

Protein Biomarkers for COPD Outcomes



Karina A. Serban, MD; Katherine A. Pratte, PhD; and Russell P. Bowler, MD, PhD

COPD is a clinically heterogeneous syndrome characterized by injury to airways, airspaces, and lung vasculature and usually caused by tobacco smoke and/or air pollution exposure. COPD is also independently associated with nonpulmonary comorbidities (eg, cardiovascular disease) and malignancies (eg, GI, bladder), suggesting a role for systemic injury. Since not all those with exposure develop COPD, there has been a search for plasma and lung biomarkers that confer increased cross-sectional and longitudinal risk. This search typically focuses on clinically relevant COPD outcomes such as FEV₁, FEV₁ decline, CT measurements of emphysema, or exacerbation frequency. The rapid advances in omics technology and the molecular phenotyping of COPD cohorts now permit large-scale evaluation of genetic, transcriptomic, proteomic, and metabolic biomarkers. This review focuses on protein biomarkers associated with clinically relevant COPD outcomes. The prototypic COPD protein biomarker is alpha-1 antitrypsin; however, this biomarker only accounts for 1% to 5% of COPD. This article reviews and summarizes the evidence for other validated biomarkers for each COPD outcome, and discusses their advantages, weaknesses, and required regulatory steps to move the biomarker from the bench into clinic. Although we highlight the emergence of many novel biomarkers (eg, fibrinogen, soluble receptor for advanced glycation, surfactant protein D, club cell secretory protein), there is increasing evidence that individual biomarkers only explain a fraction of the increased COPD risk and that multiple biomarker panels are needed to completely explain clinical variation and risk in individuals and populations.

CHEST 2021; 159(6):2244-2253

KEY WORDS: biomarker network; COPD; multiplex platforms; protein biomarker

The Need for Biomarkers in COPD

A typical patient with COPD has had decades of cigarette smoke exposure and does not manifest signs of disease until the latter stages of life. Furthermore, traditional measures of disease progression (eg, changes in postbronchodilator spirometry or CT emphysema) occur over many years. Unfortunately, there are currently no Food and Drug Administration (FDA)-

approved medications that reduce disease progression endpoints or mortality in clinical trials; therefore, there is an unmet need for surrogate endpoints (ie, biomarkers) that can be used to identify subjects at high risk of progression or that could serve as targets for particular subphenotypes or COPD.^{1,2} This contrasts with COPD exacerbator phenotype, which has intermittent and short-term manifestations and multiple

ABBREVIATIONS: AAT = alpha-1 antitrypsin; COPDGene = COPD Genetic Epidemiology; ECLIPSE = Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points; FDA = Food and Drug Administration; MESA-Lung study = Multi-Ethnic Study of Atherosclerosis; SPIROMICS = SubPopulations and Intermediate Outcome Measures In COPD Study; sRAGE = soluble receptor for advanced glycation

AFFILIATIONS: From the National Jewish Health (K. A. Serban, K. A. Pratte, and R. P. Bowler), Denver; and the University of Colorado (K. A. Serban and R. P. Bowler), Anschutz Medical Campus, Aurora, CO.

FUNDING/SUPPORT: This study was funded by National Institutes of Health/National Heart, Lung, and Blood Institute 1R01HL137995 and 1R01HL152735-01 to Dr Bowler.

CORRESPONDENCE TO: Karina A. Serban, MD; e-mail: serbank@njhealth.org

Copyright © 2021 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

DOI: <https://doi.org/10.1016/j.chest.2021.01.004>

approaches which have been shown to reduce exacerbations.^{3,4}

Definition of COPD Biomarkers

The FDA-NIH Biomarker Working Group's *BEST (Biomarkers, EndpointS and Tools)*⁵ describes a biomarker as a "characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. Molecular, histologic, radiographic, or physiologic characteristics are types of biomarkers. A biomarker is not an assessment of how an individual feels, functions, or survives."

Potentially useful COPD biomarkers might include those that could differentiate between the following: (1) healthy individuals vs individuals with COPD (diagnostic biomarkers), (2) active/progressing COPD vs stable COPD (disease activity biomarkers), and (3) treatment responders vs nonresponders (response to treatment biomarkers).

Molecular biomarkers come in a large variety of types and they serve mostly as diagnostic and disease activity markers. Next to proteins, other COPD biomarkers include the following: (1) genetic biomarkers which are linked to candidate loci and single nucleotide polymorphisms in COPD risk genes (eg, *SERPINA1* identifies individuals with alpha-1 deficiency-related COPD, chromosome 15q25: *CHRNA3/CHRNA5/IREB2* associated with smoking behavior and COPD, independent of smoking; chromosome 19q13 linked to increased susceptibility to smoking, or *PTPN6* linked to early onset familial emphysema)⁶⁻⁸; (2) epigenomic biomarkers to identify the methylation status of risk genes (eg, *HDAC6*, *PTEN*, *Nrf2*)⁹⁻¹¹ frequently associated with the clinical response to corticosteroids in individuals with COPD; (3) transcriptomic biomarkers to identify profiles of differentially expressed genes (eg, *ASAHI*, *CEBPD*, *FOXP1*, *TCF7*)⁹⁻¹¹ that are associated with low FEV₁ and/or FEV₁/FVC ratio; (4) proteomic biomarkers to identify directional changes of relevant systemic/lung-specific proteins (eg, fibrinogen, soluble receptor for advanced glycation [sRAGE])^{12,13} linked to particular COPD endotypes (eg, fibrinogen to COPD with frequent exacerbations, sRAGE to emphysema-predominant COPD); (5) metabolomic biomarkers such as sphingolipids,¹⁴ which have been linked to COPD exacerbations; and (6) microbiome signatures to identify bacterial diversity and communities (eg, *Firmicutes* phylum)¹⁵ present specifically in the airway and distal lung of individuals with COPD or sputum (eg, *Veillonella* species, *Staphylococcus* species) associated

with 1-year survival and mortality, respectively, after hospitalization for COPD exacerbation.¹⁶

In this review we discuss the protein biomarkers that are associated with COPD clinical phenotypes such as those determined by spirometry (airflow obstruction assessed by FEV₁, FEV₁/FVC ratio), imaging (% emphysema, lung densitometry as estimated by 15th percentile density, or bronchial wall thickness of the inner perimeter of 10-mm airway on CT scan),¹⁷⁻²¹ or frequency of COPD exacerbations.

