Review of recent lung biomarkers of potential harm/effect for
tobacco research [version 1; peer review: awaiting peer
review]

Erin Beattie1*, Jeffery Edmiston2*, Patrudu Makena3*, Elizabeth Mason4*,
Mike McEwan5*, Krishna Prasad5*

1Cliantha Research, 1310 Fewster Drive, Mississauga, ON L4W 1A4, Canada
2Altria Client Sercies, Richmond, Virginia, USA
3RAI Services Company, Winston-Salem, North Carolina, USA
4Imperial Brands, Bristol, UK
5British American Tobacco, Southampton, UK

* Equal contributors

Abstract
Biomarkers of potential harm (BoPH) are indicators of biological
perturbations which may contribute to the pathophysiology of
disease. In this review, we critically assessed the published data on
lung-related BoPH in human lung disease for potential use in
evaluating the effects of tobacco and nicotine products. A Scopus
literature search was conducted on lung disease biomarkers used in a
clinical setting over the last 10 years. We identified 1171 papers which
were further screened using commercial software (Sciome SWIFT-
Active Screener) giving 68 publications that met our inclusion criteria
(data on the association of the biomarker with cigarette smoking, the
impact of smoking cessation on the biomarker, and differences
between smokers and non-smokers), the majority of which
investigated chronic obstructive pulmonary disease. Several
physiological and biochemical measures were identified that are
potentially relevant for evaluating the impact of tobacco products on
lung health. Promising new candidates included blood biomarkers,
such as surfactant protein D (SP-D), soluble receptor for advanced
glycation end products (sRAGE), skin autofluorescence (SAF), and
imaging techniques. These biomarkers may provide insights into lung
disease development and progression; however, all require further
research and validation to confirm their role in the context of tobacco
and nicotine exposure, their time course of development and ability to
measure or predict disease progression.
Keywords
Biomarkers of potential harm, Lung biomarkers, Tobacco product evaluation, Nicotine, Lung disease

Corresponding author: Elizabeth Mason (elizabeth.mason@uk.imptob.com)

Author roles: Beattie E: Investigation, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; Edmiston J: Investigation, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; Makena P: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; Mason E: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; McEwan M: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; Prasad K: Investigation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: P.M. is employed by RAI Services Company. RAI Services Company is a wholly owned subsidiary of Reynolds American, Inc., which is a wholly owned subsidiary of British American Tobacco plc. M.M. and K.P. are employees of British American Tobacco (BAT) Research and Development, UK. JSE is employed by Altria Client Services LLC, Center for Research and Technology, Richmond Virginia USA. E.M. is employed by Imperial Brands plc, UK.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Beattie E et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Beattie E, Edmiston J, Makena P et al. Review of recent lung biomarkers of potential harm/effect for tobacco research [version 1; peer review: awaiting peer review] F1000Research 2021, 10:1293
https://doi.org/10.12688/f1000research.55411.1

First published: 17 Dec 2021, 10:1293 https://doi.org/10.12688/f1000research.55411.1
**Introduction**
Cigarette smoking causes many different serious lung diseases and the World Health Organization (WHO) estimated 1.5 million worldwide tobacco smoking related premature deaths from chronic respiratory disease in 2017 (World Health Organisation, 2019). Identifying and understanding biomarkers that are early indicators of changes in lung health could play an important role in evaluating the potential for new tobacco products that have a reduced risk of causing tobacco-related disease. Currently, the most prevalent non-cancer respiratory condition caused by tobacco smoking is chronic obstructive pulmonary disease (COPD). COPD is a complex disorder involving both airways and lung parenchyma with airflow limitation that is not fully reversible, generally progressive and accompanied with abnormal inflammatory response (Diaz et al., 2008).

COPD is generally diagnosed by lung function measurement using spirometry (Graham et al., 2019). Spirometry provides a method of measuring lung physiology that is relatively simple to perform and interpret results (with appropriate training), requires little equipment and is minimally invasive (Gold and Koth, 2016). Spirometry assesses lung function by measuring the volume of air that the patient can expel from the lungs after a maximal inspiration; for example, the forced expiratory volume in one second (FEV1). Expiratory airflow, which is ultimately what is measured in spirometry, is a function of muscular effort, elastic recoil of lungs and thorax, small and large airway function, and interdependence between small airway and the surrounding alveolar attachments (Gold and Koth, 2016).

Patients with suspected obstructive lung disease often present with symptoms such as persistent cough and shortness of breath (Global Initiative for Chronic Obstructive Lung Disease, 2018). Medical history and exposure to other risk factors such as cigarette smoking or air pollution, can assist with a clinical diagnosis. However, the use of spirometry is generally used to confirm a COPD diagnosis. It is a reliable method of differentiating between obstructive airways disorders (e.g. COPD, asthma) and restrictive diseases (in which the size of the lungs is reduced, e.g. fibrotic lung disease).

The Global Initiative for Obstructive Lung Disease (GOLD) guidelines provide recommendations for the spirometry readout required to confirm a COPD diagnosis and are based upon the FEV1/FVC ratio (Global Initiative for Chronic Obstructive Lung Disease, 2018). This represents the proportion of a person’s vital capacity that they are able to expire in the first second of forced expiration (FEV1) as compared to the full, forced vital capacity (FVC). A (FEV1/FVC) readout of less than 0.7 or 70 % is the current threshold to confirm the presence of persistent airflow limitation. Actual versus predicted FEV1 results can then be used to classify the severity or stage of the disease.

However, 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 (Salzman, 2012). Several clinical markers and other biomarkers have been proposed for COPD (Gonçalves et al., 2018), however, clinically useful early biomarkers, those that identify pre-symptomatic COPD condition, have yet to be identified. Biomarkers of COPD may be useful in aiding diagnosis, defining specific phenotypes of disease, monitoring exacerbations and evaluating the effects of drugs (Borrill et al., 2008). Biomarkers of COPD may also be useful when evaluating potentially reduced risk tobacco products. The purpose of this narrative review (a project by the CORESTA Biomarker Sub-Group, BMK-161-NWIP) was to examine the recent literature for biomarkers that would be useful in evaluating potential improvements in lung health as smokers switch to potentially reduced risk products. Currently there are few validated biomarkers in this area. The aim was to broadly explore potential biomarkers with specific relevance to lung disease which ranged from spirometry, imaging, to blood and lung sampling.

**Methods**
The Scopus literature search was conducted in June 2018 included the following parameters: Lung or Pulmonary, and Tobacco or Smok*, and Biomarker or Change* or Difference* or Measu*, and COPD or Emphysema or Bronchitis or Asthma or Obstruct*, and Clinical or Human. It was limited to papers published in English in the last 10 years (from January 1st 2008 to June 14th 2018) and to clinical trials. The search identified 1,171 papers to screen for the review.

A web-based collaborative review software application (SWIFT-Active Screener (Howard et al., 2020)) was used to complete the initial screening of the titles and abstracts of the 1,171 identified papers (manual screening could be used as an alternative to this software). SWIFT-Active Screener uses statistical models to continuously prioritize and estimate relevant papers left to screen. The inclusion criteria were: human studies only, data on the association of cigarette smoking, smoking cessation with the biomarker, and data on smoking versus non-smoking, preferably with data on changes with smoking cessation. Exclusion criteria were: studies of exposure to environmental tobacco smoke, studies of other non-cigarette tobacco products without smoking information, non-lung diseases, genome studies, drug-focused studies, and review articles. Once SWIFT-Active Screener predicted that we had identified 95.8 % of the relevant papers, 322/1,171 papers were screened using a two-reviewer system (i.e., both reviewers agreed the paper should be included for...
further review, based on the abstract), and finally 68 out of the predicted 71 papers were targeted for further (full-text) screening. Papers were not further filtered according to study quality in order to identify any potential biomarkers that may be of use in the future.

