Interstitial lung abnormalities and aging-related factors

Gísli Þór Axelsson

Thesis for the degree of Philosophiae Doctor

September 2022

School of Health Sciences

FACULTY OF MEDICINE

UNIVERSITY OF ICELAND

Interstitial lung abnormalities and aging-related factors

Gísli Þór Axelsson

Thesis for the degree of Philosophiae Doctor

Supervisor/s:

Gunnar Guðmundsson, Professor

Doctoral committee:

Gary Matthew Hunninghake, Thor Aspelund and Vilmundur Guðnason

September 2022

School of Health Sciences

FACULTY OF MEDICINE

UNIVERSITY OF ICELAND

Millivefslungnabreytingar og öldrunartengdir þættir

Gísli Þór Axelsson

Ritgerð til doktorsgráðu

Leiðbeinandi:

Gunnar Guðmundsson, prófessor

Doktorsnefnd:

Gary Matthew Hunninghake, Thor Aspelund og Vilmundur Guðnason

September 2022

Heilbrigðisvísindasvið

LÆKNADEILD

HÁSKÓLI ÍSLANDS

ORCID ID: 0000-0002-7156-9080

ISBN 978-9935-9657-3-8

Printing by Háskólaprent.

Reykjavik, Iceland 2022

Ágrip

Inngangur: Millivefslungnasjúkdómar eru sjúkdómar sem leggjast helst á eldra fólk, einkennast af örvefsmyndun í lungum og hafa oft slæmar horfur. Millivefslungnabreytingar (MLB) eru breytingar á myndgreiningu af lungum sem líkjast millivefslungnasjúkdómum en eru skilgreindar í þátttakendum ferilrannsókna án þekkts millivefslungnasjúkdóms. Niðurstöður fyrri rannsókna á MLB hafa sýnt að þær tengjast mörgum sömu áhættuþáttum og millivefslungnasjúkdómar, þar með talið aldri, og ályktað hefur verið að MLB séu mögulegur forveri millivefslungnasjúkdóma. Vegna sterkra tengsla MLB og millivefslungnasjúkdóma við háan aldur var markmið ritgerðarinnar að kanna tengsl MLB við öldrunartengda lífvísa og sjúkdóma sem tengjast aldri í fjórum greinum. Könnuð voru tengsl MLB við þætti sem tengjast heilsu og virkni, krabbamein innan og utan lungna, fjölda blóðpróteina og lengdar litningaenda í hvítkornum.

Aðferðir: Gögn úr þremur ferilrannsóknum, Öldrunarrannsókn Hjartaverndar (AGES-Reykjavík; greinar I-IV), Genetic epidemiology of COPD (COPDGene) rannsókninni (greinar III-IV) og Framingham hjartarannsókninni (FHS; grein IV) voru notuð. Tölvusneiðmyndir af þáttakendum höfðu áður verið lesnar m.t.t. MLB. Gögn voru fengin úr gagnasöfnum ferilrannsóknanna og rafrænni sjúkraskrá. Tengsl þriggja huglægra mælikvarða heilsu og virkni við MLB voru könnuð með tvíkosta aðhvarfsgreiningu. Samband MLB við greiningar og dánartíðni vegna lungnakrabbameina og annarra krabbameina var kannað með samanburði uppsafnaðs nýgengis með Gray's prófum og með Cox áhættulíkönum. Tengsl MLB, framþróunar MLB og einkirnabreytileika sem áður hafa verið tengd lungnatrefjun við prótein í blóði voru könnuð með leiðréttum aðhvarfslíkönum með stökum próteinum. Afbrigði LASSO aðhvarfsgreiningar í 200 handahófsúrtökum var notað til að finna hópa próteina sem spáðu fyrir um MLB og framþróun þeirra. Samband MLB við lengd litningaenda í hvítkornum var greint með aðhvarfsgreiningu og Cox lifunargreining var nýtt til að prófa hvort tengsl lifunar litningaendalengdar fyndust meðal þátttakenda með MLB.

Niðurstöður: Í AGES-Reykjavík tengdust MLB minni líkum á sjálfstæði við athafnir daglegs lífs (gagnlíkindahlutfall (OR) 0,70, 95% öryggisbil (ÖB) 0,55-0,90), minni líkum á að þátttakendum fyndust þeir vera við góða heilsu (OR 0,66, ÖB 0,52-0,82) og minni líkum á reglulegri hreyfingu (OR 0,72, ÖB 0,56-

0,91). Uppsafnað nýgengi lungnakrabbameinsgreininga og dánartíðni vegna lungnakrabbameina var hærra meðal þátttakenda með MLB en meðal þátttakenda án þeirra ($p < 0,001$) en slíkur munur fannst ekki fyrir greiningar eða dánartíðni vegna annarra krabbameina ($p = 0,31$ og $p = 0,88$, í þessari röð). Í leiðréttum Cox áhættulíkönum tengdust MLB greiningum lungnakrabbameina (hættuhlutfall (HR) 2,77, ÖB 1,76-4,36) og dánartíðni vegna lungnakrabbameina (HR 2,89, ÖB 1,80-4,66) en ekki greiningum annarra krabbameina eða dánartíðni vegna þeirra.

Í AGES-Reykjavík tengdist sermisstyrkur 287 próteina tilvist MLB í leiðréttum aðhvarfsgreiningum. Marktækust tengsl fundust fyrir Surfactant protein B (SFTPB), Secretoglobin family 3A member 1 (SCGB3A1) og WAP four-disulfide core domain protein 2 (WFDC2). Líkan með átta próteinum sem byggðist á notkun próteinanna í 200 spálíkönum í handahófsúrtökum var prófað og svæði undir ROC-kúrfu (AUROC) reyndist vera 0,880 eftir prófun með handahófsúrtökum. Niðurstöður varðandi SFTPB, SCGB3A1 og WFDC2 í eins próteins líkönum voru sannreyndar í COPDGene, sem og átta próteina líkanið (prófað AUROC 0,826). Þegar tengsl próteina við framþróun MLB voru könnuð með líkönum fyrir stök protein hafði 121 prótein tengsl við framþróun MLB og fjögurra próteina spálíkan fyrir framþróun, byggt á breytuvali í 200 handahófsúrtökum, hafði AUROC-gildið 0,824. Af einbasabreytileikum tengdum lungnatrefjun tengdist rs35705950 breytileikinn nærri *MUC5B* magni SFTPB (β 0,26, $p = 8\times10^{-18}$). Hvað varðar MLB og lengd litningaenda, reyndist styttri litningaendalengd tengjast MLB í COPDGene (OR 2,2, ÖB 1,5-3,4) og AGES-Reykjavík (OR 2,6, ÖB 1,4-4,9) og MLB tengdust einnig styttri litningaendum í FHS (β -767 basapör, ÖB -76 - -1584). Hvorki var sýnt fram á tengsl litningaendalengdar við lifun meðal þáttakenda með MLB í COPDGene né í AGES-Reykjavík.

Ályktanir: Þessar niðurstöður sýna fram á tengsl MLB við ýmsa öldrunartengda þætti og klínískar útkomur. Þetta á við um huglæga þætti er tengjast heilsu og virkni, greiningar og dánartíðni vegna lungnakrabbameina og litningaendalengdar hvítkorna. Til viðbótar benda niðurstöður til margra próteina sem nýrra lífvísa MLB í blóði, en sum þeirra hafa þekkt tengsl við öldrun. Heilt yfir auka niðurstöðurnar þekkingu á líffræðilegum eiginleikum og faraldsfræðilegri áhættu fólks með myndgreiningarbreytingar sem bent gætu til snemmbúinnar lungnatrefjunar.

Lykilorð: Millivefslungnabreytingar, öldrun, lungnakrabbamein, próteinmengjafræði, litningaendar.

Abstract

Introduction: Interstitial lung diseases (ILD), such as idiopathic pulmonary fibrosis (IPF), are diseases of elderly people that are characterised by pulmonary deposition of fibrous tissue and often have a poor prognosis. Interstitial lung abnormalities (ILA) are radiologic changes that are similar in appearance to ILD but are characterised in cohort study participants without known ILD. Prior research findings, including risk factor parallels such as advanced age, suggest that ILA are likely a precursor to ILD in some cases and are themselves associated with poor health outcomes. Due to the strong links that ILA and ILD share with advanced aging, the aims of the thesis were to assess the relationship of ILA with aging-related biological markers and outcomes in four papers. These were functional status, pulmonary and extra-pulmonary malignancies, a variety of blood proteins and leukocyte telomere length.

Methods: Data from three cohort studies, the Age/Gene-Environment Susceptibility-Reykjavik (AGES-Reykjavik) study, the Genetic epidemiology of COPD (COPDGene) study and the Framingham Heart study were used. CT images of participants had been manually read with regards to ILA status. Outcomes were ascertained from the cohorts' phenotyping data and from electronic medical records. The associations of three self-reported measures of health and functional status with ILA were assessed with logistic regression. The relationship of ILA with diagnoses of and mortality from pulmonary and non-pulmonary malignancies was assessed with Gray's tests and proportional hazards models. The associations of ILA, ILA progression and single nucleotide polymorphisms associated with pulmonary fibrosis with protein markers were assessed with adjusted regression models of single proteins. Adaptive LASSO modelling of bootstrap data samples was used to find sets of proteins predictive of ILA and ILA progression. The relations of ILA with leukocyte telomere length and telomere length with mortality among those with ILA were modelled with regression and Cox proportional hazards models.

Results: In the AGES-Reykjavik cohort, participants with ILA were less likely to be independent in activities of daily living (odds ratio (OR) 0.70, 95% confidence interval (CI) 0.55-0.90), to perceive their health as good or better (OR 0.66, CI 0.52-0.82) and to be regularly physically active (OR 0.72, CI 0.56-0.91). The cumulative incidences of lung cancer diagnoses and lung cancer

mortality were higher among participants with ILA than others ($p < 0.001$), but such differences were not found for other cancers. In adjusted Cox proportional hazards models, ILA were associated with diagnoses of lung cancer (hazard ratio (HR) 2.77, CI 1.76-4.36) and mortality from lung cancer (HR 2.89, CI 1.80-4.66) but not with diagnoses of, or mortality from, cancers excluding lung cancer. In adjusted analyses of single serum proteins and ILA in AGES-Reykjavik, 287 proteins were significantly associated with ILA. The most significant associations were for Surfactant protein B (SFTPB), Secretoglobin family 3A member 1 (SCGB3A1) and WAP four-disulfide core domain protein 2 (WFDC2). An eight-protein model of ILA status was created based on inclusion in 200 adaptive LASSO models in bootstrap data samples which had an area under the receiving operator characteristic curve (AUROC) value of 0.880 after bootstrap validation. Single-protein results for SFTPB, SCGB3A1 and WFDC2 and the eight-protein model were validated in COPDGene (validated AUROC 0.826). Similarly, 121 proteins were associated with ILA progression in single protein models in AGES-Reykjavik and multi-protein modelling yielded a four-protein model which had a validated AUROC of 0.824. In analyses of fibrosis-associated genetic polymorphisms, the rs35705950 polymorphism near the *MUC5B* gene was found to be associated with levels of SFTPB (β 0.26, $p = 8\times10^{-18}$). As for analyses of ILA and leukocyte telomere length, shorter telomere length was associated with increased odds of ILA in COPDGene (OR 2.2, CI 1.5-3.4) and AGES-Reykjavik (OR 2.6, CI 1.4-4.9) and ILA was associated with shorter telomere length in FHS (β -767 base pairs, CI -76 - -1584). Adjusted Cox proportional hazards models did not demonstrate a significant association of telomere length with mortality among those with ILA in COPDGene nor in AGES-Reykjavik.

Conclusions: These cohort study-based data demonstrate the associations of ILA with several aging-related markers and outcomes. These are subjective markers of functional status, diagnoses of and mortality from pulmonary malignancies and telomere length in leukocytes. Additionally, many novel protein biomarkers of ILA are proposed, some of them previously related to aging. These findings improve knowledge of biological markers and epidemiological outcomes associated with ILA as a possible marker of early pulmonary fibrosis.

Keywords: Interstitial lung abnormalities, aging, lung cancer, proteomics, telomeres.

Acknowledgements

First, I thank the Eimskip University Fund for providing me with grant support to work on this thesis. This support has been extremely valuable.

I would like to thank the participants of the cohort studies involved in the thesis, the AGES-Reykjavik Study, the COPDGene Study and the Framingham Heart Study. This thesis, and a great amount of incredibly important scientific work, would not have been possible without the willingness of thousands of people to donate their time and effort to participate in these studies.

I want to express my gratitude to all my instructors and co-authors at Brigham and Women's Hospital and Harvard University. First and foremost, to Matt Hunninghake, a brilliant member of my doctoral committee who has provided excellent guidance and instruction. It has been a privilege to learn from him. I also want to express my best thanks to Rachel Putman, Hiroto Hatabu, Jason Sanders, and the other scientists at Brigham and Women's Hospital that helped create the data for the study and provided guidance, review, and instructions for the presented work. My thanks also go to Katherine Pratte, Russell Bowler and their team at National Jewish Health in Denver for their help.

I am incredibly thankful to the staff at the Icelandic Heart Association. It has been an honour to work with and learn from Vilmundur Guðnason and Thor Aspelund, two of the best scientists I know, which I was lucky enough to have as part of my doctoral committee. I would also like to specifically thank Valborg Guðmundsdóttir, Elías Freyr Guðmundsson, Alexander Guðjónsson and all the others for all their help and patience.

My biggest thanks to my supervisor, Gunnar Guðmundsson, for his excellent guidance and endless encouragement and enthusiasm for the project and for my advancement as a scientist. He is an excellent role model for all aspiring clinician-scientists.

Last, a huge thanks and all my love to my wonderful wife, Auður, my parents, sisters, and family for supporting me and inspiring me every day.

Contents

List of abbreviations

ADAM9	ADAM metallopeptidase domain 9
AGES-Reykjavik	Age/Gene-Environment Susceptibility-Reykjavik
AIP	Acute interstitial pneumonia
ANXA9	Annexin A9
AUROC	Area under the receiver operating characteristic curve
CAPNS1	Calpain small subunit 1
CI	Confidence interval
COP	Cryptogenic organizing pneumonia
COPD	Chronic obstructive pulmonary disease
COPDGene	Genetic epidemiology of COPD
CSS	Cathepsin S
CT	Computed tomography
CXCL10	C-X-C Motif Chemokine Ligand 10
DIP	Desquamative interstitial pneumonia
DLCO	Diffusion capacity of carbon monoxide
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal cell transition
ER	Endoplasmic reticulum
FEV1	Forced expiratory volume in one second

FGF	Fibroblast growth factor
FHS	Framingham Heart Study
FISH	Fluorescence *in situ* hybridization
FN1	Fibronectin
FVC	Forced vital capacity
GDF-15	Growth differentiation factor 15
HAAs	High attenuation areas
HLA	Human leukocyte antigen
HNRNPDL	Heterogeneous nuclear ribonucleoprotein D-like
HR	Hazard ratio
ICD	International Classification of Diseases
IIP	Idiopathic interstitial pneumonia
ILA	Interstitial lung abnormalities
ILD	Interstitial lung disease
IPF	Idiopathic pulmonary fibrosis
LASSO	Least Absolute Shrinkage and Selection Operator
MMP	Matrix metalloproteinase
MTL	Mean relative telomere length
NSIP	Nonspecific interstitial pneumonia
OR	Odds ratio
PDGF	Platelet-derived growth factor
PLAUR	Urokinase plasminogen activator surface receptor
RB	Respiratory bronchiolitis
RNA	Ribonucleic acid
SCGB3A1	Secretoglobin family 3A member 1

SFTPB	Surfactant protein B
sICAM-1	Soluble intercellular adhesion molecule-1
SNP	Single-nucleotide polymorphism
SOMAmer	Slow-Off rate Modified Aptamer
sVCAM-1	Soluble Vascular Cell Adhesion Molecule-1
TGF	Transforming growth factor
TMPO	Lamina-associated polypeptide 2, isoforms beta/gamma
TNFRSF1B	Tumor necrosis factor receptor superfamily member 1B
UIP	Usual interstitial pneumonia
UPR	Unfolded protein response
VEGF	Vascular endothelial growth factor
WFDC2	WAP four-disulfide core domain protein 2
WFIKKN2	WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 2

List of figures

List of tables

List of original papers

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals:

I. **Axelsson GT**, Putman RK, Araki T, Sigurdsson S, Gudmundsson EF, Eiriksdottir G, Aspelund T, Miller ER, Launer LJ, Harris TB, Hatabu H, Gudnason V, Hunninghake GM, Gudmundsson G. (2018). Interstitial lung abnormalities and self-reported health and functional status. *Thorax*. Sep;73(9):884-886.

II. **Axelsson GT**, Putman RK, Aspelund T, Gudmundsson EF, Hida T, Araki T, Nishino M, Hatabu H, Gudnason V, Hunninghake GM, Gudmundsson G. (2020). The Associations of Interstitial Lung Abnormalities with Cancer Diagnoses and Mortality. *European Respiratory Journal*. Jul 9:1902154.

III. **Axelsson GT**, Gudmundsson G, Pratte KA, Aspelund T, Putman RK, Sanders JL, Gudmundsson EF, Hatabu H, Gudmundsdottir V, Gudjonsson A, Hino T, Hida T, Hobbs BD, Cho MH, Silverman EK, Bowler RP, Launer LJ, Jennings L, Hunninghake GM, Emilsson V, Gudnason V. (2022). The Proteomic Profile of Interstitial Lung Abnormalities. *American Journal of Respiratory and Critical Care Medicine*. 10.1164/rccm.202110-2296OC. Advance online publication. https://doi.org/10.1164/rccm.202110-2296OC

IV. Putman RK, **Axelsson GT**, Ash SY, Sanders JL, Menon AA, Araki T, Nishino M, Yanagawa M, Gudmundsson EF, Qiao D, San José Estépar R, Dupuis J, O'Connor GT, Rosas IO, Washko GR, El-Chemaly S, Raby BA, Gudnason V, DeMeo DL, Silverman EK, Hatabu H, DeVivo I, Cho MH, Gudmundsson G, Hunninghake GM. (2022). Interstitial lung abnormalities are associated with decreased mean telomere length. *European Respiratory Journal*. Feb 3:2101814.

All papers are reprinted by kind permission of the publishers.

Declaration of contribution

Paper I: Gisli Thor Axelsson (GTA) contributed to the planning of the study and design, in cooperation with Gunnar Gudmundsson (GG), Gary M. Hunninghake (GMH), Vilmundur Gudnason (ViG) and Hiroto Hatabu (HH). GTA performed the statistical analysis under the guidance of and in cooperation with Thor Aspelund (ThA), Rachel K. Putman (RKP), GG and GMH. GTA drafted the initial manuscript in cooperation with GG, GMH and ViG and contributed to critical revision of its content before and during the submission process along with other co-authors. ViG, GMH and GG obtained funding.

Paper II: GTA contributed to the study concept, planning and design along with GG, RKP, HH, GMH, ViG and Tetsuro Araki (TeA). GTA interpreted and prepared the data and performed the statistical analysis in cooperation with ThA, GMH, RKP and Elias F. Gudmundsson (EFG). GTA drafted the initial manuscript under guidance of GG, GMH, RKP and ViG and contributed to critical revisions with co-authers. ViG, GMH and GMH obtained funding.

Paper III: GTA contributed to the overall study conceptualization and design with GG, GMH, ViG, Russell P. Bowler (RPB), Lenore J. Launer and Lori L. Jennings (LLJ). GTA planned the statistical methodology and performed the statistical analysis in cooperation with and under the guidance of ThA, Katherine A. Pratte, Valborg Gudmundsdottir (VaG), Alexander Gudjonsson, Jason L. Sanders, Brian D. Hobbs, Michael H. Cho (MHC) and Edwin K. Silverman. GTA drafted the initial manuscript under the supervision of GG, GMH and ViG and contributed to critical revisions at all stages of the submission process in cooperation with all co-authors. RKP, ViG, VaG, Valur Emilsson, GMH, GG, LLJ, RPB and others obtained funding.

Paper IV: GG, GMH and RKP planned the study. RKP drafted the initial manuscript. GTA participated in interpretation of statistical results and writing of the manuscript and contributed to critical revision with other co-authors. GMH, MHC, RKP and GG obtained funding.

1 Introduction

1.1 Interstitial lung disease

Interstitial lung diseases (ILD), also known as diffuse parenchymal lung diseases, are nonneoplastic, non-infectious diseases mainly affecting the interstitium of the lungs, the area between the alveolar epithelial and the capillary endothelial basement membranes ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002). While many interstitial lung diseases are idiopathic, and collectively termed the idiopathic interstitial pneumonias (IIPs), several well-known exposures are recognized causes of ILD. Among these causes are connective tissue diseases (connective tissue disease related ILD, CT-ILD) (Cottin, 2016). The connective tissue diseases that most commonly cause ILD are systemic sclerosis and rheumatoid arthritis. The prognosis of patients with connective tissue disease related ILD is often better than that of patients with ILD of an unknown cause (Cottin et al., 2018). Other systemic diseases that are known to cause ILD are sarcoidosis, a multi-organ disease characterized by granulomatous inflammation (Cottin et al., 2018) and hypersensitivity pneumonitis, which is due to repeated exposure to an antigen followed by a complex immunologic response (Selman et al., 2012). While hypersensitivity pneumonitis has been associated with a myriad of causal antigens, these are primarily organic antigens, often fungal (e.g. „farmer's lung"), bacterial (e.g. „air conditioner lung") or avian (e.g. „pigeon breeder's disease" or „bird fancier's lung") (Chan et al., 2012; Selman et al., 2012). Exposure to inorganic particles, often related to occupation, can also cause interstitial lung disease. These diseases are commonly known as the occupationally related ILD or the pneumoconioses. The most common exposures causing ILD are exposure to asbestos, causing asbestosis and exposure to silica, causing silicosis (Cottin et al., 2018; Cullinan & Reid, 2013). Medications are also known to cause a subtype of ILD termed drug-induced ILD. While many medications cause this, some of the more common include cancer therapies such as bleomycin and gemcitabine, mammalian target of rapamycin inhibitors, and the anti-arrhythmic medication amiodarone (Skeoch et al., 2018). Lastly, ILD is a known, although uncommon, complication of vasculitides, most commonly of microscopic polyangiitis (Alba et al., 2017). An overview of the subtypes of interstitial lung disease is shown in Figure 1.

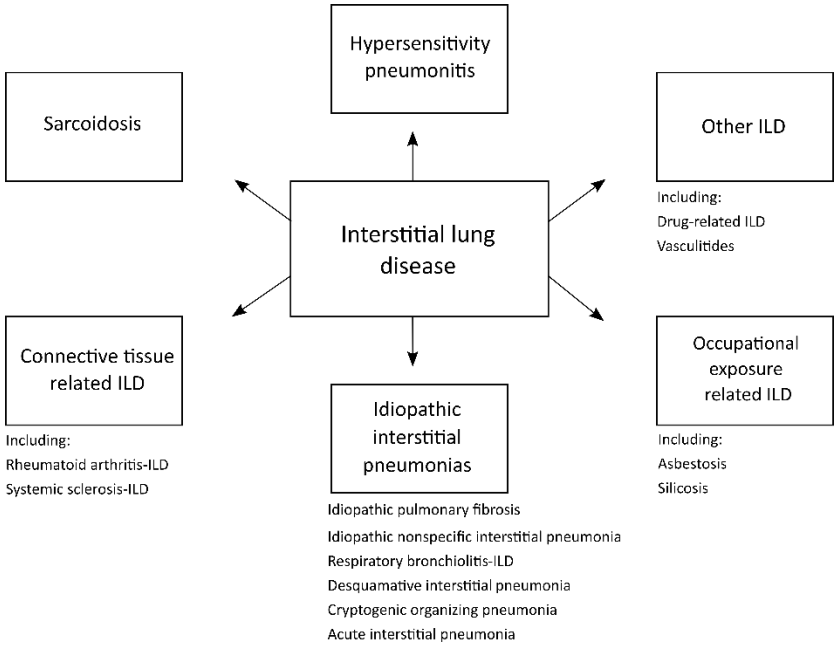

Figure 1 - Overview of interstitial lung disease
Classification is based on Cottin et al., 2018.

1.2 The idiopathic interstitial pneumonias

The idiopathic interstitial pneumonias (IIPs) are a group of interstitial lung diseases that were originally termed 'idiopathic' as none of the diseases had a known cause ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002). As understanding of their genetic risk factors and pathogenesis progresses, calls have been made for re-evaluation of the ´idiopathic´ term (Wolters et al., 2018). Six diseases are considered ´major´ diseases in this group (Travis et al., 2013). These diseases are shown in Table 1.

Table 1 - The idiopathic interstitial pneumonias

DISEASE	ACRONYM
Idiopathic pulmonary fibrosis	IPF
Nonspecific interstitial pneumonia	NSIP
Respiratory bronchiolitis-interstitial lung disease	RB-ILD
Desquamative interstitial pneumonia	DIP
Cryptogenic organizing pneumonia	COP
Acute interstitial pneumonia	AIP

Based on Travis et al., 2013.

Nonspecific interstitial pneumonia (NSIP) is a disease entity whose pathological hallmark is the NSIP histopathological pattern. It is characterised by a mild interstitial inflammation with relatively preserved lung architecture and absence or sparsity of honeycombing and fibroblastic foci with dense fibrosis ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002). On computed tomography imaging, the dominant feature of NSIP are reticular abnormalities or ground glass changes of the lower lung ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002; Belloli et al., 2016). Patients are commonly female, and a minority have a history of smoking (Belloli et al., 2016). The clinical course is variable and dependent on the extent of fibrosis, but mortality from the disease occurs in the minority of patients with the disease ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002; Travis et al., 2013; Travis et al., 2008).

Both respiratory bronchiolitis-interstitial lung disease (RB-ILD) and desquamative interstitial pneumonia (DIP) are rare diseases that are predominantly related to smoking and have been regarded to belong to a spectrum of disease collectively referred to as "smoking-related IIPs" or "smoking-related interstitial lung disease" (Margaritopoulos et al., 2016; Travis et al., 2013). While respiratory bronchiolitis, bronchiolocentric clusters of pigmented macrophages with peribronchiolar thickening of alveolar septa, can pathologically be regarded as a physiologic and asymptomatic response to smoking, RB-ILD occurs when these lesions are accompanied by clinical interstitial lung disease. This mainly occurs

in relatively young, male smokers, with clinical symptoms of cough and dyspnoea, pulmonary function tests showing a restrictive pattern and ground-glass changes with centrilobular nodules on pulmonary imaging. While the clinical presentation also consists of increasing cough and dyspnoea in smokers, DIP has more extensive filling of alveoli on pathological examinations and more diffuse ground-glass involvement of the lungs on imaging (Margaritopoulos et al., 2016).

Cryptogenic organizing pneumonia (COP), has a more acute nature than the other IIPs and, unlike other IIPs, is due to primarily intra-alveolar lesions (Cordier, 2006; Travis et al., 2013). It is also mainly a disease of adult non-smokers, causing fever, cough, dyspnoea and flu-like symptoms and has a very favourable response to corticosteroid therapy (Collard et al., 2016; Cordier, 2006). The histological hallmark of COP are intra-alveolar buds of granulation tissue made of fibroblasts and typical radiological changes are non-specific, consisting of multiple, bilateral, peripheral, often migratory opacities that can have variable appearance (Cordier, 2006).

1.3 Idiopathic pulmonary fibrosis

1.3.1 Epidemiology and risk factors

Idiopathic pulmonary fibrosis (IPF) is the most common of the IIPs, with an estimated incidence of 3-9 cases per 100,000 per year in Europe and North America, but lower in other parts of the world (Hutchinson et al., 2015). Among people aged 18-64 years in the USA, the prevalence of IPF has been found to range from 4.6-6.7 per 100,000 persons to 8.4-11.3 per 100,000 persons based on case finding criteria (Raghu et al., 2016). Meanwhile, in the US population aged 65 and older, the incidence of IPF has been estimated to be 93.7 per 100,000 per year and prevalence estimates in this age group are as high as 494.5 cases per 100,000 persons (Raghu et al., 2014). Due to these estimates being from insurance claims data, efforts have been made to recalculate these estimates adjusting for misclassification. Using this method, the total US incidence of IPF has been estimated to be 12.8 per 100,000 per year and the total US prevalence has been estimated to be 50.1 per 100,000 (Esposito et al., 2015). Other estimates of the same statistics vary however, with estimates of incidence in the range of 6.8–8.8 per 100,000 persons per year or 16.3–17.4 per 100,000 persons per year and estimates of prevalence in the range of 14-27.9 cases per 100,000 people or 42.7-63 cases per 100,000 people, with differences based on case definitions (Nalysnyk et al., 2012). This variance could stem in part from the uniform definition of IPF being relatively

recent ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002).

As the previously mentioned difference in prevalence estimates between age ranges implies, age is a cardinal risk factor for IPF diagnosis. The mean age at diagnosis has been estimated at 66 years and around two-thirds of patients are more than 60 years old at diagnosis ("American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000). IPF is also more common among males than females and cigarette smoking is the best recognised risk factor ("American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000; Baumgartner et al., 2000; Baumgartner et al., 1997). Observational studies have proposed various other external risk factors. Among such risk factors are obstructive sleep apnea (Lancaster et al., 2009; Mermigkis et al., 2010), infections from viruses, especially herpesviruses (Chioma & Drake, 2017; B. B. Moore & Moore, 2015; Tang et al., 2003), and increased burden of bacteria (Molyneaux et al., 2014) and other pathogens that lead to alterations of the lung microbiome ("American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000; Lederer & Martinez, 2018). Additional risk factors are gastroesophageal reflux (Tobin et al., 1998) and a variety of environmental and occupational exposures such as metal, wood and stone dust, sand, silica, livestock, and bird keeping ("American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000; Baumgartner et al., 2000; Sgalla et al., 2018). Expression of antimicrobial molecules is positively associated with IPF survival while decreased expression of immune response molecules is associated with worse survival, suggesting that alterations in the lung microbiome and the immune response to it could play a role in IPF pathobiology (Huang et al., 2017; Molyneaux et al., 2017). In the last decades, progress has been made in identifying genetic risk factors of IPF. The genetic risk factor accounting for the largest proportion of the population risk of IPF is the minor allele of a single-nucleotide polymorphism (SNP), rs35705950, in a promoter for the *MUC5B* gene (Allen et al., 2020; Seibold et al., 2011). This variant causes overexpression of *MUC5B*, a gene which codes for mucin 5B, a mucin glycoprotein which is a major component of airway mucus and necessary for airway clearance, macrophage function and immunity (Nakano et al., 2016; Roy et al., 2014). Additionally, mutations in telomerase genes, *TERT, TERC, RTEL1* and *PARN* that confer risk of telomere shortening over time markedly increase risk of sporadic IPF and are the cause of some familial IPF syndromes

(Armanios et al., 2007; Mushiroda et al., 2008; Stuart et al., 2015; Tsakiri et al., 2007). Telomerase related variants have been associated with shorter survival among those with pulmonary fibrosis (Borie et al., 2016). Other genetic variants associated with pulmonary fibrosis risk are *SFTPA2* and *SFTPC*, encoding for surfactant proteins, *OBFC1*, *DKC1*, *TINF2*, associated with DNA duplication, *ABCA3*, associated with surfactant secretion, two HLA alleles, *TLR3*, *IL1RN* and *IL8*, associated with immune function and *FAM13A*, *AKAP13*, *DSP*, *ATP11A*, *IVD*, *MUC2, MDGA2, MAPT, SPPL2C* and *DPP9*, associated with various other biological functions (Ahn et al., 2011; Alder et al., 2015; Allen et al., 2017; Campo et al., 2014; Cogan et al., 2015; Fingerlin et al., 2013; Fingerlin et al., 2016; Kaur et al., 2017; Korthagen et al., 2012; Kropski et al., 2014; Lederer & Martinez, 2018; Nogee et al., 2001; Noth et al., 2013; O'Dwyer et al., 2013; Sgalla et al., 2018; Stuart et al., 2015; Wang et al., 2009; Wolters et al., 2018). The relative allele frequency and effect size of some of the genetic variants associated with IPF risk is shown in Figure 2 which shows that rare genetic variants that have large effects on an individual level make smaller contributions to the population level risk compared to common variants with smaller effect sizes for each individual. The risk factors for IPF, both environmental and genetic, are summarised in Table 2.