Advantages and Disadvantages of Protein Biomarkers Compared With Other Types of Molecular Biomarkers

The first biomarker, alpha-1 antitrypsin was discovered through serum protein electrophoresis in 1963.²² More recently, there has been a focus on genetics with the discovery of > 125 single nucleotide polymorphisms in *SERPINA1* that have been associated with alpha-1 deficiency.²³ Genetic research approaches are the first to come to mind for COPD personalized medicine and have yielded many novel COPD candidate genes beyond *SERPINA1*²⁴; however, one's genotype is static throughout life and does not change with smoking or therapeutic interventions.^{2,12} Epigenetic and transcript approaches are closely related to the genome and have more potential as COPD biomarkers because they can change with time; however, most currently approved therapies in COPD and other diseases do not target the epigenome or transcriptome. Rather, they target proteins or enzymes and their products (eg, metabolites).^{12,25} Metabolites such as sphingolipids and leukotrienes could be excellent COPD biomarkers; nevertheless, there are currently few large-scale, replicated metabolomics studies in COPD.^{26,27} Technology has favored the search for protein biomarkers in COPD because of the latest development of large-scale protein arrays that can assay thousands of proteins simultaneously, cheaply, and quickly, and using low sample volume.^{28,29}

Major considerations for COPD biomarkers include the following: sample type, assays method, statistical analysis strategy, replication across similar populations in different studies, and whether a biomarker is likely to achieve regulatory approval (Fig 1). Some popular sample matrices for COPD biomarkers include plasma³⁰ or urine,³¹ but also BAL fluid,^{32,33} sputum,³⁴ exhaled breath condensate,³⁵ and lung tissue.^{36,37} For many discovery studies, there is the potential to assay multiple proteins with different detection methods (antibodies,

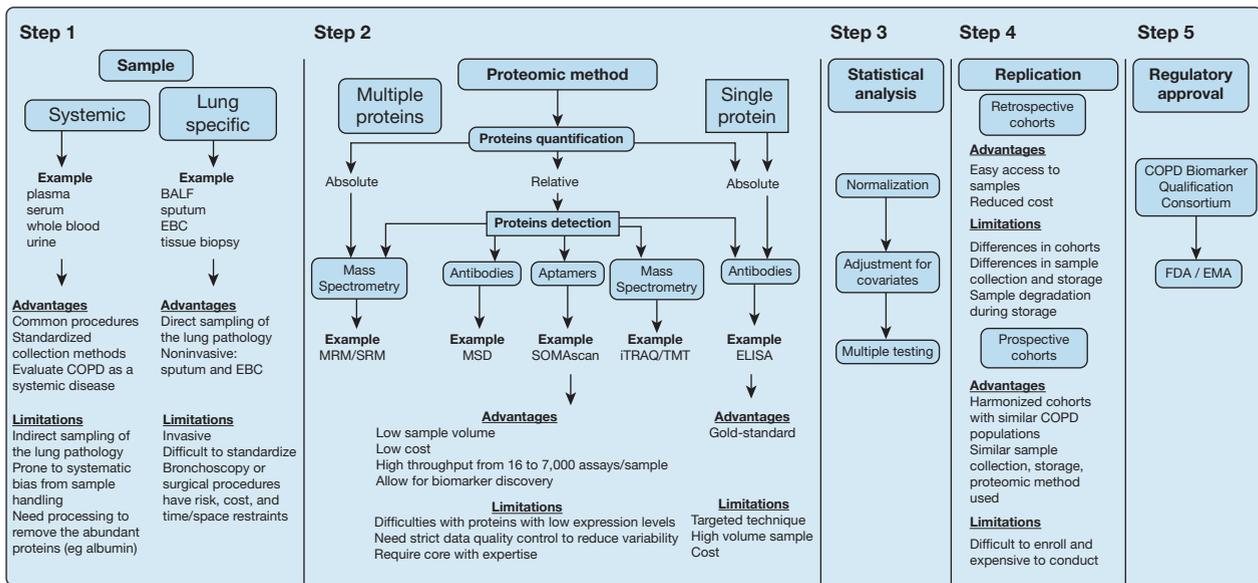


Figure 1 – Approach to judging a proteomic study. Brief description of the five steps to be considered when judging a well-designed proteomic study. Step 1: evaluation of samples used in the proteomic study with the advantages and limitations of systemic vs lung-specific samples. Step 2: evaluation of the proteomic method used in the study, including the platform (single vs multiplex), detection method (antibodies, aptamers, or mass spectrometry), and quantification (relative or absolute) with their advantages and limitations vs ELISA, the current criterion standard method. Step 3: evaluation of the statistical methods used with specific considerations for data quality control, normalization, and adjustment for covariates. Step 4: evaluation of the cohorts used for protein biomarker validation with advantages and limitations of retrospective vs prospective cohorts. Step 5: evaluation by the regulatory agencies involved in biomarker approval. BALF = BAL fluid; EBC = exhaled breath condensate; ELISA = enzyme-linked immunosorbent assay; EMA = European Medicines Agency; FDA = Food and Drug Administration; iTRAQ/TMT = isobaric tag for relative and absolute quantitation/tandem mass tags; MRM/SRM = multiple reaction monitoring/selective reaction monitoring; MSD = Meso Scale Discovery; SOMAscan = slow off-rate modified aptamers.

aptamers, or mass spectrometry) and quantification (relative or absolute). These technologic advances come with a major limitation that as the number of biomarkers (P) grows, it may significantly exceed the number of samples (N), which can lead to false discovery in small studies (where $P \gg N$). Hence, we rely on statistical analysis to ensure data quality control and normalization to adjust for covariates and to apply multiple testing methods. Replication in retrospective or prospective cohorts helps biomarker validation and the regulatory approval steps. Regulatory approval typically focuses on consistency of existing data and adding assay parameters such as upper and lower limits of detection and coefficients of variation.

Evidence for COPD Protein Biomarkers

Studies published in the last 10 years, heavily relying on plasma samples, have identified many candidate protein biomarkers. From our PubMed search performed in May 2020, we selected studies listing the terms “COPD” and “proteomic” in the key word section. These studies used a plethora of different proteomic platforms and biostatistical methods. Therefore, we decided to focus on plasma protein biomarkers that have been tested in

studies with large numbers of patients (~100 patients) and that have been replicated, defined as identifying proteins with a similar proteomic platform, in at least two independent cohorts. It is notable that many plasma protein biomarkers (Table 1) behave differently as a cross-sectional vs longitudinal biomarker. Additionally, many biomarkers are associated with multiple clinically relevant outcomes in COPD: FEV₁, FEV₁/FVC ratio, FEV₁ decline, diffusing capacity for carbon monoxide (DLCO), emphysema distribution, and CT measurements of emphysema (lung densitometry as estimated by the 15th percentile density) or exacerbation frequency. Inflammatory biomarkers (C-reactive protein, fibrinogen, IL-6, IL-8, and monocyte chemoattractant protein-1), growth factors (vascular endothelial growth factor, epidermal growth factor receptor, apolipoprotein A1, hepatocyte growth factor, and bactericidal/permeability increasing fold-containing family B member 1), proteases/antiproteases (matrix metalloproteinase 8, matrix metalloproteinase 9, myeloperoxidase, tissue inhibitor of metalloproteinases 1, and alpha-1 antitrypsin), and lung-derived proteins (surfactant protein D and club cell secretory protein) correlate positively with clinically relevant outcomes. Antiinflammatory molecules (eg, sRAGE) tend to be