Four potential sources of biomarkers were identified from the included papers: spirometry (lung function), imaging, blood and lung sampling.

An earlier version of this article can be found on coresta.org (url: https://www.coresta.org/review-recent-lung-biomarkers-potential-harmeffect-tobacco-research-34041.html).

Results

Spirometry

Of the papers screened in this study, spirometry was often used to diagnose lung function impairment; 32 of the studies included in this review which investigated non-spirometric markers of lung also included some spirometry data, reflecting the validated status of this technique in detecting changes in the lung and diagnosing COPD. Only a small number of these papers discussed spirometry in the context of lung disease and tobacco, and these are discussed here.

As noted previously, COPD is a complex and heterogenous disease; not all patients progress at the same rate and the reasons for this heterogeneity are poorly understood. This is clearly observed in the ECLIPSE study, a controlled COPD cohort that included 2,164 clinically stable COPD patients, 245 never smokers and 337 smokers with normal lung function (Agustí et al., 2010). By definition, patients with COPD had airflow limitation, whereas spirometry was normal in the two control groups; clinically stable COPD patients showed significant decreases (p < 0.01) in FEV1 (% predicted) and FEV1/FVC compared to the smoking and non-smoking controls5. Smoking controls also showed a significant reduction in spirometric measures as compared to the non-smoking controls (p < 0.05) (Agustí et al., 2010). The paper also looked at a number of other factors that may indicate the heterogeneity of the disease such as exercise tolerance and reported exacerbations.

Differences in disease progression were also seen by Lawrence et al. (2017) (Lawrence, 2017 #2997), who noted that of the 370 COPD patients assessed at baseline and then one year later, patients who declined due to FEV1 change had lower FEV1 at baseline (means 55.7 % vs 67.8 %; p = 0.0023), and larger FEV1 deterioration over 1 year (284 mls vs 44 mls; p = 0.0021), and were more likely to be current smokers (62.5 % vs 10.0 %; p = 0.04) (Lawrence et al., 2017).

Thomsen et al. (2014) assessed the relative rather than absolute changes in FEV1 of 3,218 relatively healthy heavy smokers participating in the Danish Lung Cancer Screening Trial (DLCST), a 5-year longitudinal study investigating the effect of screening on lung cancer mortality. Here, the authors reported that when measuring relative changes in FEV1, a significant acceleration of decline was observed in smokers compared to former smokers (smoker vs former smokers: −0.20 % (p = 0.027)), and those with airflow limitation (AFL) compared to those without (AFL vs no AFL: −0.94 % (p < 0.001)) (Thomsen et al., 2014). The authors suggested these factors may speed up the loss of lung function.

In terms of identifying biomarkers of harm/effect for potential use in evaluation of the effects of tobacco use, a study should track changes in spirometric values as a smoker quits tobacco cigarettes. Within the scope of this aim, only one paper attempted to do this (Drummond et al., 2012). Longitudinal data from the Lung Health Study, a clinical trial of intensive smoking cessation intervention with or without bronchodilator therapy in 5,887 smokers with mild to moderate airflow obstruction were analyzed to determine the association between spirometric measures and 5-year decline in FEV1 and 12-year mortality. When stratifying the study cohort by smoking pattern (either continuous smoker, intermittent quitter or sustained quitter), they found that the association between lower baseline lung function and accelerated decline was present in all three groups, but to a lesser extent in the sustained quitters (Drummond et al., 2012). Intermittent quitters and continuous smokers showed similar patterns when baseline lung function was normal or mildly impaired. However, at levels of more severe baseline lung impairments, they reported that intermittent quitters demonstrated less rapid lung function decline than continuous smokers (Drummond et al., 2012). All smoking patterns showed lung function decline, although this was most accelerated in the continuous smoking group as compared to the intermittent and sustained quitters (Drummond et al., 2012). Never-smokers were not part of this Lung Health Study, thus preventing a more complete comparison to lung function decline in never smokers.

While spirometry is considered the gold standard for assessing lung function, it is not without limitations (Global Initiative for Chronic Obstructive Lung Disease, 2018, Andreewa et al., 2017, Perez-Padilla et al., 2015). The technique may not be sensitive enough to detect the condition in asymptomatic patients, such as younger age smokers. There is also poor correlation between spirometry data and other clinical outcomes such as exacerbation frequency.
 Even in patients diagnosed with severe airflow obstruction, not all reported symptoms, exacerbations or showed impaired exercise tolerance, suggesting that FEV1 does not fully capture COPD complexity (Agustí et al., 2010). Therefore, the identification and validation of alternative biomarkers that can be used in the diagnosis, prognosis and treatment of COPD are still required.

**Imaging**

Our search yielded nine papers that primarily utilized lung imaging techniques to assess differences in the lungs of smokers and/or respiratory disease patients. Seven of the papers used computed tomography (CT) scans and two used magnetic resonance imaging (MRI) to evaluate the lungs of subjects in the studies. Although spirometry is the primary method used to diagnose lung function impairment, direct imaging of the lungs may add additional sensitivity to the underlying changes in the lung and provide insights into the respiratory disease mechanisms caused by smoking.

**CT scans**

In a study by Diaz et al. (2017), quantitative CT measures were used to assess ratios of bronchial lumen and adjacent artery, peak wall attenuation, wall thickness, and wall area percent in 21 smokers with bronchiectasis and 21 never smokers. They demonstrated that these measures could detect differences between the two populations in both affected and unaffected areas of the lungs. Specifically, they identified changes in pulmonary vasculature as a potential mechanism for lung functional changes during bronchiectasis.

Bodduluri et al. (2017) used CT measurements of mean lung density at the end of expiration and the end of inspiration in 8,034 subjects of which 103 were lifetime non-smokers. They evaluated the changes in lung regions with normal density and found evidence of gas trapping associated with smoking even in these normal appearing lung regions. The authors concluded that these findings represented mild small airway disease and it was associated with respiratory morbidity. Small airway disease was also investigated by Kirby et al. (2017b) using a similar technique. Inspiratory to expiratory CT measurements were captured and assessed in 133 non-smokers, 97 at risk (former or current smokers), 140 COPD GOLD stage I (early/mild), and 96 COPD GOLD stage II (moderate) patients. The authors assessed two methods for classifying gas trapping; parametric response mapping and a novel disease probability Measure. The two measures performed similarly and could distinguish between at risk populations and GOLD I and GOLD II populations, but the disease probability measure classified more areas as having gas trapping in general and it was the only method that detected more gas trapping in the COPD GOLD II group relative to the COPD GOLD I group. Additional research is needed to understand the differences between these assessment techniques.

Volume adjusted 15th percentile density (PD15) was used by two studies to assess changes in the lung. Shaker et al. (2012) studied 2,052 subjects (893 with airflow obstruction) as part of the Danish Lung Cancer Screening Trial. Increases in PD15 were associated with number of cigarettes per day and pack years in this study. PD15 was also evaluated in a study by Ashraf et al. (2011) and showed similar results with current smokers (n = 548) having higher PD15 scores than former smokers (n = 178). Additionally, this study demonstrated that decreases in PD15 continue for 2 years upon smoking cessation (n = 77). The dose response and reversibility of PD15 indicates it might be helpful in evaluating potentially reduced risk tobacco products. However, the time course for reversal took approximately two years.