Table 2 - Suggested risk factors of pulmonary fibrosis

EPIDEMIOLOGIC	
Risk factor	**Citation**
Age	"American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000
Male gender	"American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000
Cigarette smoking	"American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement.", 2000; Baumgartner et al., 1997

Obstructive sleep apnea	Lancaster et al., 2009; Mermigkis et al., 2010
Viral and bacterial infections	Chioma & Drake, 2017; Molyneaux et al., 2014; Tang et al., 2003
Gastroesophageal reflux	Tobin et al., 1998
Occupational exposure to animals, metal or stone dust	Baumgartner et al., 2000

GENETIC

SNP	**Closest gene**	**Citation**
rs35705950	*MUC5B*	Allen et al., 2020; Seibold et al., 2011
rs2736100	*TERT*	Armanios et al., 2007; Fingerlin et al., 2013; Mushiroda et al., 2008; Tsakiri et al., 2007
	STFTPA2	Wang et al., 2009
	SFTPC	Nogee et al., 2001
rs11191865	*OBFC1*	Fingerlin et al., 2013
	DKC1	Kropski et al., 2014
	TINF2	Alder et al., 2015
Various	*PARN*	Stuart et al., 2015
Various	*RTEL1*	Cogan et al., 2015; Stuart et al., 2015
2891 G > A	*ABCA3*	Campo et al., 2014
rs3775291	*TLR3*	O'Dwyer et al., 2013
rs408392, rs419598, rs2637988	*IL1RN*	Kaur et al., 2017; Korthagen et al., 2012
rs4073, rs2227307	*IL8*	Ahn et al., 2011

rs2844452, rs3020644	*C2*	Fingerlin et al., 2016
rs614549	*SLC44A4*	Fingerlin et al., 2016
rs7887	*EHMT2*	Fingerlin et al., 2016
rs2280774	*NELFE*	Fingerlin et al., 2016
rs3117116	*BTNL2*	Fingerlin et al., 2016
rs2609255	*FAM13A*	Fingerlin et al., 2013
rs62025270	*AKAP13*	Allen et al., 2017
rs2076295	*DSP*	Allen et al., 2017; Fingerlin et al., 2013
rs1278769	*ATP11A*	Fingerlin et al., 2013
rs6793295	*LRRC34*	Fingerlin et al., 2013
rs2034650	*IVD*	Peljto et al., 2013
rs7934606	*MUC2*	Fingerlin et al., 2013
rs7144383	*MDGA2*	Noth et al., 2013
rs1981997	*MAPT*	Fingerlin et al., 2013
rs17690703	*SPPL2C*	Noth et al., 2013
rs12610495	*DPP9*	Fingerlin et al., 2013

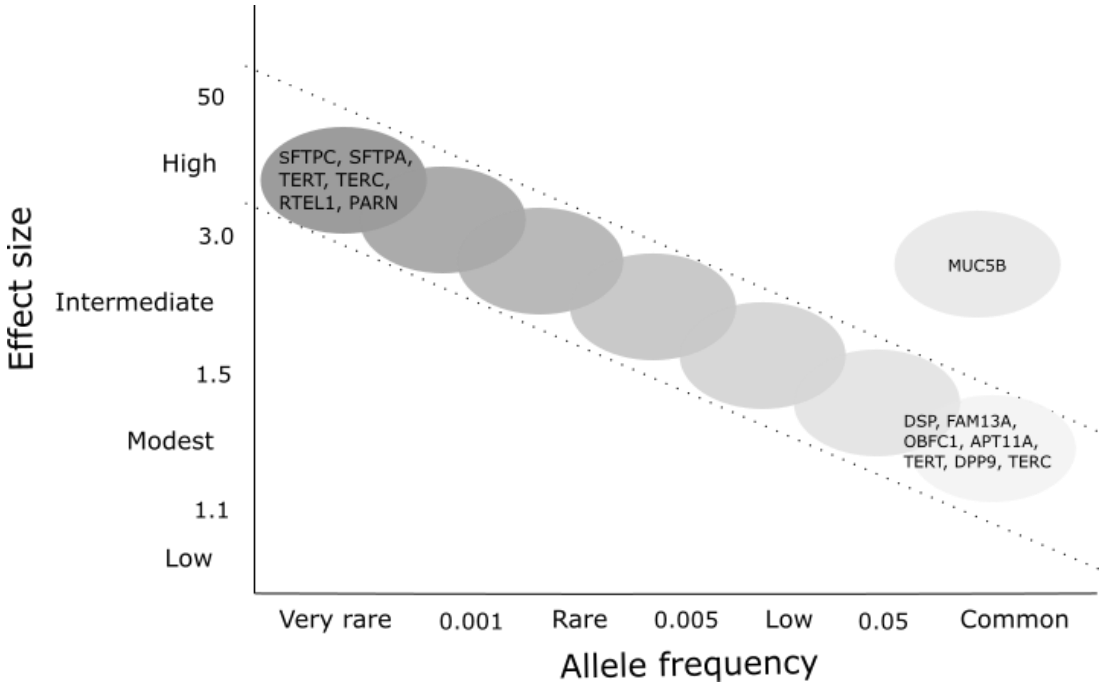

Figure 2 - The relative allele frequency and effect size of several known genetic mutations predisposing to IPF
Reproduced from Mathai et al., 2015, licensed under CC BY 4.0.

1.3.2 Clinical presentation and diagnosis

The hallmark clinical symptom, which is almost ubiquitous, of idiopathic pulmonary fibrosis is chronic exertional dyspnea in an adult, often elderly, patient. This is often accompanied with chronic, non-purulent cough and fatigue (Lederer & Martinez, 2018). These symptoms are often mistaken for chronic obstructive pulmonary disease (COPD) or heart failure at presentation, which can considerably delay diagnosis (Hewson et al., 2018). Acute exacerbations of IPF, with worsened respiratory status and fever can be seen in the disease process and can be the presenting symptoms of the disease (Richeldi et al., 2017). On clinical examination, bilateral fine ("Velcro-like") crackles are often heard on lung auscultation with clubbing and cyanosis of the extremities found in a minority of patients (Lederer & Martinez, 2018; Richeldi et al., 2017). On spirometry, reductions in forced vital capacity (FVC), forced expiratory volume in one second (FEV_1) and total lung capacity show a restrictive lung deficit and reduced diffusion capacity of carbon monoxide (DL_{CO}) is seen as a representation of reduced alveolar gas exchange (Lederer & Martinez, 2018; Richeldi et al., 2017). Hypoxemia during exercise, and hypoxemia at rest with

disease progression, can be seen (King Jr. et al., 2011). In addition to the symptoms and signs of interstitial lung disease, a thorough clinical history and examination is required to exclude other causes of ILD such as hypersensitivity pneumonitis, connective-tissue related ILD, drug-induced ILD and occupational exposure related ILD (Figure 1). Exposure to mold or birds can suggest hypersensitivity pneumonitis and various antibodies in patients' serum can suggest autoimmune disease (Lederer & Martinez, 2018).

A high-resolution computed tomography (CT) scan of the thorax is an important step in obtaining a diagnosis of IPF. The pattern seen in IPF is called the "usual interstitial pneumonia" (UIP) pattern which consists of heterogenous bilateral reticulation, more irregular than in NSIP, traction bronchiectasis (bronchial dilatation caused by surrounding fibrosis) and honeycombing in the peripheral parts of the lower lobes (Lederer & Martinez, 2018; Lynch et al., 2018; Raghu et al., 2011). Honeycombing on CT imaging, a key feature of the UIP pattern, are clustered, thick-walled cystic airspaces, most often 3-10 mm in diameter (Lynch et al., 2018; Raghu et al., 2011). The UIP pattern is diagnostic of IPF if no other cause of interstitial lung disease is found. A probable UIP pattern, bilateral reticulation of peripheral lower lobes without honeycombing but with traction bronchiectasis is also suggestive of IPF diagnosis (Lederer & Martinez, 2018). Abnormalities that are predominantly in the upper-lung or midlung or peribronchovascular, spare the peripheries of the lungs or mainly have the appearance of ground-glass opacities, consolidations, mosaic attenuations, nodules or non-honeycomb cysts suggest other forms of ILD or lung disease than IPF (Lederer & Martinez, 2018; Lynch et al., 2018; Raghu et al., 2011). Coexistent pleural plaques, calcifications or effusions also suggest a different diagnosis (Raghu et al., 2011). Still, even if the imaging pattern is more suggestive of other ILD forms such as NSIP, an IPF diagnosis should not be excluded in the right clinical setting (Lynch et al., 2018). In such cases, and in cases where a probable UIP pattern is seen, a lung biopsy may be needed. Such invasive interventions can be avoided if imaging supports the clinical diagnosis by showing a definite UIP pattern or, in some cases, a probable UIP pattern (Lederer & Martinez, 2018; Raghu et al., 2011).

The histopathological changes seen in IPF are termed the usual interstitial pneumonia (UIP) pattern like the radiological pattern. It is described as heterogenous areas of scarring and honeycomb changes, alternating with lung areas of more normal appearance. The changes are primarily found in subpleural and paraseptal areas and include hyperplasia of alveolar type II cells and bronchiolar epithelium. Inflammation in IPF is mild with patchy lymphocytic

and plasma cell infiltrates (Raghu et al., 2011). Dense collagen is also seen and foci of fibroblasts and myofibroblasts, with the appearance of tissue bulges into the airspaces, are a hallmark of the microscopic changes (Raghu et al., 2011; Wolters et al., 2014). The honeycombing described in microscopic samples of IPF patients is different from the honeycombing seen on CT as the honeycomb changes seen microscopically are too small to be seen on HRCT (Lynch et al., 2018; Raghu et al., 2011). If the UIP pattern is seen on biopsy, relatively few other diagnoses other than IPF are to be considered; among them are connective tissue-related ILD, hypersensitivity pneumonitis and occupational exposure-related ILD such as asbestosis (Raghu et al., 2011). In acute exacerbations of IPF, a histologic pattern of diffuse alveolar damage with hyaline membrane formation may be seen (Wolters et al., 2014).

1.3.3 Pathogenesis

As the aforementioned risk factors infer, predisposition to IPF is believed to be via genetic factors, aging and various environmental insults, the most common of which is cigarette smoking (Wolters et al., 2018). Accumulation of the polysaccharide chitin has also been associated with pulmonary fibrosis in animal experiments and in pulmonary fibrosis in humans (Van Dyken et al., 2017). These risk factors and predispositions affect the pulmonary epithelium where they cause intracellular changes such as shortened telomeres in alveolar type II cells and a change in epithelial cell phenotype towards senescence (Alder et al., 2008; Selman et al., 2016). The increase in IPF risk caused by short telomeres is in concordance with the fact that pulmonary fibrosis is a hallmark feature of dyskeratosis congenita and other telomere syndromes (Armanios & Blackburn, 2012) and the previously mentioned mutations of telomerase, and telomerase-associated genes known to increase risk of IPF (Alder et al., 2015; Armanios et al., 2007; Kropski et al., 2014; Stuart et al., 2015; Tsakiri et al., 2007). It has also been suggested that increased endoplasmic reticulum (ER) stress with activation of the unfolded protein response pathway and apoptosis play a role in the development of IPF. Several markers of the unfolded protein response are upregulated in the alveolar cells of IPF patients (Cha et al., 2012; Korfei et al., 2008; Lawson et al., 2008). Among the mechanisms proposed for ER stress in pulmonary fibrosis are herpesviral infections (Lawson et al., 2008) and the rs35705950 polymorphism in the *MUC5B* gene as ER stress induces expression of *MUC5B* in upper and lower airways (M. H. Kim et al., 2019). It is not known whether changes such as shortened telomeres and increased ER stress are by themselves sufficient to cause IPF (Wolters et al., 2014). Many believe that these alterations reduce the ability of the lungs to respond to injury. Thus, a continued

stimulus by an external factor such as cigarette smoking may be needed for these changes to lead to pulmonary fibrosis (Sgalla et al., 2018).

Since the best-established risk factor for pulmonary fibrosis is the rs35705950 promoter of the *MUC5B* gene (Allen et al., 2020), efforts have been made to examine the role of mucociliary dysfunction in the pathogenesis of IPF. While the polymorphism is the best-established genetic risk factor for the disease, patients with it have been proposed to have increased survival compared to patients that do not (Peljto et al., 2013). This is still disputed, and this association has been argued to be a result of index-event bias, but not a true effect (Dudbridge et al., 2019). The rs35705950 polymorphism does not only confer increased risk of IPF, but also rheumatoid arthritis-associated ILD, but not ILD associated with systemic sclerosis or autoimmune myositis (Borie et al., 2013; C. Johnson et al., 2017; Juge et al., 2018). The polymorphism confers up-regulation of expression of *MUC5B* and increased expression of *MUC5B* is associated with IPF (Nakano et al., 2016; Seibold et al., 2013; Seibold et al., 2011). This is demonstrated in the distal airways of IPF patients, additionally IPF honeycomb cysts have similar mucin expression profiles (Seibold et al., 2013). Furthermore, increased concentration of Muc5b has been implicated in impaired mucociliary clearance and the development of lung fibrosis in mice (Hancock et al., 2018). The mechanisms by which increased concentrations of Muc5b in distal airways may contribute to pulmonary fibrosis are unclear, but some hypotheses exist. It has been suggested that excessive mucin production interferes with distal airway tissue repair mechanisms after chronic lung injury, or that excessive mucus interferes with the removal of inhaled particles and pathogens due to high mucus concentration and reduced function of cilia (Evans et al., 2016).

The described alterations in pulmonary epithelial cells increase expression of mediators that lead to several changes in the pulmonary microenvironment; decreased proliferation and stem-cell failure of alveolar type II cells, the expression of mesenchymal markers by epithelial cells and activation of mesenchymal cells such as fibroblasts, causing their transition to myofibroblasts (Wolters et al., 2018; Wolters et al., 2014). The mediators that are believed to cause these changes are platelet-derived growth factor (PDGF), vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF), transforming growth factor (TGF-β), and integrin $\alpha_v\beta_6$ with TGF-β being the best recognised (Kage & Borok, 2012; Lederer & Martinez, 2018; Wolters et al., 2014). In addition to causing apoptosis of epithelial cells, fibroblast recruitment and activation and production of other cytokines, TGF-β is known to promote epithelial-to-mesenchymal cell transition (EMT) of alveolar epithelial cells which

then express many of the same markers as malignant cells do after undergoing EMT (Sgalla et al., 2018; Willis et al., 2005; Wolters et al., 2014). The significance and exact role of these cells in extracellular matrix deposition in IPF is still a topic of research (Hill et al., 2019; Kage & Borok, 2012).

No matter its relationship with the pulmonary epithelium, an undisputed hallmark of IPF pathogenesis is the deposition of ECM, including collagen (Wolters et al., 2018). The source of ECM in areas of active fibrosis are fibroblasts and myofibroblasts, a fibroblast phenotype which shares molecular features with smooth muscle cells and is resistant to apoptosis (Richeldi et al., 2017; Scotton & Chambers, 2007; Wolters et al., 2014). These myofibroblasts are believed to have three possible origins; resident fibroblasts of the lung tissue, circulating fibrocytes and/or alveolar epithelial cells that have undergone EMT (Scotton & Chambers, 2007). The excessive ECM secretion, including collagen secretion, leads to the lung tissue changes and pathologic changes seen in lung fibrosis (Wolters et al., 2014). This may be in part due to the contraction of myofibroblasts causing re-organization of collagen fibrils and leading to a stiffer ECM. The stiffer ECM in turn enhances the collagen synthesis of myofibroblasts. This mechanism could cause the myofibroblasts, activated by the aberrant pulmonary epithelium, to be driven by the pathologic, fibrotic ECM, creating a positive feedback loop which possibly plays a part in driving the disease's pathogenesis at a later stage (Parker et al., 2014; Sgalla et al., 2018; Wolters et al., 2018).

In summary, current theories of the pathogenesis of IPF, an epithelium driven disease, can be described in this way: In genetically predisposed individuals, continued exposure to exogenous risk factors causes pulmonary epithelial changes. These changes, possibly by way of cellular senescence and/or ER stress, cause the release of mediators such as TGF-β and PDGF leading to activation of fibroblasts. This in turn leads to ECM deposition and the remodelling of the IPF lung. A schematic overview of some of the features of IPF pathogenesis is shown in Figure 3.

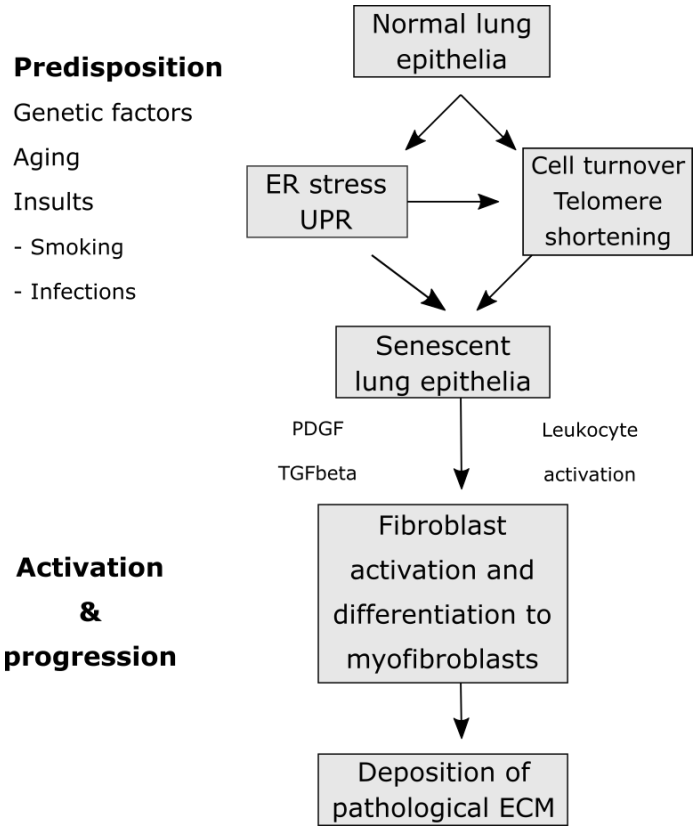

Figure 3 - A schematic model of IPF pathogenesis
Reproduced with permission from Wolters et al., 2018.

1.3.4 Prognosis and treatment

The clinical course of IPF is variable. Many patients have a slowly progressive course with an FVC reduction of 0.13-0.21 L and often present late in their disease course (King Jr. et al., 2011). A subset of patients has an accelerated variant of the disease, with faster FVC reduction and shortened survival. These patients have a distinct transcription profile and are often cigarette smokers (King Jr. et al., 2011). Overall, the median survival of patients is 3.8 years after diagnosis, but survival may be increasing (Raghu et al., 2014). A proportion of patients, estimated at 5-16% per year, have episodes of acute exacerbation which are not caused by infection or other causes. During these exacerbations, patients have worsened hypoxemic respiratory failure, with an increase in ground-glass opacities and/or consolidations on CT, and most patients die from respiratory failure (Collard et al., 2016; Lederer & Martinez, 2018).

In the last decades, two novel medications have been approved for medical treatment of IPF (Raghu et al., 2015). Pirfenidone is an anti-fibrotic, anti-inflammatory drug that blocks macrophage activation and TGF-β signalling which reduces the expression of collagen type I and smooth muscle actin (Kolb et al., 2017). It has been shown to lead to reductions in FVC decline and 6-minute walk distance decline and pooled analyses of clinical trials have shown reductions in all-cause and IPF-related mortality (King et al., 2014; Nathan et al., 2017). Nintedanib is a receptor tyrosine kinase inhibitor which inhibits platelet derived growth factor receptor, fibroblast growth factor receptor and vascular endothelial growth factor receptor which leads to attenuation of fibroblast growth and motility, myofibroblast transformation and collagen deposition (Kolb et al., 2017). It reduces FVC decline in patients with IPF (Richeldi et al., 2014) and pooled analyses of clinical trial data have shown a mortality benefit for nintedanib (Richeldi et al., 2016). In addition, such analyses have shown an increase in time to first exacerbation for patients on the drug (Richeldi et al., 2016). Both nintedanib and pirfenidone are recommended for IPF therapy (Raghu et al., 2015). Treatment with anti-inflammatory therapy such as corticosteroids is not recommended due to a lack of proven benefit and a clinical trial of prednisone, azathioprine and n-acetylcysteine that showed increased mortality and hospitalisations in the treatment group (Idiopathic Pulmonary Fibrosis Clinical Research Network, 2012; Raghu et al., 2015).

Among nonpharmacological therapies for IPF patients are smoking cessation (Lederer & Martinez, 2018), exercise training (Dowman et al., 2017), supplemental oxygen therapy for patients with resting hypoxemia (Raghu et al., 2011) and lung transplantation for selected candidates, with approximately half of lung transplants in the United States done for interstitial lung disease (Lederer & Martinez, 2018).

1.4 Interstitial lung abnormalities

This chapter is based on prior work published in Axelsson & Gudmundsson, 2021.

1.4.1 Origin and definitions

The extent of the fibrotic changes seen in interstitial lung diseases, as well as their dire prognosis, has led to efforts to find earlier stages of pulmonary fibrosis (Hunninghake, 2019). The idea that the pathobiology of pulmonary fibrosis starts before the onset of clinical symptoms was first presented decades ago (Bitterman et al., 1986). However, identifying people with abnormalities that look like early

forms of ILD was made possible by the increasing use, and lowering costs of, thoracic CT imaging. Efforts have since been made to look for such abnormalities in family members of familial pulmonary fibrosis patients and in cohorts based on the general population (Lederer et al., 2009; Rosas et al., 2007; Washko et al., 2011) using variable terminology and definitions (Gochuico et al., 2008; Guckel & Hansell, 1998; Lederer et al., 2009; Sverzellati et al., 2011; Tsushima et al., 2010).

The term 'interstitial lung abnormalities' (ILA) was coined to define abnormalities like this, first in a publication by Washko et al in 2011 (Washko et al., 2011), and its use has since become widespread in research of possible early forms of ILD (Hunninghake et al. 2013; Putman et al., 2016; C. S. Sack et al., 2017). Studies using the ILA term have still used slightly variable definitions of ILA (Buendia-Roldan et al., 2019; Hoyer et al., 2018; Jin et al., 2013; Mackintosh et al., 2019; Washko et al., 2011; Whittaker Brown et al., 2019). In the paper by Washko et al, ILA were defined as "nondependent changes that affect more than 5% of any lung zone that include ground-glass or reticular abnormalities, diffuse centrilobular nodularity, nonemphysematous cysts, honeycombing and traction bronchiectasis" (Putman et al., 2016; Washko et al., 2011) with lung areas meeting criteria for emphysema not included in these estimations. Changes "present in less than 5% of lung zones consisting of focal or unilateral ground-glass attenuations, focal or unilateral reticulation and patchy ground-glass abnormalities" were defined as indeterminate for ILA (Washko et al., 2011). Other definitions of ILA are similar, but with varying details of the magnitude and exact types of changes encompassed by the term (Buendia-Roldan et al., 2019; Hoyer et al., 2018; Jin et al., 2013; Mackintosh et al., 2019; Washko et al., 2011; Whittaker Brown et al., 2019). Various definitions of ILA in published articles using the ILA term, until the publication of the Fleischner Society consensus statement, are summarized in Table 3.

Table 3 - An overview of definitions of ILA

DEFINITION	COHORT(S)	CITATION
"nondependent changes that affect more than 5% of any lung zone that include ground-glass or reticular abnormalities, diffuse centrilobular nodularity, nonemphysematous cysts, honeycombing and traction bronchiectasis"	COPDGene Framingham Heart Study AGES-Reykjavik Study Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints Study Multi-Ethnic Atherosclerosis Study Copenhagen Comorbidity in HIV Infection cohort	Washko et al., 2011 Hunninghake et al., 2013 Putman et al., 2016 C. S. Sack et al., 2017 Ronit et al 2020
"nondependent ground-glass opacities that affected more than 5% of any lung zone, nondependent reticular abnormality, diffuse centrilobular nodularity with ground glass opacities, honeycombing, traction bronchiectasis, nonemphysematous cysts, or architectural distortion"	A sample from the National Lung Screening Trial, United States	Jin et al., 2013
"nondependent GGO that affected more than 5% of any lung zone, nondependent reticular abnormality, diffuse centrilobular nodularity with GGO, honeycombing, traction bronchiectasis, nonemphysematous cysts, or architectural distortion"	Beijing lung screening cohort	Tan et al., 2016
"Reticular abnormalities, traction bronchiectasis, bilateral independent ground-glass abnormalities, honeycombing, and nonemphysematous cysts"	Shanghai cohort of small cell lung cancer patients	Li et al., 2018

"presence of one or more of the following: (1) septal thickening, (2) honeycombing, (3) ground glass opacity (GGO), and (4) traction bronchiectasis"	United States cohort of seropositive RA	Dong et al., 2018
"ground-glass opacity, honeycombing, reticulation, pleural nodules, centrilobular nodules, paraseptal/subpleural nodules, mosaic attenuation, and mass", present or absent.	Danish Lung Cancer Screening Trial	Hoyer et al., 2018
"the presence of ground-glass abnormalities and reticulations in nondependent lung zones, with an extent above 5% of the total lung"	Aging Lung Program, Mexico	Buendia-Roldan et al., 2019
"evidence of reticular/reticulonodular opacities, honeycombing, fibrosis, or scarring", reported by study radiologists.	National Lung Screening Trial, United States	Whittaker Brown et al., 2019
"presence or absence of the following interstitial abnormalities: reticulation, ground-glass, honeycombing, consolidation, mosaicism, traction bronchiectasis, nodularity and cysts"	Queensland Lung Cancer Screening Study	Mackintosh et al., 2019
"Incidental identification of non-dependent abnormalities, including ground-glass or reticular abnormalities, lung distortion, traction bronchiectasis, honeycombing, and non-emphysematous cysts involving at least 5% of a lung zone in individuals in whom interstitial lung disease is not suspected"	Fleischner Society Position Paper	Hatabu et al., 2020

In the paper by Washko et al, ILA were also classified into four radiologic subtypes based on the location, type and extent of changes; a) centrilobular abnormalities, defined as "predominant centrilobular or peribronchial ground-glass opacities sparing the peripheral lung parenchyma", b) subpleural abnormalities, defined as "reticular, nodular or ground-glass opacities in a predominantly subpleural distribution", c) mixed abnormalities, defined as

"mixed centrilobular and subpleural abnormalities", d) radiographic ILD, defined as "extensive radiographic changes consistent with firm radiographic evidence of interstitial lung disease according to the guidelines of the American Thoracic and European Respiratory Societies" (Putman et al., 2019; Washko et al., 2011). In 2020, a Fleischner Society position paper was published in which a uniform definition for ILA was accepted. This definition is mostly in concordance with the definition used by Washko et al in 2011 (Washko et al., 2011), but the changes previously termed 'centrilobular abnormalities' have been excluded (Hatabu et al., 2020).

The definite fibrosis pattern and the usual interstitial pneumonia (UIP) pattern are specific imaging patterns that have been studied among people with ILA. In these studies, the definite fibrosis pattern has been defined as "pulmonary parenchymal architectural distortion (e.g., traction bronchiectasis, honeycombing) consistent with a fibrotic lung disease (definite fibrosis), which is not limited to those whose imaging pattern is consistent with a usual interstitial pneumonia (UIP) or probable UIP pattern". The UIP pattern definition and the definition of probable UIP are equivalent to those described in IPF diagnostic criteria by the Fleischner Society (Lynch et al., 2018; Putman et al., 2019). Thus, studies of ILA, their subtypes and imaging patterns, are based on precise definitions that stem from established diagnostic criteria for interstitial lung disease.

The prevalence of ILA has been assessed in both general population-based cohorts and cohorts of smokers. In the general population-based cohorts the mean age of participants with ILA was 70 years or older, i.e., in the Age/Gene-Environment Susceptibility-Reykjavik (AGES-Reykjavik) cohort (ILA prevalence 7%) and the Framingham Heart Study (ILA prevalence 7%) (Putman et al., 2016). ILA have also been studied in cohorts of smokers or cohorts from lung cancer screening trials in which the mean age of participants ranged from 60-66 years. In these cohorts, of the range of ILA prevalence has varied. The prevalence of ILA was 4% in the Multicentric Italian Lung Detection trial cohort (Sverzellati et al., 2011), 8% in the Genetic epidemiology of COPD (COPDGene) study (Washko et al., 2011), 8% in the Queensland Lung Cancer Screening Study (Mackintosh et al., 2019), 9% in the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints study (Putman et al., 2016), 10% in a random cohort from the National Lung Screening Trial (Jin et al., 2013), 17% in the Danish Lung Cancer Screening Trial cohort (Hoyer et al., 2020) and 20% in the National Lung Screening Trial (Whittaker Brown et al., 2019).

1.4.2 Epidemiological associations and parallels with interstitial lung disease

Interstitial lung abnormalities have been associated with many of the same features as advanced interstitial lung disease such as IPF. First and foremost, ILA is associated with age and cigarette smoking, both the amount of lifetime smoking and smoking status at the time of ILA measurement. (Putman et al., 2016; Washko et al., 2011). Differences in respiratory function have been shown between smokers with ILA and smokers without ILA, such as increased odds of having less than 80% of expected total lung capacity, decreased lung volume and less emphysema (Washko et al., 2011). Results from a different cohort of patients with COPD have suggested a higher rate of COPD exacerbations as well as significantly greater declines in FEV_1 and FVC among those with ILA compared to others (Lee et al., 2021). ILA have also been associated with possible risk factors of IPF that are not as well established as age and smoking. Among such factors are obstructive sleep apnea, air pollution, and various occupational exposures (J.S. Kim et al., 2017; Rice et al., 2019; C. Sack et al., 2017; C. S. Sack et al., 2017). Patients infected with human immunodeficiency virus are shown to have increased odds of ILA compared to uninfected controls (Ronit et al., 2020). Meanwhile, herpesviruses have received most attention with regards to IPF risk (Chioma & Drake, 2017). The genetic polymorphism that contributes the most to genetic risk of IPF is the rs35705950 promoter polymorphism of the *MUC5B* gene (Seibold et al., 2011). This polymorphism also confers a greatly increased risk of ILA among cohort study participants (Hobbs et al., 2019; Hunninghake et al., 2013). As for other genetic polymorphisms, ILA have been associated with polymorphisms near *DPP9*, *DSP*, *FAM13A*, *MAPT*, *IVD*, *IPO11*, *FCF1P3* and *HTRE1*. Of these, polymorphisms near *DPP9*, *DSP*, *FAM13A*, *MAPT* and *IVD* are previously associated with IPF (Hobbs et al., 2019). Supporting the heritability of ILA risk, family members of patients with pulmonary fibrosis are more likely to have ILA than their peers (Hunninghake et al., 2020). Histopathologic changes commonly seen in IPF, such as fibroblastic foci and honeycombing (Raghu et al., 2011), are more common in lung nodule resections from patients with ILA than in others (Miller et al., 2018). Another study of lung nodule resections showed that people with microscopic fibrotic interstitial changes were more likely to have ILA on CT (Hung et al., 2019). Radiologic features associated with ILA are increased airway thickness, also increased in IPF (Miller et al., 2019), and decreased pulmonary vessel volume on CT, termed vascular pruning (Synn et al., 2021).

In biomarker studies, parallels have been noted between IPF and ILA. Galectin-3 and matrix metalloproteinase-7 (MMP-7) are protein markers associated with IPF by elevations in measurements in bronchoalveolar lavage fluid and/or serum (Mackinnon et al., 2012; Song et al., 2013; Tzouvelekis et al., 2017). Measures of these proteins are elevated among people with ILA, and they are associated with reductions in spirometry values and/or prognosis (H.F. Armstrong et al., 2017; Buendia-Roldan et al., 2019; Ho et al., 2016). The blood levels of growth differentiation factor 15 (GDF-15), a member of the TGF-β superfamily that has elevated expression in IPF lungs (Zhang et al., 2019), are elevated in patients with ILA, as well as aging markers tumor necrosis factor α receptor II, interleukin-6, C-reactive protein, and insulin (Sanders et al., 2021). Several other markers in peripheral blood have been associated with ILA. Among these are the adhesion molecule sICAM-1 (McGroder et al., 2019), the inflammatory regulating adipokine resistin (Buendía-Roldán et al., 2021) and the rheumatoid arthritis-related autoantibodies rheumatoid factor and anti-cyclic citrullinated peptide among ever smokers (Bernstein et al., 2016). This association of resistin was not found in another study focusing on adipokines, however, an association of leptin with ILA among never smokers was found (J.S. Kim et al., 2020). As for other types of markers, several metabolite changes and upregulation of specific microRNAs, especially miR-193a-5p and miR-502-5p are associated with ILA (Ortiz-Quintero et al., 2020; Tan et al., 2016). Low levels of vitamin D are also linked with increased prevalence of ILA 10 years after vitamin D measurement (S. M. Kim et al., 2018). The use of the medication class statins (3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitors), especially hydrophilic statins (Xu et al., 2012), have also been linked with ILA prevalence. Lastly, family members of familial interstitial pneumonia patients that have ILA have shorter telomeres in mononuclear cells than family members without ILA (Salisbury et al., 2020).

1.4.3 Other definitions of ILD precursors

Automated measures have also been used to define possible precursors of interstitial lung disease. High attenuation areas (HAAs) are one such quantitative measure of interstitial changes on CT scans. HAAs are defined as areas of the lung with an attenuation value of -600 to -250 Hounsfield Units (Lederer et al., 2009). As is the case for ILA, analogous associations have been found between risk factors of HAAs and risk factors of ILD. Like the qualitatively assessed ILA, the quantitatively measured HAAs are associated with cigarette smoking (Lederer et al., 2009), nitrogen oxide exposure (a measure of air pollution), a reduction in 6-minute walk distance, lower forced vital capacity on spirometry, dyspnea on

exertion, ILD diagnoses and hospitalizations, and higher respiratory and all cause-mortality (Easthausen et al., 2020; Podolanczuk, Oelsner et al., 2017; Podolanczuk et al., 2016; C. Sack et al., 2017).