inversely correlated with emphysema and airflow limitation. The most promising plasma biomarkers have been validated, defined as identifying proteins using different proteomic platforms in multiple cohorts. In addition, these biomarkers were associated not only with cross-sectional clinically relevant outcomes, but also with longitudinal outcomes (FEV₁ decline, emphysema progression, and future exacerbations). The few biomarkers (apolipoprotein A1 and tissue inhibitor of metalloproteinases 1) found in the sputum have yet to be validated. To date, no biomarker has been validated in sputum and plasma to predict a COPD clinical outcome or disease progression. Some biomarkers have been associated with COPD diagnosis, severity, and progression; nevertheless, there is not one validated biomarker that is associated with all clinical outcomes and disease progression. Traditionally, the biomarkers associated with clinical outcomes pertinent to lung function in cross-sectional studies have been considered to reflect COPD severity (eg, sRAGE, emphysema severity); however, emerging data from longitudinal cohorts suggest that the same biomarker might reflect COPD activity considering that it was associated with emphysema progression (Table 2).

Judging Protein Biomarker Evidence

The sine qua non criterion for establishing biomarker validity is replication of associations across multiple independent populations. Similar to genetics, early biomarker studies were plagued by a lack of replication. Fortunately, major funding agencies have made coordinated efforts to build and maintain large COPD cohorts: Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE),⁵³ SubPopulations and Intermediate Outcome Measures In COPD Study (SPIROMICS), COPD Genetic Epidemiology (COPDGene),⁵⁴ Treatment Of Emphysema With A Selective Retinoid Agonist (TESRA),⁵⁵ and population cohorts with good lung phenotyping including the Framingham Heart Study (FHS), Multi-Ethnic Study of Atherosclerosis (MESA-Lung study), Cooperative Health Research in the Augsburg Region, (KORA), or those tailored toward specific COPD phenotypes (emphysema secondary to alpha-1 antitrypsin [AAT] deficiency), such as QUANTitative lung CT UnMasking emphysema progression in AATD (QUANTUM) and Genomic Research in Alpha-1 Antitrypsin Deficiency and Sarcoidosis (GRADS).⁵⁶⁻⁵⁸ The latest efforts aim to

harmonize testing platforms and data analysis to validate promising biomarkers across multiple cohorts (Table 2).⁴⁶

Other useful but not essential features that make a protein biomarker valuable include validation of the assay across multiple platforms using well-established standards. For clinical use, this may also include validating the range of the assay across concentrations that are likely to be encountered in real human populations. Stability of the assay measurements under different conditions (eg, time, method of collection, processing times) can affect assay stability and utility in real-world settings.^{59,60} Taking a long time to aliquot and rapid freezing of a sample may result in degradation of selected proteins. For example, RBC lysis from contact with ice may result in increased release of intracellular proteins, or at low centrifugation speeds, platelets and their contents may not separate from plasma.⁶¹ Less commonly, appreciated factors that can introduce bias are fasting time and time of day of sampling.⁶² There are also biological factors that can affect proteins such as exercise, age, sex, and sleep.⁶³⁻⁶⁵ All these features are important for biomarker discovery and can reduce the biomarker clinical utility by making it hard to perform consistent assays, but robust assays with the most clinical utility should not be greatly affected by these factors.

Other helpful but not essential features which improve the suitability of a biomarker include evidence for a mechanistic role (eg, AAT),^{27,57} association with disease progression or activity (eg, fibrinogen for exacerbations),⁶⁶ and a known minimal clinically important difference (eg, changes of 4 units in the St. George's Respiratory Questionnaire to detect a clinically significant change in the clinical outcome).⁶⁷ Biomarkers of COPD progression with supporting mechanistic evidence are rare for COPD because progression typically occurs over decades; however, examples of biomarkers that are associated with disease progression include serum AAT level.⁶⁸

An unmet need in the COPD biomarker field is discovery and validation of biomarkers associated with response to treatment. In monogenetic diseases such as AAT deficiency-serum levels of desmosine and isodesmosine, two markers of elastin degradation associated with clinical (FEV₁) COPD outcomes, and morphologic (emphysema CT measurements) COPD outcomes, have been modified by weekly administration of augmentation therapy, intervention recognized to

TABLE 1] Lung and Plasma Proteomics Identify Candidate Cross-Sectional and Longitudinal Biomarkers