One paper by Smith et al. (2014) investigated spatially matched airway thickness in cigarette smokers with and without COPD. This study demonstrated that spatially matching airways are thinner in the lungs of COPD patients. Previous techniques that do not match for similar airway location may result in comparisons of more proximal airways in COPD patients to more peripheral airways in the control group (Smith et al., 2014). Although this study describes a potential improvement to the assessment technique it did not report any differences due to smoking.

One study by Kimura et al. (2017) used CT scans of the sinus to investigate the extent of sinusitis and asthma severity in non-smokers (n = 130, < 10 pack years) and smokers (n = 76, ≥ 10 pack years). However, they conducted their analysis separately within each group and did not compare differences between the two populations.

**MRI**

MRI was used in two identified studies. Alamidi et al. (2016) evaluated resonance relaxation time (T1) in 24 COPD patients (current and former smokers) compared to 12 age-matched control non-smokers. The lung T1 was significantly shorter in the COPD patients and highly correlated with CT density scans and spirometry measurements. Kirby et al. (2017a) used MRI with inhaled noble gas to evaluate ventilation defect percent to assess the lungs of 34 ever-smokers, 48 COPD patients, and 42 never smokers over 30 months. Ventilation defect percent was greater in ever-smokers and...
COPD patients compared to never smoker controls. The ventilation defect percent also correlated well with spirometry and respiratory questionnaire responses. The authors concluded that “These data strongly support the use of MRI intermediate endpoints in COPD studies.”

Lung imaging as an evaluation tool for potentially reduced risk tobacco products is appealing because it allows the visualization of the primarily exposed tissue during cigarette smoking. The imaging techniques are evolving and refinements of the techniques as well as additional data from longitudinal studies will prove useful in the future. The use of current imaging techniques is limited as both CT and MRI systems are expensive and require specialized facilities and training to conduct and interpret the results.

**Blood**

Persistent systemic inflammation is thought to play a significant pathogenic role in COPD (De Martinis et al., 2005). Levels of circulating fibrinogen, C-reactive protein (CRP), interleukins 6 (IL-6) and 8 (IL-8) are biomarkers that have been most often studied in COPD (Agustí et al., 2012).

**Fibrinogen**

Fibrinogen is a glycoprotein found in blood, which is converted to fibrin by the enzyme thrombin during tissue and vascular injury leading to fibrin-based blood clots. Elevated levels of fibrinogen have been implicated in cardiovascular disease (Cook and Ubben, 1990). In addition, both CRP and fibrinogen have been associated with the severity of COPD and the risk of cardiovascular disease (van Dijk et al., 2013).

A number of the papers selected through the screening process identified fibrinogen as a potential biomarker for lung disease, especially COPD, with both targeted and untargeted methods being used. Generally, these papers investigated potential biomarkers and their association with severity of COPD or airflow limitation. In the six papers that investigated fibrinogen in COPD patients versus control subjects (made up of current smokers, former smokers and never smokers), elevated levels were found in the COPD groups, irrespective of smoking status (Agustí et al., 2012, Dickens et al., 2011, Merali et al., 2014, York et al., 2010). Devanarayan et al. (2010) and Garcia-Rio et al. (2010) both showed increases in fibrinogen in COPD subjects versus healthy subjects, but these were not significant. However, the Devanarayan et al. (2010) study did identify fibrinogen as one of nine markers capable of distinguishing COPD “rapid decliners” from “slow decliners”. This is supported by the study of Rana et al. (2010) which identified fibrinogen among 33 proteins that were differentially expressed between COPD fast and slow decliners. In addition, both Garcia-Rio et al. (2010) and Dickens et al. (2011) indicated an association between elevated fibrinogen levels and reduced exercise tolerance as measured by the 6-minute walk test.

Very few of the identified papers investigated the effects of smoking on fibrinogen levels. One study by van Dijk et al. (2013) investigating the acute effects of cigarette smoking in COPD patients, showed that fibrinogen levels significantly increased directly after smoking and remained elevated at 35 minutes. While Kalhan et al. (2010) reported a longitudinal study, which followed subjects over 20 years (CARDIA study) and showed that higher fibrinogen was associated with smoking and higher body mass index (BMI). This study also showed an association between higher fibrinogen and 15-year decline in both FVC and FEV1 and that higher levels of fibrinogen at year 7 was associated with greater risk of having abnormal FVC and FEV1 in middle age (Kalhan et al., 2010).

As these papers generally focused on biomarkers that were associated with COPD once it had been diagnosed, there was scant information on potential early biomarkers associated with COPD development that could be used in the context of potential risk reduction with potentially reduced-risk products (PRRPs). However, the study by Agustí et al. (2012) does show data which indicates a potentially significant increase in levels of fibrinogen in healthy smokers compared with healthy non-smokers, while Dickens et al. (2011) indicate that there was no significant difference between smokers and non-smokers for plasma fibrinogen levels.

In terms of use of fibrinogen as a biomarker for COPD in PRRP switching studies, Scherer (2018) indicated that it may be a useful marker in larger field studies as power calculations indicate a sample size of a minimum 216 subjects would be required to achieve a statistical difference. This review also estimated that from cessation studies, it would take one year to achieve non-significant levels of fibrinogen in plasma. However, it should be noted that in a cross-sectional analysis, Haswell et al. (2014) noted there was no significant difference in levels of fibrinogen between smokers and never-smokers.
**Interleukins**

Interleukins are a subset of a larger group of cellular messenger molecules called cytokines, which are modulators of cellular behavior, particularly immune signaling. Like other cytokines, interleukins are produced in response to a stimulus, such as an infectious agent. Increased levels of interleukins in blood serum have been associated with COPD.

The papers identified through the screening process mainly investigated levels of interleukins in COPD patients and healthy subjects which were both smokers and never smokers using targeted analyses. Of the different interleukins analyzed, IL-6 and IL-8 were the most common. Table 1 shows the levels of IL-6 and IL-8 reported in the papers identified through the screening process for this review.

Two studies identified IL-6 as being significantly higher in subjects with COPD than healthy subjects (Agustí et al., 2012, Dickens et al., 2011), whereas Garcia-Rio et al. (2010) reported IL-6 as not significantly different between COPD patients and healthy subjects after adjustment for gender, age, BMI and pack-years, although there was a significant difference before adjustment. In addition, a study by Hackett et al. (2011) investigating single nucleotide polymorphisms (SNPs) identified a SNP associated with IL-6 that had a significant interaction with cigarette smoking and lung function and that subjects with this SNP would be at risk of developing both COPD and cardiovascular disease. IL-8 was also significantly higher in COPD patients versus healthy subjects in two studies (Dickens et al., 2011, García-Río et al., 2010). Furthermore, García-Río et al. (2010) reported that IL-6 and IL-8 were related to severity of COPD and both were inversely related to exercise tolerance as measured by the 6-minute walk test, whereas Agustí et al. (2012) indicated that IL-6 tended to increase with severity of airflow limitation while IL-8 did not.

Other interleukins were reported in the identified publications, but these were only incorporated in single papers. Dickens et al. (2011) investigated levels of IL-12p40, but this showed no significant change between COPD patients and healthy subjects. Devanarayan et al. (2010) studied a number of interleukins in COPD patients and healthy controls. This study showed that IL-4 and IL-7 had the largest increase in the COPD patients versus healthy controls. In addition, a study on elderly subjects in Uppsala, Sweden showed no association between lung function and IL-6 levels, even though this study did show an association between CRP and FEV1 (Kuhlmann et al., 2013).