HAAs have been found to be associated with various other biological markers and measurements. In a genome-wide association study, polymorphisms near several genes (*GNPDA2*, *ZNF664*, *FAM101A*, *DAAM1, PFKP*, *FOXP4*) were associated with HAAs (Manichaikul et al., 2017). Among the biological markers associated with HAAs are MMP-7, interleukin-6, the adhesion molecules sICAM-1, sVCAM-1 and P-selectin, collagen biomarkers, higher levels of the adipokine resistin and lower levels of the adipokine adiponectin (J.S. Kim et al., 2020; Madahar et al., 2018; McGroder et al., 2019; Podolanczuk et al., 2016). Autoantibodies that are markers of rheumatoid arthritis, i.e., rheumatoid factor IgM and IgA and anti-cyclic citrullinated peptide, are also associated with HAAs among ever-smokers (Bernstein et al., 2016). Vitamin-D deficiency was associated with HAAs and HAA progression in the same study that linked it to ILA prevalence (S. M. Kim et al., 2018). HAAs have also been associated with lipoproteins, mainly high-density lipoprotein-C and ApoA-1, but associations with ILA were not tested (Podolanczuk, Raghu et al., 2017). When the relations of arsenic exposure, as measured by the surrogate measure that is amount of rice in diet and blood levels of D-dimer with HAAs and ILA were assessed, an association was found with HAAs but not with ILA (J.S. Kim et al., 2021; Sanchez et al., 2019). Lastly, HAAs are significantly associated with ILA, although not strongly predictive of the visually assessed phenotype as shown by positive predictive values of 19% and 13% in two cohorts. (Kliment et al., 2015; Podolanczuk et al., 2016). HAAs were not associated with the *MUC5B* polymorphism (rs35705950) strongly related with IPF and ILA (Kliment et al., 2015).

Other methods of quantitative detection of early interstitial lung changes exist. One such method, in which the lung is divided into areas and local histogram analysis of image density in addition to distance from the pleura is used to quantify such changes, correlates fairly well with ILA (area under the receiver operating characteristic curve [AUROC] 0.82, sensitivity 87.8%, specificity 57.5%) and fibrotic ILA (AUROC 0.89) (Ash, Harmouche, Ross, et al., 2017). Another machine-learning method using an ensemble of convolutional neural networks predicted ILA with an AUROC of 0.86 (Bermejo-Peláez et al., 2020). Interstitial features assessed by the prior method are associated with the rs35705950 polymorphism near *MUC5B*, spirometry measures, a measure of quality of life and mortality (Ash, Harmouche, Putman, et al., 2017). Among

patients that underwent lung cancer resection, those that had fibrosis-like changes detected using yet another method had worse disease-free survival (Iwasawa et al., 2019) and finally, a study of another machine-learning algorithm designed to identify normal-appearing high attenuation areas showed that such areas were associated with spirometry impairments, shorter distance on a 6-minute walking test, C-reactive protein and sICAM1 and mortality during follow-up (Harmouche et al., 2019).

1.4.4 Subtypes and imaging patterns

ILA have been classified into several subtypes, as previously described, among which some of the key associations of interstitial lung abnormalities vary. The most common ILA subtype among the four subtypes described in Washko et al. (2011) is subpleural abnormalities. The reduction in lung volumes seen among people with ILA is greatest among people with the radiographic ILD subtype, followed by subpleural abnormalities and mixed abnormalities and the centrilobular subtype has shown the least reduction in lung volumes (Washko et al., 2011). Associations with the *MUC5B* promoter polymorphism known to confer IPF risk (rs35705950) are also shown to vary among ILA subtypes. The subtype that is most significantly associated with the *MUC5B* variant is radiographic ILD (odds ratio [OR] 4.4), with subpleural and then mixed abnormalities following (ORs 2.6 and 1.5). Centrilobular abnormalities meanwhile, have not been shown to be associated with said *MUC5B* polymorphism (Putman, Gudmundsson et al., 2017). As the centrilobular abnormalities do not consistently associate with ILD risk factors and symptoms, they were excluded from the consensus definition of ILA presented by the Fleischner Society (Hatabu et al., 2020).

Associations of ILA have also been found to vary based on the presence of specific imaging patterns, i.e., the 'definite fibrosis' pattern and the UIP pattern. Both these features conferred increased risk of mortality in a longitudinal study (Putman et al., 2019). The rs35705950 polymorphism near the *MUC5B* gene is more strongly associated with fibrotic ILA than with non-fibrotic ILA (respective ORs 3.0 vs 1.8) and more strongly with the possible UIP or UIP patterns (ORs 2.7 and 4.1) than with ILA without the UIP pattern (OR 1.4) (Putman, Gudmundsson et al., 2017).

1.4.5 Clinical outcomes and progression

Efforts have been made to study the progression of ILA in longitudinal studies. In these studies, progression has been based on visual comparison of participant

images. In the AGES-Reykjavik study and the Framingham Heart Study, images taken five years apart were simultaneously assessed for progression, with progression graded on the following five-point scale: definite regression, probable regression, no change, probable progression, and definite progression. Progression was defined as "increase in lung areas affected with nondependent ground-glass, reticular abnormalities, diffuse centrilobular nodularity, nonemphysematous cysts, honeycombing, or traction bronchiectasis, or a new appearance of at least one such abnormality" (Araki et al., 2016; Putman et al., 2019; Rice et al., 2019; Synn et al., 2021). Studies of participants of the National Lung Screening Trial and the Queensland Lung Cancer Screening study defined progression by simultaneous consensus comparison of images taken two years apart, grading images from resolution or improvement to progression, with no further definition of progression provided (Jin et al., 2013; Mackintosh et al., 2018). Increases in symptoms or reductions in spirometric measurements have only been included in the definition of ILA progression in one cohort, a lung aging study at the Mexican National Institute of Respiratory Diseases (Buendía-Roldan et al., 2021). In these studies, 73-76% of people who had or developed ILA had progressed over five years of follow-up (Araki et al., 2016; Putman et al., 2019). While studies assessing follow-up over two years have shown progression in a much lower percentage of participants, it must be considered that the ILA definitions in these studies are not perfectly identical (Jin et al., 2013; Mackintosh et al., 2019). In studies of the risk factors of ILA progression, the rs35705950 polymorphism near *MUC5B*, smoking at study entry, increased age, exposure to elemental carbon, and an increased level of pulmonary vessel pruning on CT have all been associated with progression while the suggested association of body mass index has been inconsistent (Araki et al., 2016; Putman et al., 2019; Rice et al., 2019; Synn et al., 2021). Among subtypes and imaging patterns, subpleural ILA, fibrotic ILA, and the UIP pattern have been related to progression of ILA (Jin et al., 2013; Putman et al., 2019). Additionally, specific imaging features (e.g., nonemphysematous cysts, subpleural reticular markings, traction bronchiectasis) confer increased progression risk but centrilobular nodules confer decreased risk of progression (Putman et al., 2019). In studies examining outcomes of progressors, progression is associated with increased FVC decline and mortality (Araki et al., 2016; Putman et al., 2019).

The presence of ILA has been associated with adverse outcomes. Of most relevance is the increased all-cause mortality that ILA have been linked with in multiple studies and cohorts (Hoyer et al., 2018; Putman et al., 2016). In the Icelandic AGES-Reykjavik cohort, ILA were found to be specifically associated

with mortality from a respiratory cause and with mortality from pulmonary fibrosis (identified at the time of death) (Putman et al., 2016). An increased risk of ILD diagnosis was found among participants in a lung cancer screening study that had ILA (Hoyer et al., 2020). Associations of ILA with COPD have varied between cohorts. One cohort study has found ILA to associate with decreased odds of COPD and its severity (Washko et al., 2011), while later studies have either not shown an association with COPD or associated ILA with increased risk of COPD or group B of the Global Initiative for Chronic Obstructive Lung Disease classification of COPD specifically (Bozzetti et al., 2016; Hoyer et al., 2020; Ohgiya et al., 2017). Prevalence of paraseptal emphysema, a subtype of emphysema whose clinical significance is largely unknown, correlates with ILA (Araki et al., 2015). Emphysema aside, a study has shown that critically ill patients with ILA are at increased risk of acute respiratory distress syndrome compared to those without ILA (Putman, Hunninghake et al., 2017). Participants in lung cancer screening studies or lung cancer patients with ILA also have been found to be at greater risk of lung cancer and/or mortality from lung cancer compared to those without ILA and ILA are associated with increased complications after lung cancer surgery (Araki et al., 2019; Chubachi et al., 2017; Gu et al., 2019; Hoyer et al., 2020; Hoyer et al., 2018; Im et al., 2019; Nishino et al., 2015; Whittaker Brown et al., 2019; Wille et al., 2016). Additionally, patients with small cell lung cancer and ILA are at increased risk of severe radiation pneumonitis after radiation therapy (Li et al., 2018) and those with non-small cell lung cancer and ILA are at increased risk of immune checkpoint inhibitor-induced interstitial lung disease (Nakanishi et al., 2019). Finally, the prevalence of ILA and their progression has been shown to be high among patients with rheumatoid arthritis. Among these patients, ILA have been shown to be associated with increased respiratory symptoms, restriction on spirometry, decreased exercise capacity and increased rheumatoid arthritis severity (Dong et al., 2018; Doyle et al., 2014). ILA have additionally been associated with respiratory failure, pneumonia, lung abscesses or pleural empyema as well as decreased 6-minute walking distance among smokers (Doyle et al., 2012; Hoyer et al., 2020). The associations of ILA and HAAs are summarized in Table 4 and associations of ILA progression are listed in Table 5.

Table 4 - Associations of interstitial lung abnormalities and high attenuation areas

	ASSOCIATION WITH ILA	ASSOCIATION WITH HAA
MAJOR RISK FACTORS AND COMORBIDITIES OF IPF (LEDERER & MARTINEZ, 2018)		
Age	Putman et al., 2016; Washko et al., 2011	-
Pack-years of smoking	Putman et al., 2016; Washko et al., 2011	Lederer et al., 2009
Airway pollution and exposures	Rice et al., 2019; C. Sack et al., 2017; C. S. Sack et al., 2017	C. Sack et al., 2017
Obstructive sleep apnea	J. S. Kim et al., 2017	J. S. Kim et al., 2017
Gastroesophageal reflux	-	-
The rs35705950 promoter polymorphism of the MUC5B gene	Hobbs et al., 2019; Hunninghake et al., 2013	-
Polymorphisms in *DPP9*, *DSP*, *FAM13A*, *MAPT* and *IVD*	Hobbs et al., 2019	-
Matrix metalloproteinase-7	H. F. Armstrong et al., 2017; Buendia-Roldan et al., 2019	Choi et al., 2020; Podolanczuk et al., 2016
Galectin-3	Ho et al., 2016	-
Lung cancer (Le Jeune et al., 2007)	Chubachi et al., 2017; Whittaker Brown et al., 2019; Wille et al., 2016	-
OTHER FEATURES		

Genetic polymorphisms and biomarkers		
Polymorphisms near *IPO11*, *FCF1P3* and *HTRE1*	Hobbs et al., 2019	-
Polymorphisms near *GNPDA2*, *ZNF664*, *FAM101A*, *PFKP*, *SAMD4A, GYPC*, *FUT10*, *GNPDA2*, *PFKP*, *SLC45A*, *FOXP4, ALCAM*	-	Manichaikul et al., 2017
Tumour necrosis factor-α receptor II	Sanders et al., 2021	-
Growth differentiation factor 15 (Zhang et al., 2019)	Sanders et al., 2021	-
Interleukin-6	Sanders et al., 2021	Choi et al., 2020; Podolanczuk et al., 2016
C-reactive protein	Sanders et al., 2021	-
Insulin	Sanders et al., 2021	-
Resistin	Buendía-Roldán et al., 2021	J.S. Kim et al., 2020
Soluble intracellular adhesion molecules	McGroder et al., 2019	McGroder et al., 2019
Metabolite alterations	Tan et al., 2016	-
Upregulation of microRNAs	Ortiz-Quintero et al., 2020	-
Leptin	J.S. Kim et al., 2020	-
Collagen biomarkers	-	Madahar et al., 2018
Lower levels of high-density lipoprotein-C and ApoA-1	-	Podolanczuk, Raghu et al., 2017
Lower levels of adiponectin	-	J.S. Kim et al., 2020
Higher levels of D-dimer	-	J.S. Kim et al., 2021

Daily rice consumption	-	Sanchez et al., 2019
Clinical features		
Reduced total lung capacity	Washko et al., 2011	Kliment et al., 2015
Lesser amount of emphysema	Washko et al., 2011	Podolanczuk, Oelsner et al., 2017
Reduced exercise capacity or exertional dyspnea	Doyle et al., 2012	Easthausen et al., 2020; Podolanczuk et al., 2016
Paraseptal emphysema	Araki et al., 2015	-
Increased airway wall thickness	Miller et al., 2019	-
CT vascular pruning	Synn et al., 2021	-
Vitamin D deficiency	S. M. Kim et al., 2018	S. M. Kim et al., 2018
Statin use	Xu et al., 2012	-
Human immunodeficiency virus infection	Ronit et al., 2020	-
Human immunodeficiency virus viral load	-	Leader et al., 2016
Rheumatoid arthritis or associated autoantibodies	Bernstein et al., 2016; Dong et al., 2018; Doyle et al., 2014	Bernstein et al., 2016
Clinical outcomes		
All-cause mortality	Hoyer et al., 2020; Hoyer et al., 2018; Putman et al., 2016	Choi et al., 2020; Podolanczuk et al., 2016
Death from a respiratory cause or ILD	Putman et al., 2016	Choi et al., 2020; Podolanczuk, Oelsner et al., 2017
ILD diagnosis or hospitalization	Hoyer et al., 2020	Podolanczuk, Oelsner et al., 2017

Acute respiratory distress syndrome	Putman, Hunninghake et al., 2017	-
Mortality due to cancer and shorter survival of cancer patients	Araki et al., 2019; Gu et al., 2019; Hoyer et al., 2020; Hoyer et al., 2018; Nishino et al., 2015; Whittaker Brown et al., 2019	-
Treatment complications among lung cancer patients	Im et al., 2019; Li et al., 2018; Nakanishi et al., 2019	-
Respiratory infections (pneumonia, pleural empyema, lung abscesses, respiratory failure among smokers, COPD exacerbations)	Hoyer et al., 2018; Lee et al., 2021	-

Based on prior work published in Axelsson & Gudmundsson, 2021.

Table 5 - Associations with progression of interstitial lung abnormalities

FEATURE ASSOCIATED WITH ILA PROGRESSION	CITATION(S)
Age	Araki et al., 2016; Putman et al., 2019
The rs35705950 promoter polymorphism of the MUC5B gene	Araki et al., 2016; Putman et al., 2019
Smoking at the time of study	Araki et al., 2016
Subpleural ILA subtype	Putman et al., 2019
Imaging features (subpleural reticular markings, nonemphysematous cysts, traction bronchiectasis, lower lobe predominance, less centrilobular nodules)	Putman et al., 2019
The definite fibrosis and usual interstitial pneumonia imaging patterns	Putman et al., 2019
CT vascular pruning	Synn et al., 2021
A decline in forced vital capacity (FVC)	Araki et al., 2016
All-cause mortality	Araki et al., 2016; Putman et al., 2019

Based on prior work published in Axelsson & Gudmundsson, 2021.

1.4.6 Relevance and future directions

So, as detailed in previous chapters, interstitial lung abnormalities are analogous with interstitial lung disease in many ways. They associate with the main epidemiologic, environmental, and genetic risk factors of ILD, some histopathological characteristics and many biomarkers (H. F. Armstrong et al., 2017; Ho et al., 2016; Hobbs et al., 2019; Miller et al., 2018; Rice et al., 2019; Washko et al., 2011). These findings are different between subtypes and imaging patterns, with the main theme being that stronger associations with risk factors and greater odds of dire outcomes are seen with more extensive abnormalities (Washko et al., 2011). Increased risk of ILA progression is also noted with more extensive abnormalities and the main genetic ILD risk factor, rs35705950 (Araki et al., 2016; Putman et al., 2019). Also, the most extensive abnormalities included in the ILA term are of a magnitude that can be consistent with the existence of ILD, such as the probable UIP and the definite UIP patterns. Overall, this supports the proposal that at least some of the changes that ILA term includes are precursors or early stages of ILD (Hunninghake, 2019). Despite these similarities, the vast difference in prevalence leads to the inevitable conclusion that only a fraction of people with ILA will ultimately develop ILD (Hutchinson et al., 2015; Lederer et al., 2009; Nalysnyk et al., 2012; Putman et al., 2016; Raghu et al., 2016; Raghu et al., 2014). So, while individuals with ILA have changes that are likely to, at least in some cases, be related to ILD, it is important to not equate ILA with ILD. This is because only a minority of changes termed ILA will ever progress to a level that can be called ILD, with self-sustaining, widespread disruption of lung tissue structure. Therefore, it can also be presumed that the disease processes that maintain ILD are not fully active in those with ILA, or at least only in a minority of the group.

However, ILA are associated with other clinically relevant endpoints such as a considerable increase in risk of all-cause mortality (Putman et al., 2016) and other lung diseases with high mortality such as lung cancer and acute respiratory distress syndrome (Hoyer et al., 2020; Putman, Hunninghake et al., 2017). This has added to the current understanding of the presumed outcomes and of the possible clinical relevance of ILA. A graphic depicting the current understanding of the clinical endpoints of ILA is shown in Figure 4.

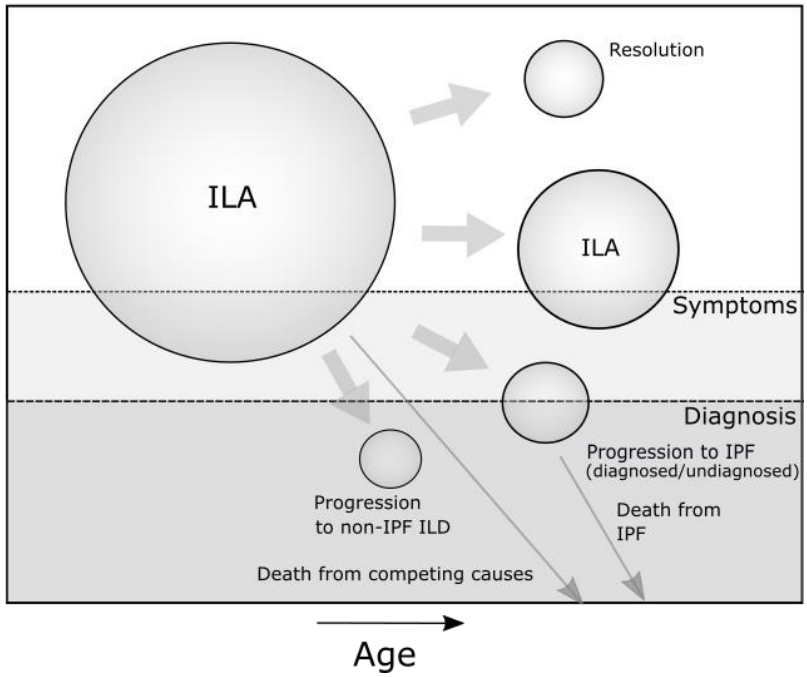

Figure 4 - Potential outcomes of ILA
Reproduced with permission from Hunninghake, 2019, licensed under CC BY-NC 4.0.

Studies of ILA could be valuable in multiple ways. First, they could lead to earlier detection, and possibly earlier treatment, of patients with ILD. This has been an aim of ILA research since the term was coined (Hunninghake, 2019). However, it is not yet known whether and in which populations such earlier intervention would improve outcomes as, as discussed earlier, only a fraction of those with ILA will ever develop ILD and require anti-fibrotic therapy according to current treatment indications. Therefore, finding those with ILA at greatest risk of ILD by finding markers or predictive factors of such progression that add to the current knowledge of radiologic patterns that predict ILA progression (Putman et al., 2019), will be of paramount importance for this aim and for the utility of the ILA term to reduce the morbidity related to ILD. Second, the study of imaging features of ILA that likely represent a precursor to, or the least severe end of a spectrum of, ILD represents an opportunity for researchers to learn about the many unknowns in ILD pathogenesis, especially in its early stages. The study of epidemiological and genetic risk factors of ILA could thus increase understanding of early ILD. For studies of ILA to reach their potential, a uniform definition such as the one proposed by the Fleischner Society (Hatabu et al., 2020) needs to become standard so researchers can compare and meta-analyse findings from different studies. This viewpoint can be extended to subtypes and

imaging patterns of ILA that are studied in different cohorts. Lastly, automated methods to detect ILD precursors on CT could greatly increase the utility of the study of ILD precursors as they aid detection of such changes in large groups of patients.

1.5 Aging and aging-related disease

Interstitial lung diseases, especially the idiopathic interstitial pneumonias, are diseases of middle-aged and elderly people ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002) and ILA have strong epidemiological associations with age (Lederer et al., 2009; Washko et al., 2011). These epidemiological associations raise the question of whether fibrotic changes of the lung could share biological mechanisms with aging, bringing the biology of aging into focus.

Aging is a complex process on the cellular, molecular, and biochemical levels that leads to declining physiological capability and increased risk of various diseases and ultimately, death. On the cellular and molecular levels, several key characteristics or hallmarks of aging have been described. Among these are epigenetic alterations such as altered histone and DNA methylation, impaired protein homeostasis via altered function of chaperone proteins such as heat shock proteins, deregulation of hormonal response to nutrition, altered cellular communication with an increase in inflammatory mediators, a lack of tissue regeneration capacity as marked by a deficiency in tissue stem cells and finally a decreased integrity of the genome, by shortening telomeres and other changes (López-Otín et al., 2013).

Telomere shortening is a hallmark of the aging process of cells (López-Otín et al., 2013). Telomeres are guanin-rich DNA-protein complexes that are found at the ends of eukaryotic chromosomes and protect them from degradation. During cell division, these DNA rows shorten due to their incomplete replication during DNA synthesis. This in turn causes the cells' demise. However, in immortal cell lines such as embryonic stem cells, telomeres are maintained by a specialised reverse transcriptase complex, telomerase, that elongates telomeres and thus induces their immortality (Hiyama & Hiyama, 2007). Mutations in genes encoding for telomerase cause a distinct group of diseases, termed the short telomere syndromes, in which dyskeratosis congenita is the best-known disorder. Some of the common pathological phenotypes across the telomere syndromes are premature hair greying, aplastic anaemia, liver cirrhosis, various types of cancer and pulmonary fibrosis (Armanios & Blackburn, 2012). Even in

individuals without mutations in telomere genes, telomere length is inversely associated with mortality (Boonekamp et al., 2013) and animal experiments have shown that changes to telomere length incur changes in the animals' lifespan (de Jesus et al., 2012). From these data, it is evident that the function and appropriate renewal of telomeres is integral to cells' anti-aging properties. Therefore, their length has been suggested as a biomarker of aging. Still, the epidemiological associations of telomere length with diseases and functional measures sensitive to aging have been inconsistent in prior studies (Mather et al., 2011; Sanders & Newman, 2013).

Age-related cellular changes have been associated with several diseases that are strongly correlated with aging. Therefore, they could partly explain the demographic distribution of these illnesses. As previously stated, robust links exist between telomerase dysfunction and a handful of diseases to the extent that they have been called "short telomere syndromes". Among these diseases are dyskeratosis congenita, pulmonary fibrosis, aplastic anaemia and rarer diseases such as the Hoyeraal-Hreidarsson syndrome (Armanios & Blackburn, 2012). Absence of protein stability as manifested by the presence of misfolded or aggregated proteins is a major part of the pathogenesis of neurodegenerative diseases strongly associated with aging, such as Parkinson's disease, Alzheimer's disease and Huntington's chorea (Powers et al., 2009) while increased genomic instability due to ineffective DNA repair mechanisms or defects in the nuclear lamina causes various rare progeroid syndromes (Hoeijmakers, 2009; Liu et al., 2005; López-Otín et al., 2013). As for the association of cellular senescence with aging-related disease, mutations in the *INK4a/ARF* locus, encoding for proteins implicated in cellular senescence, have an extraordinarily strong association with aging-related disease (Jeck et al., 2012). Inflammation, meanwhile, is a component in the pathogenesis of several aging-related diseases and conditions, perhaps most notably atherosclerosis and the metabolic syndrome (Barzilai et al., 2012; Tabas, 2010). Deregulated nutrient sensing and hormonal changes are pillars of a metabolic syndrome that is associated with diseases such as diabetes and cardiovascular disease, as well as aging. These diseases are strongly associated with aging, however, the pathways by which systemic inflammation due to aging ("inflammaging") and systemic inflammation due to over-intake of nutrients (termed "metaflammation") contribute to the metabolic syndrome are largely similar and interconnected (Barzilai et al., 2012; Franceschi et al., 2018). The molecular hallmarks of aging are therefore tied with many of the major aging-associated diseases.

Neoplasms represent a group of diseases that are strongly associated with aging as a whole and cancer and aging share many molecular hallmarks, albeit with diverging results at the cellular level. For cancer, the mechanisms make the cells hyperactive and increase mutation rate while the aging process causes hypoactivity and slowed cellular mutation rate (Aunan et al., 2017). Genomic instability is a hallmark of both aging and cancer with many progeroid syndromes having much increased rates of cancer at a young age. Still, the mutations that make a cell a cancer cell are mostly advantageous to the cell's growth. This is contrary to normal aging, in which genomic instability is harmful to the cell's survival (Aunan et al., 2017). Telomeres also feature differently in the biology of aging and cancer. The progressive shortening of telomeres with cellular division contributes to cellular senescence in aging but the constitutional expression of telomerase plays a vital role in the "immortality" of cancer cells (Shay, 2016). Believed to play a part in the biology of aging, loss of proteostasis, autophagy and mitochondrial dysfunction causing oxidative stress has additionally been implicated in carcinogenesis by several mechanisms (Dou & Zonder, 2014; Kudryavtseva et al., 2016; Rappa et al., 2012; White, 2015). Lastly, a hypothesis regarding cellular senescence and the senescence-associated secretory phenotype suggests that it can both drive degenerative aging-related disease and protect against tumorigenesis as well as incite some components of neoplastic transformation such as EMT and tumour invasion and angiogenesis, depending on context (Campisi, 2013).

1.6 Aging and interstitial lung disease

Some data exist on the relationship between the mechanisms of aging and the pathobiology of interstitial lung disease. Firstly, regarding telomeres, mutations in telomerase genes are associated with familial and sporadic IPF, even when patients do not show other symptoms of dyskeratosis congenita or other telomerase syndromes (Armanios et al., 2007; Tsakiri et al., 2007). Furthermore, patients with IPF without telomerase mutations have disproportionately short telomeres when compared to the normal population (Alder et al., 2008; Cronkhite et al., 2008). This applies both to telomere length in peripheral blood leukocytes and in the alveolar epithelium (Alder et al., 2008). A recent study exploiting Mendelian randomisation has implied that a causal relationship exists between telomere shortening and IPF (Duckworth et al., 2021).

Other aging mechanisms have also been implicated in pulmonary fibrosis (Selman et al., 2019). Cellular senescence is suggested as a major driver of pulmonary fibrosis. Markers of cellular senescence are elevated in fibrotic lung disease and senescent cells are believed to contribute to the pathobiology of the

disease as the senescent cellular phenotype is suggested as one of the drivers of interstitial lung disease (Schafer et al., 2017). Some of the drivers of cellular senescence are telomere attrition and oxidative stress, possibly in part via their effect on the regulation of microRNAs, and some of the molecules secreted by senescent cells are important mediators of fibrosis, e.g., VEGF, TGF-β and matrix metalloproteinases (Barnes, et al., 2019; Disayabutr et al., 2016; Schafer et al., 2017). Some evidence exists for increased oxidative stress in fibrotic lungs, as oxidative DNA damage is increased in alveolar epithelial cells of patients with pulmonary fibrosis (Kuwano et al., 2003) and mitochondrial dysfunction has been noted in the same cells (Bueno et al., 2015). The lungs of ILD patients have also been suggested to have defective autophagy (Patel et al., 2012) and to have a distinct genetic methylation pattern (Yang et al., 2014). These data suggest that some of the mechanisms of aging are associated with interstitial lung disease and its pathogenesis.

2 Aims

The strong epidemiological relationship between aging and both interstitial lung abnormalities and advanced interstitial lung disease supports the theory that aging itself or aging related processes could be causally implicated in the pathobiology of interstitial lung disease. This is further supported by parallels that exist between the biology of aging and the pathobiology of advanced interstitial lung disease. However, knowledge is lacking whether these parallels extend to earlier, less extensive forms of pulmonary fibrosis, such as ILA.

One of the most pressing needs in research of interstitial lung abnormalities and related conditions is the discovery of biomarkers that could aid in the stratification of individuals with regards to the progression of the radiographic changes as well as the risk of mortality (Hatabu et al., 2020). The importance of such markers stems from the difference in prevalence between early forms of pulmonary fibrosis such as ILA (Putman et al., 2016) and advanced fibrotic lung diseases such as IPF (Hutchinson et al., 2015). Aging-related markers could be valuable as such biomarkers.

Lastly, the aging process is associated with frailty, functional decline, and a multitude of diseases, including degenerative diseases and malignancies (López-Otín et al., 2013). With the ever-increasing use of CT imaging in clinical practice (Brenner & Hall, 2007) and a recently published unifying definition of ILA for clinical radiologists (Hatabu et al., 2020), it is likely that the number of reports of ILA to practicing clinicians will increase. Therefore, it is important to clarify whether they confer increased risk of aging-related impairment and disease.

To approach these questions, the aims of the thesis were as following:

- Aim 1: To explore the relationship of interstitial lung abnormalities with functional status (Paper I).
- Aim 2: To explore the relationship of interstitial lung abnormalities with pulmonary and non-pulmonary malignancies (Paper II).
- Aim 3: To explore the relationship of interstitial lung abnormalities with markers of cellular senescence (Paper III).
- Aim 4: To explore the relationship of interstitial lung abnormalities with telomere length in leukocytes (Paper IV).

3 Materials and methods

3.1 Study population

3.1.1 The AGES-Reykjavik Study

3.1.1.1 Study design

The roots of the Age/Gene-Environment Susceptibility-Reykjavik (AGES-Reykjavik) Study lie in its forerunner, the Reykjavik Study. The Reykjavik Study was a study of 30,795 inhabitants of Reykjavik, Iceland, initiated in 1967 with the overarching aim of discovering risk factors of coronary heart disease. The participants were born in 1907-1934 and were invited for study participation in five groups in five phases in the years 1967 through 1996 with a subgroup of participants invited often, in multiple or all study phases, to investigate the effect of repeated health surveying. The focus of the health evaluations in the Reykjavik Study was health questionnaires and evaluation of cardiac health with electrocardiograms as well as laboratory analysis of blood samples, lung spirometry and anthropometric measurements. Some of the key findings from the Reykjavik Study concern the incidence, risk factors for and prognosis of unrecognized myocardial infarction, blood-based markers that are predictive of myocardial infarction and the association of smoking with cancer incidence (G. Bjornsson et al., 1968; O. Bjornsson et al., 1979).

In 2002, there were 11,549 members of the Reykjavik Study research cohort alive. From this group, the 5,764 participants of the AGES-Reykjavik Study were sampled, prioritising those who had the most extensive phenotyping in the Reykjavik Study. The majority (58 percent) of these participants were female. The phenotyping in AGES-Reykjavik took place in a single wave from 2002-2006, with each participant undergoing examinations in three clinic visits within a four-to-six-week time frame.

In these three clinic visits, participants were subject to extensive examinations. These included detailed questionnaires on prior medical history and history of medical symptoms, occupational, residential, and social history as well as participants' exercise and leisure habits and medication history. Participants had anthropometric measurements of height, weight, and abdominal circumference, tests of total body fat, measurements of isometric muscle strength as well as measurements of gait speed and proprioception. They also underwent

measurements of pulse and blood pressure, an electrocardiogram, neuropsychological testing, a three-step evaluation for dementia, depression screening, visual acuity testing, assessment of intraocular pressure and audiometry. Participants had various laboratory measurements of blood, urine, and salivary samples. Samples of plasma, serum, urine, and white blood cells were then stored for the possibility of future laboratory analyses and genotyping. Additionally, participants underwent imaging studies which included ultrasonography of their carotids, CT imaging of coronary and aortic calcium, retinal photographs, magnetic resonance imaging of the brain, CT imaging of the lumbar spine and thighs for evaluation of bone quality and muscle mass and photographs of hands for osteoarthritis evaluation. Some but not all participants in the cohort underwent arterial tonometry, echocardiography, magnetic resonance imaging of the heart and pulmonary function testing. A small minority of patients had a home examination, due to inability or unwillingness to come to clinic. Lastly, information on AGES-Reykjavik participants has also been collected through registries. Mainly, data on vital status and causes of death is collected via the Icelandic Directorate of Health and data on disease status is collected via hospital records using the International Classification of Diseases coding system (Harris et al., 2007).

