis of COPD studies (Study I, II and III) and the Finnish discovery cohort and the Japanese replication cohort for ACOS studies (Study IV and V). We used the strict exclusion criteria for both cohorts. In the 6-year longitudinal
cohort of adult smokers, none of the participants had any chronic pulmonary or other diseases requiring regular medication. In addition, they had no allergies, no risk factors for other pulmonary diseases (e.g. known exposures, bronchiectasis, malignancies, and also previous lung tuberculosis) and no any lung infections during the last two months before entering the study (Toljamo et al. 2015). In the Finnish discovery cohort, the patients with COPD and ACOS had no history of α1-antitrypsin deficiency in their family, and no patients were either receiving oral steroid treatment or experienced an exacerbation during the month before inclusion (Iwamoto et al. 2014). In the early COPD project, we detected proteomic signatures related to smoking, COPD, AATD, IPF, overlapping and unspecific changes, and identified specific proteins for lung diseases by the longitudinal studies. The correlation between specific COPD biomarkers, disease severity, smoking and other clinical characters were further evaluated, particularly in mild and moderate COPD. We established a functional ELISA for quantitative detection of BPIFB1. Our study firstly evaluated the relationship between plasma sRAGE and longitudinal airway obstruction. In the ACOS study, our project was the first study to investigate plasma or sputum biomarkers in patients with ACOS using unbiased comparisons. We used the discovery cohort and replication cohort in ACOS and our results were validated in a different population. Our controls had normal lung function without other exposures.

Aside from the strengths, there were some limitations in our study. In the early phase of the project, our lung tissue samples were evaluated and the severity of COPD assessed according to the previous GOLD 2010 and classified into four stages (I-IV), but this seems not to be sufficient to reflect the heterogeneity of COPD (Toljamo et al. 2015, Gao et al. 2016). On the other hand, the follow-up period of 4 years was relatively short, but this was due to the study design of a longitudinal analysis among apparently healthy middle-aged to elderly individuals (Gao et al. 2015). In the ACOS project, the diagnosis criteria for the discovery and replication cohorts were similar but not identical. Accordingly, diagnosis differences may influence the comparison of results between the two cohorts. Meanwhile, the sample size in both cohorts were limited. Despite several sample types (lung tissue, sputum and plasma) were used, all three biological samples do not come from the same subjects. On the other hand, the limitation would be a lack of BALF samples that could reflect the inflammation on a level of small airways. Our selected biomarkers were measured by quantitative ELISA after the proteomic studies, that may contribute to understanding COPD pathology, but further COPD research is still needed via other methods (e.g. genomics, transcriptomics and metabolics studies), as proteomics may not adequately verify the specificity of the identified changes. Furthermore, multivariate stepwise regression analysis was performed for obtaining an independent predictor for biomarker or lung function in our studies, but the other statistical methods may be needed in future, such as cluster analyses and principal component analysis.
CONCLUSIONS

We identified the 18 proteins in a well-defined cohort in the study I. CTSD, DPYSL2 and TPP1 seemed to be specific biomarkers for COPD, and these biomarkers are related to smoking. TGM2 was another COPD-specific biomarker, and TGM2 was not associated with smoking and may thus be associated with the disease process itself. Our results suggest that TGM2 may be associated with COPD severity, and may also be used in the future in COPD diagnosis, monitoring and targeted therapies.

According to the results received with the ELISA method established in the study, elevated sputum BPIFB1 levels were shown in smokers with COPD compared to smokers without COPD. Importantly, sputum BPIFB1 levels were significantly associated with longitudinal changes of airflow limitation during a 4-year follow-up period, especially in current smokers with COPD. Our immunohistochemistry results showed that BPIFB1 localised in regions of goblet cell metaplasia and its staining was restricted to the airway epithelium, confirming the known function of BPIFB1 in pulmonary immunity. These results suggest that BPIFB1 may be involved in the pathogenesis of smoking-related lung diseases.

In the longitudinal study set-up, plasma sRAGE levels were significantly lower in smokers with and without COPD compared with controls. Our study is the first to longitudinally evaluate the relationship between plasma sRAGE and airway obstruction. Our results showed that reduced plasma sRAGE levels were significantly correlated with the progression of airflow limitation in all subjects and in smokers with COPD. Therefore, our findings suggest that RAGE may have a protective role against the progression of smoking-induced lung damage.