Phenotype	Clinical Parameter	Biomarker	
Cross-sectional			
Lung function	Case/control	<ul style="list-style-type: none"> • 13 proteins (TIMP1, APOA1, AHSG, C6orf58, KRT19, ATIII, ALB, HRG, TF, BPIFB1, CNDP1)^{38,a} • Apolipoprotein A1, Lipocalin-1^{39,a} • 2-macroglobulin, haptoglobin, ceruloplasmin, hemopexin^{40,b,c} • 43 proteins (eotaxin-2, MCP-1, IL-8, IFNγ, TNFα, VEGF, MMP9, TIMP1, PAI-II)^{41,b} • 21 proteins (CRP, eotaxin-2, IL-10, tenascin)^{30,b} • CRP^{42,b} • B defensin-2, CCL2, CCL18, CRP, CXCL10, fibrinogen, HGF, IL-8, IL-12, MMP8, MMP9, MPO^{43,b} • 19 proteins (cathepsin B, ATP synthase, chaperonin)^{44,a} 	
	FEV ₁	<ul style="list-style-type: none"> • GRP78, sCD163^{45,b,c} • 13 proteins (TIMP1, APOA1, AHSG, C6orf58, KRT19, ATIII, ALB, HRG, TF, BPIFB1, CNDP1)^{38,a} • 25 proteins (eotaxin-2, MCP-1, IL-8, IFNγ, TNFα, VEGF, MMP9, TIMP1, PAI-II, prolactin)^{41,b} • IL-6, P-selectin, CRP, ICAM-1^{42,b} • sRAGE, CC16, SP-D, CRP, fibrinogen^{46,b,c} • Eotaxin, IL-10, IL-2, IL-6, IL-8, TNFα^{47,b,c} • 26 proteins (APOA1, C5a, ATIII)^{48,b} 	
	FEV ₁ /FVC	<ul style="list-style-type: none"> • Lipocalin-1^{39,a} • sRAGE,^{46,b} eotaxin^{47,b} 	
	Dl _{CO}	<ul style="list-style-type: none"> • 13 proteins (TIMP1, APOA1, C6orf58, KRT19, ATIII, ALB, APOA1, HRG, TF, BPIFB1, CNDP1)^{38,a} • 25 proteins (eotaxin-2, MCP-1, IL-8, IFNγ, TNFα, VEGF, MMP9, TIMP1, PAI-II, prolactin)^{41,b} • 9 proteins (sRAGE, MMP10, EGFR, MDC, IL-12P40, APOA1, TS1, fibrinogen)^{13,b} 	
	GOLD stages	<ul style="list-style-type: none"> • GRP78, sCD163, IL-1AP, MSTP9^{45,b,c} • Apolipoprotein A1, lipocalin-1^{39,a} • sRAGE^{13,b} • b-defensin-2, IL-8, MMP8, MMP9, adiponectin, HGF^{30,b} 	
	Imaging	Emphysema	<ul style="list-style-type: none"> • 24 proteins (eotaxin-2, MCP-1, VEGF, TIMP1, PAI-II)^{41,b} • 21 proteins (CRP, eotaxin-2, IL-10, tenascin)^{43,b} • RAGE, ICAM-1, CCL20, SERPINA7^{49,b,c}
		LAA-950	<ul style="list-style-type: none"> • GRP78, sCD163^{45,b,c} • sRAGE^{49,50,b,c} • RAGE, ICAM-1, CCL20, SERPINA7^{49,b,c} • MMP9^{43,b}
PD15		<ul style="list-style-type: none"> • 11 proteins (sRAGE, ICAM-1, EGFR, IL-2RA, IL-1A, IL-7, IL-16, IL-3, MIP3A, TS1, IL-2)^{13,b} • sRAGE, SP-D, CCL18^{50,b} • Adiponectin, CXCL10, SOD1, FAS, SP-D, CCL20^{49,b,c} 	
Clinical phenotypes	Pi10	<ul style="list-style-type: none"> • sRAGE^{13,b,c} 	
	Frequent exacerbations	<ul style="list-style-type: none"> • SAA^{51,b} • 25 proteins (eotaxin-2, MCP-1, IL-8, IFNγ, TNFα, VEGF, MMP9, TIMP1, PAI-II, prolactin)^{41,b} • Fibrinogen^{43,b} 	
Longitudinal			
Lung function	FEV ₁ decline	<ul style="list-style-type: none"> • sRAGE, SP-D, fibrinogen, IL-6, CRP^{46,b,c} 	
Imaging	Emphysema progression	<ul style="list-style-type: none"> • sRAGE, SP-D, fibrinogen, IL-6, IL-8, CRP^{46,47,50,b,c} 	
Clinical phenotypes	Future exacerbations	<ul style="list-style-type: none"> • sRAGE, CC16, SP-D, CRP, fibrinogen^{46,b,c} • Decorin, α2 macroglobulin, AAT, ferritin, CXCL10, CXCL9, CCL11^{52,b,c} 	

TABLE 2] Protein Biomarker Associated With Cross-sectional and Longitudinal COPD Outcomes

Biomarker	Lung Function			Imaging		Exacerbations	
	FEV ₁ or FEV ₁ /FVC	Decline FEV ₁	D _{lco}	Emphysema	Emphysema Progression	Frequent and Previous	Future
sRAGE	↑ ^b	↑ ^b	↓ ^a	↓ ^b	↓ ^b	↑ ^b	↑ ^a
IL-6	↑ ^b	↑ ^a	NS	↑ ^a	↑ ^a	NS	NS
IL-8	↑ ^b	NS	↑ ^a	NS	↑ ^a	↑ ^b	NS
Fibrinogen	↓ ^b	↑ ^b	NS	↑ ^b	↓ ^b	↑ ^b	↑ ^a
CRP	↓ ^b	↑ ^a	NS	↑ ^b	↑ ^a	↑ ^b	↑ ^a
SPD	↓ ^b	↑ ^a	NS	↓ ^b	↑ ^a	↑ ^b	NS
CC16	↑ ^b	↓ ^b	NS	↑ ^a	↓ ^a	NS	↑ ^a
APOA1	↑ ^a	NS	↑ ^a	NS	NS	NS	NS

↑ = positive biomarker - clinical outcome association; ↓ = negative biomarker - clinical outcome association; APOA1 = apolipoprotein A1; CC16 = club cell secretory protein; CRP = C-reactive protein; D_{lco} = diffusing capacity for carbon monoxide; NS = not significant biomarker - clinical outcome association; SPD = surfactant protein D; sRAGE = soluble receptor for advanced glycation.

^aNo replication.

^bReplication in two cohorts.

modify COPD progression in individuals with AAT deficiency.⁶⁹ In smoking-related COPD, this task is not readily explored. First, there are few disease-modifying therapies in COPD to study their effect on biomarkers identified to correlate with COPD severity and activity. Second, individuals enrolled in COPD cohorts, both in the control or diseased arms, may be on non-COPD therapies (eg, statins, thiazolidinediones, angiotensin converting enzyme inhibitors, angiotensin II type 1 receptor antagonists) that have been shown to increase serum levels of COPD relevant biomarkers (eg, sRAGE).⁷⁰ Stringent inclusion criteria and biostatistical methods can estimate the additive effect of these therapies to the biomarker variability.

Final considerations for a biomarker should include the target population. For the COPD population, these may include older age and current or former smokers, but also features such as sex and race, both of which are strongly associated with differential expression of blood biomarkers.⁶³ Clear delineation of the population used to assess biomarker performance is essential when

interpreting a biomarker test, but is often not considered (eg, D-dimer performs better in the diagnosis of pulmonary embolism in a high pretest probability population⁷¹ rather than all patients evaluated for chest pain).

Why Do Protein Biomarkers Fail to Validate Across All Cohorts?

Many of the discovery biomarkers presented in Table 1 may not be generalizable to all COPD populations because of inherent differences between the cohorts. Inflammatory (C-reactive protein, fibrinogen, IL-6, and IL-8) and distal lung injury (sRAGE, surfactant protein D, club cell secretory protein, and apolipoprotein A1) biomarkers associated with clinical outcomes across multiple cohorts are shown in Table 2.^{13,46,47} The difficulties in replicating some biomarkers may be related to differences in the individuals recruited in these cohorts.