Follow-up examinations were organised, starting in 2007. Of 5,245 AGES-Reykjavik participants, 3,411 participated in the follow-up. These follow-up examinations finished in 2011. The time between the first examination and the follow-up examination was around five years for each participant. In the follow-up examination, participants underwent a similar range of clinical examinations as in the first round of examinations. They had anthropometric measurements, measurements of their physical capacity, electrocardiograms, a coronary CT, neuropsychological testing and dementia evaluation, CT scanning with regards to bone and muscle mass and magnetic resonance imaging of their brain, as well as extensive laboratory profiling of blood and urine samples. Therefore, comparisons of a wide range of phenotypes are possible for a majority of AGES participants (Harris et al., 2007).

3.1.1.2 Study-specific phenotyping

A large majority of participants that were evaluated during the original AGES-Reykjavik examination as well as the follow-up examination underwent CT scans of their thorax, with the aim to obtain an assessment of participants' calcium load in their coronaries and thoracic aorta. The scans were separate with a four-row CT scanner (Sensation, Siemens Medical System, Erlangen, Germany) using 2.5 mm thick slices, 140 kVp, 50 mAs and 0.361 second scan time. The scans were

acquired during maximum inhalation and covered an estimated 95% of participants' lung fields. The only lung parts omitted were the superior parts of the apexes.

These two CT scans were evaluated jointly when ILA status was assessed. For the original assessment ILA were defined according to the definition first published in Washko et al (2011), i.e., nondependent changes affecting more than 5% of any lung zone, including reticular or ground-glass abnormalities, diffuse centrilobular nodularity, nonemphysematous cysts, honeycombing or traction bronchiectasis, with images with focal ground-glass attenuation or reticulation under 5% of any lung zone considered indeterminate. The assessment of CT scans for ILA status was done at the Brigham and Women's Hospital, Boston, Mass. AZE VirtualPlace Fujin Raijin workstations (AZE, Tokyo, Japan) were used to read axial images with a window level of -700 Hounsfield Units and a window width of 1500 Hounsfield Units. The readers of the images were three doctors, two chest radiologists and a pulmonologist that were blinded to participant specific information. This was done using a sequential reading method. First, all scans were classified as showing no evidence of ILA, being indeterminate for ILA or showing ILA. Next, another reader reviewed all scans labelled as ILA, indeterminate scans and 20% of the normal scans. Lastly, the third reader added a final vote to those scans discordantly scored (Putman et al., 2016).

CT scans with ILA were then characterized as belonging to one of four specific subtypes; centrilobular abnormalities defined as predominantly centrilobular or peribronchial ground glass opacities or nodules that spare the peripheral lung, subpleural abnormalities defined as predominantly subpleural abnormalities that had nodular, reticular or ground-glass appearance, mixed abnormalities defined as a mix of centrilobular and subpleural abnormalities and the radiographic ILD subtype, defined as changes that meet criteria from the American Thoracic and European Respiratory Societies for radiographic evidence of ILD ("American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias.", 2002). For scans with ILA the presence or absence of certain imaging features was registered. These were ground-glass opacities, subpleural reticulation, centrilobular nodularity, honeycombing, traction bronchiectasis and nonemphysematous cysts. The location and distribution of these features was registered as well. In addition to this, images from the participants with ILA were classified based on the presence of two imaging patterns. These patterns were the definite fibrosis pattern and the usual interstitial pneumonia (UIP) pattern. The definite fibrosis pattern is defined as a scan with architectural distortion of the

pulmonary parenchyma that is consistent with a fibrotic lung disease, irrespective of whether a UIP pattern is present. The classification with regards to the UIP imaging pattern was done based on official guidelines for IPF diagnosis (Lynch et al., 2018; Raghu, Remy-Jardin, et al., 2018) with participants classified as having no UIP, being indeterminate for UIP, having probable UIP or definite UIP. This characterising of images from participants with ILA was done by the consensus of three readers that were blinded from participant information (Putman et al., 2019). Lastly, for participants that had CT scans available at a five-year follow-up (3,167 participants), ILA progression was evaluated. The follow-up CT scans were evaluated for ILA and all participants that had ILA at baseline, at follow-up or both had both their CT scans simultaneously compared. In that direct comparison the scans were scored on a scale with the possible scores being definite regression, probable regression, no change, probable progression and definite progression. Progression was defined as "increase in lung areas affected with nondependent ground-glass, reticular abnormalities, diffuse centrilobular nodularity, nonemphysematous cysts, honeycombing, or traction bronchiectasis, or a new appearance of at least one such abnormality". These direct comparisons were done by a consensus of at least three readers (Putman et al., 2019).

These data were used in all four papers in which this thesis is based on. Papers I and II use the data with the ILA definition mentioned above (Washko et al., 2011). Papers III and IV were submitted after the publication of the 2020 Fleischner Society Position paper (Hatabu et al., 2020) which excludes abnormalities only consisting of centrilobular nodules from a uniform definition of the ILA term. Therefore, in these papers, participants with ILA of the centrilobular abnormalities subtype were excluded from analyses to comply with the updated definition.

In Paper III, measurements of multiple serum proteins were used. The measurements of these serum proteins were done using an emerging technology for proteomic profiling – the SOMAscan (SomaLogic, Inc. Boulder, CO, USA) which utilises Slow-Off rate Modified Aptamers (SOMAmers) to measure amounts of multiple blood proteins with high throughput and sensitivity. SOMAmers are synthetic, single-stranded DNA aptamers with modified nucleic acids that selectively bind three-dimensional protein targets with high specificity. In multiplex arrays of such aptamers, multiple SOMAmers with fluorescent tags bind to their target proteins in a biological sample and unbound SOMAmers as well as unbound proteins are then washed away. After removal of the SOMAmers from their target proteins, and the removal of these proteins from the sample, the

amount of remaining SOMAmers, i.e., the SOMAmers that bound to their target protein, are quantified. This quantification is done by fluorescence measurement utilizing the SOMAmers' fluorescent tags, after hybridization to a microarray which contains single stranded DNA probes complementary to the SOMAmer. This is done in three dilution sets of the original samples (40%, 1% and 0.005%) to be able to quantify proteins across a great scale of abundance. The SOMAscan method allows for proteomic studies that use simultaneous, consistent measurements of much higher numbers of proteins from a single biological sample than previously possible (Candia et al., 2017; Gold et al., 2010). In the AGES-Reykjavik study, a custom version of the SOMAscan 5K platform with 5,034 SOMAmers was used. Thus, 5,034 protein analytes were targeted by these SOMAmers in serum samples from 5,457 AGES-Reykjavik participants. Of these, 4,783 SOMAmers targeted 4,137 individual human proteins which are annotated to a total of 4,115 or 4,118 Entrez gene symbols (depending on the version of Entrez database used for annotation). For protein measurement in the AGES-Reykjavik cohort, both sample collection and sample processing were randomized, and all samples run as a single set. This was done to avoid measurement bias by batches or time of processing. The 5,034 SOMAmers used had median intra-assay and inter assay coefficient of variation <5%, meaning that the variability of protein measurements in the AGES-Reykjavik cohort was similar to previously reported variabilities of protein measurements with the SOMAscan method (Candia et al., 2017; Emilsson et al., 2018). For validation of the assay, the specificity of 779 SOMAmers for their target was performed with mass spectrometry and measurements of proteins measured with standard immunoassays as well as the SOMAscan in the AGES-Reykjavik cohort (i.e., insulin, C-reactive protein, and natriuretic peptide B) were compared and found to be highly correlated. In addition, inferential support for the SOMAmer specificity was also established by validating previously known associations of proteins with phenotypic measures or disease outcomes and with cis-acting SNPs in the AGES-Reykjavik cohort using the SOMAscan measurements (Emilsson et al., 2018; Gudjonsson et al., 2022).

To achieve aim 4, discussed in paper IV, acquisition of data on telomere length were necessary. To measure telomere length in the AGES-Reykjavik cohort, genomic DNA was extracted from participants' peripheral blood leukocytes. Telomere length measurements were done using a high throughput assay based on quantitative real time polymerase chain reaction. This method has been previously described and validated versus other methods of telomere measurement (Cawthon, 2002; McGrath, et al., 2007). This method allows for the calculating of the mean relative telomere length (MTL) by calculating the

exponentiated telomere repeat copy number to a single gene copy number (T/S) ratio, i.e., by comparing the amount of telomere product amplified by polymerase chain reaction to the quantity of a single copy gene amplified by polymerase chain reaction. Each sample was run three times and the mean of those values was reported as the MTL. This ratio is proportional to the average telomere length when adjusted for a reference sample (Cawthon, 2002). For statistical analyses of the MTL Z-scores, obtained by dividing the MTL by the mean MTL per 384 well plate, were used. This was done to account for potential batch effect.

3.1.1.3 Participant inclusion

Of the 5,764 participants of the AGES-Reykjavik Study, the 5,320 participants with ILA phenotyping available (92%) were included in paper I. Included in Paper II were 5,270 participants (91% of total number) that had data on cancer diagnoses and ILA phenotyping available. Included in the main analyses in Paper III were 5,259 participants (91%) that had protein measurements and ILA phenotyping available. For analyses of pulmonary fibrosis related SNPs and proteins 5,368 participants (93%) were included, i.e., those that had genotyping data and protein measurements available. Included in paper IV were 498 participants (9%) that had data available on telomere length. This sample included a higher proportion of participants with ILA than in the whole cohort.

3.1.2 The COPDGene Study

3.1.2.1 Study design

Data from the Genetic epidemiology of COPD (COPDGene) Study were used in Paper III and Paper IV. The COPDGene Study is a bi-racial, multi-centre, cohort study designed to discover genetic factors associated with chronic obstructive pulmonary disease among smokers. The study was designed to enroll a total number of 10,000 participants, two-thirds non-Hispanic whites and one-third African American with even distribution of genders and COPD severity. Ultimately, 10,192 participants were enrolled, thereof 3,408 were African American (Castaldi et al., 2014). Participants were to be between 45 and 80 years of age, with a history of minimum 10 pack-years of smoking, except for controls without a history of smoking. The enrolment took place in 21 clinical centres in the United States. Participants with a history of other lung disease (except asthma but including pulmonary fibrosis), history of lobectomy or suspected lung cancer, active cancer and several other medical conditions were excluded. Participants were included in four phases to be able to confirm genetic signals in independent participant groups. Participants underwent CT scans of the lungs in full inhalation end at the end of expiration, standardized

spirometry, measurements of height, weight, blood pressure, oxygen saturation, completed a six-minute walk test and questionnaires on their respiratory symptoms and medical history and had blood samples collected (Regan et al., 2010).

3.1.2.2 Phenotyping

The evaluation of COPDGene participants' CT scans for ILA took place in an identical manner to the previously described ILA evaluation in AGES-Reykjavik. Scans were sequentially evaluated by up to three readers using the definition of ILA described above (Washko et al., 2011). For statistical analyses, as in analyses of the AGES-Reykjavik cohort, participants with centrilobular abnormalities were excluded to comply with the updated definition of ILA published in a Fleischner Society Position Paper (Hatabu et al., 2020).

Interstitial features, identified in an automated, objective manner, on CT scans of participants of COPDGene were analysed as part of Paper IV. This machine-learning based method splits the lung parenchyma into regions and uses the distribution of the density in the region and its distance from the pleura to classify portions of the lungs as normal, emphysematous or as an interstitial lung feature. The method is thus able to quantify the proportion of CT scans occupied by interstitial lung features. The interstitial features that were used to train the classification tool were reticular features, honeycombing, centrilobular nodules, linear scars, nodular features, subpleural lines and ground glass abnormalities. Details of the method are previously described in detail (Ash, Harmouche, Putman, et al., 2017; Ash, Harmouche, Ross, et al., 2017). COPD was also modelled in paper IV using data from the COPDGene study. It was defined by spirometry as a Global Initiative for Chronic Obstructive Lung Disease stage II and higher.

To achieve aim 3, measurements of the same blood proteins previously measured in the AGES-Reykjavik cohort were obtained for the COPDGene cohort. These measurements took place using the same proteomic method as for AGES-Reykjavik, i.e., the SOMAScan (SomaLogic, Inc.) using DNA aptamers as described above. The measurements in COPDGene were made using EDTA plasma samples from participants, with a version of the SOMAScan platform with 5,285 SOMAmers targeting 4,776 unique human proteins. The standardization of data, including plate scaling and calibration to control for inter-assay variation and batch differences was done per SomaLogic protocol. After normalization and cleaning of data the median coefficient of variation was 5.0%.

The methods for measurement of telomere length, necessary to achieve aim 4 and discussed in paper IV, were identical to those used for measuring telomere length in AGES-Reykjavik. Briefly, after extraction of genomic DNA from blood leukocytes, telomere length measurements were based on quantitative real time PCR (McGrath et al., 2007). Then, the MTL was calculated by comparison of the telomere repeat copy number to a single gene copy number.

3.1.2.3 Participant inclusion

In Paper III, the 4,899 participants of Phase 2 of the COPDGene study that had SOMAmer measurements and CT data available were included. In Paper IV, 4,452 participants that had data on MTL measurements as well as ILA phenotyping available in Phase 1 were included in the study.

3.1.3 The Framingham Heart Study

3.1.3.1 Study design

The Framingham Heart Study (FHS) was the first study of its kind, originally launched in 1948 in Framingham, Massachusetts with the goal of finding risk factors for cardiovascular disease. Residents of Framingham were originally recruited between 1948 and 1952. The offspring of the original cohort were recruited to the study in 1971 and with the recruitment of their offspring in 2002 recruitment, FHS extended to three generations (Mahmood, et al., 2014). Participants in the Offspring Study were first examined in 1971. The 5,135 participants originally underwent interviews on medical history, anthropometric and blood pressure measurements, electrocardiography, and the collection of a fasting blood sample (Kannel, et al., 1979). After the baseline examination, participants in the Offspring cohort were re-examined after eight years and once every four years after that (Djoussé, et al., 2002).

3.1.3.2 Phenotyping

A subset (n = 2,633) of participants of the FHS Offspring and Third Generation cohorts participated in the FHS-Multidetector Computed Tomography 2 study carried out in the years between 2008 and 2011. As part of that study, they underwent a volumetric inspiratory chest CT that covered all parts of the lungs. These images have been previously evaluated for ILA, with the methodology of evaluation for ILA as previously described, with the sequential evaluation of up to three readers (Araki et al., 2016; Hunninghake et al., 2013). The updated ILA definition published by the Fleischner Society has been used for recent analyses (Hatabu et al., 2020).

Of all participants of the FHS Offspring Study, 3,532 participated in the sixth examination conducted between 1995 and 1998. Of these, 1,589 participants, selected to be unrelated, had DNA samples taken for telomere analysis, although 345 were of insufficient quality for telomere measurement. The telomere measurement was done by Southern blot analyses using telomere-specific DNA probes with the telomere length defined as the mean of the leukocyte DNA terminal restriction lengths (Vasan et al., 2008).

3.1.3.3 Participant inclusion

Included in Paper IV were the 439 participants of the FHS Offspring Study that had undergone a chest CT with ILA phenotyping as part of the FHS-Multidetector Computed Tomography 2 study as well as telomere length measurements as part of their sixth Offspring Study examination.

3.2 Study design and statistical analyses

3.2.1 The relationship of interstitial lung abnormalities with functional status (Paper I)

In Paper I, the associations of three self-reported measures of health and functional status with ILA status were assessed. These three measures, all obtained from questionnaires, were self-reported health, independence in activities of daily living and physical activity in the 12 months before participating. Details of outcome definition and variable dichotomisation are found in the Methods section of Paper I. The association of each of these variables with ILA status was evaluated after dichotomisation using logistic regression models, both unadjusted models and models adjusted for age, sex, body mass index, pack-years of smoking and smoking status at time of study participation. Participants with indeterminate ILA status were excluded from analyses (n=1,726, 32%) to provide comparison of participants with and without ILA. Analyses for each outcome were done using complete-case analyses. The number of participants with missing data for independence in activities of daily living was 32 (0.6%), 7 participants (0.1%) had missing data on perception of health status and 52 (1%) had missing data on physical activity participation. Seven participants (0.1%) were excluded from adjusted analyses due to missing covariate data.

3.2.2 The relationship of interstitial lung abnormalities with pulmonary and non-pulmonary malignancies (Paper II)

In Paper II, the aim was to assess the association of ILA with diagnoses of, and mortality from, malignancies. This was assessed for both all malignancies and

pulmonary malignancies. The observation period was from participants' study entry (between 2002-2006) to August 31, 2016. Details of outcome definitions and statistical analyses are found in the Methods section of Paper II. Data on cancer diagnoses, both prior to and during the observation period, was obtained from medical records from hospital visits or admissions to Landspitali, Iceland's only tertiary care hospital. Cancer diagnoses were defined using the International Classification of Diseases, Tenth Revision (ICD-10). The date of diagnosis was defined as the date of the first hospital visit with code registration. Data on mortality and cause of death were ascertained from the Icelandic Directorate of Health.

Baseline characteristics and outcome data were summarised by ILA status. Further statistical analyses were done comparing participants without ILA and participants with ILA, i.e., participants with indeterminate ILA were excluded. The cumulative incidence of lung cancer diagnosis and cancer diagnosis of other causes among participants with and without ILA were calculated with mortality regarded as a competing risk. The null hypotheses of equality of these cumulative incidences among participants with and without ILA were tested using Gray's tests. Likewise, the cumulative incidences of mortality with lung cancer registered as the cause of death, mortality with other cancers registered as the cause of death and mortality with other registered causes of death was calculated and equality between participants with and without ILA tested with Gray's tests. Cox proportional hazards models were used to assess the associations of ILA with study outcomes. All models were adjusted for age, sex, smoking at study entry and pack years of smoking as covariates.

Additional models were created, testing the sensitivity of the results to modifications to the outcome definition and the analysis methodology. Models of the association of ILA with main outcomes were replicated with participants with indeterminate ILA included, i.e., comparing participants with ILA to participants that either had no ILA or indeterminate ILA. Differences in the hazard of lung cancer based on the presence of the definite fibrosis pattern were assessed. Finally, the association of ILA with main study outcomes was assessed using proportional subdistribution hazards models that account for the competing risk of mortality from other causes (Fine & Gray, 1999).

3.2.3 The relationship of interstitial lung abnormalities with markers of cellular senescence (Paper III)

The aim of Paper III was to assess the relationship of various protein markers, including cellular senescence markers, with ILA. Prior to analyses, protein

measurements were transformed using the Yeo-Johnson transformation, a method of data transformation that is a part of the Box-Cox family of data transformations (Yeo & Johnson, 2000) and extreme data outliers, defined as values above the 99.5th percentile of the distribution of 99th percentile cutoffs across all proteins after scaling (0.2% of values in total), were removed and imputed prior to analysis using K-Nearest Neighbour imputing. Participants with indeterminate ILA status were excluded from analyses. Analyses were performed at the protein level, meaning that when more than one SOMAmer per genetic target was available, the SOMAmer that had a stronger association with ILA at baseline was used. The Bonferroni correction was applied to adjust for multiple testing as appropriate, with alpha defined *a priori* as 0.05. Details of statistical analyses are found in the Methods and Supplement methods of Paper III.

For analyses of proteins and ILA in the AGES-Reykjavik cohort, logistic regression models of the association of each protein with ILA, adjusted for age, sex, pack-years of smoking and smoking status at study entry, were fitted. These results were then used to find a set of proteins that predicts ILA in multivariate models by application of adaptive LASSO (Least Absolute Shrinkage and Selection Operator) models to 200 bootstrap data samples. The measurements of the proteins that, in single protein models, were associated with ILA without Bonferroni correction (an unadjusted p-value < 0.05) were used to create 200 bootstrap data samples. An adaptive LASSO model was fitted with 10-fold cross validation to maximize AUROC in each sample. All proteins were then ranked based on the number of bootstrap samples in which they were included in the adaptive LASSO models. The proteins used in all 200 bootstrap samples were then jointly assessed in an adjusted logistic regression model with ILA as the outcome. The AUROC of this model was calculated as well as the variance inflation factor. To validate the AUROC, or rather estimate the value in the population from which the AGES-Reykjavik cohort is sampled, 200 bootstrap samples of the data were created and the mean AUROC across all samples presented as the validated AUROC (Harrell Jr., 2020). The analyses assessing proteins' relationship with ILA were replicated in the COPDGene cohort. These were the single protein logistic regression models of the three proteins with the strongest association with ILA and the multi-protein model, based on adaptive LASSO of bootstrap samples.

The modelling process for ILA progression was analogous to that for ILA at baseline. Included in it were the 223 participants that had ILA progression at the five-year follow-up, defined as an imaging score of probable or definite progression, and they were compared with the 1,425 participants that had no

ILA neither at baseline nor at follow-up. The associations of single proteins with ILA progression were assessed using adjusted logistic regression with ILA progression as the outcome. These results were then used to find a set of proteins that predicted ILA progression in a multivariate model with an identical methodology to that used for ILA at baseline.

To assess whether the associations of proteins with ILA differed based on specific ILA patterns, the proteins that were associated with ILA overall were evaluated for associations with imaging patterns. The patterns were the definite fibrosis pattern and the UIP pattern. To assess whether proteins were associated with genetic risk factors for ILA or IPF, data on SNPs that have previously published associations with ILA or IPF was retrieved (Hobbs et al., 2019). The associations of these SNPs, using AGES-Reykjavik genotype data (Gudjonsson et al., 2022; Hobbs et al., 2019), with all human SOMAmers were tested using linear regression. Results that met the genome wide significance threshold of $P < 5\times10^{-8}$ were considered significant and reported in the results. Non-genome wide significant results for rs35705950, the best-established genetic risk factor for ILA and IPF, were reported as well.

The proteins that were associated with ILA were assessed for functional enrichment, by assessing the enrichment of gene ontology terms based on data from the Gene Ontology project (Gene Ontology Consortium, 2021), biological pathways based on the Kyoto Encyclopedia of Genes and Genomes (Kanehisa & Goto, 2000), Reactome database (Fabregat et al., 2017) and WikiPathways (Martens et al., 2020), DNA transcription factors and regulatory motifs based on the TRANScription FACtor database (Matys et al., 2006), protein complexes based on the Comprehensive Resource of Mammalian protein complexes database (Giurgiu et al., 2019) and phenotypic abnormalities based on the Human Phenotype Ontology database (Köhler et al., 2020). This was done with a cumulative hypergeometric test using the g:Profiler R package, a publicly available tool for such analysis (Raudvere et al., 2019). Furthermore, the enrichment of these genes for tissue specificity was tested using data from the Genotype-Tissue Expression database (GTEx Consortium, 2015). For the corresponding genes of the 25 proteins most significantly associated with ILA, the median expression values for a range of tissues were retrieved from the Genotype-Tissue Expression database and shown graphically.

3.2.4 The relationship of interstitial lung abnormalities with telomere length in leukocytes (Paper IV)

The aim of Paper IV was to assess the relationship of ILA with leukocyte telomere length in three cohorts, COPDGene, AGES-Reykjavik and the FHS. Adjusted logistic regression was used to model the association of ILA with telomere length in COPDGene and AGES-Reykjavik. Participants indeterminate for ILA were excluded from analyses. The association of ILA and telomere length was modelled both with telomere length as a continuous variable and by comparing participants by telomere length quartiles. In the FHS, linear regression of the association of ILA status with telomere length was used, with telomere length as the outcome. Further details are provided in the Methods section of Paper IV.

Additionally, the associations between telomere length and specific imaging patterns and features of ILA were assessed in COPDGene and AGES-Reykjavik. Associations with the definite fibrosis pattern were explored, both with comparison to participants without ILA and to participants with non-fibrotic ILA. Similar analyses were done for the UIP pattern. The association of centrilobular nodules on chest CT, not included in the current definition of ILA (Hatabu et al., 2020), with telomere length was explored in COPDGene. The association between mortality and telomere length was assessed in the COPDGene and AGES-Reykjavik cohorts among those with ILA. This was done using adjusted Cox proportional hazards models in which telomere length was modelled both by comparison of the shortest quartile with the longest quartile and of the shortest decile with the longest decile. Additionally, to test whether the UIP pattern affected this relationship, Cox models modelling the association of telomere length were created using only data from AGES-Reykjavik and COPDGene participants with probable or definite UIP patterns. Furthermore, it was assessed whether quantitatively assessed interstitial lung features were associated with telomere length using data from COPDGene by using adjusted linear regression models. Lastly, the association between COPD and telomere length was evaluated in the COPDGene cohort by means of logistic regression, with the same adjustments as models of ILA.

4 Results and discussion

4.1 The relationship of interstitial lung abnormalities with functional status (Paper I)

An overview of study participants is shown in Table 1, Paper I. It is seen that the mean age of participants is around 77 years of age, that most participants are female and have a history of smoking while a minority were smokers at the start of the study. Participants with ILA were on average older and were more often male than participants without ILA. They more often had a history of smoking, more often smoked at study entry, and had smoked a greater number of pack-years. In Table 2, Paper I, the associations of subjective measures of functional status and health with ILA are shown. Independence in the basic activities of daily living, a self-perception of health as good or better and participation in physical activity at least weekly were all associated with lower odds of ILA in both unadjusted models and models adjusted for covariates.

These results show that ILA were consistently associated with deficits in a range of subjective measures of health and functional status. The associations were consistent across the range of measures assessed. Independence in activities of daily living is a predictor of morbidity and mortality in the geriatric population (Storeng et al., 2020) and participation in physical activity and self-reported health assessment are associated with mortality and chronic disease (Banerjee, et al., 2010; Bijnen et al., 1999; McGee, et al., 1999; Sundquist, et al., 2004). It can therefore be concluded that these markers found to associate with ILA are significant in the context of the health and function of elderly people and that they are implicated in aging.

These results must be interpreted considering several limitations and considerations. First, causality cannot be inferred in a cross-sectional study. This is especially true since biological mechanisms that could explain these mechanisms are not obvious and cannot be elucidated from the results. Second, unmeasured confounders, such as other chronic diseases, or residual confounding related to factors such as aging, could, at least partly, explain the associations seen in the results. Third, analyses were done using complete-case analysis, introducing the possibility of bias. Therefore, imputation of missing

values could have been a more appropriate way of handling missing data (Harrell Jr., 2016). However, the proportions of missing values were low for all variables and under such conditions complete-case analyses are more acceptable (Harrell Jr., 2016). Fourth, the classification of functional status measures to dichotomous variables and their analyses as such can be criticised. It particularly hinders the possibility to find a dose-response effect between predictors and outcomes. However, the dichotomisation was deemed necessary as variables were very unevenly distributed with low numbers of participants with some levels. For example, there were few participants that were independent in none or only one or two activities of daily living. This made the power to detect an association of these small groups with ILA low. In addition, there are no widely acknowledged ways to dichotomise these measures, posing the danger that the dichotomisation may be over-reliant on the shape of the data. Still, the cut-offs applied in this study were established trying to reflect real-world relevance and applicability while maintaining simplicity, for example whether participants were independent in all activities of daily living or regular participants in physical activity. Studies using these cut-offs to associate the functional measures with important outcomes are previously published (McGee et al., 1999; Storeng et al., 2020).

In conclusion, taking these limitations into account, interstitial lung abnormalities were found to associate with multiple subjective functional measures that are of importance to the geriatric population.

4.2 The relationship of interstitial lung abnormalities with pulmonary and non-pulmonary malignancies (Paper II)

An overview of the study participants is shown in Table 1, Paper II. A higher proportion of participants with ILA received any cancer diagnosis or a lung cancer diagnosis during the observation period than participants without ILA, which was also the case for mortality due to cancer and mortality due to lung cancer.

When cumulative incidences of lung cancer diagnoses and other cancer diagnoses were compared among participants with and without ILA, the cumulative incidence of lung cancer diagnoses was significantly higher among participants with ILA (Figure 1, Paper II). In Cox proportional hazards models and a proportional subdistribution hazards model assessing the association of ILA with lung cancer diagnoses, the hazard of lung cancer diagnosis was significantly higher among participants with ILA than without (Table 2, Paper II; Table 6). Comparing the cumulative incidences of other cancer diagnoses

among participants with and without ILA, no significant difference was found (Figure 1, Paper II). ILA was associated with increased hazard of all cancer diagnoses. However, when lung cancer diagnoses were excluded from that outcome, ILA remained associated in an unadjusted model, but in the adjusted model the association was attenuated and could not be regarded as significant (Table 2, Paper II). In an adjusted proportional subdistribution hazards model, ILA did not have a significant association with non-pulmonary cancer diagnoses (Table 6).

When cumulative incidences of mortality from lung cancer and other cancers were calculated and compared between participants with and without ILA, the cumulative incidence of lung cancer mortality was significantly higher among those with ILA but the mortality from other cancers was not (Figure 2, Paper II). Cox proportional hazards and proportional subdistribution hazards models showed a significant association of ILA with lung cancer mortality (Table 4, Paper II; Table 6). In similar models, ILA was associated with mortality from all cancers. However, in models in which the outcome was mortality from cancers excluding mortality from lung cancers the association was not significant (Table 4, Paper II; Table 6).

Table 6 - Results of proportional subdistribution hazards models of the association of ILA with cancer-related outcomes

MODEL	HR (95% CI)	P-VALUE
Diagnoses		
Lung cancer	2.63 (1.58-4.38)	1.9×10^{-4}
Cancer excluding lung cancer	0.97 (0.76-1.25)	0.84
Mortality		
Lung cancer	2.55 (1.56-4.18)	2.1×10^{-4}
Cancer excluding lung cancer	0.91 (0.65-1.29)	0.60

Models are proportional subdistribution hazards models. The model of diagnoses had lung cancer diagnosis, diagnosis of other cancers and mortality as possible outcomes. The model of mortality had mortality from lung cancer, mortality from other cancers and other mortality as possible outcomes. Models were adjusted for age, sex, pack-years, smoking at study entry. HR = hazard ratio. 95% CI = 95% confidence interval.

In analyses in which associations of specific imaging patterns with lung cancer diagnoses were assessed, both ILA without fibrosis and ILA with definite fibrosis were associated with lung cancer diagnoses in unadjusted models and adjusted models. The adjusted HR was higher in analyses of ILA with definite fibrosis than in analyses of non-fibrotic ILA. As for associations of the same patterns with lung cancer mortality, ILA with definite fibrosis was associated with the outcome in an adjusted model, while ILA without fibrosis was not (Table 3, Paper II).

Cox proportional hazards models of the association of ILA with lung cancer diagnoses, all cancer diagnoses, cancer diagnoses excluding lung cancer diagnoses, mortality from lung cancer, mortality from all cancers and mortality from all cancers excluding lung cancers were repeated in which participants with indeterminate ILA were included and grouped with those without ILA. In these models ILA was associated with lung cancer diagnoses, all cancer diagnoses, mortality from lung cancer and mortality from all cancers. ILA was not significantly associated with cancer diagnoses excluding lung cancer nor mortality from all cancers excluding lung cancers (Tables S1-S2, Paper II).

In summary, ILA were associated with an increased hazard of lung cancer diagnoses and mortality due to lung cancer in an elderly population. The associations were more reliable and with higher hazard ratios for fibrotic ILA than non-fibrotic ILA, suggesting that participants with more advanced imaging abnormalities were at greater risk of lung cancer related outcomes. The associations of ILA with diagnoses of, and mortality from, all cancers did not remain significant after exclusion of those with lung cancer from the outcome which suggests that the associations of ILA with all-cause cancer diagnoses and mortality were driven by the associations of lung cancer diagnoses and lung cancer mortality. It can therefore not be reliably concluded that ILA are associated with cancer diagnoses other than that of lung cancer or mortality from non-pulmonary cancers.

Overall, these results support the previously well documented association between IPF and lung cancer and extends that association to include earlier forms of lung fibrosis, such as ILA. Patients with IPF have a well-established, repeatedly documented, increased risk of pulmonary malignancies (Hubbard et al., 2000; Le Jeune et al., 2007; Turner-Warwick, et al., 1980), with data suggesting that the cumulative incidence of lung cancer among IPF patients reaches 15% after 5 years (Ozawa et al., 2009). Biological mechanisms that may explain this association have been explored. Rare mutations in *SFTPA2*, encoding for an isoform of surfactant protein A, cause a familial syndrome of IPF and lung cancer (Wang et al., 2009). Among patients with non-familial

pulmonary fibrosis, it has been shown that methylation profiles in IPF lungs bear similarities to that of lung cancers (Rabinovich et al., 2012), genetic abnormalities commonly linked to lung cancer have been found in IPF patients (Demopoulos, et al., 2002), IPF patients have abnormalities in gap junction cell-to-cell communication also seen in lung cancer (Trovato-Salinaro et al., 2006), cellular transformations with epithelial-to-mesenchymal transition are believed to play a part in IPF and lung cancer (Ballester et al., 2019; Kasai et al., 2005) and activation of signalling pathways via growth factors such as VDGF and FGF are implicated in both diseases (Vancheri, 2013). This last point is supported by the efficacy of the growth factor inhibitor nintedanib in treating both diseases (Reck et al., 2014; Richeldi et al., 2014; Vancheri, 2013). Other similarities in expression of biological molecules and pathways have also been noted (Ballester et al., 2019; Vancheri, 2013). The results presented here imply that at least some of these biological similarities may also apply to earlier forms of pulmonary fibrosis than IPF. In addition, because of the vast difference in prevalence between ILA and IPF, the results suggest that ILA, or at least extensive or fibrotic ILA, may be a significant risk factor for lung cancer at the population level. As such, if these results are corroborated in future studies, studies assessing whether incorporating incidentally detected ILA into lung cancer screening protocols may be warranted.