Few of these studies investigated the levels of interleukins in smokers and non-smokers, however, Agustí et al. (2012) found that IL-8 was higher in smokers versus non-smokers. In concordance with this, Dickens et al. (2011) reported that IL-8 was significantly higher in smokers versus non-smokers, but that IL-6 in smokers was not significantly different compared to non-smokers. A study by Hoonhorst et al. (2014b) investigated the acute effects of smoking three cigarettes in one hour. This study showed an increase in activation of neutrophils in young subjects who were deemed susceptible to COPD, but that levels of IL-6 and IL-8 decreased in older healthy controls, whereas IL-6 decreased significantly in young healthy controls compared with COPD patients.

A review on biomarkers of biological effect by Scherer (2018) also looked at IL-6 for use in PRRP switching studies and identified this marker in blood as having limited suitability for use in field studies. However, for IL-8 there were few studies carried out to investigate this biomarker in this context. In addition, a cross-sectional analysis of smokers versus non-smokers showed that plasma levels for both IL-6 and IL-8 were all below the limit of detection (Haswell et al., 2014).

**C-reactive protein**

CRP is produced in response to inflammatory conditions such as bacterial, viral and fungal infections and is synthesized in the liver and released into the circulation. It is part of the innate immune system and is associated with the activation of the complement cascade that has been associated with atherosclerosis as well as COPD and asthma (Anderson, 2006, Peisajovich et al., 2008, Thompson et al., 1999).

Of the papers selected through the screening process that investigated plasma biomarkers associated with COPD, eight included CRP through either targeted or untargeted analyses. In most cases, the main objective of these papers was to identify potential biomarkers and their association with COPD disease severity. Of these papers, three identified CRP as being significantly elevated in COPD patients versus healthy subjects (including both smokers and non-smokers) (Agustí et al., 2012, Dickens et al., 2011, Garcia-Río et al., 2010). In addition, one study using a targeted analysis of plasma CRP on an elderly population of 888 subjects showed CRP and leucocyte count were independently associated with FEV1 (Kuhlmann et al., 2013). Leeming et al. (2017) investigated matrix metalloprotease generated collagen type 1 and type IV fragments (C1M and C6M) to investigate predictors of FEV1 change in COPD patients. This study indicated that C1M and C6M along with CRP and emphysema were significant predictors of lung function decline.
Table 1. IL-6 and IL-8 levels in blood serum of chronic obstructive pulmonary disease (COPD) subjects, smoker and non-smoker controls. IQR = interquartile range.

| Reference                  | Bio-marker | COPD subjects | Smoker controls | COPD subjects vs Smoker controls | COPD subjects vs Non-smoker controls |
|----------------------------|------------|---------------|----------------|----------------------------------|-------------------------------------|
|                            |            | N | Median (pg/mL) | IQR | N | Median (pg/mL) | IQR | P value | N | Median (pg/mL) | IQR | P value |
| Dickens et al. (2011)      | IL-6       | 194 | 0.2           | 2.8 | 37 | 0.2           | 0.0 | 0.004    | 37 | 0.2           | 0.4 | 0.072    |
| Garcia-Rio et al. (2010)   | IL-6       | 324 | 4.5           | -   | 110 | 3.5           | -   | 0.095    | -   | -           | -   | -        |
| Agustí et al. (2012)       | IL-6       | 1755 | 1.5          | 2.3 | 297 | 0.6           | 1.0 | <0.001   | 202 | 0.4           | 0.7 | <0.001   |
| Dickens et al. (2011)      | IL-8       | 190 | 8.8           | 5.2 | 36  | 7.7           | 4.8 | 0.051    | 34  | 6.7           | 3.0 | 0.009    |
| Garcia-Rio et al. (2010)   | IL-8       | 324 | 7.3           | -   | 110 | 3.0           | -   | 0.003    | -   | -           | -   | -        |
| Agustí et al. (2012)       | IL-8       | 1755 | 6.9          | 10  | 297 | 7.8           | 10  | 0.013    | 202 | 4.3           | 4.9 | <0.001   |
One study on untargeted analyses on the plasma proteome of COPD patients and matched ex-smokers with no airflow obstruction identified CRP among 31 different proteins that were differentially expressed (Merali et al., 2014). However, two other papers identified through the screening process also looked at untargeted analysis of the plasma proteome of COPD and did not identify CRP as differentially expressed (Rana et al., 2010, York et al., 2010). Of the proteins that were differentially expressed, the majority of these proteins belonged to the complement and coagulation cascades in both studies.

On the question of the effects of smoking on plasma CRP levels, only two papers addressed the question of the effects of smoking on plasma CRP levels. van Dijk et al. (2013) investigated the acute effects of smoking in a cohort of 31 COPD subjects and showed that CRP increased directly after smoking and returning to normalized levels after 35 minutes. Whereas Kalhan et al. (2010) analyzed plasma samples from a longitudinal study on healthy young adults at time of enrolment over a 20-year period and showed that subjects with higher levels of CRP at year 7 were associated with a greater risk of airflow decline in middle age. A greater risk of having obstructive lung disease by year 20 of the study was noted in subjects that had elevated CRP levels at year 7 and had 10 or more pack years of smoking history.

Other observations from the screened papers were that BMI was associated with CRP (Agustí et al., 2012, Garcia-Rio et al., 2010, Kalhan et al., 2010, Kuhlmann et al., 2013) and that exacerbations in COPD subjects are associated with further elevation of CRP (Dickens et al., 2011).

As most of the papers are aimed at looking for discriminatory biomarkers for decline in lung function for subjects with COPD, there is little information on how these biomarkers behave in healthy subjects and whether they are predictive of COPD development. A study by Kalhan et al. (2010) does indicate there is a possibility that CRP may be predictive of obstructive lung disease from analysis of CRP levels in young adulthood. However, Scherer (2018) indicated that CRP is less suitable for use in such studies. This is supported by Ogden et al. (2015) and Haswell et al. (2014) which both reported that there was no significant difference between smokers and non-smokers for CRP.

In addition, CRP is also associated with cardiovascular events and as this affects the cardiovascular system, can potentially be a major confounder for studies investigating effects on the lung from tobacco smoking (Penson et al., 2018, Ridker et al., 2002). In addition, CRP is an acute phase protein produced through microbial infection, and this may also be a confounding source in clinical studies. These challenges may be alleviated through longitudinal studies with large subject numbers, and thus CRP may be relevant for use in longer term studies, such as post-market surveillance.

**SP-D**

Several of the screened publications identified surfactant protein D (SP-D) as a potential biomarker related to the development of COPD. SP-D (43kDa) is a local inflammatory biomarker belonging to the collecting sub-group of the C-type lectin superfamily. The protein can form multimers assembling into trimers and can then form higher multimers such as dodecamers (formed from four trimers) via disulfide crosslinking. It is principally produced in and secreted from type II pneumocytes in the lung and found in alveoli, distal airways and in the blood and plays an important role in regulating immune and inflammatory responses in the lungs (Ito et al., 2015).

Generally, this biomarker was investigated using targeted analysis and was found to be elevated in the serum of subjects with COPD compared with healthy subjects (Winkler et al., 2011). Dickens et al. (2011) observed that serum SP-D levels were elevated in COPD subjects with unresolved exacerbations and reproducible over a three-month period. Whereas Carolan et al. (2014) found a potential association of SP-D with emphysema and Kim et al. (2012) reported a significant association between SNPs (rs10836312 and rs12793173) and serum SP-D. Contrary to this, Ito et al. (2015) found that total serum SP-D was not significantly different between subjects with smokers with COPD, non-COPD smokers and never smokers. However, this study also investigated fucosylated SP-D in serum and found levels to be significantly higher in COPD patients compared with non-COPD smokers and was directionally higher but not significantly different when compared to never smokers. This study also investigated levels of fucosylated SP-D in non-COPD smokers and never smokers and saw no significant difference between these groups, suggesting fucosylation levels were not affected by smoking status (Ito et al., 2015). Only one other of the screened manuscripts investigated levels of SP-D in smokers and never smokers and this indicated that smoking increased levels of serum SP-D and reduced levels in bronchioalveolar lavage (BAL) fluid from the lungs (Winkler et al., 2011). In addition, Winkler et al. (2011) also found that smoking led to changes in the quaternary structure of SP-D in BAL compared to never smokers with increased levels of these smaller subunits of pulmonary SP-D being found in smokers and subjects with COPD. In support of this, one other publication which was not identified through the screening process, showed that serum SP-D levels were significantly higher in smokers compared to never smokers and former smokers. This study carried out on 756 twin pairs in Denmark also
showed there was no significant difference in serum SP-D levels between never smokers and ex-smokers (Johansson et al., 2014).