Even with a balanced number of men and women recruited in the cohorts, we are able to pin out protein biomarkers that might be sex-specific (eg, IL-16, VEGF).⁶³

AAT = alpha-1 antitrypsin; AHSG = alpha 2-HS glycoprotein; ALB = albumin; APOA1 = apolipoprotein A1; ATIII = antithrombin III; ATP = adenosine triphosphate; BPIFB1 = bactericidal/permeability-increasing-fold-containing family B member 1; C5a = complement C5a protein; C6orf58 = chromosome 6 open reading frame 58; CC16 = club cell secretory protein; CCL2 = C-C motif chemokine ligand 2; CCL11 = C-C motif chemokine ligand 11; CCL18 = C-C motif chemokine ligand 18; CCL20 = C-C motif chemokine ligand 20; CNDP1 = carnosine dipeptidase 1; CRP = C-reactive protein; CXCL9 = C-X-C motif chemokine ligand 9; CXCL10 = C-X-C motif chemokine ligand 10; D_{lco} = diffusing capacity for carbon monoxide; EGFR = epidermal growth factor receptor; FAS = Fas cell surface death receptor; GOLD = Global Initiative for Chronic Obstructive Lung Disease; GRP78 = 78-kDa glucose-regulated protein; HGF = hepatocyte growth factor; HRG = histidine-rich glycoprotein; ICAM-1 = intercellular adhesion molecule 1; IFN γ = interferon gamma; KRT19 = keratin 19; LAA-950 = low attenuation areas less than a threshold of -950 Hounsfield units; MCP-1 = monocyte chemoattractant protein-1; MDC = macrophage-derived chemokine; MIP3A = macrophage inflammatory protein 3 alpha; MMP8 = matrix metalloproteinase 8; MMP9 = matrix metalloproteinase 9; MMP10 = matrix metalloproteinase 10; MPO = myeloperoxidase; MSTP9 = putative macrophage-stimulating protein; PAI-II = plasminogen activator inhibitor-2; PD15 = 15th percentile lung density; Pi10 = inner perimeter of 10-mm airway; RAGE = receptor for advanced glycation endproducts; SAA = serum amyloid A; sCD163 = cluster of differentiation 163 protein; SERPINA7 = serpin family A member 7; SOD1 = superoxide dismutase; SP-D = surfactant protein D; sRAGE = soluble receptor for advanced glycation endproducts; TF = transferrin; TIMP1 = tissue inhibitor of metalloproteinases 1; TNF α = tumor necrosis factor alpha; TS1 = thymidylate synthase; VEGF = vascular endothelial growth factor.

^aLung proteomics.

^bPlasma proteomics.

^cValidated using at least two methods in two cohorts.

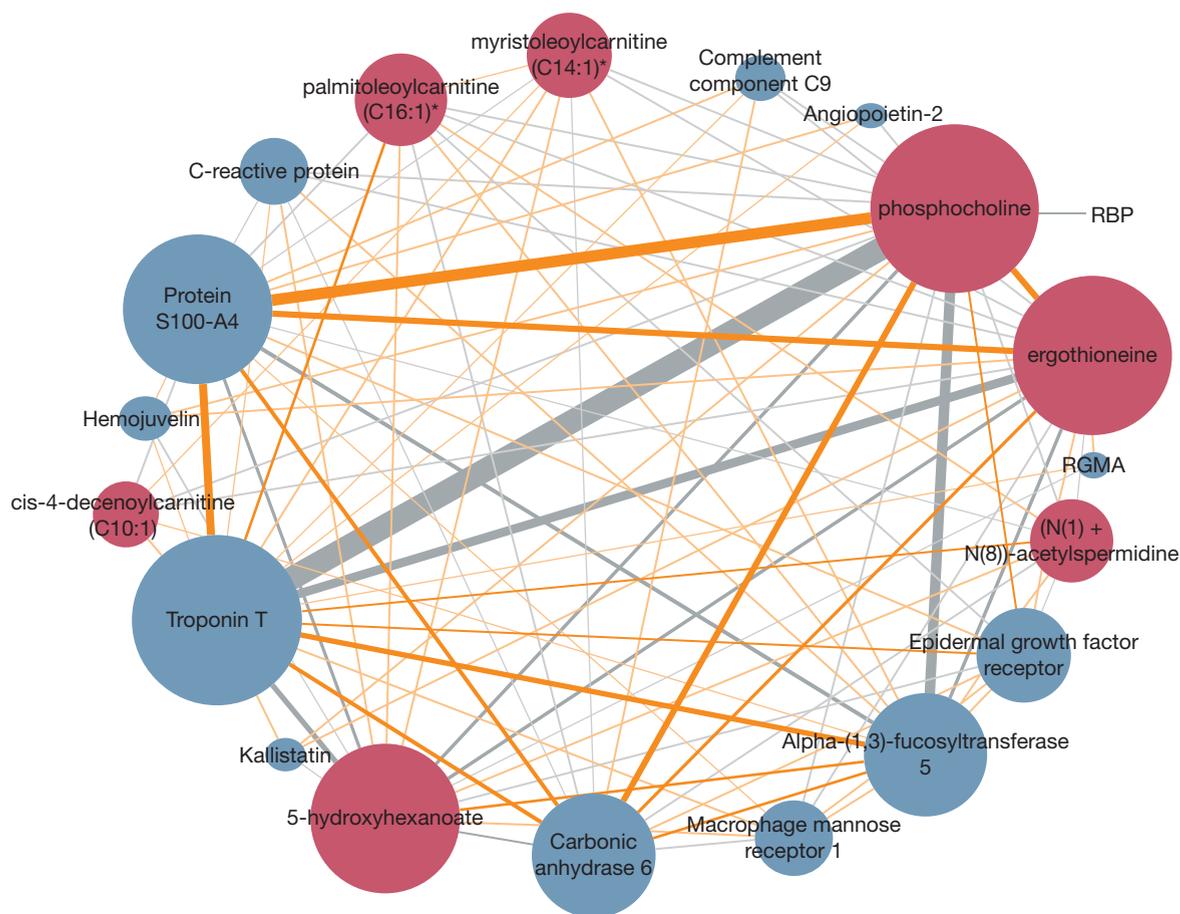


Figure 2 – Protein-metabolite network associated with FEV₁. The network was identified after applying sparse multiple canonical correlation network (SmCCNet) to adjust proteomic and metabolomic data from the blood of 1,008 participants in the COPD Genetic Epidemiology study to study omics data association with FEV₁. Edges thickness represents the level of association between metabolite-protein pairs relative to percent FEV₁. The size of the network hubs and the nodes correspond to the highest connectivity (ie, number of edges connected to the node). The 13 proteins (blue) and seven metabolites (red) included in the network were individually associated with percent FEV₁, but within the network the pair phosphocholine-troponin T has the highest pairwise correlation with percent FEV₁, suggestive of a strong negative (gray) link between systemic inflammation (phosphocholine)-heart muscle strain (troponin T)-percent FEV₁. Gray edges indicate negative correlation between the nodes and orange edges indicate positive correlation between the nodes. RBP = retinol-binding protein; RGMA = repulsive guidance molecule A. (Reprinted with permission from Mastej et al.⁷⁵)