The presented results are subject to some limitations. First, while associations of ILA with diagnoses of and mortality from non-pulmonary malignancies did not reach the α value of 0.05, such associations cannot be entirely excluded using these data. This is especially true given the small numbers of some specific cancers other than lung cancers which theoretically could drive such an association or be independently associated with ILA. Second, cancer diagnoses and causes of mortality were obtained from hospital records and Health Directorate data. Since these data are registered for clinical purposes, a potential for misclassification exists. However, the concordance between results using diagnoses and causes of mortality, obtained from different sources, supports the reliability of the data. Additionally, there is nothing that implies bias in registry of malignancies based on ILA status during the study follow-up period. Third, histopathological diagnoses of malignancies were not available, but such information would be of interest, especially for subtyping of lung cancers. A larger study would however be needed to draw rigid conclusions regarding specific types of lung cancer. Fourth, although models were adjusted for confounding variables, residual confounding cannot be excluded entirely nor can the possible effect of unmeasured confounders. This holds especially true for analyses of all cancers which have many different risk factors that were not all

adjusted for. Fifth, the study cohort is comprised of elderly people and is racially homogenous and of limited genetic diversity. The study findings would benefit from replication in a more diverse cohort, especially regarding age, drawn from the general population. Sixth, the study findings could be limited to extensive abnormalities or distinct subtypes of ILA, or at least vary based on abnormality severity and/or type. This is shown in part by the strength of associations of ILA with definite fibrosis as opposed to those of non-fibrotic ILA, but analyses of further subtypes or imaging patterns were not feasible due to concerns with statistical power. Seventh, the validity of conventional survival analysis methods, i.e., Kaplan-Meier estimators and Cox proportional hazards models, can be questioned due to the competing risk of death due to other causes than cancer. Using conventional survival analysis methods, the censored subjects, that for example drop out of a study, are assumed to have the same risk of the studied event as those remaining under follow-up (D. F. Moore, 2016). If subjects' follow-up times are "censored" due to mortality of other causes than the studied event, this assumption does not hold. This concern led to the calculation of cumulative incidences, taking the competing risk of mortality into account (Austin et al., 2016). In general, Cox proportional hazards models that do not have all-cause mortality as an endpoint assume no competing risks and noninformative censoring (D. F. Moore, 2016). The competing risk of death of another might therefore lead to biased hazard estimates. Hence, the data were also modelled using Fine and Gray proportional subdistribution hazards models that account for the competing risks of mortality (Fine & Gray, 1999). As the results were similar, it can be concluded that the presence of competing risks does not undermine the results of the Cox proportional hazards models.

In conclusion, ILA were associated with lung cancer diagnoses and lung cancer mortality, but reliable associations with diagnoses of or mortality from other cancers were not found.

4.3 The relationship of interstitial lung abnormalities with markers of cellular senescence (Paper III)

Participant characteristics in AGES-Reykjavik and in COPDGene are shown in Table 1, Paper III. In both cohorts, participants with ILA were older and had smoked more than participants without ILA.

The 287 proteins that were associated with ILA after Bonferroni adjustment for multiple statistical comparisons in adjusted logistic regression models are shown in Table E1, Paper III and graphically in Figure 1A, Paper III. Notably, the most significant association was for Surfactant protein B (SFTPB), followed by

Secretoglobin family 3A member 1 (SCGB3A1) and WAP four-disulfide core domain protein 2 (WFDC2). The validated areas under the receiving operator characteristic curve (AUROC) by 200-fold bootstrapping of these models, compared to a "baseline" model of ILA with age, sex, pack-years and smoking at study entry are shown in Figure 2A, Paper III. The associations of three proteins, SFTPB, WFDC2 and SCGB3A1 were replicated in COPDGene and they were all significantly associated with ILA, as shown in Table E7, Paper III and graphically in comparison with results from AGES-Reykjavik in Figure 1A, Paper III. Validated AUROC values for SFTPB, SCGB3A1 and WFDC2 are also shown (Figure 2B, Paper III). The associations of proteins that are previously suggested biomarkers of ILA (H. F. Armstrong et al., 2017; Ho et al., 2016; McGroder et al., 2019; Sanders et al., 2021) with ILA are shown in Table E4, Paper III.

The proteins that most frequently were a part of adaptive LASSO models in 200 bootstrap samples of the data are shown in Table E5, Paper III. Eight proteins were featured in all 200 bootstrap samples. In a multi-protein logistic regression model of associations with ILA adjusted for covariates, seven of these proteins were associated with ILA. Positive associations were found for SFTPB, WFDC2, SCGB3A1 and CBLN4, while WFIKKN2, ADAM9 and ANXA9 were negatively associated (Table E6 and Figure 2A, Paper III). The validated AUROC by 200-fold bootstrapping was 0.880, compared to 0.670 for a model with only covariates. The multi-protein model was replicated in COPDGene, with four proteins significantly associated with ILA in that model (Figure 2B and Table E7, Paper III). After validation by 200-fold bootstrap sampling, the AUROC was 0.826.

After Bonferroni adjustment, 121 proteins were associated with ILA progression in single protein logistic regression models. These proteins are shown in Table E12, Paper III and graphically in Figure 1B, Paper III. The most significantly associated proteins were SFTPB, WFDC2 and GDF15. Four proteins were a part of adaptive LASSO model in all 200 bootstrap models created (Table E14, Paper III). They were all associated with ILA in a multiprotein logistic regression model (Figure 2C and Table E15, Paper III) for which the AUROC value was 0.824 after validation (Figure 3C, Paper III).

The associations of ILA related proteins with ILA with fibrosis and ILA without fibrosis, as well as ILA with the UIP pattern and ILA without the UIP pattern are shown in Tables E10 and E11, Paper III. The 20 proteins most significantly associated with ILA were associated with ILA with and without fibrosis and ILA with and without UIP. Still, for all these 20 proteins, the odds ratios associated with definite fibrosis were higher than those associated with non-fibrotic ILA and

the odds ratios for ILA with the UIP pattern were higher than the odds ratios for ILA without the UIP pattern (Table E10, Paper III).

The associations of SNPs previously associated with ILA and/or IPF with proteins that reached the threshold for genome-wide significance, the associations of these proteins with ILA and previously reported associations of these proteins with ILA (Hobbs et al., 2019) are shown in Table 2, Paper III. The only SNP that in addition to being associated with an ILA-associated protein that was itself associated with ILA was rs35705950 near *MUC5B*, associated with SFTPB. Due to these associations, the association of rs35705950 with SFTPB was evaluated with stratification based on ILA status and was found to be significant regardless of ILA status (Table E8, Paper III).

The most significantly enriched gene ontology terms for proteins associated with ILA after Bonferroni correction were terms involving the extracellular region (GO:0005576, p = 0.0069), chemokine activity (GO:0008009, p = 0.013) and chemokine receptor binding (GO:0042379, p = 0.013). All terms and pathways for which ILA-associated proteins were significantly enriched are shown in Table E3, Paper III and enrichment for tissue-specificity is shown in Figure E4, Paper III.

To summarise, proteins with highly significant associations with ILA, most notably SFTPB, WFDC2, SCGB3A1 and WFIKKN2, were found and replicated in an unrelated cohort. A multivariate model that effectively discriminated participants with and without ILA was created by way of machine learning and replicated. When analyses were done for ILA progression, some of the same proteins were highly associated with ILA progression and were featured in a multi-protein predictive model for progression. The association of SFTPB with the best-established genetic risk factor of ILA was demonstrated and associations of the most heavily featured proteins were found to be greater for more advanced imaging abnormalities. Lastly, functional analyses highlight possible roles of the proteins in the complex functions that are cytokine signalling and mRNA processing.

These findings highlight several proteins, such as SFTPB and WFDC2, as possible biomarkers of early fibrotic lung changes and indicate a need for studying a selection of these proteins in advanced ILD, in addition to further validating their relationship with ILA. Not only do these proteins confer increased odds of ILA, their ability to predict ILA and ILA progression, alone or as parts of multi-protein models, is also demonstrated. These data may also provide insights into the biology of ILA and similar changes. Some experimental data support

such notions. SFTPB is a surfactant protein and genetic mutations in these have been implicated in IPF (Ley et al., 2014), WFDC2 has been found to be elevated in IPF (Raghu, Richeldi, et al., 2018) and WFIKKN2 is a part of the TGF-beta superfamily that is central to IPF pathogenesis (Szláma et al., 2010). Some of the associations presented substantiate the proposed link of ILA with biological aging. As an example, Growth Differentiation Factor 15 (GDF-15), Calpain small subunit 1 (CAPNS1), Urokinase plasminogen activator surface receptor (PLAUR), Lamina-associated polypeptide 2, isoforms beta/gamma (TMPO), Tumor necrosis factor receptor superfamily member 1B (TNFRSF1B), Heterogeneous nuclear ribonucleoprotein D-like (HNRNPDL) and C-X-C Motif Chemokine Ligand 10 (CXCL10), here associated with ILA and/or ILA progression, are all proposed biomarkers of aging in multiple studies (A. A. Johnson et al., 2020). Still, other proposed biomarkers of aging, such as albumin, annexins A1 and A2, complement C3, cathepsin S (CSS), epidermal growth factor receptor (EGFR), fibronectin (FN1) and haptoglobin, were not associated with ILA (Table E2) (A. A. Johnson et al., 2020). The association of GDF-15, specifically, could be related to its upregulation in type 2 alveolar epithelial cells in response to telomere dysfunction and stress (Zhang et al., 2019). Protein biomarkers previously associated with ILA (H. F. Armstrong et al., 2017; Ho et al., 2016; McGroder et al., 2019; Sanders et al., 2021) were all associated with ILA based on an unadjusted p-value and all, except interleukin-6 and insulin, were associated with ILA or ILA progression after Bonferroni adjustment (Table E4, Paper III).

Several methodological considerations must be discussed further. Statistical significance for single protein models in this study was based on Bonferroni adjustments for multiple comparisons which, while simple and interpretable, tests the null hypothesis that no proteins are associated with the outcome which may be considered unrealistic for the presented work (R. A. Armstrong, 2014; Perneger, 1998). Therefore, these stringent adjustments may invite the risk of type II error for some proteins measured, i.e., that the null hypothesis is wrongly accepted and that associations of proteins with ILA are therefore underreported. However, the explorative nature of the single protein analyses with proteins selected for analysis based on availability rather than an *a priori* hypothesis of ILA association supports the use of stringent adjustments for multiple testing. Another methodological consideration of the paper is the method used for machine-learning based multiprotein models for prediction of ILA and ILA progression. Here, the method was selected with several aims in mind. The selected model was to have the maximum ability to discriminate participants with ILA or ILA progression from participants without these outcomes while not

overfitting the data at hand, i.e., the prediction was to be maximized in the population from which the cohort was sampled rather than in the cohort itself. The model was preferred to use a low number of proteins for the model to be more practical and the model's results had to be comprehensible. Lastly, since part of the research question revolved around discovering which proteins could predict the ILA status of participants, the model had to be interpretable with regards to which proteins composed or were most important for the model's classification. Due to this last point, the use of popular supervised machine learning methods such as support vector machines or neural networks were refrained from as such models are poorly interpretable although they can provide excellent classification performance (Stafford et al., 2020). LASSO regression and ridge regression are regression methods that add a penalty to the regression equation introducing stable bias to reduce variance. This is done to reduce overfitting to the data in which the model is trained. For the ridge regression approach, this penalty is applied to the sum of the squares of the regression coefficients, with the optimal size of the penalty often obtained by cross-validation. For the LASSO the approach is to penalize the sum of the absolute values of the regression coefficients. This causes variables with low coefficients to be excluded from the model. This allows for the selection of a set of variables while maximizing prediction (Keller & Rice, 2018; Ranstam & Cook, 2018; Tibshirani, 2011). The adaptive LASSO is a variant of the LASSO model that uses weights for predictor variables, that can e.g., be obtained from a ridge regression model. This is shown to increase the consistency of the model (Zou, 2006). A bootstrap data sample is a resampling method that randomly resamples observations from a given set of observations, allowing for repeated sampling of the same single observation. The rationale of the method is that multiple bootstrap samples of a set of data reflect the population from which the data set is drawn (Harrell Jr., 2016). Therefore, by selecting proteins that were used in adaptive LASSO models in all bootstrap samples, we aimed to find proteins that consistently jointly predicted the outcome in a robust manner, not only in the cohort but in the population from which it is drawn. The bootstrap method, with different samples created, was then utilized to validate AUROC calculations. The aim of that was to adjust for optimism and provide a AUROC value reflecting the population from which the cohort is sampled. This was done instead of splitting of the data due to the power concerns introduced by data splitting for model validation (Harrell Jr., 2016; Steyerberg, 2018). This method of feature selection has potential drawbacks. It is driven by statistical correlations in the data but not based on biological relationships. The dropping of variables via LASSO may mean that a selected variable for the model, from a group of highly correlated variables, is not the one with the most biological relevance. The variables

selected for the model were selected based on their ability to jointly predict ILA, as noted by the low collinearity of model features. Therefore, some proteins may be selected because they predict ILA in tandem with a stronger predictor, not because of their individual importance. While the use of the bootstrap is intended to limit the reliance of feature selection on the cohort and extend it to the population from which the cohort is drawn, the AGES-Reykjavik cohort is drawn from the elderly inhabitants of Iceland. This population is of advanced age and is homogenic with regards to race. Its history with regards to environmental exposures is likely similar as well. Therefore, the protein selection might have had different results in a cohort sampled from another population, even with identical methods. This might also explain the imperfectness of the replication of the multi-protein model of ILA in COPDGene, i.e., why some model proteins are not significant predictors of ILA in that cohort.

Some limitations of the results presented in Paper III must be mentioned. In the definition of ILA progression, participants with progression are compared to participants without ILA on either scan. This definition, used to obtain sufficient statistical power for meaningful analyses of progression, is identical to that used in previous articles exploring associations with ILA progression (Putman et al., 2019). Still, it encompasses incident ILA cases as well as progression during the follow-up period and can therefore more precisely be described as the 'development and progression' of ILA. The comparison to participants without ILA means that the findings may not be identical to a comparison with participants with ILA that have non-progression or resolution. Additionally, the definition of progression used in this paper and most other papers studying ILA entirely relies on radiologic criteria (Araki et al., 2016). This not consistent with definitions of the progression of ILD such as IPF in research and in clinical practice, in which a decline in results of lung function tests and an increase of clinical symptoms are a part of the concept of progression (King et al., 2014; Raghu et al., 2011; Richeldi et al., 2014). It is still unknown, and should be a topic of further research, whether the radiologic progression of ILA correlates with clinical and spirometric parameters. Therefore, it should be emphasized that the term "ILA progression" remains distinct from progression to ILD or progression of ILD. Second, the statistical associations presented in the paper, both with ILA and ILA progression are not sufficient to infer causal relations of the featured proteins with these outcomes. Other mechanisms could explain the associations, for instance reverse causality, i.e., that changes in the serum levels of proteins represent a response to the outcome, either directly or to comorbid conditions or biological differences associated with the outcome. Such conditions could e.g., be comorbid disease not adjusted for or biological

processes, such as aging processes in the lung not fully adjusted for. This discussion highlights the possible importance of residual or unmeasured confounders that could be associated with this large number of protein markers. Third, the specificity of the serum level changes of the proteins for ILA is not guaranteed. It is a possibility that the serum level changes of some proteins are a reaction to non-specific lung pathology, possibly not limited to only fibrotic or interstitial pathology. Fourth, the measurement of protein levels by way of DNA-aptamers may be subject to some limitations. The specificity of the aptamers for single proteins or the protein isoforms is not perfect, causing aptamers to bind to another protein or isoform in a small minority of cases (Sun et al., 2018). It is also presumed that changes in protein structure by way of single amino acid changes due to genetic polymorphisms, posttranslational modification of proteins, binding to carrier proteins or forming of protein complexes effects the affinity of aptamers for their targets, but unclear how much it does so (Joshi & Mayr, 2018; Pietzner et al., 2021). Fifth, while the present study is large compared to many others of its kind, it is possible that a larger number of participants with ILA is needed to detect further associations with proteins or pathways, especially given the stringent adjustment used for multiple comparisons, as discussed above.

Despite these limitations and considerations, it can be concluded that replicable associations with several proteins were found in the study, underlining these proteins as potential biomarkers of early pulmonary fibrosis.

4.4 The relationship of interstitial lung abnormalities with telomere length in leukocytes (Paper IV)

The baseline characteristics of participants included in analyses of telomere length are provided for all three cohorts, with stratification by ILA status in Table 1, Paper IV and for COPDGene and AGES-Reykjavik with stratification by quartiles of telomere length in Supplementary Table 1, Paper IV.

Measures of telomere length were associated with ILA status in all three cohorts after adjustment for covariates as shown in Table 2, Paper IV. When evaluated continuously, MTL z-score was associated with an increase in ILA odds in AGES-Reykjavik and in COPDGene. In the FHS, the telomere length of participants with ILA was estimated to be shorter than that of participants without ILA. When analysed by quartiles of telomere length, comparing participants in the shortest quartile of telomere length to participants in the longest quartile, the shortest quartile was associated with ILA in AGES-Reykjavik and in COPDGene.

Relations between telomere length and mortality among those with ILA were evaluated in COPDGene and in AGES-Reykjavik. In COPDGene, an association was not found between MTL and mortality. In AGES-Reykjavik, while no association with mortality was found in comparison between the lowest and highest quartile, participants in the lowest decile of MTL were found to be at increased risk of mortality compared to those in the highest decile. When analyses were restricted to those with probable UIP or UIP pattern, no association of MTL with mortality was found in COPDGene while such an association was noted in AGES-Reykjavik. These results are shown in full in Table 7.

Table 7 - Associations of MTL with mortality among those with ILA

	HR	**95% CI**	**P-VALUE**
ILA			
COPDGene			
Lowest vs highest quartile	0.82	0.4-1.7	0.6
Lowest vs highest decile	1.3	0.9-1.8	0.14
AGES-Reykjavik			
Lowest vs highest quartile	1.2	0.6-2.2	0.5
Lowest vs highest decile	2.0	1.2-3.4	0.007
ILA WITH UIP PATTERN			
COPDGene			
Lowest vs highest decile	1.2	0.8-1.9	0.3
AGES-Reykjavik			
Lowest vs highest decile	3.5	1.7-7.4	0.0009

Shown are results of Cox proportional hazards models evaluating the associations of MTL with mortality among those with ILA or ILA with UIP as specified. Analyses were adjusted for age, sex, body mass index, pack-years of smoking and smoking status. Additional adjustments for race were made in COPDGene.

Analyses of MTL were also made based on the presence of the definite fibrosis pattern of ILA in COPDGene and AGES-Reykjavik. Both fibrotic ILA and non-

fibrotic ILA were associated with MTL in COPDGene when compared with those without ILA (OR = 11.6, 95% CI = 2.7-50, p = 0.009 and OR = 1.7, 95% CI = 1.1-2.6, p = 0.02, respectively), although the hazard was greater for fibrotic ILA. When participants with fibrotic ILA were compared to those with non-fibrotic ILA, a significant association was found with MTL in COPDGene but not in AGES-Reykjavik (Table 8). A significant association of MTL with the UIP pattern, compared to participants indeterminate for or probable UIP, was likewise found among participants of COPDGene but not participants of AGES-Reykjavik. Additionally, associations of MTL with centrilobular nodules, now excluded from the ILA term (Hatabu et al., 2020), were evaluated in COPDGene and not found to be significant.

Table 8 - Associations of MTL with subtypes of ILA

	OR	**95% CI**	**P-VALUE**
COPDGene			
Fibrotic vs non-fibrotic ILA	6.8	1.4-33.3	0.02
UIP vs indeterminate for UIP	2.2	1.1-4.2	0.02
UIP vs probable UIP	20	2.7-150	0.004
AGES-Reykjavik			
Fibrotic vs non-fibrotic ILA	1.2	0.4-3.3	0.8
UIP vs indeterminate for UIP	3.0	0.2-42	0.4
UIP vs probable UIP	1.4	0.2-8.2	0.7

Shown are results of logistic regression evaluating the associations of quartiles of MTL with ILA imaging patterns among those with ILA. Analyses were adjusted for age, sex, body mass index, pack-years of smoking and smoking status. Additional adjustments for race were made in COPDGene.

Data from the COPDGene cohort were also used to assess the association of MTL with a quantitative measure of interstitial features on chest CT. Shorter MTL was associated with an increase in interstitial features whether modelled

continuously or modelled comparing the lowest MTL quartile to the longest. The association between MTL and COPD was also modelled in COPDGene by comparing the shortest MTL quartile to the longest. A positive association was found with a COPD diagnosis and removing those with ILA from this analysis did not substantially alter the results.

These findings from three independent cohorts suggest that shorter telomere length is associated with increased odds of ILA. This is in concordance with prior studies that associate telomere length with more advanced forms of pulmonary fibrosis (Dai et al., 2015; Duckworth et al., 2021) and ILA in relatives of people with advanced pulmonary fibrosis (Hunninghake et al., 2020). However, associations with mortality among those with ILA were not found which somewhat contradicts previously demonstrated associations of telomere length with impaired survival among those with idiopathic pulmonary fibrosis (Stuart et al., 2014). Additionally, telomere shortening, a driver of cellular senescence, has been suggested as a biomarker of aging (Barnes et al., 2019; Bekaert et al., 2005), as discussed in detail in the Introduction, and causes the so-called short telomere syndromes that are suggested to be syndromes of "accelerated aging" (Armanios & Blackburn, 2012). Therefore, the association of telomere length with ILA supports the suggestion that aging of the lungs may play a role in the development of ILA. However, as associations of telomere length with aging-related biological changes and diseases are not consistent (Mather et al., 2011; Sanders & Newman, 2013), the shortening of telomeres may be a specific driver of pulmonary aging and pulmonary pathology rather than of all aging-related diseases. The finding that pulmonary fibrosis is the most common disease manifestation of telomere attrition supports this notion (Armanios, 2012). The role of telomere shortening in invoking cellular senescence, and the pro-fibrotic mediators secreted by senescent cells, could partly explain a specific link between telomere attrition and pulmonary fibrosis (Barnes et al., 2019).

Some limitations must be considered when interpreting the results of Paper IV. A large part of the results is obtained by comparing quartiles or deciles of the continuous measure that is telomere length. This approach can be problematic in some ways. First and foremost, it assumes that the risk of ILA modelled is proportional between these cut-offs in the data, which are data-driven and arbitrary. This approach also is prone to reduction of statistical power. Second, the usage of data-driven cut-offs that are not numerically identical between cohorts, can bias comparisons between studies or even possibly comparisons between cohorts presented in the same study (Bennette & Vickers, 2012). While this is a limitation of Paper IV, it can be argued that the effect on the association

of ILA with telomere length is limited as results from models in which telomere length is modelled continuously were concordant to quartile analyses in two cohorts. Additionally, when the association of MTL with ILA odds is examined graphically in the AGES-Reykjavik cohort, it is close to linear, supporting the validity of comparisons between quartiles (Figure 5A-B). However, this may be more of a concern for the association of MTL with mortality among those with ILA, as this association in AGES-Reykjavik cannot be shown to be clearly linear when examined in the same manner, although the low number of participants might contribute to the erratic shape of the association (Figure 5C-D). Second, only a small proportion of the AGES-Reykjavik cohort had telomere measurement data. This group of participants was oversampled for the presence of ILA, and therefore not as representative of the elderly population as the whole cohort. Such sampling also invites the risk of selection bias. However, the demographic characteristics of both participants with and without ILA were similar for the sample used in Paper IV as in the whole cohort (Table 1, Paper IV, Table 1, Paper I). The low number of participants also creates issues of statistical power for analyses of the AGES-Reykjavik cohort, especially for analyses of subtypes and imaging patterns and analyses of mortality. Third, the time between blood sampling for telomere length measurement and ILA phenotyping differed between COPDGene and AGES-Reykjavik on one hand (blood sampling and CT imaging done synchronously) and the FHS (blood sampling performed 13 years before CT imaging in which ILA were evaluated) on the other hand. Additionally, the lack of statistical power in the FHS cohort precluded the analysis of mortality or ILA subtypes. This limits ability to estimate whether data from the FHS cohort are comparable to the other cohorts. Fourth, telomere length was measured in cells from the peripheral blood, but telomere shortening in some IPF patients may be specific to the lung tissue, or even distinct cell types within the lung (van Batenburg et al., 2020). However, such measurements cannot be considered feasible for large cohort studies. Fifth, the selection of participants into the AGES-Reykjavik could introduce a certain survival bias to analyses of mortality among those with ILA, as the cohort is comprised of participants of the Reykjavik study that survived to advanced age. However, results of mortality analyses in the AGES-Reykjavik cohort were largely in concordance with results from COPDGene, apart from analyses using the extremes of MTL.

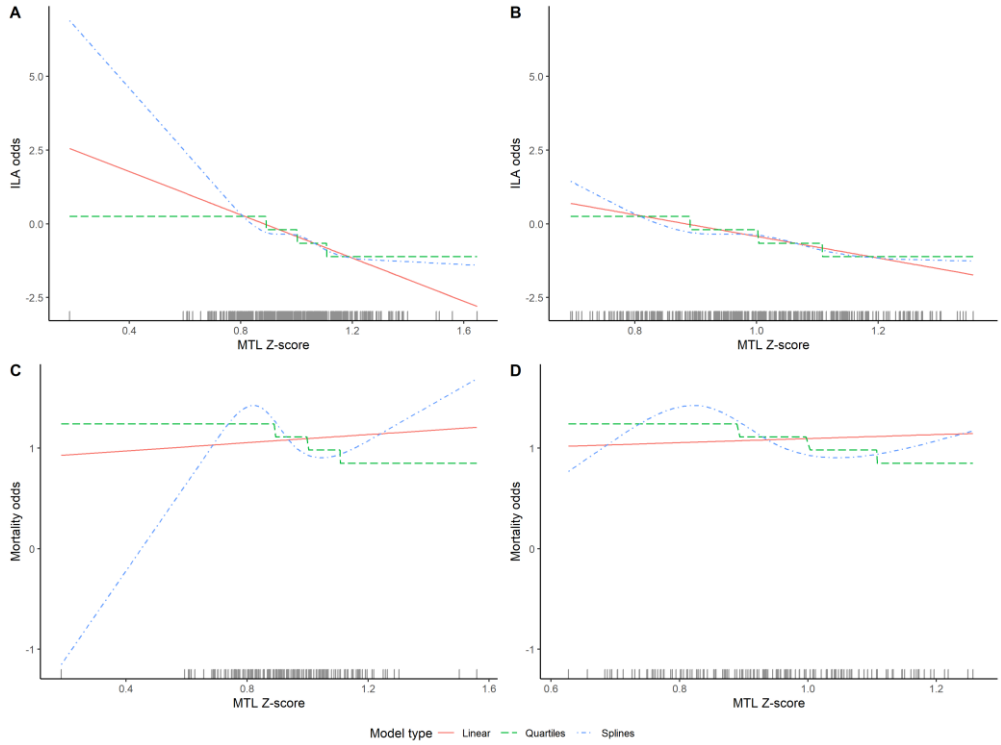

Figure 5 - Predictions of ILA odds and mortality based on MTL z-scores from linear, quartile-based and spline models among those with ILA in AGES-Reykjavik Panel A: Predictions of ILA odds across the whole range of MTL. Panel B: Predictions of ILA odds across the middle 95% of MTL range. Panel C: Predictions of mortality among those with ILA across the whole range of MTL. Panel D: Predictions of mortality among those with ILA across the middle 95% of MTL range.

Finally, different methods of telomere measurements must be discussed. The bulk of the findings were based on MTL as a measure of telomere length, a measure based on quantitative real time polymerase chain reaction. The measure was converted to z-scores to avoid batch effect, this makes it more difficult to compare and impossible to pool results from the AGES-Reykjavik and COPDGene cohorts. A different method of telomere measurement altogether, the Southern blot, was used in the FHS. These methods have their strengths and disadvantages (Aubert et al., 2012). The Southern blot analysis of terminal restriction fragments used in the FHS can accurately estimate average telomere length in absolute values. It was the first method used in research of telomere length and was the reference against which other methods were developed. Its main disadvantages are that it is labor-intensive and therefore expensive, requires a considerable amount of DNA for analysis and can be insensitive to

very short telomeres (Aubert et al. 2012; Kimura et al., 2010). The method used in COPDGene and AGES-Reykjavik was polymerase chain reaction-based amplification of telomeres which were then compared to a single copy gene amplification of an equivalent sample. The obtained estimate of telomere length is therefore not a direct measurement but a ratio of telomere to single copy gene. This method is popular as it is relatively fast, inexpensive, and does not require large amounts of DNA. Still, as there is fairly high variability within and between assays, careful quality control is needed and larger numbers of participants may be needed to detect small differences between groups of participants in epidemiological research (Cawthon, 2002; Kimura et al., 2010). Other methods, such as flow fluorescence *in situ* hybridization (FISH), are available. That method is based on fluorescently labeled probes that hybridize to telomere DNA and are quantified by flow cytometry. This method has the benefit of being able to measure average telomere length in specific cell populations and is highly accurate. While it requires a rather large amount of sample, is not as efficient for a large number of samples as polymerase chain reaction-based techniques and is sensitive to variations in the complex sample processing required (Aubert et al., 2012; Baerlocher et al., 2006). In summary, most of the presented data were based on a polymerase chain reaction-based method chosen for its high throughput, but the smallest cohort had telomere measurements done by the gold standard method that is the Southern blot. While this limits direct comparisons and meta-analyses between cohorts, it provides validation of the association of telomere length with ILA between different methods of measurement.

From these data, despite their limitations, it can be concluded that measures of telomere length in leukocytes were associated with ILA in multiple cohorts. A consistent association of telomere length with mortality among those with ILA was not found.

4.5 Future directions

The findings presented suggest several directions for future research of ILA. Overall, the validation of findings in other, more diverse, cohorts would be valuable due to the distinct characteristics and racial and genetic homogeneity of the AGES-Reykjavik population. As for findings that regard the associations of ILA, health, and functional status, it would be interesting to examine further whether people with ILA have pulmonary physiologic impairments, such as changes in oxygen saturation with exercise. Such studies would help establish whether a causal relationship exists between ILA and the impairments associated with them. For research of ILA and lung cancer, the association between these

two conditions suggests that research exploring common mechanisms of pulmonary fibrosis and malignancies could also focus on mechanisms implicated in early pulmonary fibrosis. As for future epidemiological studies, an inquiry into whether the associations of lung cancers with ILA are stronger for specific pathological subtypes would be valuable. Similarly, as tumour genetics become more important in oncology, larger studies exploring whether ILA is associated with specific genetic mutations of lung cancers would be an important addition to the literature. Studies like these could bring possible common biological mechanisms of early pulmonary fibrosis and lung cancer into focus and elucidate pathways that are causally implicated in both conditions. Mendelian randomisation studies could possibly be useful in such efforts. Lastly, if future studies continue to corroborate the link between ILA and lung cancer, studies exploring the utility of ILA as a risk factor in lung cancer screening could impact management of patients with ILA.

In Paper IV, leukocyte telomere length is associated with ILA, adding to previously known associations of leukocyte telomere length with IPF. To advance knowledge in that field, it would be important to discern whether short telomeres confer increased risk of progression of ILA. Such insights would again help answer whether telomere length is causally implicated in progression of ILA and the evolution of pulmonary fibrosis. This knowledge would be important to assess whether telomere lengths could be a biomarker of ILA progression, or even a target of therapeutic interventions aimed at hindering such progression. Lastly, for the findings of Paper III, further validation of the suggested protein biomarkers of ILA is very important, both in other cohorts with different demographic characteristics and with different methods of measurement. This would help answer whether any, and which, of the highlighted proteins could be useful as blood biomarkers that could clinically predict the existence of ILA and progression to ILD. Also, Mendelian randomisation studies using genetic polymorphisms as instruments in calculations to eliminate the effects of confounding, and laboratory work comparing the amounts of these proteins in healthy and fibrotic lung tissue could help approach the question of whether any, and which, of the proposed protein biomarkers are implicated in the pathophysiology of ILA and pulmonary fibrosis.