**AGE and RAGE**

Accelerated formation and accumulation of advanced glycation end products (AGEs) is believed to occur as a result of oxidative stress and smoking.

AGEs are a heterogeneous and complex group of compounds formed by non-enzymatic glycation and oxidation of proteins and lipids. These AGEs then accumulate in tissues with aging. Furthermore, under oxidative stress and inflammatory conditions their formation and accumulation increase. Examples of AGEs are \( \text{N}^\varepsilon-(\text{carboxymethyl})-\text{lysine} \) (CML) and \( \text{N}^\varepsilon-(\text{carboxyethyl})-\text{lysine} \) (CEL) and pentosidine (Hoonhorst et al., 2016).

The receptor for AGE (RAGE) is an immunoglobulin family member that is highly expressed in human lung (Carolan et al., 2014). It has been associated with several inflammatory diseases such as diabetes mellitus, cardiovascular and respiratory diseases (Carpagnano et al., 2016, Kim et al., 2012). This membrane bound protein can bind to AGE, triggering inflammatory responses, oxidative stress, and RAGE over-expression (Hoonhorst et al., 2016).

In addition, there is a soluble form of RAGE (sRAGE) which can be found in the circulation and is believed to act as a decoy receptor for clearance of circulating AGEs, preventing ligation of membrane bound RAGE and therefore may act as a 'protective' mechanism (Cheng et al., 2013, Hoonhorst et al., 2016).

In the three manuscripts that investigate AGEs and RAGEs there was a consensus that levels of sRAGE decrease in COPD patients compared to healthy controls (Carolan et al., 2014, Cheng et al., 2013, Hoonhorst et al., 2016). Both Carolan et al. (2014) and Cheng et al. (2013) indicated that sRAGE was negatively associated with severity of emphysema, whereas Hoonhorst et al. (2016) indicated sRAGE was lower in GOLD stage III (Severe COPD) patients compared to GOLD stage I.

Only Hoonhorst et al. (2016) investigated the levels of AGEs in the circulation and concluded that levels of CML and CEL were comparable in COPD and healthy subjects after adjustment for confounding variables.

For comparisons between smokers and non-smokers, Hoonhorst et al. (2016) showed higher levels of pentosidine in young healthy never smokers compared to young healthy smokers. However, this is reversed in old healthy never smokers compared to old healthy smokers. In the case of plasma sRAGE levels, there appears to be higher levels in both young and healthy never smokers compared to the smoker cohorts, but this was not significant. This is supported by Cheng et al. (2013) who found there was no significant difference between sRAGE levels in smokers versus non-smokers and no association with pack years of smoking.

Interestingly, AGEs have been shown to accumulate differently in body compartments and has been shown to accumulate in the skin but not plasma, sputum or bronchial biopsies. This accumulation in skin can be measured by skin autofluorescence (SAF) (Hoonhorst et al., 2014a, Hoonhorst et al., 2016).

Although not a plasma biomarker, SAF was found at higher levels along with lower plasma sRAGE and this was associated with decreased lung function (Hoonhorst et al., 2014a). In addition, this study shows higher SAF values in both old and young smokers compared to old and young never smokers (Hoonhorst et al., 2016). Another study reported by Hoonhorst et al. (2014a) also showed that with smoking history, a higher number of pack-years smoking was associated with higher SAF levels in the young and old healthy groups analysed. In addition, it showed that within the old healthy subjects, current smoking was associated with higher SAF levels. This is in agreement with a further study carried out by Isami et al. (2018) which shows that SAF was also correlated to pack year history of smoking. It should be noted that in subjects with cardiovascular disease and diabetes SAF levels were significantly higher than those without (Hoonhorst et al., 2014a), and this should be a consideration for this biomarker if applied in clinical studies.

**Lung sampling**

As previously discussed, COPD is a complex disorder involving both airways and lung parenchyma with airflow limitation that is not fully reversible, generally progressive and accompanied with abnormal inflammatory response (Diaz et al., 2008). 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. Several clinical markers and other
biomarkers have been proposed for COPD (Gonçalves et al., 2018), however, clinically useful early biomarkers that identify pre-symptomatic COPD condition have yet to be identified. Biomarkers of COPD may be useful in aiding diagnosis, defining specific phenotypes of disease, monitoring exacerbations and evaluating the effects of drugs (Borrill et al., 2008).

A variety of pathological changes at different locations of the lung were reported in COPD. Increased numbers of macrophages and CD8 T lymphocytes were observed in proximal airways (>2mm), peripheral airways (<2mm), lung parenchyma and pulmonary vasculature (MacNee, 2005). Goblet cell and squamous cell metaplasia is detected with mucus hyper secretion in both proximal and peripheral airways (Ramos et al., 2014). Enlarged smooth muscle and connective tissue, altered ciliary function, and increased neutrophils and lymphocytes infiltration of bronchial glands are observed in proximal airways. In the peripheral airways, an early-stage bronchiolitis with luminal and inflammatory exudates, peribronchial fibrosis and airway narrowing is noticed (MacNee, 2006). Alveolar wall destruction from loss of epithelial and endothelial cells and abnormal enlargement of airspaces distal to terminal bronchioles with micro- and macroscopic emphysematous changes are key pathological features observed in lung parenchyma (MacNee, 2006). In pulmonary vasculature, intimal thickening, endothelial dysfunction is seen in early stage and hypertrophy of vascular smooth muscle, collagen deposition, destruction of capillary bed, development of pulmonary hypertension and cor pulmonale are observed in late stage (MacNee, 2006).

There is an urgent need for discovery and validation of reliable disease biomarkers to better characterize the phenotypic heterogeneity and the prognosis of COPD. To that end, early measurable changes in molecular, cellular, structural or functional level are the important targets for biomarker development. These alterations progressively accumulate and contribute to the pathophysiology and eventually to the adverse disease outcomes. Therefore, collection and processing of high-quality biological specimens from different locations of the lung are critical to the development of COPD biomarkers. This section will focus on some of the COPD biomarkers from lung specific samples, 1) induced sputum, 2) bronchial biopsies, 3) bronchoalveolar lavage fluid, 4) exhaled breath condensate, 5) exhaled breath temperature, and present results from selected publications.

**Induced sputum**

A number of studies have suggested a pathogenetic role for airway inflammation in the induction of both chronic sputum production and chronic airflow obstruction in smokers (Pauwels, 2001). Sputum originates from both proximal and distal airways, which can be induced by inhalation of hypertonic saline. Induction of sputum is a safe, reliable, and relatively non-invasive method. Induced sputum provides information about both inflammatory cells and mediators present in the airways, potentially relevant for phenotypic characterization of COPD and asthma. Induced sputum differs from spontaneous sputum by having a higher number of viable cells and less squamous cell contamination (Pauwels, 2001). Several sputum induction methods were previously described (Efthimiadis et al., 2002, Paggiaro et al., 2002, Pavord et al., 1997, Scheicher et al., 2003, Weizschar and Horvath, 2013). The few limitations of this sample type include 1) the difficulty in obtaining healthy control samples, 2) the lack of general standardization, and 3) the presence of highly charged mucins that require usage of detergents which may interfere with some of the biomarkers of interest (Woolhouse et al., 2002). Several COPD biomarkers in induced sputum were previously described (Barnes et al., 2006, Comandini et al., 2009, Iwamoto et al., 2014, Keatings and Barnes, 1997, Kleniewska et al., 2016, Rutgers et al., 2000).