Because many cohorts are enriched for white, European-descent individuals, fewer publications have reported on racially diversified COPD cohorts. COPDGene included approximately 33.5% non-Hispanic African Americans, but the plasma biomarkers identified within this subgroup have yet to be replicated and validated in other cohorts with admixed backgrounds, such as SPIROMICS, MESA-Lung study, or Jackson Heart Study (JHS).⁴⁸ Plasma endothelin-1 surfaced as a possible biomarker for heart failure and mortality in African Americans enrolled in the racially diverse Jackson Heart Study (JHS) cohort.⁷²

SPIROMICS and COPDGene included smokers and COPD of all severity, with fewer healthy nonsmoker individuals; ECLIPSE included primarily white,

European ancestry individuals with more severe COPD and rather fewer active smokers or patients with mild stage 1 COPD. Neither COPDGene nor ECLIPSE are population-based studies, but the Framingham Heart Study (FHS) and MESA-Lung study have been designed as population-based studies; their caveat is that they are not enriched for the population at risk for COPD and the effect of smoking or COPD development on any biomarker in these cohorts is skewed by the variability of the biomarker in the nondiseased subjects because of age, sex, race, and sample biases. The positive predictive value of a biomarker drops in a population whose disease prevalence is low (eg, general population) vs a population enriched for COPD (eg, patients in the pulmonary clinic). Ideally, matching individuals enrolled for age, sex, race, and baseline lung function

may overcome this conundrum, but this is difficult to achieve in longitudinal COPD studies where inclusion and follow-up of healthy, nonsmoker individuals are hindered by individuals' cooperation, age-related storage material degradation, and costs usually underrecognized by the funding agencies.

Moving a Candidate Biomarker From Research to Clinical Practice

The COPD biomarkers described in [Table 2](#) have made it far on the roadmap for novel biomarkers regulatory approval. Importantly, the COPD Biomarker Qualification Consortium, after reviewing the strong preclinical and clinical data, and the validation studies in multiple cohorts deemed plasma fibrinogen appropriate for FDA evaluation and approval as a biomarker to identify and predict the risk of future COPD exacerbations. Plasma fibrinogen benefits from a reproducible and widely available detection method; the availability of the testing method outside big research centers where most of the validation studies have been conducted is one big technical challenge. Subsequently, the regulatory agencies FDA and European Medicines Agency agreed that when plasma fibrinogen passes the last two relevant hurdles, interventional and prospectively designed studies, it will be ready for clinical use. The FDA/European Medicines Agency approved plasma fibrinogen as a biomarker of high risk for COPD exacerbation and all-cause mortality in COPD. Plasma fibrinogen dossier initiation, consultation, review, and approval stages took approximately 4 years.⁶⁶ Next on the approvable biomarkers list are sRAGE⁴⁶ and blood eosinophil count.

The Future of COPD Biomarkers: Where Do We Go From Here?

Phenotypes and Progression

Although there are now several encouraging biomarkers (eg, AAT, fibrinogen, sRAGE), there still remains significant gaps in our knowledge. For instance, there are few large studies that have investigated and replicated biomarkers of chronic bronchitis or disease progression, comorbidities, and death from COPD. These biomarkers will be crucial for identifying subjects who are at high risk and may benefit from an intervention. Furthermore, there are also no large studies which identify biomarkers

of treatment response. These biomarkers will be essential in evaluating disease-modifying interventions because traditional metrics (FEV₁, mortality, or emphysema progression) typically change over decades rather than months.

Single Biomarkers vs Biomarker Scores

Current evidence suggests that single biomarkers typically explain < 10% of the risk of COPD phenotypes.^{27,46} Similar to genetic risk scores, in which multiple genetic markers are weighted for a single total score, evidence suggests that multiple biomarker-derived scores significantly explain > 10% of risk. Examples of a multiple biomarker panel might be club cell secretory protein, fibrinogen, sRAGE, C-reactive protein, and surfactant protein D for FEV₁.⁴⁶ As with genetic risk scores, the disadvantage of protein risk scores is that they are likely to be sex-, phenotype-, and population-specific.

Integration of Protein Biomarkers With Other Omics

In biological systems, proteins do not act independently of other proteins or other classes of molecules (eg, nucleic acids, metabolites). For instance, we know that expression or measurement of most proteins are at least partially associated with genotype.^{73,74} Therefore, it is reasonable to assume that the best predictive models will include knowledge of multiple classes of molecules. Additional multiple proteins can be represented as networks and these networks can be integrated with other omics networks to better understand COPD. For instance, by integrating proteomics networks with metabolomics networks, we have been able to deconstruct several pathways that are associated with distinct COPD clinical phenotypes: oxidative phosphorylation with emphysema progression, antigen processing and presentation with exacerbation frequency, and glycerophospholipid metabolism with FEV₁ and FEV₁/FVC ratio ([Fig 2](#)).⁷⁵ A network-based interactome facilitates a better understanding of the interplay between interconnected causative factors in a manner compounding the effects of any one factor. The interactome can lead to the identification of key hub molecules that might not be even evaluated by discovery platforms.

We are optimistic that harmonization of standard-of-practice procedures, storage methods, and data

integration will pave the way for promising protein biomarkers and networks to become this decade's contribution to personalized medicine in COPD.

Acknowledgments

Author contributions: K. A. S., K. A. P., and R. P. B. contributed substantially to the data collection and interpretation and the writing of the manuscript. K. A. S. and R. P. B. are responsible for the content of this manuscript. The manuscript has not, either in part or whole, presented or been reviewed before accepted for publication in *CHEST*.

Financial/nonfinancial disclosures: None declared.

Role of sponsors: The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

Other contributions: We thank Dr Maneesh Bhargava, MD, PhD (University of Minnesota) for his insightful comments.