In addition to the specific questions prompted by the presented data, studies of ILA must address some key questions. First, because of the chasm in prevalence between ILA and ILD, risk factors, biomarkers or radiologic subtypes that predict the progression of ILA would be important findings. In this endeavour, there must be clear differentiation of different endpoints, i.e., radiologic progression

of ILA to more extensive ILA, progression to an increase in symptoms or a change in lung function, and progression to clinical ILD. Understanding of the natural history of progressing ILA and their clinical outcomes should be a focus of future studies. Rigorous validation of biomarkers currently suggested in the literature is needed as well as studies that assess whether proposed biomarkers are a marker of ILA only or are related to progression, either of radiologic abnormalities or to symptomatic and progressive ILD. Integrating respiratory symptoms and lung function test results into future longitudinal studies of ILA could help this aim. Validation of associations of genetic and radiologic findings that have been associated with radiologic development and progression of ILA (Araki et al., 2016; Putman et al., 2019) in more populations and with more outcomes would also prove valuable and it is possible that properly validated automated measures of interstitial findings could help predict progression. The summarized results of such studies will hopefully become sufficient to risk-stratify patients with regards to the risk of progression to ILD, providing a better evidence basis for the follow-up of those with ILA. If sufficiently high-risk patient groups can be defined, trials of anti-fibrotic treatment of patients with ILA could be warranted, which would be a step towards earlier intervention into ILD by screening for, follow-up and treatment of those with ILA, a long-standing aim of ILA research (Hunninghake et al., 2019). The high prevalence of ILA and undiagnosed ILD in close relatives of pulmonary fibrosis patients suggests that this group could be a suitable group for trials of screening and early intervention (Hunninghake et al., 2020). However, it must be mentioned that the possible benefits of early intervention would have to be weighed against the considerable adverse effects of currently available therapies (King et al., 2014; Richeldi et al., 2014), again emphazising the need for selection of patient groups at high risk of negative outcomes. Also, such efforts would have to take the possible psychological burden related to screening and long-term follow-up into account.

Second, it must be mentioned that future studies of ILA would benefit from the use of a standardized definition of ILA in research. An attempt at a unifiying definition has already been made in a Fleischner Society Position Paper (Hatabu et al., 2020). It would greatly aid the comparison and meta-analysis of research findings and the development of clinical guidelines if future studies use a standardized definition such as that one. It must also be mentioned that screening studies of high-risk groups would likely lead to the diagnosis of undiagnosed ILD (Hunninghake et al., 2020). As knowledge and clinical awareness of ILA and early ILD increases, it will be important to have clear terminology and guidelines regarding which abnormalities, in conjunction with

defined clinical and physiologic findings, should be termed a disease and likewise, at which point progressive ILA become early ILD.

Third, as ILA are more common than is ILD, the excess mortality of those with ILA is greater than can be explained by progression to overt ILD (Putman et al., 2016). This raises the question of which other causes of death among those with ILA could explain this. The presented studies suggest that increased risk of lung cancer may play a role in this excess mortality which agrees with some prior findings (Hoyer et al., 2020; Whittaker Brown et al., 2019). While it is possible that this is due to a diverse relationship with aging processes, other diseases, and epidemiological risk factors such as smoking, follow-up studies of people with ILA may establish an association with specific diseases or pathologies. Due to the relatively high prevalence of ILA, this could be useful epidemiologic information if such pathologies account for a substantial part of excess mortality of those with ILA.

Last, some studies of the pathobiology and epidemiology of ILA could be easier to do than studies of the pathobiology of ILD due to the high prevalence of ILA. Such studies could, especially if combined with some risk-stratification so that the group studied represents ILA with high risk of progression, could shed light on common epidemiological risk factors, e.g., those that cause epithelial injury to the lung in the early stages of ILD pathogenesis. Studies of ILA could be for example be well suited to assess the proposed roles of infective agents (Chioma & Drake, 2017) and occupational exposures (Baumgartner et al. 2000) associated with IPF in earlier stages of ILD and to assess whether common comorbidities of IPF such as gastroesophageal reflux (Tobin et al., 1998) are also frequent in those with ILA. As mentioned before, further studies of biomarkers with replicable associations with ILA at high risk of progression could be the basis of new knowledge of the pathogenesis of ILA and early ILD. Studies of biomarkers, comorbidities and epidemiologic risk factors could benefit from novel epidemiological methods that allow for causal inference, such as mendelian randomisation. Any possible causal mechanisms found could then be examined further in tissue specimens from patients with fibrotic lung disease or in animal models of pulmonary fibrosis.

5 Summary and conclusions

Several conclusions can be drawn from the presented findings. ILA are associated with worse self-reported health among the elderly. People with ILA have an increased hazard of lung cancer diagnoses and mortality but are not at increased risk of other types of cancers. Measures of telomere length, a marker of cellular senescence, are inversely associated with ILA odds and lastly, ILA are associated with levels of serum proteins that can be proposed as serum markers of ILA. These findings are mainly from the AGES-Reykjavik cohort, a population of elderly Icelanders, except for findings for telomere length that are mainly based on the COPDGene Study. A recurring pattern in the findings was that associations were stronger for more extensive abnormalities. Overall, major findings were concordant with results of previous studies, either for ILA in other cohorts such as the findings of Paper II and III (Whittaker Brown et al., 2019) (Table E4, Paper III), or for IPF such as the findings of Papers II and IV (Alder et al., 2008; Le Jeune et al., 2007).

Functional status is an important marker of general health among the elderly. Cancer has pathobiologic parallels with aging and telomere length is a driver of cellular senescence, although epidemiological associations with age-related diseases vary. Many aging-associated proteins were associated with ILA but that was not true for all aging markers. Overall, the findings provide some support to the notion that aging-related processes are associated with ILA, a theory that has often been proposed for more advanced forms of pulmonary fibrosis. However, it is not possible to draw a definite conclusion that ILA is caused by accelerated aging from epidemiological studies such as this one due to the complexity of the pathobiology of ILA and ILD. It is also difficult, with findings such as these, to disentangle whether associations apply to accelerated lung aging or accelerated biological aging overall.

However, the findings are useful and a significant addition to the state-of-the-art knowledge of ILA. Notably, the findings highlight adverse health outcomes that are more common among the relatively large proportion of elderly people with ILA. The increased risk of lung cancer and lung cancer mortality is especially relevant in this matter. The results regarding lung cancer and telomere length add to the known associations of advanced pulmonary fibrosis that are now shown to apply to ILA as well. This further demonstrates that at least a part of ILA

cases may belong on a spectrum of ILD and that pathobiological features of pulmonary fibrosis may extend to earlier forms of ILD than previously known. Lastly, candidate biomarkers of ILA are uncovered that could prove valuable in aiding recognition of early forms of ILD and their progression. So, examining ILA in the light of aging-related processes has proved valuable.

The findings presented in the thesis are subject to some limitations overall. First, the characteristics of the main study cohort limit the assumptions that can be made from the findings. Mainly, any extrapolation of the results is limited to the elderly population in Western countries. Additionally, the homogeneity of the Icelandic population in terms of race and genetics suggest that results may not be identical in more diverse populations. Second, while results of all studies are adjusted for important confounders, the role of residual confounding in specific analyses cannot be excluded. Third, although the AGES-Reykjavik cohort is relatively large, the total number of participants with ILA is relatively small. This limits the statistical power of analyses of subgroups of those with ILA, i.e., analyses of progression and imaging patterns. Fourth, because the presented conclusions are based on epidemiological associations, biological causality cannot be inferred from such findings. Further research is needed to elucidate the biological mechanisms underlying these findings.

In conclusion, ILA are associated with several aging-related phenotypes and biological markers. While the results do not suffice to state whether ILA is a direct result of advanced or accelerated aging, approaching this hypothesis has generated knowledge that improves understanding of ILA and has implications for further research.

References

Ahn, M.H., Park, B.L., Lee, S.H., Park, S.W., Park, J.S., Kim, D.J., . . . Shim, Y.-S. (2011). A promoter SNP rs4073T>A in the common allele of the interleukin 8 gene is associated with the development of idiopathic pulmonary fibrosis via the IL-8 protein enhancing mode. *Respiratory Research, 12*(1), 73-73. doi:10.1186/1465-9921-12-73

Alba, M. A., Flores-Suárez, L. F., Henderson, A. G., Xiao, H., Hu, P., Nachman, P. H., . . . Charles Jennette, J. (2017). Interstital lung disease in ANCA vasculitis. *Autoimmunity reviews, 16*(7), 722-729. doi:10.1016/j.autrev.2017.05.008

Alder, J. K., Chen, J. J.-L., Lancaster, L., Danoff, S., Su, S.-c., Cogan, J. D., . . . Armanios, M. Y. (2008). Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proceedings of the National Academy of Sciences, 105*(35), 13051-13056. doi:10.1073/pnas.0804280105

Alder, J. K., Stanley, S. E., Wagner, C. L., Hamilton, M., Hanumanthu, V. S., & Armanios, M. (2015). Exome sequencing identifies mutant TINF2 in a family with pulmonary fibrosis. *Chest, 147*(5), 1361-1368. doi:10.1378/chest.14-1947

Allen, R. J., Guillen-Guio, B., Oldham, J. M., Ma, S. F., Dressen, A., Paynton, M. L., . . . Wain, L. V. (2020). Genome-Wide Association Study of Susceptibility to Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 201*(5), 564-574. doi:10.1164/rccm.201905-1017OC

Allen, R. J., Porte, J., Braybrooke, R., Flores, C., Fingerlin, T. E., Oldham, J. M., . . . Jenkins, R. G. (2017). Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *The Lancet Respiratory Medicine, 5*(11), 869-880. doi:https://doi.org/10.1016/S2213-2600(17)30387-9

American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). (2000). *American Journal of Respiratory and Critical Care Medicine, 161*(2 Pt 1), 646-664. doi:10.1164/ajrccm.161.2.ats3-00

American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. (2002). *American Journal of Respiratory and Critical Care Medicine, 165*(2), 277-304. doi:10.1164/ajrccm.165.2.ats01

Araki, T., Dahlberg, S. E., Hida, T., Lydon, C. A., Rabin, M. S., Hatabu, H., . . . Nishino, M. (2019). Interstitial lung abnormality in stage IV non-small cell lung cancer: A validation study for the association with poor clinical outcome. *Eur J Radiol Open, 6*, 128-131. doi:10.1016/j.ejro.2019.03.003

Araki, T., Nishino, M., Zazueta, O. E., Gao, W., Dupuis, J., Okajima, Y., . . . Hatabu, H. (2015). Paraseptal emphysema: Prevalence and distribution on CT and association with interstitial lung abnormalities. *European Journal of Radiology, 84*(7), 1413-1418. doi:10.1016/j.ejrad.2015.03.010

Araki, T., Putman, R. K., Hatabu, H., Gao, W., Dupuis, J., Latourelle, J. C., . . . Hunninghake, G. M. (2016). Development and Progression of Interstitial Lung Abnormalities in the Framingham Heart Study. *American Journal of Respiratory and Critical Care Medicine, 194*(12), 1514-1522. doi:10.1164/rccm.201512-2523OC

Armanios, M. (2012). Telomerase and idiopathic pulmonary fibrosis. *Mutation Research, 730*(1-2), 52-58. doi:10.1016/j.mrfmmm.2011.10.013

Armanios, M., & Blackburn, E. H. (2012). The telomere syndromes. *Nature reviews. Genetics, 13*(10), 693-704. doi:10.1038/nrg3246

Armanios, M. Y., Chen, J. J., Cogan, J. D., Alder, J. K., Ingersoll, R. G., Markin, C., . . . Loyd, J. E. (2007). Telomerase mutations in families with idiopathic pulmonary fibrosis. *New England Journal of Medicine, 356*(13), 1317-1326. doi:10.1056/NEJMoa066157

Armstrong, H. F., Podolanczuk, A. J., Barr, R. G., Oelsner, E. C., Kawut, S. M., Hoffman, E. A., . . . Lederer, D. J. (2017). Serum Matrix Metalloproteinase-7, Respiratory Symptoms, and Mortality in Community-Dwelling Adults. MESA (Multi-Ethnic Study of Atherosclerosis). *American Journal of Respiratory and Critical Care Medicine, 196*(10), 1311-1317. doi:10.1164/rccm.201701-0254OC

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics, 34*(5), 502-508. doi:10.1111/opo.12131

Ash, S. Y., Harmouche, R., Putman, R. K., Ross, J. C., Diaz, A. A., Hunninghake, G. M., . . . Investigators, C. O. (2017). Clinical and Genetic Associations of Objectively Identified Interstitial Changes in Smokers. *Chest, 152*(4), 780-791. doi:10.1016/j.chest.2017.04.185

Ash, S. Y., Harmouche, R., Ross, J. C., Diaz, A. A., Hunninghake, G. M., Putman, R. K., . . . Washko, G. R. (2017). The Objective Identification and Quantification of Interstitial Lung Abnormalities in Smokers. *Academic Radiology, 24*(8), 941-946. doi:10.1016/j.acra.2016.08.023

Aubert, G., Hills, M., & Lansdorp, P. M. (2012). Telomere length measurement-caveats and a critical assessment of the available technologies and tools. *Mutation Research*, 730(1-2), 59-67. doi:10.1016/j.mrfmmm.2011.04.003

Aunan, J. R., Cho, W. C., & Søreide, K. (2017). The Biology of Aging and Cancer: A Brief Overview of Shared and Divergent Molecular Hallmarks. *Aging and Disease, 8*(5), 628-642. doi:10.14336/AD.2017.0103

Austin, P. C., Lee, D. S., & Fine, J. P. (2016). Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation, 133*(6), 601-609. doi:10.1161/CIRCULATIONAHA.115.017719

Axelsson, G. T., & Gudmundsson, G. (2021). Interstitial lung abnormalities - current knowledge and future directions. *Eur Clin Respir J, 8*(1), 1994178. doi:10.1080/20018525.2021.1994178

Ballester, B., Milara, J., & Cortijo, J. (2019). Idiopathic Pulmonary Fibrosis and Lung Cancer: Mechanisms and Molecular Targets. *International Journal of Molecular Sciences, 20*(3). doi:10.3390/ijms20030593

Baerlocher, G. M., Vulto, I., de Jong, G., & Lansdorp, P. M. (2006). Flow cytometry and FISH to measure the average length of telomeres (flow FISH). *Nature Protocols*, 1(5), 2365-2376. doi:10.1038/nprot.2006.263

Banerjee, D., Perry, M., Tran, D., & Arafat, R. (2010). Self-reported health, functional status and chronic disease in community dwelling older adults: untangling the role of demographics. *Journal of Community Health, 35*(2), 135-141.

Barnes, P. J., Baker, J., & Donnelly, L. E. (2019). Cellular Senescence as a Mechanism and Target in Chronic Lung Diseases. *American Journal of Respiratory and Critical Care Medicine, 200*(5), 556-564. doi:10.1164/rccm.201810-1975TR

Barzilai, N., Huffman, D. M., Muzumdar, R. H., & Bartke, A. (2012). The critical role of metabolic pathways in aging. *Diabetes, 61*(6), 1315-1322. doi:10.2337/db11-1300

Baumgartner, K. B., Samet, J. M., Coultas, D. B., Stidley, C. A., Hunt, W. C., Colby, T. V., . . . Centers, C. (2000). Occupational and Environmental Risk Factors for Idiopathic Pulmonary Fibrosis: A Multicenter Case-Control Study. *American Journal of Epidemiology, 152*(4), 307-315. doi:10.1093/aje/152.4.307 %J American Journal of Epidemiology

Baumgartner, K. B., Samet, J. M., Stidley, C. A., Colby, T. V., & Waldron, J. A. (1997). Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine, 155*(1), 242-248. doi:10.1164/ajrccm.155.1.9001319

Bekaert, S., De Meyer, T. I. M., & Van Oostveldt, P. (2005). Telomere Attrition as Ageing Biomarker. *Anticancer Research, 25*(4), 3011. Retrieved from http://ar.iiarjournals.org/content/25/4/3011.abstract

Belloli, E. A., Beckford, R., Hadley, R., & Flaherty, K. R. (2016). Idiopathic non-specific interstitial pneumonia. *Respirology, 21*(2), 259-268. doi:10.1111/resp.12674

Bennette, C., & Vickers, A. (2012). Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. *BMC Medical Research Methodology, 12*(1), 21. doi:10.1186/1471-2288-12-21

Bermejo-Peláez, D., Ash, S. Y., Washko, G. R., San José Estépar, R., & Ledesma-Carbayo, M. J. (2020). Classification of Interstitial Lung Abnormality Patterns with an Ensemble of Deep Convolutional Neural Networks. *Scientific Reports, 10*(1), 338-338. doi:10.1038/s41598-019-56989-5

Bernardes de Jesus, B., Vera, E., Schneeberger, K., Tejera, A. M., Ayuso, E., Bosch, F., & Blasco, M. A. (2012). Telomerase gene therapy in adult and old mice delays aging and increases longevity without increasing cancer. *EMBO Molecular Medicine, 4*(8), 691-704. doi:10.1002/emmm.201200245

Bernstein, E. J., Barr, R. G., Austin, J. H., Kawut, S. M., Raghu, G., Sell, J. L., . . . Lederer, D. J. (2016). Rheumatoid arthritis-associated autoantibodies and subclinical interstitial lung disease: the Multi-Ethnic Study of Atherosclerosis. *Thorax*. doi:10.1136/thoraxjnl-2016-208932

Bijnen, F. C. H., Feskens, E. J. M., Caspersen, C. J., Nagelkerke, N., Mosterd, W. L., & Kromhout, D. (1999). Baseline and Previous Physical Activity in Relation to Mortality in Elderly Men: The Zutphen Elderly Study. *American Journal of Epidemiology, 150*(12), 1289-1296. doi:10.1093/oxfordjournals.aje.a009960

Bitterman, P. B., Rennard, S. I., Keogh, B. A., Wewers, M. D., Adelberg, S., & Crystal, R. G. (1986). Familial idiopathic pulmonary fibrosis. Evidence of lung inflammation in unaffected family members. *New England Journal of Medicine, 314*(21), 1343-1347. doi:10.1056/nejm198605223142103

Bjornsson, G., Bjornsson, O., Davidsson, D., Kristjansson, B., Olafsson, O., Sigfusson, N., & Thorsteinsson, T. (1968). Report abc XXIV. *Health survey in the Reykjavik area—women. Stages I-III, 1969*, 1971-1972.

Bjornsson, O., Davidsson, D., Olafsson, H., Olafsson, O., Sigfusson, N., & Thorsteinsson, T. (1979). Report ABC XVIII Health Survey in the Reykjavik Area—Men Stages I–III 1967–1969 1970–1971 and 1974–1976 Participants Invitation Response etc. *Reykjavik: The Icelandic Heart Association*.

Boonekamp, J. J., Simons, M. J. P., Hemerik, L., & Verhulst, S. (2013). Telomere length behaves as biomarker of somatic redundancy rather than biological age. *Aging Cell, 12*(2), 330-332. doi:https://doi.org/10.1111/acel.12050

Borie, R., Crestani, B., Dieude, P., Nunes, H., Allanore, Y., Kannengiesser, C., . . . Grandchamp, B. (2013). The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population. *PloS One, 8*(8), e70621-e70621. doi:10.1371/journal.pone.0070621

Borie, R., Tabèze, L., Thabut, G., Nunes, H., Cottin, V., Marchand-Adam, S., . . . Crestani, B. (2016). Prevalence and characteristics of TERT and TERC mutations in suspected genetic pulmonary fibrosis. *European Respiratory Journal, 48*(6), 1721. doi:10.1183/13993003.02115-2015

Bozzetti, F., Paladini, I., Rabaiotti, E., Franceschini, A., Alfieri, V., Chetta, A., . . . Sverzellati, N. (2016). Are interstitial lung abnormalities associated with COPD? A nested case-control study. *International Journal of Chronic Obstructive Pulmonary Disease, 11*, 1087-1096. doi:10.2147/COPD.S103256

Brenner, D. J., & Hall, E. J. (2007). Computed Tomography — An Increasing Source of Radiation Exposure. *New England Journal of Medicine, 357*(22), 2277-2284. doi:10.1056/NEJMra072149

Buendía-Roldán, I., Fernandez, R., Mejía, M., Juarez, F., Ramirez-Martinez, G., Montes, E., . . . Selman, M. (2021). Risk factors associated with the development of interstitial lung abnormalities. *European Respiratory Journal*, 2003005. doi:10.1183/13993003.03005-2020

Buendia-Roldan, I., Machuca, N., Mejía, M., Maldonado, M., Pardo, A., & Selman, M. (2019). Lower levels of α-Klotho in serum are associated with decreased lung function in individuals with interstitial lung abnormalities. *Scientific Reports, 9*(1), 10801-10801. doi:10.1038/s41598-019-47199-0

Bueno, M., Lai, Y.-C., Romero, Y., Brands, J., St. Croix, C. M., Kamga, C., . . . Mora, A. L. (2015). PINK1 deficiency impairs mitochondrial homeostasis and promotes lung fibrosis. *The Journal of Clinical Investigation, 125*(2), 521-538. doi:10.1172/JCI74942

Campisi, J. (2013). Aging, cellular senescence, and cancer. *Annual Review of Physiology, 75*, 685-705. doi:10.1146/annurev-physiol-030212-183653

Campo, I., Zorzetto, M., Mariani, F., Kadija, Z., Morbini, P., Dore, R., . . . Luisetti, M. (2014). A large kindred of pulmonary fibrosis associated with a novel ABCA3 gene variant. *Respiratory Research, 15*(1), 43-43. doi:10.1186/1465-9921-15-43

Candia, J., Cheung, F., Kotliarov, Y., Fantoni, G., Sellers, B., Griesman, T., . . . Biancotto, A. (2017). Assessment of Variability in the SOMAscan Assay. *Scientific Reports, 7*(1), 14248. doi:10.1038/s41598-017-14755-5

Castaldi, P. J., Dy, J., Ross, J., Chang, Y., Washko, G. R., Curran-Everett, D., . . . Cho, M. H. (2014). Cluster analysis in the COPDGene study identifies subtypes of smokers with distinct patterns of airway disease and emphysema. *Thorax, 69*(5), 415-422. doi:10.1136/thoraxjnl-2013-203601

Cawthon, R. M. (2002). Telomere measurement by quantitative PCR. *Nucleic Acids Research, 30*(10), e47-e47. doi:10.1093/nar/30.10.e47

Cha, S. I., Ryerson, C. J., Lee, J. S., Kukreja, J., Barry, S. S., Jones, K. D., . . . Wolters, P. J. (2012). Cleaved cytokeratin-18 is a mechanistically informative biomarker in idiopathic pulmonary fibrosis. *Respiratory Research, 13*(1), 105. doi:10.1186/1465-9921-13-105

Chan, A. L., Juarez, M. M., Leslie, K. O., Ismail, H. A., & Albertson, T. E. (2012). Bird fancier's lung: a state-of-the-art review. *Clinical Reviews in Allergy and Immunology, 43*(1-2), 69-83. doi:10.1007/s12016-011-8282-y

Chioma, O. S., & Drake, W. P. (2017). Role of Microbial Agents in Pulmonary Fibrosis. *The Yale journal of biology and medicine, 90*(2), 219-227. Retrieved from https://pubmed.ncbi.nlm.nih.gov/28656009

Choi, B., Kawut, S. M., Raghu, G., Hoffman, E., Tracy, R., Madahar, P., . . . Podolanczuk, A. (2020). Regional distribution of high-attenuation areas on chest computed tomography in the Multi-Ethnic Study of Atherosclerosis. *ERJ Open Res, 6*(1), 00115-02019. doi:10.1183/23120541.00115-2019

Chubachi, S., Takahashi, S., Tsutsumi, A., Kameyama, N., Sasaki, M., Naoki, K., . . . Betsuyaku, T. (2017). Radiologic features of precancerous areas of the lungs in chronic obstructive pulmonary disease. *International Journal of Chronic Obstructive Pulmonary Disease, 12*, 1613-1624. doi:10.2147/COPD.S132709

Cogan, J. D., Kropski, J. A., Zhao, M., Mitchell, D. B., Rives, L., Markin, C., . . . Blackwell, T. S. (2015). Rare variants in RTEL1 are associated with familial interstitial pneumonia. *American Journal of Respiratory and Critical Care Medicine, 191*(6), 646-655. doi:10.1164/rccm.201408-1510OC

Collard, H. R., Ryerson, C. J., Corte, T. J., Jenkins, G., Kondoh, Y., Lederer, D. J., . . . Martinez, F. J. (2016). Acute Exacerbation of Idiopathic Pulmonary Fibrosis. An International Working Group Report. *American Journal of Respiratory and Critical Care Medicine, 194*(3), 265-275. doi:10.1164/rccm.201604-0801CI

Cordier, J. F. (2006). Cryptogenic organising pneumonia. *European Respiratory Journal, 28*(2), 422. doi:10.1183/09031936.06.00013505

Cottin, V. (2016). Idiopathic interstitial pneumonias with connective tissue diseases features: A review. *Respirology, 21*(2), 245-258. doi:10.1111/resp.12588

Cottin, V., Hirani, N. A., Hotchkin, D. L., Nambiar, A. M., Ogura, T., Otaola, M., . . . Wells, A. U. (2018). Presentation, diagnosis and clinical course of the spectrum of progressive-fibrosing interstitial lung diseases. *European Respiratory Review, 27*(150), 180076. doi:10.1183/16000617.0076-2018

Cronkhite, J. T., Xing, C., Raghu, G., Chin, K. M., Torres, F., Rosenblatt, R. L., & Garcia, C. K. (2008). Telomere shortening in familial and sporadic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine, 178*(7), 729-737. doi:10.1164/rccm.200804-550OC

Cullinan, P., & Reid, P. (2013). Pneumoconiosis. *Primary Care Respiratory Journal, 22*(2), 249-252. doi:10.4104/pcrj.2013.00055

Dai, J., Cai, H., Li, H., Zhuang, Y., Min, H., Wen, Y., . . . Yi, L. (2015). Association between telomere length and survival in patients with idiopathic pulmonary fibrosis. *Respirology, 20*(6), 947-952. doi:10.1111/resp.12566

Demopoulos, K., Arvanitis, D. A., Vassilakis, D. A., Siafakas, N. M., & Spandidos, D. A. (2002). MYCL1, FHIT, SPARC, p16INK4and TP53 genes associated to lung cancer in idiopathic pulmonary fibrosis. *Journal of Cellular and Molecular Medicine, 6*(2), 215-222. doi:10.1111/j.1582-4934.2002.tb00188.x

Disayabutr, S., Kim, E. K., Cha, S.-I., Green, G., Naikawadi, R. P., Jones, K. D., . . . Wolters, P. J. (2016). miR-34 miRNAs Regulate Cellular Senescence in Type II Alveolar Epithelial Cells of Patients with Idiopathic Pulmonary Fibrosis. *PloS One, 11*(6), e0158367. doi:10.1371/journal.pone.0158367

Djoussé, L., Rothman, K. J., Cupples, L. A., Levy, D., & Ellison, R. C. (2002). Serum Albumin and Risk of Myocardial Infarction and All-Cause Mortality in the Framingham Offspring Study. *Circulation, 106*(23), 2919-2924. doi:10.1161/01.CIR.0000042673.07632.76

Dong, H., Julien, P. J., Demoruelle, M. K., Deane, K. D., & Weisman, M. H. (2018). Interstitial lung abnormalities in patients with early rheumatoid arthritis: A pilot study evaluating prevalence and progression. *European journal of rheumatology, 6*(4), 193-198. doi:10.5152/eurjrheum.2019.19044

Dou, Q. P., & Zonder, J. A. (2014). Overview of proteasome inhibitor-based anti-cancer therapies: perspective on bortezomib and second generation proteasome inhibitors versus future generation inhibitors of ubiquitin-proteasome system. *Current Cancer Drug Targets, 14*(6), 517-536. doi:10.2174/1568009614666140804154511

Dowman, L. M., McDonald, C. F., Hill, C. J., Lee, A. L., Barker, K., Boote, C., . . . Holland, A. E. (2017). The evidence of benefits of exercise training in interstitial lung disease: a randomised controlled trial. *Thorax, 72*(7), 610. doi:10.1136/thoraxjnl-2016-208638

Doyle, T. J., Dellaripa, P. F., Batra, K., Frits, M. L., Iannaccone, C. K., Hatabu, H., . . . Rosas, I. O. (2014). Functional impact of a spectrum of interstitial lung abnormalities in rheumatoid arthritis. *Chest, 146*(1), 41-50. doi:10.1378/chest.13-1394

Doyle, T. J., Washko, G. R., Fernandez, I. E., Nishino, M., Okajima, Y., Yamashiro, T., . . . Hunninghake, G. M. (2012). Interstitial lung abnormalities and reduced exercise capacity. *American Journal of Respiratory and Critical Care Medicine, 185*(7), 756-762. doi:10.1164/rccm.201109-1618OC

Duckworth, A., Gibbons, M. A., Allen, R. J., Almond, H., Beaumont, R. N., Wood, A. R., . . . Scotton, C. J. (2021). Telomere length and risk of idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease: a mendelian randomisation study. *The Lancet Respiratory Medicine, 9*(3), 285-294. doi:https://doi.org/10.1016/S2213-2600(20)30364-7

Dudbridge, F., Allen, R. J., Sheehan, N. A., Schmidt, A. F., Lee, J. C., Jenkins, R. G., . . . Patel, R. S. (2019). Adjustment for index event bias in genome-wide association studies of subsequent events. *Nature Communications, 10*(1), 1561-1561. doi:10.1038/s41467-019-09381-w

Easthausen, I., Podolanczuk, A., Hoffman, E., Kawut, S., Oelsner, E., Kim, J. S., . . . Lederer, D. J. (2020). Reference values for high attenuation areas on chest CT in a healthy, never-smoker, multi-ethnic sample: The MESA study. *Respirology*. doi:10.1111/resp.13783

Emilsson, V., Ilkov, M., Lamb, J. R., Finkel, N., Gudmundsson, E. F., Pitts, R., . . . Gudnason, V. (2018). Co-regulatory networks of human serum proteins link genetics to disease. *Science, 361*(6404), 769-773. doi:10.1126/science.aaq1327

Esposito, D. B., Lanes, S., Donneyong, M., Holick, C. N., Lasky, J. A., Lederer, D., . . . Tran, T. N. (2015). Idiopathic Pulmonary Fibrosis in United States Automated Claims. Incidence, Prevalence, and Algorithm Validation. *American Journal of Respiratory and Critical Care Medicine, 192*(10), 1200-1207. doi:10.1164/rccm.201504-0818OC

Evans, C. M., Fingerlin, T. E., Schwarz, M. I., Lynch, D., Kurche, J., Warg, L., . . . Schwartz, D. A. (2016). Idiopathic Pulmonary Fibrosis: A Genetic Disease That Involves Mucociliary Dysfunction of the Peripheral Airways. *Physiological Reviews, 96*(4), 1567-1591. doi:10.1152/physrev.00004.2016

Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnau, V., . . . Hermjakob, H. (2017). Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics, 18*(1), 142. doi:10.1186/s12859-017-1559-2

Fine, J. P., & Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association, 94*(446), 496-509. doi:10.1080/01621459.1999.10474144

Fingerlin, T. E., Murphy, E., Zhang, W., Peljto, A. L., Brown, K. K., Steele, M. P., . . . Schwartz, D. A. (2013). Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nature Genetics, 45*(6), 613-620. doi:10.1038/ng.2609

Fingerlin, T. E., Zhang, W., Yang, I. V., Ainsworth, H. C., Russell, P. H., Blumhagen, R. Z., . . . Schwartz, D. A. (2016). Genome-wide imputation study identifies novel HLA locus for pulmonary fibrosis and potential role for auto-immunity in fibrotic idiopathic interstitial pneumonia. *BMC Genetics, 17*(1), 74-74. doi:10.1186/s12863-016-0377-2

Franceschi, C., Garagnani, P., Parini, P., Giuliani, C., & Santoro, A. (2018). Inflammaging: a new immune–metabolic viewpoint for age-related diseases. *Nature Reviews Endocrinology, 14*(10), 576-590. doi:10.1038/s41574-018-0059-4

Gene Ontology Consortium. (2021). The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Research, 49*(D1), D325-d334. doi:10.1093/nar/gkaa1113

Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., . . . Ruepp, A. (2019). CORUM: the comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Research, 47*(D1), D559-d563. doi:10.1093/nar/gky973

Gochuico, B. R., Avila, N. A., Chow, C. K., Novero, L. J., Wu, H. P., Ren, P., . . . Rosas, I. O. (2008). Progressive preclinical interstitial lung disease in rheumatoid arthritis. *Archives of Internal Medicine, 168*(2), 159-166. doi:10.1001/archinternmed.2007.59

Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E. N., . . . Zichi, D. (2010). Aptamer-based multiplexed proteomic technology for biomarker discovery. *PloS One, 5*(12), e15004-e15004. doi:10.1371/journal.pone.0015004

GTEx Consortium. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science, 348*(6235), 648-660. doi:10.1126/science.1262110

Gu, K., Lee, H. Y., Lee, K., Choi, J. Y., Woo, S. Y., Sohn, I., . . . Shim, Y. M. (2019). Integrated evaluation of clinical, pathological and radiological prognostic factors in squamous cell carcinoma of the lung. *PloS One, 14*(10), e0223298-e0223298. doi:10.1371/journal.pone.0223298

Guckel, C., & Hansell, D. M. (1998). Imaging the 'dirty lung'–has high resolution computed tomography cleared the smoke? *Clinical Radiology, 53*(10), 717-722.