Titz et al. (2015) compared proteomic and transcriptomic changes in the sputum of asymptomatic smokers, COPD smokers, former smokers, and never smokers. In this study, smokers exhibited the proteomic and transcriptomic changes in mucin/trefoil proteins and a prominent xenobiotic/oxidative stress response. Proteomic analysis revealed that TIMP1, APOA1, C6orf58, and BPIFB1 (LPLUNC1), KRT19, PP1B, TF, AHSG, SERPINC1, AFM, ALB, HRG, and CNDP1 are differentially abundant proteins between the COPD and the asymptomatic smoker group. Aldehyde dehydrogenase 3A1 (ALDH3A1) showed the strongest and most significant increase in the asymptomatic and COPD smokers compared to non-smokers and former smokers (Table 2). In this study, except S100A6, there were no differentially expressed transcripts between the former smoker and never smoker groups (Titz et al., 2015).

Nicholas et al. (2010) used a proteomic approach to compare novel noninvasive biomarkers using induced sputum in GOLD stage 2 COPD and healthy smoker control subjects. In this study, the authors used a sequential approach of identifying 1,325 individual protein spots in which 37 were quantitatively and 3 qualitatively different between the two groups. Using this approach, apolipoprotein A1 and lipocalin-1 were potential biomarkers and reduced in patients with COPD when compared with healthy smokers (Table 2). Immunohistochemistry revealed that these two markers were localized to bronchial mucosa (Nicholas et al., 2010). Hoonhorst et al. (2016) assessed the association of AGEs and sRAGE in plasma, sputum, bronchial biopsies, and skin and found no differences in sputum sRAGE between smokers and COPD groups.
Table 2. Biomarkers in bronchial biopsies, induced sputum, bronchioalveolar lavage (BAL) and exhaled breath. COPD = chronic obstructive pulmonary disease; SP-D = surfactant protein D; EBT = exhaled breath temperature.

| Reference            | Sample type   | Biomarkers                                                                 | Study population                                      | Analytical method | Comments                                                                                                                                                                                                 |
|----------------------|---------------|----------------------------------------------------------------------------|-------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Kunz et al. (2013)   | Bronchial Biopsies | Versican, Collagen III, Collagen I                                         | COPD Smokers (40), COPD Ex-smokers (24)               | Immunohistochemistry | Long-term inhaled corticosteroids treatment is associated with increased lung function, suggesting that long-term inhaled steroids modulate airway remodelling thereby potentially preventing airway collapse in COPD. Smoking status is not associated with bronchial extracellular matrix proteins. |
| Hoonhorst et al. (2014b) | Bronchial Biopsies | Inflammatory cells count                                                   | (Young) Non-susceptible (29), (Young) Susceptible (21), (Old) COPD (13), Healthy controls (27) | Flow Cytometry     | COPD susceptibility at young age associates with an increased systemic innate immune response to cigarette smoking and suggests systemic inflammation in the early induction phase of COPD.                                      |
| Hoonhorst et al. (2016) | Bronchial Biopsies | AGER (Nε-(carboxymethyl)-lysine (CML), Nε-(carboxyethyl) lysine (CEL) and pentosidine) RAGE | Healthy smokers (Young (72), old (55)), Never-smokers (Young (36), old (28)) COPD (GOLD I-IV) (97) | Immunohistochemistry | No differences were found between young and old healthy subjects. RAGE levels in sputum and RAGE immunopositivity in whole sections from bronchial biopsies did not differ between groups.                          |
| Freeman et al. (2013) | Lung Tissue    | CD8+ T-Cells (TLR1, TLR2, TLR4, TLR6 and TLR2/1), CD4+ T cells (TLR2/1), CD8+ NK T cells (TLR2) | Healthy smokers, COPD smokers                        | Flow Cytometry     | Low levels of specific TLR expression, but the percentage of CD8+ T cells expressing TLR1, TLR2, TLR4, TLR6 and TLR2/1 was significantly increased in COPD subjects relative to those without COPD. |
| Titz et al. (2015)   | Induced Sputum | mucin/trefoil proteins (MUC5AC and TFF1/3), xenobiotic metabolism enzymes (NQO1, and GSTA1), peptidase regulators (TIMP1 and SERPINB1) proteins involved in redox processes (TXN, PRDX1, and SOD1). | Asymptomatic smokers, Smokers with early stage COPD, Former smokers, Never smokers | Proteomics         | Current smokers dearly reflected the common physiological responses, including alterations in mucin/trefoil proteins and a prominent xenobiotic/oxidative stress response. This study demonstrated that sputum profiling can capture the complex and reversible physiological response to cigarette smoke exposure. |
| Röpcke et al. (2012) | Induced Sputum | α1-Antitrypsin albumin (HSA), MMP3                                           | Healthy smokers (23), COPD Smokers (24)               | Immunoassays (Luminex® or ELISA) | This study describes a lung function independent inflammatory phenotype with an improved repeatability compared to sputum neutrophils alone.                                                                  |
| Reference          | Sample type | Biomarkers                      | Study population                      | Analytical method                  | Comments                                                                                                                                 |
|--------------------|-------------|---------------------------------|---------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Nicholas et al. (2010) | Induced Sputum | Apolipoprotein A1 Lipocalin-1 | Healthy smoker COPD GOLD II | Proteomics and Mass spectrometry | Apolipoprotein A1 and Lipocalin-1 were found to be significantly reduced in patients with COPD when compared with healthy smokers. Their levels correlated with FEV1/FVC, indicating their relationship to disease severity. |
| Röpcke et al. (2012) | BAL         | Total cell counts CD14+ MONO α1-Antitrypsin albumin (HSA) EGFR Calprotectin TIMP-1 IL-8 | Healthy smokers (23) COPD Smokers (24) | Immunoassays (Luminex® or ELISA) | A novel finding was the upregulation of Calprotectin in BAL fluid.                                                                        |
| Winkler et al. (2011) | BAL         | SP-D                            | Non-smokers (10) Healthy smokers (20) COPD Smokers (20) | ELISA                              | SP-D levels in BAL were decreased in smokers and patients with COPD. Pulmonary and serum SP-D levels are stable markers influenced by smoking and related to airflow obstruction and disease state. |
| Holz et al. (2014)   | BAL/Induced sputum | Calprotectin IL-8 MMP3          | Healthy smokers (23) COPD Smokers (24) | ELISA                              | Higher levels of BAL or induced sputum markers associated with neutrophilic airway inflammation (e.g. IL8, calprotectin or matrix metalloproteases, MMPs) in 2006/2007 were associated with a larger decline in lung function over a 5 year period. |
| Labor et al. (2016)  | Exhaled Breath | EBT                             | Current smokers (140 Woman (76); Men (64)) | X-Halo"                          | Change in EBT after smoking a cigarette at initial visit was significantly predictive for disease progression after 2 years. EBT shows potential for predicting the future development of COPD in current smokers. |
Bronchial biopsies

Bronchial biopsy is an invasive sample collection technique and specimens obtained by bronchoscopy provide a very useful tool to study cellular, immunological and molecular abnormalities of the airway mucosa (Azzawi et al., 1990, Beasley et al., 1989, Elston et al., 2004). Bronchial biopsies have been useful for documenting the structural changes, cellular patterns, and expression of inflammatory proteins in patients with COPD (Barnes et al., 2006). An accurate method of analysis for endobronchial biopsies in health and disease was previously described in the literature (Jeffery et al., 2003). The advantage of bronchial biopsies is that they represent airway tissue, maintaining the spatial relationships of structural components that may be important to functional changes (Jeffery et al., 2003). Bronchial biopsies can be cultured and used for *ex vivo* evaluations to assess their responses to experimental treatments. However, due to the invasive nature of the sample collection, it is difficult to recruit subjects in clinical trials. The biopsy of proximal airways may not closely reflect all the pathologic changes present in peripheral airways and lung parenchyma, which are the sites responsible for airflow limitation in COPD (Barnes et al., 2006).