References

- Han MK, Martinez CH, Au DH, et al. Meeting the challenge of COPD care delivery in the USA: a multiprovider perspective. *Lancet Respir Med*. 2016;4(6):473-526.
- Regan EA, Hersh CP, Castaldi PJ, et al. Omics and the search for blood biomarkers in chronic obstructive pulmonary disease. Insights from COPDGene. *Am J Respir Cell Mol Biol*. 2019;61(2):143-149.
- Wedzicha JA, Brill SE, Allinson JP, Donaldson GC. Mechanisms and impact of the frequent exacerbator phenotype in chronic obstructive pulmonary disease. *BMC Med*. 2013;11:181.
- Wedzicha JA, Rabe KF, Martinez FJ, et al. Efficacy of roflumilast in the COPD frequent exacerbator phenotype. *Chest*. 2013;143(5):1302-1311.
- FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource*. MD: Silver Spring; 2016.
- Hobbs BD, Hersh CP. Integrative genomics of chronic obstructive pulmonary disease. *Biochem Biophys Res Commun*. 2014;452(2):276-286.
- Sakornsakolpat P, Prokopenko D, Lamontagne M, et al. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet*. 2019;51(3):494-505.
- Bosse Y, Lamontagne M, Gaudreault N, et al. Early-onset emphysema in a large French-Canadian family: a genetic investigation. *Lancet Respir Med*. 2019;7(5):427-436.
- Barnes PJ. Reduced histone deacetylase in COPD: clinical implications. *Chest*. 2006;129(1):151-155.
- Lam HC, Cloonan SM, Bhashyam AR, et al. Histone deacetylase 6-mediated selective autophagy regulates COPD-associated cilia dysfunction. *J Clin Invest*. 2013;123(12):5212-5230.
- Vucic EA, Chari R, Thu KL, et al. DNA methylation is globally disrupted and associated with expression changes in chronic obstructive pulmonary disease small airways. *Am J Respir Cell Mol Biol*. 2014;50(5):912-922.
- Bowler RP, Wendt CH, Fessler MB, et al. New strategies and challenges in lung proteomics and metabolomics. An Official American Thoracic Society Workshop Report. *Ann Am Thorac Soc*. 2017;14(12):1721-1743.
- Cheng DT, Kim DK, Cockayne DA, et al. Systemic soluble receptor for advanced glycation endproducts is a biomarker of emphysema and associated with AGER genetic variants in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2013;188(8):948-957.
- Cruikshank-Quinn CI, Jacobson S, Hughes G, et al. Metabolomics and transcriptomics pathway approach reveals outcome-specific perturbations in COPD. *Sci Rep*. 2018;8(1):17132.
- Martinez FJ, Erb-Downward JR, Huffnagle GB. Significance of the microbiome in chronic obstructive pulmonary disease. *Ann Am Thorac Soc*. 2013;10(suppl):S170-S179.
- Leitao Filho FS, Alotaibi NM, Ngan D, et al. Sputum microbiome is associated with 1-year mortality after chronic obstructive pulmonary disease hospitalizations. *Am J Respir Crit Care Med*. 2019;199(10):1205-1213.
- Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med*. 2004;350(10):1005-1012.
- Meguro M, Barley EA, Spencer S, Jones PW. Development and validation of an improved, COPD-specific version of the St. George Respiratory Questionnaire. *Chest*. 2007;132(2):456-463.
- Troosters T, Vilaro J, Rabinovich R, et al. Physiological responses to the 6-min walk test in patients with chronic obstructive pulmonary disease. *Eur Respir J*. 2002;20(3):564-569.
- Schroeder JD, McKenzie AS, Zach JA, et al. Relationships between airflow obstruction and quantitative CT measurements of emphysema, air trapping, and airways in subjects with and without chronic obstructive pulmonary disease. *AJR Am J Roentgenol*. 2013;201(3):W460-W470.
- Celli BR, MacNee W, Force AET. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J*. 2004;23(6):932-946.
- Laurell CB, Eriksson S. Hypo-alpha-1-antitrypsinemia [in German]. *Verh Dtsch Ges Inn Med*. 1964;70:537-539.
- DeMeo DL, Silverman EK. Alpha1-antitrypsin deficiency. 2: genetic aspects of alpha(1)-antitrypsin deficiency: phenotypes and genetic modifiers of emphysema risk. *Thorax*. 2004;59(3):259-264.
- Shrine N, Guyatt AL, Erzurumluoglu AM, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet*. 2019;51(3):481-493.
- Agusti A. The path to personalised medicine in COPD. *Thorax*. 2014;69(9):857-864.
- Kan M, Shumyatcher M, Himes BE. Using omics approaches to understand pulmonary diseases. *Respir Res*. 2017;18(1):149.
- Stockley RA, Halpin DMG, Celli BR, Singh D. Chronic obstructive pulmonary disease biomarkers and their interpretation. *Am J Respir Crit Care Med*. 2019;199(10):1195-1204.
- Raffield LM, Dang H, Pratte KA, et al. Comparison of proteomic assessment methods in multiple cohort studies. *Proteomics*. 2020;20(12):e1900278.
- Candia J, Cheung F, Kotliarov Y, et al. Assessment of variability in the SOMAScan assay. *Sci Rep*. 2017;7(1):14248.
- Zarei S, Mirtar A, Morrow JD, Castaldi PJ, Belloni P, Hersh CP. Subtyping chronic obstructive pulmonary disease using peripheral blood proteomics. *Chronic Obstr Pulm Dis*. 2017;4(2):97-108.
- Oelsner EC, Balte PP, Grams ME, et al. Albuminuria, lung function decline, and risk of incident chronic obstructive pulmonary disease. The NHLBI Pooled Cohorts Study. *Am J Respir Crit Care Med*. 2019;199(3):321-332.
- Ropcke S, Holz O, Lauer G, et al. Repeatability of and relationship between potential COPD biomarkers in bronchoalveolar lavage, bronchial biopsies, serum, and induced sputum. *PLoS One*. 2012;7(10):e46207.
- Leiten EO, Eagan TML, Martinsen EMH, et al. Complications and discomfort after research bronchoscopy in the MicroCOPD study. *BMJ Open Respir Res*. 2020;7(1):e000449.
- Baraniuk JN, Casado B, Pannell LK, et al. Protein networks in induced sputum from smokers and COPD patients. *Int J Chron Obstruct Pulmon Dis*. 2015;10(18):1957-1975.
- Franciosi L, Govorukhina N, Fusetti F, et al. Proteomic analysis of human epithelial lining fluid by microfluidics-based nanoLC-MS/MS: a feasibility study. *Electrophoresis*. 2013;34(18):2683-2694.
- Lee EJ, In KH, Kim JH, et al. Proteomic analysis in lung tissue of smokers and COPD patients. *Chest*. 2009;135(2):344-352.
- Steiling K, Kadar AY, Bergerat A, et al. Comparison of proteomic and transcriptomic profiles in the bronchial airway epithelium of current and never smokers. *PLoS One*. 2009;4(4):e5043.