Gudjonsson, A., Gudmundsdottir, V., Axelsson, G. T., Gudmundsson, E. F., Jonsson, B. G., Launer, L. J., . . . Gudnason, V. (2022). A genome-wide association study of serum proteins reveals shared loci with common diseases. *Nat Commun, 13*(1), 480. doi:10.1038/s41467-021-27850-z

Hancock, L. A., Hennessy, C. E., Solomon, G. M., Dobrinskikh, E., Estrella, A., Hara, N., . . . Schwartz, D. A. (2018). Muc5b overexpression causes mucociliary dysfunction and enhances lung fibrosis in mice. *Nature Communications, 9*(1), 5363. doi:10.1038/s41467-018-07768-9

Harmouche, R., Ash, S. Y., Putman, R. K., Hunninghake, G. M., San Jose Estepar, R., Martinez, F. J., . . . San Jose Estepar, R. (2019). Objectively Measured Chronic Lung Injury on Chest CT. *Chest, 156*(6), 1149-1159. doi:https://doi.org/10.1016/j.chest.2019.05.020

Harrell Jr., F. E. (2016). *Regression modelling strategies*. Cham, Switzerland: Springer International Publishing.

Harrell Jr., F. E. (2020). rms: Regression Modeling Strategies. R package retrieved from https://CRAN.R-project.org/package=rms

Harris, T. B., Launer, L. J., Eiriksdottir, G., Kjartansson, O., Jonsson, P. V., Sigurdsson, G., . . . Gudnason, V. (2007). Age, Gene/Environment Susceptibility – Reykjavik Study: Multidisciplinary Applied Phenomics. *American Journal of Epidemiology, 165*(9), 1076-1087. doi:10.1093/aje/kwk115

Hatabu, H., Hunninghake, G. M., Richeldi, L., Brown, K. K., Wells, A. U., Remy-Jardin, M., . . . Lynch, D. A. (2020). Interstitial lung abnormalities detected incidentally on CT: a Position Paper from the Fleischner Society. *The Lancet Respiratory Medicine, 8*(7), 726-737. doi:10.1016/S2213-2600(20)30168-5

Hewson, T., McKeever, T. M., Gibson, J. E., Navaratnam, V., Hubbard, R. B., & Hutchinson, J. P. (2018). Timing of onset of symptoms in people with idiopathic pulmonary fibrosis. *Thorax, 73*(7), 683. doi:10.1136/thoraxjnl-2017-210177

Hill, C., Jones, M. G., Davies, D. E., & Wang, Y. (2019). Epithelial-mesenchymal transition contributes to pulmonary fibrosis via aberrant epithelial/fibroblastic cross-talk. *Journal of lung health and diseases, 3*(2), 31-35. Retrieved from https://pubmed.ncbi.nlm.nih.gov/31032489

Hiyama, E., & Hiyama, K. (2007). Telomere and telomerase in stem cells. *British Journal of Cancer, 96*(7), 1020-1024. doi:10.1038/sj.bjc.6603671

Ho, J. E., Gao, W., Levy, D., Santhanakrishnan, R., Araki, T., Rosas, I. O., . . . Hunninghake, G. M. (2016). Galectin-3 Is Associated with Restrictive Lung Disease and Interstitial Lung Abnormalities. *American Journal of Respiratory and Critical Care Medicine, 194*(1), 77-83. doi:10.1164/rccm.201509-1753OC

Hobbs, B. D., Putman, R. K., Araki, T., Nishino, M., Gudmundsson, G., Gudnason, V., . . . Hunninghake, G. M. (2019). Overlap of Genetic Risk between Interstitial Lung Abnormalities and Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 200*(11), 1402-1413. doi:10.1164/rccm.201903-0511OC

Hoeijmakers, J. H. J. (2009). DNA Damage, Aging, and Cancer. *New England Journal of Medicine, 361*(15), 1475-1485. doi:10.1056/NEJMra0804615

Hoyer, N., Thomsen, L. H., Wille, M. M. W., Wilcke, T., Dirksen, A., Pedersen, J. H., . . . Shaker, S. B. (2020). Increased respiratory morbidity in individuals with interstitial lung abnormalities. *BMC Pulmonary Medicine, 20*(1), 67-67. doi:10.1186/s12890-020-1107-0

Hoyer, N., Wille, M. M. W., Thomsen, L. H., Wilcke, T., Dirksen, A., Pedersen, J. H., . . . Shaker, S. B. (2018). Interstitial lung abnormalities are associated with increased mortality in smokers. *Respiratory Medicine, 136*, 77-82. doi:10.1016/j.rmed.2018.02.001

Huang, Y., Ma, S.F., Espindola, M. S., Vij, R., Oldham, J. M., Huffnagle, G. B., . . . Noth, I. (2017). Microbes Are Associated with Host Innate Immune Response in Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 196*(2), 208-219. doi:10.1164/rccm.201607-1525OC

Hubbard, R., Venn, A., Lewis, S., & Britton, J. (2000). Lung cancer and cryptogenic fibrosing alveolitis. A population-based cohort study. *American Journal of Respiratory and Critical Care Medicine, 161*(1), 5-8. doi:10.1164/ajrccm.161.1.9906062

Hung, Y. P., Hunninghake, G. M., Miller, E. R., Putman, R., Nishino, M., Araki, T., . . . Vivero, M. (2019). Incidental nonneoplastic parenchymal findings in patients undergoing lung resection for mass lesions. *Human Pathology, 86*, 93-101. doi:https://doi.org/10.1016/j.humpath.2019.01.002

Hunninghake, G. M. (2019). Interstitial lung abnormalities: erecting fences in the path towards advanced pulmonary fibrosis. *Thorax*, thoraxjnl-2018-212446. doi:10.1136/thoraxjnl-2018-212446

Hunninghake, G. M., Hatabu, H., Okajima, Y., Gao, W., Dupuis, J., Latourelle, J. C., . . . Schwartz, D. A. (2013). MUC5B promoter polymorphism and interstitial lung abnormalities. *New England Journal of Medicine, 368*(23), 2192-2200. doi:10.1056/NEJMoa1216076

Hunninghake, G. M., Quesada-Arias, L. D., Carmichael, N. E., Martinez Manzano, J. M., Poli De Frias, S., Alvarez Baumgartner, M., . . . Rosas, I. O. (2020). Interstitial Lung Disease in Relatives of Patients with Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine*. doi:10.1164/rccm.201908-1571OC

Hutchinson, J., Fogarty, A., Hubbard, R., & McKeever, T. (2015). Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *European Respiratory Journal, 46*(3), 795. doi:10.1183/09031936.00185114

Idiopathic Pulmonary Fibrosis Clinical Research Network, N., Raghu, G., Anstrom, K. J., King, T. E., Jr., Lasky, J. A., & Martinez, F. J. (2012). Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *The New England Journal of Medicine, 366*(21), 1968-1977. doi:10.1056/NEJMoa1113354

Im, Y., Park, H. Y., Shin, S., Shin, S. H., Lee, H., Ahn, J. H., . . . Kim, J. (2019). Prevalence of and risk factors for pulmonary complications after curative resection in otherwise healthy elderly patients with early stage lung cancer. *Respiratory Research, 20*(1), 136-136. doi:10.1186/s12931-019-1087-x

Iwasawa, T., Okudela, K., Takemura, T., Fukuda, T., Matsushita, S., Baba, T., . . . Yoshizawa, A. (2019). Computer-aided Quantification of Pulmonary Fibrosis in Patients with Lung Cancer: Relationship to Disease-free Survival. *Radiology, 292*(2), 489-498. doi:10.1148/radiol.2019182466

Jeck, W. R., Siebold, A. P., & Sharpless, N. E. (2012). Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell, 11*(5), 727-731. doi:10.1111/j.1474-9726.2012.00871.x

Jin, G. Y., Lynch, D., Chawla, A., Garg, K., Tammemagi, M. C., Sahin, H., . . . Kwon, K. S. (2013). Interstitial lung abnormalities in a CT lung cancer screening population: prevalence and progression rate. *Radiology, 268*(2), 563-571. doi:10.1148/radiol.13120816

Johnson, A. A., Shokhirev, M. N., Wyss-Coray, T., & Lehallier, B. (2020). Systematic review and analysis of human proteomics aging studies unveils a novel proteomic aging clock and identifies key processes that change with age. *Ageing Res Rev, 60*, 101070. doi:https://doi.org/10.1016/j.arr.2020.101070

Johnson, C., Rosen, P., Lloyd, T., Horton, M., Christopher-Stine, L., Oddis, C. V., . . . Danoff, S. K. (2017). Exploration of the MUC5B promoter variant and ILD risk in patients with autoimmune myositis. *Respiratory Medicine, 130*, 52-54. doi:https://doi.org/10.1016/j.rmed.2017.07.010

Joshi, A., & Mayr, M. (2018). In Aptamers They Trust. *Circulation, 138*(22), 2482-2485. doi:10.1161/CIRCULATIONAHA.118.036823

Juge, P. A., Lee, J. S., Ebstein, E., Furukawa, H., Dobrinskikh, E., Gazal, S., . . . Dieudé, P. (2018). MUC5B Promoter Variant and Rheumatoid Arthritis with Interstitial Lung Disease. *New England Journal of Medicine, 379*(23), 2209-2219. doi:10.1056/NEJMoa1801562

Kage, H., & Borok, Z. (2012). EMT and interstitial lung disease: a mysterious relationship. *Current Opinion in Pulmonary Medicine, 18*(5), 517-523. doi:10.1097/MCP.0b013e3283566721

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research, 28*(1), 27-30. doi:10.1093/nar/28.1.27

Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J., & Castelli, W. P. (1979). AN INVESTIGATION OF CORONARY HEART DISEASE IN FAMILIES: THE FRAMINGHAM OFFSPRING STUDY. *American Journal of Epidemiology, 110*(3), 281-290. doi:10.1093/oxfordjournals.aje.a112813

Kasai, H., Allen, J. T., Mason, R. M., Kamimura, T., & Zhang, Z. (2005). TGF-beta1 induces human alveolar epithelial to mesenchymal cell transition (EMT). *Respiratory Research, 6*(1), 56-56. doi:10.1186/1465-9921-6-56

Kaur, A., Mathai, S. K., & Schwartz, D. A. (2017). Genetics in Idiopathic Pulmonary Fibrosis Pathogenesis, Prognosis, and Treatment. *Front Med (Lausanne), 4*, 154. doi:10.3389/fmed.2017.00154

Keller, J. P., & Rice, K. M. (2018). Selecting Shrinkage Parameters for Effect Estimation: The Multi-Ethnic Study of Atherosclerosis. *American Journal of Epidemiology, 187*(2), 358-365. doi:10.1093/aje/kwx225

Kim, J. S., Anderson, M. R., Bernstein, E. J., Oelsner, E. C., Raghu, G., Noth, I., . . . Podolanczuk, A. J. (2021). Associations of D-dimer with CT Lung Abnormalities, Serum Biomarkers of Lung Injury, and Forced Vital Capacity: MESA Lung Study. *Ann Am Thorac Soc, 0*(ja), null. doi:10.1513/AnnalsATS.202012-1557OC

Kim, J. S., Anderson, M. R., Podolanczuk, A. J., Kawut, S. M., Allison, M. A., Raghu, G., . . . Giles, J. T. (2020). Associations of Serum Adipokines With Subclinical Interstitial Lung Disease Among Community-Dwelling Adults: The Multi-Ethnic Study of Atherosclerosis (MESA). *Chest, 157*(3), 580-589. doi:https://doi.org/10.1016/j.chest.2019.10.011

Kim, J. S., Podolanczuk, A. J., Borker, P., Kawut, S. M., Raghu, G., Kaufman, J. D., . . . Lederer, D. J. (2017). Obstructive Sleep Apnea and Subclinical Interstitial Lung Disease in the Multi-Ethnic Study of Atherosclerosis (MESA). *Ann Am Thorac Soc, 14*(12), 1786-1795. doi:10.1513/AnnalsATS.201701-091OC

Kim, M. H., Bae, C. H., Choi, Y. S., Na, H. G., Song, S. Y., & Kim, Y. D. (2019). Endoplasmic Reticulum Stress Induces MUC5AC and MUC5B Expression in Human Nasal Airway Epithelial Cells. *Clinical and Experimental Otorhinolaryngology, 12*(2), 181-189. doi:10.21053/ceo.2018.00493

Kim, S. M., Zhao, D., Podolanczuk, A. J., Lutsey, P. L., Guallar, E., Kawut, S. M., . . . Michos, E. D. (2018). Serum 25-Hydroxyvitamin D Concentrations Are Associated with Computed Tomography Markers of Subclinical Interstitial Lung Disease among Community-Dwelling Adults in the Multi-Ethnic Study of Atherosclerosis (MESA). *Journal of Nutrition, 148*(7), 1126-1134. doi:10.1093/jn/nxy066

Kimura, M., Stone, R. C., Hunt, S. C., Skurnick, J., Lu, X., Cao, X., . . . Aviv, A. (2010). Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. *Nature Protocols*, *5*(9), 1596-1607. doi:10.1038/nprot.2010.124

King, T. E., Bradford, W. Z., Castro-Bernardini, S., Fagan, E. A., Glaspole, I., Glassberg, M. K., . . . Noble, P. W. (2014). A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine, 370*(22), 2083-2092. doi:10.1056/NEJMoa1402582

King, T. E., Jr., Pardo, A., & Selman, M. (2011). Idiopathic pulmonary fibrosis. *Lancet, 378*(9807), 1949-1961. doi:10.1016/s0140-6736(11)60052-4

Kliment, C. R., Araki, T., Doyle, T. J., Gao, W., Dupuis, J., Latourelle, J. C., . . . Hunninghake, G. M. (2015). A comparison of visual and quantitative methods to identify interstitial lung abnormalities. *BMC Pulmonary Medicine, 15*, 134. doi:10.1186/s12890-015-0124-x

Köhler, S., Gargano, M., Matentzoglu, N., Carmody, L. C., Lewis-Smith, D., Vasilevsky, N. A., . . . Robinson, P. N. (2020). The Human Phenotype Ontology in 2021. *Nucleic Acids Research, 49*(D1), D1207-D1217. doi:10.1093/nar/gkaa1043

Kolb, M., Bonella, F., & Wollin, L. (2017). Therapeutic targets in idiopathic pulmonary fibrosis. *Respiratory Medicine, 131*, 49-57. doi:https://doi.org/10.1016/j.rmed.2017.07.062

Korfei, M., Ruppert, C., Mahavadi, P., Henneke, I., Markart, P., Koch, M., . . . Guenther, A. (2008). Epithelial endoplasmic reticulum stress and apoptosis in sporadic idiopathic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine, 178*(8), 838-846. doi:10.1164/rccm.200802-313OC

Korthagen, N. M., van Moorsel, C. H. M., Kazemier, K. M., Ruven, H. J. T., & Grutters, J. C. (2012). IL1RN genetic variations and risk of IPF: a meta-analysis and mRNA expression study. *Immunogenetics, 64*(5), 371-377. doi:10.1007/s00251-012-0604-6

Kropski, J. A., Mitchell, D. B., Markin, C., Polosukhin, V. V., Choi, L., Johnson, J. E., . . . Loyd, J. E. (2014). A novel dyskerin (DKC1) mutation is associated with familial interstitial pneumonia. *Chest, 146*(1), e1-e7. doi:10.1378/chest.13-2224

Kudryavtseva, A. V., Krasnov, G. S., Dmitriev, A. A., Alekseev, B. Y., Kardymon, O. L., Sadritdinova, A. F., . . . Snezhkina, A. V. (2016). Mitochondrial dysfunction and oxidative stress in aging and cancer. *Oncotarget, 7*(29), 44879-44905. doi:10.18632/oncotarget.9821

Kuwano, K., Nakashima, N., Inoshima, I., Hagimoto, N., Fujita, M., Yoshimi, M., . . . Hara, N. (2003). Oxidative stress in lung epithelial cells from patients with idiopathic interstitial pneumonias. *European Respiratory Journal, 21*(2), 232. doi:10.1183/09031936.03.00063203

Lancaster, L. H., Mason, W. R., Parnell, J. A., Rice, T. W., Loyd, J. E., Milstone, A. P., . . . Malow, B. A. (2009). Obstructive Sleep Apnea Is Common in Idiopathic Pulmonary Fibrosis. *Chest, 136*(3), 772-778. doi:https://doi.org/10.1378/chest.08-2776

Lawson, W. E., Crossno, P. F., Polosukhin, V. V., Roldan, J., Cheng, D.-S., Lane, K. B., . . . Blackwell, T. S. (2008). Endoplasmic reticulum stress in alveolar epithelial cells is prominent in IPF: association with altered surfactant protein processing and herpesvirus infection. *American Journal of Physiology-Lung Cellular and Molecular Physiology, 294*(6), L1119-L1126. doi:10.1152/ajplung.00382.2007

Le Jeune, I., Gribbin, J., West, J., Smith, C., Cullinan, P., & Hubbard, R. (2007). The incidence of cancer in patients with idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Respiratory Medicine, 101*(12), 2534-2540. doi:https://doi.org/10.1016/j.rmed.2007.07.012

Leader, J. K., Crothers, K., Huang, L., King, M. A., Morris, A., Thompson, B. W., . . . Diaz, P. T. (2016). Risk Factors Associated With Quantitative Evidence of Lung Emphysema and Fibrosis in an HIV-Infected Cohort. *Journal of Acquired Immune Deficiency Syndromes*, 71(4), 420-427. doi:10.1097/qai.0000000000000894

Lederer, D. J., Enright, P. L., Kawut, S. M., Hoffman, E. A., Hunninghake, G., van Beek, E. J. R., . . . Barr, R. G. (2009). Cigarette Smoking Is Associated with Subclinical Parenchymal Lung Disease: The Multi-Ethnic Study of Atherosclerosis (MESA)—Lung Study. *American Journal of Respiratory and Critical Care Medicine, 180*(5), 407-414. doi:10.1164/rccm.200812-1966OC

Lederer, D. J., & Martinez, F. J. (2018). Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine, 378*(19), 1811-1823. doi:10.1056/NEJMra1705751

Lee, T. S., Jin, K. N., Lee, H. W., Yoon, S.Y., Park, T. Y., Heo, E. Y., . . . Lee, J.-K. (2021). Interstitial Lung Abnormalities and the Clinical Course in Patients With COPD. *Chest, 159*(1), 128-137. doi:https://doi.org/10.1016/j.chest.2020.08.017

Ley, B., Brown, K. K., & Collard, H. R. (2014). Molecular biomarkers in idiopathic pulmonary fibrosis. *American journal of physiology. Lung cellular and molecular physiology, 307*(9), L681-L691. doi:10.1152/ajplung.00014.2014

Li, F., Zhou, Z., Wu, A., Cai, Y., Wu, H., Chen, M., & Liang, S. (2018). Preexisting radiological interstitial lung abnormalities are a risk factor for severe radiation pneumonitis in patients with small-cell lung cancer after thoracic radiation therapy. *Radiation Oncology (London, England), 13*(1), 82-82. doi:10.1186/s13014-018-1030-1

Liu, B., Wang, J., Chan, K. M., Tjia, W. M., Deng, W., Guan, X., . . . Zhou, Z. (2005). Genomic instability in laminopathy-based premature aging. *Nature Medicine, 11*(7), 780-785. doi:10.1038/nm1266

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The Hallmarks of Aging. *Cell, 153*(6), 1194-1217. doi:http://dx.doi.org/10.1016/j.cell.2013.05.039

Lynch, D. A., Sverzellati, N., Travis, W. D., Brown, K. K., Colby, T. V., Galvin, J. R., . . . Wells, A. U. (2018). Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. *The Lancet Respiratory Medicine, 6*(2), 138-153. doi:https://doi.org/10.1016/S2213-2600(17)30433-2

Mackinnon, A. C., Gibbons, M. A., Farnworth, S. L., Leffler, H., Nilsson, U. J., Delaine, T., . . . Sethi, T. (2012). Regulation of transforming growth factor-beta1-driven lung fibrosis by galectin-3. *American Journal of Respiratory and Critical Care Medicine, 185*(5), 537-546. doi:10.1164/rccm.201106-0965OC

Mackintosh, J. A., Marshall, H. M., Slaughter, R., Reddy, T., Yang, I. A., Bowman, R. V., & Fong, K. M. (2019). Interstitial lung abnormalities in the Queensland Lung Cancer Screening Study: prevalence and progression over 2 years of surveillance. *49*(7), 843-849. doi:10.1111/imj.14148

Madahar, P., Duprez, D. A., Podolanczuk, A. J., Bernstein, E. J., Kawut, S. M., Raghu, G., . . . Lederer, D. J. (2018). Collagen biomarkers and subclinical interstitial lung disease: The Multi-Ethnic Study of Atherosclerosis. *Respiratory Medicine, 140*, 108-114. doi:10.1016/j.rmed.2018.06.001

Mahmood, S. S., Levy, D., Vasan, R. S., & Wang, T. J. (2014). The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet (London, England), 383*(9921), 999-1008. doi:10.1016/S0140-6736(13)61752-3

Manichaikul, A., Wang, X.-Q., Sun, L., Dupuis, J., Borczuk, A. C., Nguyen, J. N., . . . Lederer, D. J. (2017). Genome-wide association study of subclinical interstitial lung disease in MESA. *Respiratory Research, 18*(1), 97-97. doi:10.1186/s12931-017-0581-2

Margaritopoulos, G. A., Harari, S., Caminati, A., & Antoniou, K. M. (2016). Smoking-related idiopathic interstitial pneumonia: A review. *Respirology, 21*(1), 57-64. doi:10.1111/resp.12576

Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, Denise N., Hanspers, K., . . . Kutmon, M. (2020). WikiPathways: connecting communities. *Nucleic Acids Research, 49*(D1), D613-D621. doi:10.1093/nar/gkaa1024

Mathai, S. K., Yang, I. V., Schwarz, M. I., & Schwartz, D. A. (2015). Incorporating genetics into the identification and treatment of Idiopathic Pulmonary Fibrosis. *BMC Medicine, 13*(1), 191. doi:10.1186/s12916-015-0434-0

Mather, K. A., Jorm, A. F., Parslow, R. A., & Christensen, H. (2011). Is Telomere Length a Biomarker of Aging? A Review. *The Journals of Gerontology: Series A, 66A*(2), 202-213. doi:10.1093/gerona/glq180

Matys, V., Kel-Margoulis, O. V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., . . . Wingender, E. (2006). TRANSFAC® and its module TRANSCompel®: transcriptional gene regulation in eukaryotes. *Nucleic Acids Research, 34*(suppl_1), D108-D110. doi:10.1093/nar/gkj143

McGee, D. L., Liao, Y., Cao, G., & Cooper, R. S. (1999). Self-reported Health Status and Mortality in a Multiethnic US Cohort. *American Journal of Epidemiology, 149*(1), 41-46. doi:10.1093/oxfordjournals.aje.a009725

McGrath, M., Wong, J. Y. Y., Michaud, D., Hunter, D. J., & De Vivo, I. (2007). Telomere Length, Cigarette Smoking, and Bladder Cancer Risk in Men and Women. *Cancer Epidemiology Biomarkers & Prevention, 16*(4), 815. doi:10.1158/1055-9965.EPI-06-0961

McGroder, C. F., Aaron, C. P., Bielinski, S. J., Kawut, S. M., Tracy, R. P., Raghu, G., . . . Podolanczuk, A. J. (2019). Circulating Adhesion Molecules and Subclinical Interstitial Lung Disease: The Multi-Ethnic Study of Atherosclerosis. *European Respiratory Journal*. doi:10.1183/13993003.00295-2019

Mermigkis, C., Stagaki, E., Tryfon, S., Schiza, S., Amfilochiou, A., Polychronopoulos, V., . . . Bouros, D. (2010). How common is sleep-disordered breathing in patients with idiopathic pulmonary fibrosis? *Sleep and Breathing, 14*(4), 387-390. doi:10.1007/s11325-010-0336-5

Miller, E. R., Putman, R. K., Diaz, A. A., Xu, H., San Jose Estepar, R., Araki, T., . . . Hunninghake, G. M. (2019). Increased Airway Wall Thickness in Interstitial Lung Abnormalities and Idiopathic Pulmonary Fibrosis. *Ann Am Thorac Soc, 16*(4), 447-454. doi:10.1513/AnnalsATS.201806-424OC

Miller, E. R., Putman, R. K., Vivero, M., Hung, Y., Araki, T., Nishino, M., . . . Hunninghake, G. M. (2018). Histopathology of Interstitial Lung Abnormalities in the Context of Lung Nodule Resections. *American Journal of Respiratory and Critical Care Medicine, 197*(7), 955-958. doi:10.1164/rccm.201708-1679LE

Molyneaux, P. L., Cox, M. J., Willis-Owen, S. A., Mallia, P., Russell, K. E., Russell, A. M., . . . Moffatt, M. F. (2014). The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine, 190*(8), 906-913. doi:10.1164/rccm.201403-0541OC

Molyneaux, P. L., Willis-Owen, S. A. G., Cox, M. J., James, P., Cowman, S., Loebinger, M., . . . Maher, T. M. (2017). Host–Microbial Interactions in Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 195*(12), 1640-1650. doi:10.1164/rccm.201607-1408OC

Moore, B. B., & Moore, T. A. (2015). Viruses in Idiopathic Pulmonary Fibrosis. Etiology and Exacerbation. *Ann Am Thorac Soc, 12 Suppl 2*(Suppl 2), S186-S192. doi:10.1513/AnnalsATS.201502-088AW

Moore, D. F. (2016). *Applied Survival Analysis Using R*. Switzerland: Springer International Publishing.

Mushiroda, T., Wattanapokayakit, S., Takahashi, A., Nukiwa, T., Kudoh, S., Ogura, T., . . . Nakamura, Y. (2008). A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. *Journal of Medical Genetics, 45*(10), 654-656. doi:10.1136/jmg.2008.057356

Nakanishi, Y., Masuda, T., Yamaguchi, K., Sakamoto, S., Horimasu, Y., Nakashima, T., . . . Hattori, N. (2019). Pre-existing interstitial lung abnormalities are risk factors for immune checkpoint inhibitor-induced interstitial lung disease in non-small cell lung cancer. *Respiratory Investigation, 57*(5), 451-459. doi:https://doi.org/10.1016/j.resinv.2019.05.002

Nakano, Y., Yang, I. V., Walts, A. D., Watson, A. M., Helling, B. A., Fletcher, A. A., . . . Schwartz, D. A. (2016). MUC5B Promoter Variant rs35705950 Affects MUC5B Expression in the Distal Airways in Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 193*(4), 464-466. doi:10.1164/rccm.201509-1872LE

Nalysnyk, L., Cid-Ruzafa, J., Rotella, P., & Esser, D. (2012). Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. *European Respiratory Review: An Official Journal of the European Respiratory Society, 21*(126), 355-361. doi:10.1183/09059180.00002512

Nathan, S. D., Albera, C., Bradford, W. Z., Costabel, U., Glaspole, I., Glassberg, M. K., . . . Noble, P. W. (2017). Effect of pirfenidone on mortality: pooled analyses and meta-analyses of clinical trials in idiopathic pulmonary fibrosis. *The Lancet Respiratory Medicine, 5*(1), 33-41. doi:https://doi.org/10.1016/S2213-2600(16)30326-5

Nishino, M., Cardarella, S., Dahlberg, S. E., Araki, T., Lydon, C., Jackman, D. M., . . . Johnson, B. E. (2015). Interstitial lung abnormalities in treatment-naïve advanced non-small-cell lung cancer patients are associated with shorter survival. *European Journal of Radiology, 84*(5), 998-1004. doi:10.1016/j.ejrad.2015.01.021

Nogee, L. M., Dunbar, A. E., Wert, S. E., Askin, F., Hamvas, A., & Whitsett, J. A. (2001). A Mutation in the Surfactant Protein C Gene Associated with Familial Interstitial Lung Disease. *344*(8), 573-579. doi:10.1056/nejm200102223440805

Noth, I., Zhang, Y., Ma, S. F., Flores, C., Barber, M., Huang, Y., . . . Garcia, J. G. (2013). Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med, 1*(4), 309-317. doi:10.1016/s2213-2600(13)70045-6

O'Dwyer, D. N., Armstrong, M. E., Trujillo, G., Cooke, G., Keane, M. P., Fallon, P. G., . . . Donnelly, S. C. (2013). The Toll-like receptor 3 L412F polymorphism and disease progression in idiopathic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine, 188*(12), 1442-1450. doi:10.1164/rccm.201304-0760OC

Ohgiya, M., Matsui, H., Tamura, A., Kato, T., Akagawa, S., & Ohta, K. (2017). The Evaluation of Interstitial Abnormalities in Group B of the 2011 Global Initiative for Chronic Obstructive Lung Disease (GOLD) Classification of Chronic Obstructive Pulmonary Disease (COPD). *Internal medicine (Tokyo, Japan), 56*(20), 2711-2717. doi:10.2169/internalmedicine.8406-16

Ortiz-Quintero, B., Buendía-Roldán, I., Ramírez-Salazar, E. G., Balderas-Martínez, Y. I., Ramírez-Rodríguez, S. L., Martínez-Espinosa, K., & Selman, M. (2020). Circulating microRNA Signature Associated to Interstitial Lung Abnormalities in Respiratory Asymptomatic Subjects. *Cells, 9*(6), 1556. doi:10.3390/cells9061556

Ozawa, Y., Suda, T., Naito, T., Enomoto, N., Hashimoto, D., Fujisawa, T., . . . Chida, K. (2009). Cumulative incidence of and predictive factors for lung cancer in IPF. *Respirology, 14*(5), 723-728. doi:10.1111/j.1440-1843.2009.01547.x

Parker, M. W., Rossi, D., Peterson, M., Smith, K., Sikström, K., White, E. S., . . . Bitterman, P. B. (2014). Fibrotic extracellular matrix activates a profibrotic positive feedback loop. *The Journal of Clinical Investigation, 124*(4), 1622-1635. doi:10.1172/JCI71386

Patel, A. S., Lin, L., Geyer, A., Haspel, J. A., An, C. H., Cao, J., . . . Morse, D. (2012). Autophagy in Idiopathic Pulmonary Fibrosis. *PloS One, 7*(7), e41394. doi:10.1371/journal.pone.0041394

Peljto, A. L., Zhang, Y., Fingerlin, T. E., Ma, S. F., Garcia, J. G., Richards, T. J., . . . Schwartz, D. A. (2013). Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA, 309*(21), 2232-2239. doi:10.1001/jama.2013.5827

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ, 316*(7139), 1236. doi:10.1136/bmj.316.7139.1236

Pietzner, M., Wheeler, E., Carrasco-Zanini, J., Kerrison, N. D., Oerton, E., Koprulu, M., . . . Langenberg, C. (2021). Synergistic insights into human health from aptamer- and antibody-based proteomic profiling. *Nature Communications, 12*(1), 6822. doi:10.1038/s41467-021-27164-0

Podolanczuk, A. J., Oelsner, E. C., Barr, R. G., Bernstein, E. J., Hoffman, E. A., Easthausen, I. J., . . . Lederer, D. J. (2017). High-Attenuation Areas on Chest Computed Tomography and Clinical Respiratory Outcomes in Community-Dwelling Adults. *American Journal of Respiratory and Critical Care Medicine, 196*(11), 1434-1442. doi:10.1164/rccm.201703-0555OC

Podolanczuk, A. J., Oelsner, E. C., Barr, R. G., Hoffman, E. A., Armstrong, H. F., Austin, J. H. M., . . . Lederer, D. J. (2016). High attenuation areas on chest computed tomography in community-dwelling adults: the MESA study. *European Respiratory Journal, 48*(5), 1442. Retrieved from http://erj.ersjournals.com/content/48/5/1442.abstract

Podolanczuk, A. J., Raghu, G., Tsai, M. Y., Kawut, S. M., Peterson, E., Sonti, R., . . . Lederer, D. J. (2017). Cholesterol, lipoproteins and subclinical interstitial lung disease: the MESA study. *Thorax*. doi:10.1136/thoraxjnl-2016-209568

Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., & Balch, W. E. (2009). Biological and Chemical Approaches to Diseases of Proteostasis Deficiency. *Annual Review of Biochemistry, 78*(1), 959-991. doi:10.1146/annurev.biochem.052308.114844

Putman, R. K., Gudmundsson, G., Araki, T., Nishino, M., Sigurdsson, S., Gudmundsson, E. F., . . . Hunninghake, G. M. (2017). The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *European Respiratory Journal, 50*(3). doi:10.1183/13993003.00537-2017

Putman, R. K., Gudmundsson, G., Axelsson, G. T., Hida, T., Honda, O., Araki, T., . . . Hunninghake, G. M. (2019). Imaging Patterns are Associated with Interstitial Lung Abnormality Progression and Mortality. *American Journal of Respiratory and Critical Care Medicine*. doi:10.1164/rccm.201809-1652OC

Putman, R. K., Hatabu, H., Araki, T., Gudmundsson, G., Gao, W., Nishino, M., . . . Hunninghake, G. M. (2016). Association Between Interstitial Lung Abnormalities and All-Cause Mortality. *JAMA, 315*(7), 672-681. doi:10.1001/jama.2016.0518