In our first ACOS study on the basis of the Finnish single cohort, our results showed that the levels of plasma SP-A, plasma sRAGE and sputum MPO that distinguished ACOS from asthma, whereas those levels did not distinguish ACOS from COPD. However, only sputum NGAL levels were significantly elevated in ACOS compared with those in COPD and asthma and was associated with airflow obstruction independent of sputum neutrophil level. These results suggest that NGAL may be a characteristic feature of ACOS, reflecting increased neutrophilic airway inflammation, airway epithelial injury or both.

In the second ACOS study using the two well-defined independent cohorts, levels of sputum IL-6, YKL-40 and IL-13 in the replication cohort were not significantly different between groups, which was inconsistent with the results from the discovery cohort. The difference in sputum IL-13 levels in the replication cohort was statistically significant between asthma and COPD, whereas no difference was observed in the discovery cohort. Consistent with the discovery cohort and the previous study, sputum NGAL had diagnostic value in differentiating...
ACOS from asthma and COPD and independently correlated with declined lung function. Together with our previous study IV, our results suggest a relationship between sputum biomarkers and ACOS, and NGAL especially may be a potential biomarker for ACOS.
FUTURE RESEARCH

On the basis of our results, no biomarker accurately clarified COPD phenotypes, and biomarkers solely specific for one disease might be very rare. For our early COPD project, it would be very useful to verify the suggested smoking-independent mechanisms of TGM2 action in the future (Ohlmeier et al. 2016). Although our results suggest that BPIFB1 and sRAGE may play important roles in COPD, future large-scale studies are needed to clarify these mechanisms and to specify the roles of modulating lung inflammation in COPD (Gao et al. 2015, Iwamoto et al. 2014). More studies are needed to investigate the RAGE/HMGB1 pathway in COPD. For our ACOS project, future studies are needed to determine the diagnostic value of these biomarkers in differentiating ACOS from asthma and COPD. Therefore, future research of COPD may include genomic, transcriptomic, proteomic and metabolic studies and use of other statistical methods (e.g. cluster analyses and principal component analysis). By investigating the multi-level biological network in COPD, we could further clarify the endotypes based on the different biological mechanisms and identify relevant biomarkers.
ACKNOWLEDGMENTS

This thesis work was performed at the Heart and Lung Center, Division of Pulmonary Medicine, University of Helsinki and Helsinki University Hospital. I would like to express my deep gratitude to everyone who helped me in this study.

This work was supported financially by the SalWe Research Program for IMO (Tekes-the Finnish Funding Agency for Technology and Innovation grant 648/10), a governmental subsidy from the Ministry of Social Affairs and Health, a governmental subsidy for health science research (EVO) in Rovaniemi, the EVO funding of the Helsinki University Central Hospital, Research Funds of University of Helsinki, Sigrid Jusélius Foundation, Finnish Anti-Tuberculosis Association Foundation, Jalmari and Rauha Ahokas Foundation, Finnish Cultural Foundation, the Research Foundation of the Pulmonary Diseases, Ida Montin Foundation, China Scholarship Council, CIMO, and Väinö and Laina Kivi Foundation. B. Araujo was supported by an ERS Long Term Research Fellowship.

I owe my warmest sincere gratitude to my supervisor, Docent Ville Pulkkinen PhD, for guiding me with his excellent scientific knowledge, for teaching me with his rigorous attitude, for supporting me with his powerful energy, and for revising the final version of this thesis. I am very grateful to him giving me great help in difficult times.

I sincerely thank my other supervisor Docent Witold Mazur PhD, MD, for always giving me his endless support, for teaching me his knowledge in the pulmonary field, for sharing his professional skills with me, and for revising this thesis. I also wish to thank him for helping me to contact with Docent Ville Pulkkinen. I shall never forget his warmly help for me and my family.

I give my special thanks to my late Professor Vuokko Kinnula PhD, MD, for opening a scientific door to me in the pulmonology field and providing me a chance to study in Finland. She and her family gave me warmly help for my study and life. I shall always remember her kindness and cleverness.