38. Titz B, Sewer A, Schneider T, et al. Alterations in the sputum proteome and transcriptome in smokers and early-stage COPD subjects. *J Proteomics*. 2015;128:306-320.
39. Nicholas BL, Skipp P, Barton S, et al. Identification of lipocalin and apolipoprotein AI as biomarkers of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2010;181(10):1049-1060.
40. Verrills NM, Irwin JA, He XY, et al. Identification of novel diagnostic biomarkers for asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2011;183(12):1633-1643.
41. Pinto-Plata V, Toso J, Lee K, et al. Profiling serum biomarkers in patients with COPD: associations with clinical parameters. *Thorax*. 2007;62(7):595-601.
42. Walter RE, Wilk JB, Larson MG, et al. Systemic inflammation and COPD: the Framingham Heart Study. *Chest*. 2008;133(1):19-25.
43. Dickens JA, Miller BE, Edwards LD, et al. COPD association and repeatability of blood biomarkers in the ECLIPSE cohort. *Respir Res*. 2011;12:146.
44. Kohler M, Sandberg A, Kjellqvist S, et al. Gender differences in the bronchoalveolar lavage cell proteome of patients with chronic obstructive pulmonary disease. *J Allergy Clin Immunol*. 2013;131(3):743-751.
45. Merali S, Barrero CA, Bowler RP, et al. Analysis of the plasma proteome in COPD: novel low abundance proteins reflect the severity of lung remodeling. *COPD*. 2014;11(2):177-189.
46. Zemans RL, Jacobson S, Keene J, et al. Multiple biomarkers predict disease severity, progression and mortality in COPD. *Respir Res*. 2017;18(1):117.
47. Bradford E, Jacobson S, Varasteh J, et al. The value of blood cytokines and chemokines in assessing COPD. *Respir Res*. 2017;18(1):180.
48. Ngo D, Peterson B, Keyes M, Gao Y, et al. The proteomic profile associated with lung function and COPD in the Jackson Heart Study. *Am J Respir Crit Care Med*. 2020;201:A6136.
49. Carolan BJ, Hughes G, Morrow J, et al. The association of plasma biomarkers with computed tomography-assessed emphysema phenotypes. *Respir Res*. 2014;15:127.
50. Coxson HO, Dirksen A, Edwards LD, et al. The presence and progression of emphysema in COPD as determined by CT scanning and biomarker expression: a prospective analysis from the ECLIPSE study. *Lancet Respir Med*. 2013;1(2):129-136.
51. Bozinovski S, Hutchinson A, Thompson M, et al. Serum amyloid A is a biomarker of acute exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2008;177(3):269-278.
52. Keene JD, Jacobson S, Kechris K, et al. Biomarkers predictive of exacerbations in the SPIROMICS and COPDGene cohorts. *Am J Respir Crit Care Med*. 2017;195(4):473-481.
53. Vestbo J, Anderson W, Coxson HO, et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *Eur Respir J*. 2008;31(4):869-873.
54. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. *COPD*. 2010;7(1):32-43.
55. Jones PW, Rames AD. Tesra (treatment Of Emphysema With A Selective Retinoid Agonist) Study Results. *Am J Respir Crit Care Med*. 2011;183:2011. A641.
56. Rennard SI. The Promise of Observational Studies (ECLIPSE, SPIROMICS, and COPDGene) in achieving the goal of personalized treatment of chronic obstructive pulmonary disease. *Semin Respir Crit Care Med*. 2015;36(4):478-490.
57. Beiko T, Janech MG, Alekseyenko AV, et al. Serum proteins associated with emphysema progression in severe alpha-1 antitrypsin deficiency. *Chronic Obstr Pulm Dis*. 2017;4(3):204-216.
58. Strange C, Senior RM, Sciruba F, et al. Rationale and design of the genomic research in alpha-1 antitrypsin deficiency and sarcoidosis study. Alpha-1 protocol. *Ann Am Thorac Soc*. 2015;12(10):1551-1560.
59. Tuck MK, Chan DW, Chia D, et al. Standard operating procedures for serum and plasma collection: early detection research network consensus statement standard operating procedure integration working group. *J Proteome Res*. 2009;8(1):113-117.
60. Geyer PE, Voytik E, Treit PV, et al. Plasma proteome profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol Med*. 2019;11(11):e10427.
61. Rikkert LG, Coumans FAW, Hau CM, Terstappen L, Nieuwland R. Platelet removal by single-step centrifugation [published online ahead of print June 17, 2020]. *Platelets*. <https://doi.org/10.1080/09537104.2020.1779924>.
62. Depner CM, Melanson EL, McHill AW, Wright KP Jr. Mistimed food intake and sleep alters 24-hour time-of-day patterns of the human plasma proteome. *Proc Natl Acad Sci U S A*. 2018;115(23):E5390-E5399.
63. de Torres JP, Casanova C, Pinto-Plata V, et al. Gender differences in plasma biomarker levels in a cohort of COPD patients: a pilot study. *PLoS One*. 2011;6(1):e16021.
64. Jammes Y, Steinberg JG, Ba A, Delliaux S, Bregeon F. Enhanced exercise-induced plasma cytokine response and oxidative stress in COPD patients depend on blood oxygenation. *Clin Physiol Funct Imaging*. 2008;28(3):182-188.
65. Agusti A, Hedner J, Marin JM, Barbe F, Cazzola M, Rennard S. Night-time symptoms: a forgotten dimension of COPD. *Eur Respir Rev*. 2011;20(121):183-194.
66. Miller BE, Tal-Singer R, Rennard SI, et al. Plasma fibrinogen qualification as a drug development tool in chronic obstructive pulmonary disease. Perspective of the Chronic Obstructive Pulmonary Disease Biomarker Qualification Consortium. *Am J Respir Crit Care Med*. 2016;193(6):607-613.
67. Jones PW. St. George's Respiratory Questionnaire: MCID. *COPD*. 2005;2(1):75-79.
68. Campbell EJ, Campbell MA, Boukedes SS, Owen CA. Quantum proteolysis by neutrophils: implications for pulmonary emphysema in alpha 1-antitrypsin deficiency. *J Clin Invest*. 1999;104(3):337-344.
69. Ma S, Lin YY, Turino GM. Measurements of desmosine and isodesmosine by mass spectrometry in COPD. *Chest*. 2007;131(5):1363-1371.
70. Sukkar MB, Ullah MA, Gan WJ, et al. RAGE: a new frontier in chronic airways disease. *Br J Pharmacol*. 2012;167(6):1161-1176.
71. Kearon C, de Wit K, Parpia S, et al. Diagnosis of pulmonary embolism with d-Dimer adjusted to clinical probability. *N Engl J Med*. 2019;381(22):2125-2134.
72. Jankowich MD, Wu WC, Choudhary G. Association of elevated plasma endothelin-1 levels with pulmonary hypertension, mortality, and heart failure in african american individuals: the Jackson Heart Study. *JAMA Cardiol*. 2016;1(4):461-469.
73. Sun W, Kechris K, Jacobson S, et al. Common genetic polymorphisms influence blood biomarker measurements in COPD. *PLoS Genet*. 2016;12(8):e1006011.
74. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558(7708):73-79.
75. Mastej E, Gillenwater L, Zhuang Y, Pratte KA, Bowler RP, Kechris K. Identifying protein-metabolite networks associated with COPD phenotypes. *Metabolites*. 2020;10(4):124.