The selected papers on bronchial biopsies or lung tissue evaluated several biomarkers of effect.

Kunz et al. (2013) compared several potential markers, including elastic fibers; proteoglycans (versican, decorin); collagens type I and III in bronchial biopsies after 30-months of inhaled steroids treatment or placebo in COPD current and ex-smokers. This study shows that there is no significant difference in the selected biomarkers in current and ex-smokers with COPD (Kunz et al., 2013). Freeman et al. (2013) evaluated subsets of specific TLRs on lung CD8+ T cells, CD4+ T cells, CD56+ NK cells, IFN-γ, and TNFα by CD8+ T cells in lung tissues of healthy smokers and COPD smokers. In this study, the percentage of CD8+ T cells expressing TLR1, TLR2, TLR4, TLR6 and TLR2/1 and only TLR2/1 and TLR2 on lung CD4+ T cells and CD8+ NKT cells was significantly increased in COPD subjects relative to those without COPD (Freeman et al., 2013). In separate experiments, IFN-γ and TNFα release significantly increased CD8+ T cells of COPD smokers when co-stimulation by Pam3CSK4, a specific TLR2/1 ligand.

Bronchoalveolar lavage fluid

Bronchoalveolar lavage (BAL), performed during flexible bronchoscopy, has gained widespread acceptance and provides important information about immunologic, inflammatory, and infectious processes taking place at the alveolar level (Meyer et al., 2012). BAL is an invasive procedure in which a bronchoscope with light and a small camera is passed through the mouth or nose into the lungs and epithelial lining fluid is collected by flushing into a small part of the lung and then recollected for examination. A flexible bronchoscope is longer and thinner than a rigid bronchoscope. Flexible bronchoscopy is more common than rigid bronchoscopy, and flexible bronchoscopy usually does not require general anesthesia and contains a fiberoptic system that transmits an image from the tip of the instrument to an eyepiece or video camera at the opposite end (Mehta et al., 2018). Generally, BAL has been performed safely on patients with stable COPD (Paradis et al., 2016). However, it is unsafe to perform BAL on patients with moderate/severe lung disease, or during acute exacerbations (Louhelainen et al., 2008). One of the limitations of BAL collection in COPD patients is often inadequate and not representative of the situation in the bronchioles due to airway collapse and reduced fluid recovery (Barnes et al., 2006).

BAL is the most common clinical procedure to sample the components of the pulmonary airways. BAL samples can be analyzed for cytological, hematologic, biochemical, and microbiological examination. A variety of lung specific biomarkers can be evaluated in BAL including surfactant proteins (Barnes et al., 2006, Dahl, 2008, Hartl and Griese, 2006, Moré et al., 2010, Winkler et al., 2011), proinflammatory cytokines (Barnes, 2008, Barnes, 2009), inflammatory responses/cells (Löfdahl et al., 2006, O’Donnell et al., 2006, Rovina et al., 2013, Wang et al., 2018, Wen et al., 2019), matrix metalloproteins (Molet et al., 2005, Ostridge et al., 2016, Papakonstantinou et al., 2015), ‘omic’ markers (including proteomic, metabolomic and microRNA) (Adamko et al., 2015, Chen et al., 2010, Cruickshank-Quinn et al., 2017, Fujii et al., 2017, Molina-Pinelo et al., 2014, Tu et al., 2014, Wendt et al., 2016), and oxidative stress markers (Drost et al., 2005, Morrison et al., 1999).

Some of the selected papers in this review compared biomarkers in smokers, non-smokers and COPD smokers in BAL and the majority of them were inflammatory markers. Röpcke et al. (2012) analyzed a large panel of markers in bronchoalveolar lavage, bronchial biopsies, serum and induced sputum of healthy smokers and COPD smokers (GOLD II). Samples were collected twice within a period of 6 weeks and over 100 different markers were validated for the respective matrices prior to analysis. Several biomarkers were reported, including total cell counts, CD14+ monocytes, α1-antitripsin, EGFR, HAS, TIMP1, calprotectin, and IL-8 in BAL and α1-antitripsin, HAS and MMP3 in induced sputum. (Holz et al., 2014) reported that potential prognostic value of biomarkers in BALF, sputum and serum in a
five-year follow-up of smokers with and without COPD. The lung function decline was larger in smoking COPD patients than in healthy smokers. This study demonstrated that higher levels of BAL or induced sputum markers, IL-8, calprotectin or, MMPs (neutrophilic airway inflammation markers) in the baseline were associated with a larger decline in lung function over a 5-year period. Winkler et al. (2011) compared BAL and serum SP-D in young and elderly smokers, COPD smokers and non-smokers. SP-D levels decreased in an order of young, smokers, and COPD smokers compared to non-smokers.

**Exhaled Breath**

**Condensate.** Collection of exhaled breath condensate (EBC) for sampling the lower airways involves tidal breathing into a chilled collection device where water vapor in exhaled breath is condensed. EBC unlike BAL, is a noninvasive, inexpensive, and easy to collect method and repeat without causing airway injury or inflammation (Kharitonov and Barnes, 2002, Mutlu et al., 2001, O’Reilly and Bailey, 2007) and has the potential for screening and diagnosis of COPD, especially during exacerbations and treatment response monitoring (Borrill et al., 2008, Holz, 2005). Several COPD biomarkers in EBC were reported, including pH (Borrill et al., 2005, Kostikas et al., 2002), prostaglandins (Antczak et al., 2012, Montuschi et al., 2003), oxidative stress markers (Dekhuijzen et al., 1996, Nowak et al., 1999), proinflammatory cytokines (Bucchioni et al., 2003, He et al., 2014, Lin et al., 2017, Liu et al., 2014), and matrix metalloproteinases (Hao et al., 2019, Mercer et al., 2005).

**Temperature.** Exhaled breath temperature (EBT), a component of exhaled breath, is regarded as a promising noninvasive diagnostic tool to evaluate respiratory and other diseases (Popov, 2011). Most studies on EBT have been performed in asthmatic patients to assess changes in the degree of airway inflammation (Paredi et al., 2002, Piacentini et al., 2002, Piacentini et al., 2007). Some studies also evaluated EBT as a potential biomarker in smoking and COPD (Carpagnano et al., 2016, Hoffmeyer et al., 2009, Lázár et al., 2014, Popov et al., 2017, Vribica et al., 2017).

The selected paper described EBT as a novel and early biomarker of COPD in the pre-symptomatic stage (Labor et al., 2016). In a two year follow up study, Labor and coworkers evaluated EBT and lung function markers in smokers with no prior diagnosis of COPD until the subjects developed COPD (GOLD 1) (Labor et al., 2016). This study demonstrated a change in EBT after smoking a cigarette at an initial visit was significantly predictive for disease progression. Carpagnano et al. (2016) showed the sensitivity of EBT to cigarette smoke and the potential to predict future development of COPD in current smokers.