I am very grateful to Professor Terttu Harju PhD, MD, for accepting the invitation to be my opponent.

I would like to thank my thesis revisers Docent Anne Pietinalho MD, and Docent Siri Lehtonen PhD, for their constructive comments on my thesis.
I thank Professor Marjukka Myllärniemi PhD, MD, for her warmly help and friendly support during my study.

I am most grateful to my co-author and collaborator Docent Hiroshi Iwamoto PhD, MD, for giving me much valuable advice throughout my study, for providing insightful comments in my manuscripts, for helping me solve the research problems, and for his friendship and invaluable help.

I also wish to thank my other collaborators Docent Pentti Nieminen PhD, for sharing his statistics skills with me and for providing excellent statistics support in our project; Docent Tuula Toljamo PhD, MD, for giving us support in sample collection and for giving me a lot of help during the time I worked in Rovaniemi; Docent Steffen Ohlmeier PhD, for sharing his experience and providing inspiration; Professor Tarja Laitinen PhD, MD, for giving us great support in the thesis projects; Professor Colin D Bingle PhD, for his excellent help and sharing his expertise with us; Jukka Koskela, PhD, MD for his advice and friendly attitude during my project, and for sharing his experience.

I send my warmest thanks to my excellent colleagues, Tiina Marjomaa, Sari Nummikoski, Tanja Kanerva, Outi Leppäranta, Tanja Törölä, Eva Sutinen, Hely Ollila and Merita Kouri for conducting the analysis, for providing effective assistance throughout the whole collaboration period, and for their friendship. I would like to express my gratitude to my co-authors: Professor Harri Alenius PhD, MD; Docent Lynne Bingle PhD; Professor Hideo Kobayashi PhD; Professor Noboru Hattor PhD, MD; Professor Nobuoki Kohno PhD, MD; Mikko Rönty PhD; Ulrich Bergmann PhD; Bianca Araujo PhD; Matti Höyhtyä PhD, and Sari Tiitinen.

I owe profound gratitude to all the participants who volunteered to take part in my project during all these years.

I am truly thankful to the language centre of University of Helsinki proving the linguistic revision for my thesis. I wish to thank Derek. Ho PhD, for linguistic revision of my thesis.

Finally, I deeply thank my family for endless support. I thank my parents, Yonghong and Guangmian for encouraging me to study abroad and for giving me great support in my life. I would like to thank my dear husband Xia Han, for helping me in my life and for giving me powerful support. I truly appreciate him for understanding me in the difficult times.
REFERENCE


28 Bel EH, Ten Brinke A: New anti-eosinophil drugs for asthma and copd: Targeting the trait! Chest 2017


48  Buckley ST, Ehrhardt C: The receptor for advanced glycation end products (rage) and the lung. Journal of Biomedicine and Biotechnology 2010
63 Carolan BJ, Sutherland ER: Clinical phenotypes of chronic obstructive pulmonary disease and asthma: Recent advances. J Allergy Clin Immunol 2013;131:627-634; quiz 635.


106 Evans RA, Steiner MC: Pulmonary rehabilitation: The lead singer of copd therapy but not a "one-man band". Chest 2017;152:1103-1105.


155 Ishikawa N, Ohlmeier S, Salmenkivi K, Myllarniemi M, Rahman I, Mazur W, Kinnula VL: Hemoglobin alpha and beta are ubiquitous in the human lung, decline in idiopathic pulmonary fibrosis but not in copd. Respir Res 2010;11


177 Kobayashi H, Kanoh S, Motoyoshi K: Serum surfactant protein-a, but not surfactant protein-d or kl-6, can predict preclinical lung damage induced by smoking. Biomarkers 2008;13:385-392.


221 Nashida T, Shimomura-Kuroki J, Mizuhashi F, Haga-Tsujimura M, Yoshimura K, Hayashi-Sakai S: Presence of bpifb1 in saliva from non-obese diabetic mice. Odontology 2017
226 Network BTSSIG: British guideline on the management of asthma - a national clinical guideline - introduction. Thorax 2008;63