**Discussion**

Several prospective biomarkers of potential harm have been described in the last decade, including molecular, cellular, structural and functional endpoints that are implicated in different stages of COPD with or without exacerbations. Clinically relevant early biomarkers of pre-symptomatic COPD would be important for early diagnosis and effective treatment. In this review, potential biomarkers for future validation and key limitations were identified from the 68 publications that met our preset criteria and were subdivided into 4 categories based on the technique and location of sample type (spirometry, imaging, blood and lung sampling).

Spirometry is widely accepted as the gold standard for assessing lung function and diagnosis of COPD. However, it has some limitations, for example an accurate result depends on collective and active participation from both the subject and the operator. In addition, the technique may lack the required sensitivity to detect the condition in asymptomatic patients, such as healthy smokers. There is also poor correlation between spirometry data and clinical outcomes such as exacerbation frequency. Even in patients diagnosed with severe airflow obstruction, not all reported symptoms and exacerbations show impaired exercise tolerance, suggesting that FEV1 does not fully capture COPD complexity. Therefore, the development and validation of alternative biomarkers that can be used in the diagnosis, prognosis and treatment of COPD are still required.

Understanding of the pathophysiological mechanisms in COPD is growing with the utility of improved imaging tools that are now more widely available. Advanced imaging techniques allow thorough anatomical, structural and functional abnormalities to be acquired in vivo. In combination, the acquired imaging data demonstrate a remarkable degree of regional variation in lung function. Different imaging modalities vary significantly in terms of temporal and spatial resolution, and each has its own advantages and disadvantages. The major limitation of ionizing radiation exposure is being overcome by new technologies, which reduce radiation dose while enhancing image quality. However, the utility of current imaging techniques is limited as both CT and MRI systems are expensive and require specialized facilities and training to conduct and interpret the results. Lung imaging as an evolving tool for the assessment of potentially reduced risk tobacco products is appealing because it allows the visualization of the primarily exposed tissue during cigarette smoking. The imaging techniques are evolving and refinements of the techniques as well as additional data from
longitudinal studies will prove useful in the future. These modalities are likely to become increasingly important in early diagnosis of COPD.

Blood biomarkers associated with the development of smoking-related lung disease have been a major area of research. However, as these biomarkers may come from other organs in the body and not necessarily the lungs, this creates a level of difficulty in interpreting their specific association with lung disease. In addition, most of these papers investigated blood biomarkers in patients with COPD with a view to identifying biomarkers that could be indicative of disease progression. Nevertheless, these studies present little evidence to establish the utility of these biomarkers in demonstrating the potential for reduced risk of lung disease if a subject switches from tobacco cigarettes to a PRRP.

Of the blood biomarkers reviewed here, only IL-8 was reported as significantly higher in smokers versus non-smokers; IL-6, sRAGE and SP-D levels showed some promise, but may require a larger sample size to differentiate between smokers and non-smokers. These biomarkers may have potential for use in PRRP switching studies, however further research is required to investigate their utility for use in this context. In addition, further understanding of the role these biomarkers may play in the pathological changes involved in the development of COPD would be required. Fibrinogen and CRP show conflicting data on comparisons between smokers and non-smokers and may be confounded by other factors such as BMI. Finally, although not a plasma biomarker, SAF warrants further investigation as the available evidence, whilst limited, shows some promise in its utility for use in potential reduced risk tobacco product studies context.

Several pathophysiological alterations at different locations of the lung were reported in all stages of COPD. The selected papers in lung sampling yielded several biomarkers that were previously well described; for example: IL-8, α1-Antitrypsin, Apolipoprotein A1, and MUC5AC. Novel biomarkers that were identified include AGTER, RAGE, SP-D and exhaled breath temperature. Interestingly, RAGE, SP-D and IL-8 in lung sampling are also found in plasma. Therefore, validation of these biomarkers would provide additional confidence. However, obtaining lung sample quality and consistency in COPD subjects is a difficult task. Development and validation of noninvasive techniques such as measuring exhaled breath condensate and temperature would be useful for local biomarker readouts.

Conclusion

Overall, in the past decade, researchers have identified many potential biomarkers of smoking related lung diseases, including COPD. The existing tools such as spirometry may not address all the current gaps, while other cutting-edge imaging tools are rapidly developing and promising for the identification of novel biomarkers. This review identified existing and new biomarkers that have potential for being predictive biomarkers in smokers and COPD subjects, as well as provide insights into disease development and progression. All these identified biomarkers require further development and fit-for-purpose assessments including validation of the biomarkers in tobacco and nicotine context and understanding their potential role in the development of disease.

Data availability

No data are associated with this article.

Acknowledgments

This publication is the work product of a cross-industry work item commissioned by the Cooperation Centre for Scientific Research Relative to Tobacco (CORESTA) Biomarker Sub-Group. CORESTA is a non-profit association, the purpose being to promote international cooperation in scientific research relative to tobacco and its derived products. We would like to express our gratitude to members of this Sub-Group who provided their diverse perspective and technical expertise during the drafting of this review.

References

Adamko DJ, Nair P, Mayers I, et al.: Metabolomic profiling of asthma and chronic obstructive pulmonary disease: A pilot study differentiating diseases. J. Allergy Clin. Immunol. 2015; 136: 571-580.e3. Publisher Full Text

Agustí A, Calverley PM, Celli B, et al.: Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respir. Res. 2010; 11: 122. PubMed Abstract | Publisher Full Text | Free Full Text

Agustí A, Edwards LD, Rennard SI, et al.: Persistent systemic inflammation is associated with poor clinical outcomes in COPD: a novel phenotype. PLoS One. 2012; 7: e37483. PubMed Abstract | Publisher Full Text | Free Full Text

Alamidi DF, Morgan AR, Hubbard Cristinacce PL, et al.: COPD Patients Have Short Lung Magnetic Resonance T1 Relaxation Time. COPD. 2016; 13: 153-159. PubMed Abstract | Publisher Full Text
Weiszhar Z, Horvath I: Induced sputum analysis: step by step. Breathe. 2013; 9: 300–306. PubMed Full Text

Wen Y, Reid DW, Zhang D, et al.: Assessment of airway inflammation using sputum, BAL, and endobronchial biopsies in current and ex-smokers with established COPD. Int. J. Chron. Obstruct. Pulmon. Dis. 2010; 5: 327–334. PubMed Abstract | Publisher Full Text | Free Full Text

Wendt CH, Nelsestuen G, Harvey S, et al.: Peptides in Bronchoalveolar Lavage in Chronic Obstructive Pulmonary Disease. PLoS One. 2016; 11: e0155724. PubMed Abstract | Publisher Full Text | Free Full Text

Winkler C, Atochina-Vasserman EN, Holz O, et al.: Comprehensive characterisation of pulmonary and serum surfactant protein D in COPD. Respir. Res. 2011; 12: 29. PubMed Abstract | Publisher Full Text | Free Full Text

Woolhouse IS, Bayley DL, Stockley RA: Effect of sputum processing with dithiothreitol on the detection of inflammatory mediators in chronic bronchitis and bronchiectasis. Thorax. 2002; 57: 667–671. PubMed Abstract | Publisher Full Text | Free Full Text

World Health Organisation: WHO highlights huge scale of tobacco-related lung disease deaths. World Health Organisation; 2019.

York TP, Van Den Oord EJ, Langston TB, et al.: High-resolution mass spectrometry proteomics for the identification of candidate plasma protein biomarkers for chronic obstructive pulmonary disease. Biomarkers. 2010; 15: 367–377. PubMed Abstract | Publisher Full Text
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com