Putman, R. K., Hunninghake, G. M., Dieffenbach, P. B., Barragan-Bradford, D., Serhan, K., Adams, U., . . . Englert, J. A. (2017). Interstitial Lung Abnormalities Are Associated with Acute Respiratory Distress Syndrome. *American Journal of Respiratory and Critical Care Medicine, 195*(1), 138-141. doi:10.1164/rccm.201604-0818LE

Rabinovich, E. I., Kapetanaki, M. G., Steinfeld, I., Gibson, K. F., Pandit, K. V., Yu, G., . . . Kaminski, N. (2012). Global methylation patterns in idiopathic pulmonary fibrosis. *PloS One, 7*(4), e33770. doi:10.1371/journal.pone.0033770

Raghu, G., Chen, S. Y., Hou, Q., Yeh, W. S., & Collard, H. R. (2016). Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18-64 years old. *European Respiratory Journal, 48*(1), 179-186. doi:10.1183/13993003.01653-2015

Raghu, G., Chen, S. Y., Yeh, W. S., Maroni, B., Li, Q., Lee, Y. C., & Collard, H. R. (2014). Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001-11. *Lancet Respir Med, 2*(7), 566-572. doi:10.1016/s2213-2600(14)70101-8

Raghu, G., Collard, H. R., Egan, J. J., Martinez, F. J., Behr, J., Brown, K. K., . . . Schunemann, H. J. (2011). An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *American Journal of Respiratory and Critical Care Medicine, 183*(6), 788-824. doi:10.1164/rccm.2009-040GL

Raghu, G., Remy-Jardin, M., Myers, J. L., Richeldi, L., Ryerson, C. J., Lederer, D. J., . . . Wilson, K. C. (2018). Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *American Journal of Respiratory and Critical Care Medicine, 198*(5), e44-e68. doi:10.1164/rccm.201807-1255ST

Raghu, G., Richeldi, L., Jagerschmidt, A., Martin, V., Subramaniam, A., Ozoux, M.-L., . . . Soubrane, C. (2018). Idiopathic Pulmonary Fibrosis: Prospective, Case-Controlled Study of Natural History and Circulating Biomarkers. *Chest, 154*(6), 1359-1370. doi:https://doi.org/10.1016/j.chest.2018.08.1083

Raghu, G., Rochwerg, B., Zhang, Y., Garcia, C. A. C., Azuma, A., Behr, J., . . . Schünemann, H. J. (2015). An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *American Journal of Respiratory and Critical Care Medicine, 192*(2), e3-e19. doi:10.1164/rccm.201506-1063ST

Ranstam, J., & Cook, J. A. (2018). LASSO regression. *British Journal of Surgery, 105*(10), 1348-1348. doi:10.1002/bjs.10895

Rappa, F., Farina, F., Zummo, G., David, S., Campanella, C., Carini, F., . . . Macario, A. J. L. (2012). HSP-Molecular Chaperones in Cancer Biogenesis and Tumor Therapy: An Overview. *Anticancer Research, 32*(12), 5139. Retrieved from http://ar.iiarjournals.org/content/32/12/5139.abstract

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research, 47*(W1), W191-W198. doi:10.1093/nar/gkz369

Reck, M., Kaiser, R., Mellemgaard, A., Douillard, J.-Y., Orlov, S., Krzakowski, M., . . . Novello, S. (2014). Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *The Lancet Oncology, 15*(2), 143-155. doi:https://doi.org/10.1016/S1470-2045(13)70586-2

Regan, E. A., Hokanson, J. E., Murphy, J. R., Make, B., Lynch, D. A., Beaty, T. H., . . . Crapo, J. D. (2010). Genetic epidemiology of COPD (COPDGene) study design. *COPD, 7*(1), 32-43. doi:10.3109/15412550903499522

Rice, M. B., Li, W., Schwartz, J., Di, Q., Kloog, I., Koutrakis, P., . . . Mittleman, M. A. (2019). Ambient air pollution exposure and risk and progression of interstitial lung abnormalities: the Framingham Heart Study. *Thorax*. doi:10.1136/thoraxjnl-2018-212877

Richeldi, L., Collard, H. R., & Jones, M. G. (2017). Idiopathic pulmonary fibrosis. *Lancet, 389*(10082), 1941-1952. doi:10.1016/s0140-6736(17)30866-8

Richeldi, L., Cottin, V., du Bois, R. M., Selman, M., Kimura, T., Bailes, Z., . . . Brown, K. K. (2016). Nintedanib in patients with idiopathic pulmonary fibrosis: Combined evidence from the TOMORROW and INPULSIS® trials. *Respiratory Medicine, 113*, 74-79. doi:https://doi.org/10.1016/j.rmed.2016.02.001

Richeldi, L., du Bois, R. M., Raghu, G., Azuma, A., Brown, K. K., Costabel, U., . . . Collard, H. R. (2014). Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine, 370*(22), 2071-2082. doi:10.1056/NEJMoa1402584

Ronit, A., Benfield, T., Lundgren, J., Vestbo, J., Afzal, S., Nordestgaard, B. G., . . . Kristensen, T. (2020). Interstitial Lung Abnormalities in People With HIV Infection and Uninfected Controls. *Journal of Infectious Diseases, 221*(12), 1973-1977. doi:10.1093/infdis/jiaa047

Rosas, I. O., Ren, P., Avila, N. A., Chow, C. K., Franks, T. J., Travis, W. D., . . . Gochuico, B. R. (2007). Early Interstitial Lung Disease in Familial Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 176*(7), 698-705. doi:10.1164/rccm.200702-254OC

Roy, M. G., Livraghi-Butrico, A., Fletcher, A. A., McElwee, M. M., Evans, S. E., Boerner, R. M., . . . Evans, C. M. (2014). Muc5b is required for airway defence. *Nature, 505*(7483), 412-416. doi:10.1038/nature12807

Sack, C., Vedal, S., Sheppard, L., Raghu, G., Barr, R. G., Podolanczuk, A., . . . Kaufman, J. D. (2017). Air pollution and subclinical interstitial lung disease: the Multi-Ethnic Study of Atherosclerosis (MESA) air-lung study. *The European respiratory journal, 50*(6), 1700559. doi:10.1183/13993003.00559-2017

Sack, C. S., Doney, B. C., Podolanczuk, A. J., Hooper, L. G., Seixas, N. S., Hoffman, E. A., . . . Kaufman, J. D. (2017). Occupational Exposures and Subclinical Interstitial Lung Disease. The MESA (Multi-Ethnic Study of Atherosclerosis) Air and Lung Studies. *American Journal of Respiratory and Critical Care Medicine, 196*(8), 1031-1039. doi:10.1164/rccm.201612-2431OC

Salisbury, M. L., Hewlett, J. C., Ding, G., Markin, C. R., Douglas, K., Mason, W., . . . Blackwell, T. S. (2020). Development and Progression of Radiologic Abnormalities in Individuals at Risk for Familial ILD. *American Journal of Respiratory and Critical Care Medicine*. doi:10.1164/rccm.201909-1834OC

Sanchez, T. R., Oelsner, E. C., Lederer, D. J., Lo Cascio, C. M., Jones, M. R., Grau-Perez, M., . . . Navas-Acien, A. (2019). Rice Consumption and Subclinical Lung Disease in US Adults: Observational Evidence From the Multi-Ethnic Study of Atherosclerosis. *American Journal of Epidemiology, 188*(9), 1655-1665. doi:10.1093/aje/kwz137

Sanders, J. L., & Newman, A. B. (2013). Telomere Length in Epidemiology: A Biomarker of Aging, Age-Related Disease, Both, or Neither? *Epidemiologic Reviews, 35*(1), 112-131. doi:10.1093/epirev/mxs008

Sanders, J. L., Putman, R. K., Dupuis, J., Xu, H., Murabito, J. M., Araki, T., . . . Hunninghake, G. M. (2021). The Association of Aging Biomarkers, Interstitial Lung Abnormalities, and Mortality. *American Journal of Respiratory and Critical Care Medicine, 203*(9), 1149-1157. doi:10.1164/rccm.202007-2993OC

Schafer, M. J., White, T. A., Iijima, K., Haak, A. J., Ligresti, G., Atkinson, E. J., . . . LeBrasseur, N. K. (2017). Cellular senescence mediates fibrotic pulmonary disease. *Nature Communications, 8*, 14532. doi:10.1038/ncomms14532

Scotton, C. J., & Chambers, R. C. (2007). Molecular Targets in Pulmonary Fibrosis: The Myofibroblast in Focus. *Chest, 132*(4), 1311-1321. doi:https://doi.org/10.1378/chest.06-2568

Seibold, M. A., Smith, R. W., Urbanek, C., Groshong, S. D., Cosgrove, G. P., Brown, K. K., . . . Reynolds, S. D. (2013). The idiopathic pulmonary fibrosis honeycomb cyst contains a mucocilary pseudostratified epithelium. *PloS One, 8*(3), e58658-e58658. doi:10.1371/journal.pone.0058658

Seibold, M. A., Wise, A. L., Speer, M. C., Steele, M. P., Brown, K. K., Loyd, J. E., . . . Schwartz, D. A. (2011). A common MUC5B promoter polymorphism and pulmonary fibrosis. *New England Journal of Medicine, 364*(16), 1503-1512. doi:10.1056/NEJMoa1013660

Selman, M., López-Otín, C., & Pardo, A. (2016). Age-driven developmental drift in the pathogenesis of idiopathic pulmonary fibrosis. *European Respiratory Journal, 48*(2), 538. doi:10.1183/13993003.00398-2016

Selman, M., Martinez, F. J., & Pardo, A. (2019). Why Does an Aging Smoker's Lung Develop Idiopathic Pulmonary Fibrosis and Not Chronic Obstructive Pulmonary Disease? *American Journal of Respiratory and Critical Care Medicine, 199*(3), 279-285. doi:10.1164/rccm.201806-1166PP

Selman, M., Pardo, A., & King, T. E. (2012). Hypersensitivity Pneumonitis. *American Journal of Respiratory and Critical Care Medicine, 186*(4), 314-324. doi:10.1164/rccm.201203-0513CI

Sgalla, G., Iovene, B., Calvello, M., Ori, M., Varone, F., & Richeldi, L. (2018). Idiopathic pulmonary fibrosis: pathogenesis and management. *Respiratory Research, 19*(1), 32-32. doi:10.1186/s12931-018-0730-2

Shay, J. W. (2016). Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discovery, 6*(6), 584-593. doi:10.1158/2159-8290.CD-16-0062

Skeoch, S., Weatherley, N., Swift, A. J., Oldroyd, A., Johns, C., Hayton, C., . . . Chaudhuri, N. (2018). Drug-Induced Interstitial Lung Disease: A Systematic Review. *Journal of clinical medicine, 7*(10), 356. doi:10.3390/jcm7100356

Song, J. W., Do, K. H., Jang, S. J., Colby, T. V., Han, S., & Kim, D. S. (2013). Blood Biomarkers MMP-7 and SP-A: Predictors of Outcome in Idiopathic Pulmonary Fibrosis. *Chest, 143*(5), 1422-1429. doi:https://doi.org/10.1378/chest.11-2735

Stafford, I. S., Kellermann, M., Mossotto, E., Beattie, R. M., MacArthur, B. D., & Ennis, S. (2020). A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ digital medicine, 3*, 30-30. doi:10.1038/s41746-020-0229-3

Steyerberg, E. W. (2018). Validation in prediction research: the waste by data splitting. *Journal of Clinical Epidemiology, 103*, 131-133. doi:10.1016/j.jclinepi.2018.07.010

Storeng, S. H., Vinjerui, K. H., Sund, E. R., & Krokstad, S. (2020). Associations between complex multimorbidity, activities of daily living and mortality among older Norwegians. A prospective cohort study: the HUNT Study, Norway. *BMC Geriatrics, 20*(1), 21. doi:10.1186/s12877-020-1425-3

Stuart, B. D., Choi, J., Zaidi, S., Xing, C., Holohan, B., Chen, R., . . . Garcia, C. K. (2015). Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nature Genetics, 47*(5), 512-517. doi:10.1038/ng.3278

Stuart, B. D., Lee, J. S., Kozlitina, J., Noth, I., Devine, M. S., Glazer, C. S., . . . Garcia, C. K. (2014). Effect of telomere length on survival in patients with idiopathic pulmonary fibrosis: an observational cohort study with independent validation. *Lancet Respir Med, 2*(7), 557-565. doi:10.1016/s2213-2600(14)70124-9

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., . . . Butterworth, A. S. (2018). Genomic atlas of the human plasma proteome. *Nature, 558*(7708), 73-79. doi:10.1038/s41586-018-0175-2

Sundquist, K., Qvist, J., Sundquist, J., & Johansson, S.-E. (2004). Frequent and occasional physical activity in the elderly: A 12-year follow-up study of mortality. *American Journal of Preventive Medicine, 27*(1), 22-27. doi:https://doi.org/10.1016/j.amepre.2004.03.011

Sverzellati, N., Guerci, L., Randi, G., Calabro, E., La Vecchia, C., Marchiano, A., . . . Pastorino, U. (2011). Interstitial lung diseases in a lung cancer screening trial. *European Respiratory Journal, 38*(2), 392-400. doi:10.1183/09031936.00201809

Synn, A. J., Li, W., Hunninghake, G. M., Washko, G. R., San José Estépar, R., O'Connor, G. T., . . . Rice, M. B. (2021). Vascular Pruning on CT and Interstitial Lung Abnormalities in the Framingham Heart Study. *Chest, 159*(2), 663-672. doi:10.1016/j.chest.2020.07.082

Szláma, G., Kondás, K., Trexler, M., & Patthy, L. (2010). WFIKKN1 and WFIKKN2 bind growth factors TGFβ1, BMP2 and BMP4 but do not inhibit their signalling activity. *The FEBS Journal, 277*(24), 5040-5050. doi:https://doi.org/10.1111/j.1742-4658.2010.07909.x

Tabas, I. (2010). Macrophage death and defective inflammation resolution in atherosclerosis. *Nature reviews. Immunology, 10*(1), 36-46. doi:10.1038/nri2675

Tan, Y., Jia, D., Lin, Z., Guo, B., He, B., Lu, C., . . . Lu, A. (2016). Potential Metabolic Biomarkers to Identify Interstitial Lung Abnormalities. *International Journal of Molecular Sciences, 17*(7), 1148. doi:10.3390/ijms17071148

Tang, Y.-W., Johnson, J. E., Browning, P. J., Cruz-Gervis, R. A., Davis, A., Graham, B. S., . . . Stecenko, A. A. (2003). Herpesvirus DNA Is Consistently Detected in Lungs of Patients with Idiopathic Pulmonary Fibrosis. *Journal of Clinical Microbiology, 41*(6), 2633. doi:10.1128/JCM.41.6.2633-2640.2003

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73*(3), 273-282. doi:https://doi.org/10.1111/j.1467-9868.2011.00771.x

Tobin, R. W., Charles E. Pope, I., Pellegrini, C. A., Emond, M. J., Sillery, J., & Raghu, G. (1998). Increased Prevalence of Gastroesophageal Reflux in Patients with Idiopathic Pulmonary Fibrosis. *158*(6), 1804-1808. doi:10.1164/ajrccm.158.6.9804105

Travis, W. D., Costabel, U., Hansell, D. M., King, T. E., Jr., Lynch, D. A., Nicholson, A. G., . . . Valeyre, D. (2013). An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *American Journal of Respiratory and Critical Care Medicine, 188*(6), 733-748. doi:10.1164/rccm.201308-1483ST

Travis, W. D., Hunninghake, G., King, T. E., Lynch, D. A., Colby, T. V., Galvin, J. R., . . . Wells, A. (2008). Idiopathic Nonspecific Interstitial Pneumonia. *American Journal of Respiratory and Critical Care Medicine, 177*(12), 1338-1347. doi:10.1164/rccm.200611-1685OC

Trovato-Salinaro, A., Trovato-Salinaro, E., Failla, M., Mastruzzo, C., Tomaselli, V., Gili, E., . . . Vancheri, C. (2006). Altered intercellular communication in lung fibroblast cultures from patients with idiopathic pulmonary fibrosis. *Respiratory Research, 7*(1), 122-122. doi:10.1186/1465-9921-7-122

Tsakiri, K. D., Cronkhite, J. T., Kuan, P. J., Xing, C., Raghu, G., Weissler, J. C., . . . Garcia, C. K. (2007). Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proceedings of the National Academy of Sciences of the United States of America, 104*(18), 7552-7557. doi:10.1073/pnas.0701009104

Tsushima, K., Sone, S., Yoshikawa, S., Yokoyama, T., Suzuki, T., & Kubo, K. (2010). The radiological patterns of interstitial change at an early phase: over a 4-year follow-up. *Respiratory Medicine, 104*(11), 1712-1721. doi:10.1016/j.rmed.2010.05.014

Turner-Warwick, M., Lebowitz, M., Burrows, B., & Johnson, A. (1980). Cryptogenic fibrosing alveolitis and lung cancer. *Thorax, 35*(7), 496-499. doi:10.1136/thx.35.7.496

Tzouvelekis, A., Herazo-Maya, J. D., Slade, M., Chu, J.-H., Deiuliis, G., Ryu, C., . . . Kaminski, N. (2017). Validation of the prognostic value of MMP-7 in idiopathic pulmonary fibrosis. *Respirology (Carlton, Vic.), 22*(3), 486-493. doi:10.1111/resp.12920

van Batenburg, A. A., Kazemier, K. M., van Oosterhout, M. F. M., van der Vis, J. J., van Es, H. W., Grutters, J. C., . . . van Moorsel, C. H. M. (2020). From organ to cell: Multi-level telomere length assessment in patients with idiopathic pulmonary fibrosis. *PloS One, 15*(1), e0226785-e0226785. doi:10.1371/journal.pone.0226785

Van Dyken, S. J., Liang, H.-E., Naikawadi, R. P., Woodruff, P. G., Wolters, P. J., Erle, D. J., & Locksley, R. M. (2017). Spontaneous Chitin Accumulation in Airways and Age-Related Fibrotic Lung Disease. *Cell, 169*(3), 497-509.e413. doi:10.1016/j.cell.2017.03.044

Vancheri, C. (2013). Common pathways in idiopathic pulmonary fibrosis and cancer. *European Respiratory Review, 22*(129), 265. doi:10.1183/09059180.00003613

Vasan, R. S., Demissie, S., Kimura, M., Cupples, L. A., Rifai, N., White, C., . . . Aviv, A. (2008). Association of Leukocyte Telomere Length With Circulating Biomarkers of the Renin-Angiotensin-Aldosterone System. *Circulation, 117*(9), 1138-1144. doi:10.1161/CIRCULATIONAHA.107.731794

Wang, Y., Kuan, P. J., Xing, C., Cronkhite, J. T., Torres, F., Rosenblatt, R. L., . . . Garcia, C. K. (2009). Genetic defects in surfactant protein A2 are associated with pulmonary fibrosis and lung cancer. *American Journal of Human Genetics, 84*(1), 52-59. doi:10.1016/j.ajhg.2008.11.010

Washko, G. R., Hunninghake, G. M., Fernandez, I. E., Nishino, M., Okajima, Y., Yamashiro, T., . . . Rosas, I. O. (2011). Lung volumes and emphysema in smokers with interstitial lung abnormalities. *New England Journal of Medicine, 364*(10), 897-906. doi:10.1056/NEJMoa1007285

White, E. (2015). The role for autophagy in cancer. *The Journal of Clinical Investigation, 125*(1), 42-46. doi:10.1172/JCI73941

Whittaker Brown, S.-A., Padilla, M., Mhango, G., Powell, C., Salvatore, M., Henschke, C., . . . Wisnivesky, J. (2019). Interstitial Lung Abnormalities and Lung Cancer Risk in the National Lung Screening Trial. *Chest*. doi:https://doi.org/10.1016/j.chest.2019.06.041

Wille, M. M., Thomsen, L. H., Petersen, J., de Bruijne, M., Dirksen, A., Pedersen, J. H., & Shaker, S. B. (2016). Visual assessment of early emphysema and interstitial abnormalities on CT is useful in lung cancer risk analysis. *European Radiology, 26*(2), 487-494. doi:10.1007/s00330-015-3826-9

Willis, B. C., Liebler, J. M., Luby-Phelps, K., Nicholson, A. G., Crandall, E. D., du Bois, R. M., & Borok, Z. (2005). Induction of epithelial-mesenchymal transition in alveolar epithelial cells by transforming growth factor-beta1: potential role in idiopathic pulmonary fibrosis. *The American Journal of Pathology, 166*(5), 1321-1332. doi:10.1016/s0002-9440(10)62351-6

Wolters, P. J., Blackwell, T. S., Eickelberg, O., Loyd, J. E., Kaminski, N., Jenkins, G., . . . Schwartz, D. A. (2018). Time for a change: is idiopathic pulmonary fibrosis still idiopathic and only fibrotic? *Lancet Respir Med, 6*(2), 154-160. doi:10.1016/S2213-2600(18)30007-9

Wolters, P. J., Collard, H. R., & Jones, K. D. (2014). Pathogenesis of idiopathic pulmonary fibrosis. *Annual Review of Pathology, 9*, 157-179. doi:10.1146/annurev-pathol-012513-104706

Xu, J. F., Washko, G. R., Nakahira, K., Hatabu, H., Patel, A. S., Fernandez, I. E., . . . Hunninghake, G. M. (2012). Statins and pulmonary fibrosis: the potential role of NLRP3 inflammasome activation. *American Journal of Respiratory and Critical Care Medicine, 185*(5), 547-556. doi:10.1164/rccm.201108-1574OC

Yang, I. V., Pedersen, B. S., Rabinovich, E., Hennessy, C. E., Davidson, E. J., Murphy, E., . . . Schwartz, D. A. (2014). Relationship of DNA methylation and gene expression in idiopathic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine, 190*(11), 1263-1272. doi:10.1164/rccm.201408-1452OC

Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika, 87*(4), 954-959.

Zhang, Y., Jiang, M., Nouraie, M., Roth, M. G., Tabib, T., Winters, S., . . . Alder, J. K. (2019). GDF15 is an epithelial-derived biomarker of idiopathic pulmonary fibrosis. *American Journal of Physiology: Lung Cellular and Molecular Physiology, 317*(4), L510-l521. doi:10.1152/ajplung.00062.2019

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association, 101*(476), 1418-1429. doi:10.1198/016214506000000735

Original publications

Paper I

Interstitial lung abnormalities and self-reported health and functional status

ABSTRACT

We investigated the association between interstitial lung abnormalities (ILA) and self-reported measures of health and functional status in 5764 participants from the Age, Gene/Environment Susceptibility-Reykjavik study. The associations of ILA to activities of daily living (ADLs), general health status and physical activity were explored using logistic regression models. Participants with ILA were less likely to be independent in ADLs (OR 0.70; 95% CI 0.55 to 0.90) to have good or better self-reported health (OR 0.66; 95% CI 0.52 to 0.82) and to participate in physical activity (OR 0.72; CI 0.56 to 0.91). The results demonstrate ILA's association with worsening self-reported health and functional status.

INTRODUCTION

Interstitial lung abnormalities (ILA) includes a set of precisely defined CT changes, similar to those used to detect interstitial lung disease (ILD), but are often of a lesser magnitude.[1–3] The argument for the relation of ILA to ILD has been strengthened by the association of ILA to several factors, including reductions in total lung capacity and an increase in mortality.[1 3 4] Other notable associations of ILA include increased age, decreased diffusion capacity, in addition to shortness of breath with exertion, and chronic cough.[1 2 5] Despite these associations, little is known about the relationship between ILA and self-reported measures of health and functional status.

METHODS

To explore these associations, data were obtained from the Age, Gene/Environment Susceptibility-Reykjavik (AGES-Reykjavik) study. Details of the study design have been previously published.[6] Informed consent was obtained from all participants. Two separate CT scans of the thorax were obtained with a four-row CT scanner (Sensation, Siemens Medical System, Erlangen, Germany) that covered about 95% of the lungs, omitting the superior portion of the apices. The axial images were evaluated by up to three readers (pulmonologists and chest radiologists) blinded to participant-specific information, using a sequential reading method as previously described.[1] Scans labelled as ILA, indeterminate and 20% of normal scans were read by two readers and the third reader provided majority opinion on discordantly labelled scans. ILA was defined as 'non-dependent changes affecting more than 5% of any lung zone, including reticular or ground-glass abnormalities, diffuse centrilobular nodularity, non-emphysematous cysts, honeycombing or traction bronchiectasis' while 'focal or unilateral ground-glass attenuation, focal or unilateral reticulation or patchy ground-glass abnormalities (<5% of any lung zone)' were defined as indeterminate changes. The data on smoking status, health status, activities of daily living (ADLs) and physical activity were obtained from participant questionnaires.[6] The question regarding health status was 'In general, how would you say your health is?', with the answers ranging from 1 ('excellent') to 5 ('poor'). For statistical analyses, participants were categorised into two groups based on whether they perceived their health as 'good' or better. Thus, one group was composed of participants that perceived their health as 'good', 'very good' or 'excellent', while participants that perceived their health as 'poor' or 'fair' made up the other group. ADLs were assessed by asking participants whether they had difficulties performing the following five activities: bathing, dressing, walking from room to room, transferring out of bed/chair or eating.[7] For statistical analyses, participants were grouped into two groups: people independent in all ADLs and people

Table 1 Baseline characteristics of participants

Participants	No ILA (n=3216, 61%)	Indeterminate ILA status (n=1726, 32%)	ILA (n=378, 7%)	P value
Mean age (95% CI)	75.9 (75.7 to 76.1)	77.4 (77.1 to 77.6)	77.8 (77.2 to 78.3)	3.4×10^{-9}
Women, n (%)	1910 (59)	962 (56)	172 (46)	3.1×10^{-7}
Mean BMI (95% CI)	27.2 (27.0 to 27.3)	26.8 (26.6 to 27.0)	27.1 (26.6 to 27.5)	0.60
History of smoking, n (%)	1750 (54)	1021 (59)	271 (72)	1.1×10^{-10}
Median pack-years (IQR)	0 (0–16)	2.5 (0–22.5)	11.0 (0–28.5)	3.5×10^{-16}
Current smoker, n (%)	374 (12)	205 (12)	69 (18)	0.0003
Activities of daily living				
Independence in all five ADLs, n (%)	2422 (75)	1191 (69)	251 (66)	
Independence in four ADLs, n (%)	434 (13)	253 (15)	65 (17)	
Independence in three or less ADLs, n (%)	336 (10)	259 (15)	54 (14)	
Health status				
Health described as 'excellent', n (%)	742 (23)	368 (21)	70 (19)	
Health described as 'very good', n (%)	480 (15)	225 (13)	38 (10)	
Health described as 'good', n (%)	1010 (31)	512 (30)	109 (29)	
Health described as 'fair', n (%)	813 (25)	503 (29)	128 (34)	
Health described as 'poor', n (%)	166 (5)	114 (7)	31 (8)	
Physical activity in the past 12 months				
No or rare physical activity, n (%)	1914 (60)	1115 (65)	258 (68)	
Weekly physical activity, n (%)	1263 (39)	580 (34)	108 (29)	

P values comparing participants without ILA and participants with ILA.
ADLs, activities of daily living; BMI, body mass index; ILA, interstitial lung abnormalities.

Table 2 Associations of ILA with measures of health

	Unadjusted OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Independence in activities of daily living				
No ILA	1.00	0.0008	1.00	0.0051
ILA	0.67 (0.53 to 0.85)		0.70 (0.55 to 0.90)	
Health status perceived as good or better*				
No ILA	1.00	4.0×10^{-6}	1.00	0.0003
ILA	0.60 (0.48 to 0.75)		0.66 (0.52 to 0.82)	
Participation in physical activity				
No ILA	1.00	0.0002	1.00	0.0076
ILA	0.63 (0.50 to 0.80)		0.72 (0.56 to 0.91)	

Associations were estimated with logistic regression models. The adjusted models are adjusted for sex, body mass index, age, pack-years and current smoking status as covariates.
*This includes self-reports of health status perceived as 'good', 'very good' and 'excellent', as opposed to 'poor' or 'fair'.
ILA, interstitial lung abnormalities.

dependent in one or more ADLs. Physical activity was estimated by asking participants how many hours per week, during the last 12 months, they participated in moderate-vigorous intensity physical activity. The possible answers were never, rarely, every week but less than 1 hour, 1–3 hours per week, 4–7 hours per week and more than 7 hours per week. For statistical analyses, people were categorised into two groups: people who never or rarely participated in physical activity and those who did so at least weekly. Comparisons of baseline characteristics were done using X^2 tests, t-tests and Wilcoxon rank-sum tests as appropriate. The analysis of the association of ILA with measures of health and functional status was done using logistic regression modelling. Participants with indeterminate ILA status were excluded from these analyses. The measures of health were assigned to multivariable models adjusted for covariates including age, sex, body mass index (BMI), pack-years of smoking and current smoking status. All statistical analyses were done using R V.3.3.2.

RESULTS

The baseline characteristics of the participants in the AGES-Reykjavik cohort for which CT imaging data were available (5320 of 5764 or 92%) are shown in table 1. Similar to previous reports,[3] people with ILA were significantly older and more often male than people without ILA. Participants with ILA were more likely to have a history of smoking, had a higher number of pack-years and were more often smokers at the time of data collection. Results regarding the associations between measures of health with ILA are shown in table 2. Associations with ILA were observed for all metrics of self-reported health and functional status. Odds of independence in ADLs were decreased among participants with ILA (OR 0.70; 95% CI 0.55 to 0.90, P=0.0052) compared with people free of ILA (table 2). The same applied to worse health; people with ILA were significantly less likely to have good or better self-reported health status than people without ILA (OR 0.66; 95% CI 0.52 to 0.82, P=0.0003). People with ILA were less likely to participate in physical activity (OR 0.72; 95% CI 0.56 to 0.91, P=0.0076) as shown in table 2.

DISCUSSION

These data demonstrate that ILA is associated with worse self-reported health and functional status in the AGES-Reykjavik study. This study extends previous findings[1–3 8] demonstrating that, even though not diagnosed, research participants with ILA may, in some cases, share similar health outcomes as patients with ILD.

More than one reason is possible for these associations. First, it is possible that the reductions in self-reported health and functional status are being driven by the underlying, interstitial abnormalities. Another possibility is that similar underlying biological factors can cause both ILA and reduced measures of health and functional status. Lastly, it is possible that unmeasured confounders could, in part, explain some of these findings of association. These findings demonstrate that in the AGES-Reykjavik cohort, people with ILA have poorer self-reported health and reduced functional status. These results add to previous suggestions that research programmes aimed at improving treatment and prevention of ILD could benefit from directing their efforts to early stages of ILD.[3 9] In conclusion, ILA are associated with measures of decreased health among elderly people.

Gisli Thor Axelsson,[1,2] Rachel K Putman,[3] Tetsuro Araki,[4,5] Sigurdur Sigurdsson,[1] Elias Freyr Gudmundsson,[1] Gudny Eiriksdottir,[1] Thor Aspelund,[1,2] Ezra R Miller,[3] Lenore J Launer,[6] Tamara B Harris,[6] Hiroto Hatabu,[4,5] Vilmundur Gudnason,[1,2] Gary Matt Hunninghake,[3,5] Gunnar Gudmundsson[2,7]

GTA and RKP contributed equally.

[1]Icelandic Heart Association, Kopavogur, Iceland
[2]Faculty of Medicine, University of Iceland, Reykjavik, Iceland
[3]Pulmonary and Critical Care Division, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[4]Department of Radiology, Brigham and Women's Hospital, Boston, Massachusetts, USA
[5]Center for Pulmonary Functional Imaging, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[6]Intramural Research Program, National Institute on Aging, Bethesda, Maryland, USA
[7]Department of Respiratory Medicine, Landspitali University Hospital, Reykjavik, Iceland

Correspondence to Dr Gunnar Gudmundsson, Department of Respiratory Medicine, Landspitali University Hospital, Reykjavik IS-108, Iceland; ggudmund@landspitali.is

Contributors Study design: HH, VG, GMH and GG. Acquisition, analysis or interpretation of the data: GTA, RKP, TAs, SS, EFG, GE, TAr, HH, VG, GMH and GG. Statistical analysis: GTA, RKP, TAs and GMH. Obtained funding: VG, GMH and GG. All authors contributed to the critical revision of the manuscript for important intellectual content.

Funding RKP is supported by a National Institutes of Health (NIH) grant (T32 HL007633). GG is supported by a project grant from the Icelandic Research Fund (141513-051) and from the Landspitali Scientific Fund (A-2015-030, A-2016-023). The Age, Gene/Environment Susceptibility-Reykjavik study was supported by a National Institute on Aging (NIA) grant (27120120022C), two NIH contracts (N01-AG-1-2100 and HHSN27120120022C), the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association) and the Althingi (the Icelandic Parliament). VG is supported by an NIA grant (27120120022C) and an Icelandic Research Fund project grant (141513-051). GMH is supported by NIH grants (P01 HL114501, R01 HL111024, R01 HL135142, RO1 HL130974) and a project grant from

the Icelandic Research Fund (141513-051).

Competing interests None declared.

Ethics approval National Bioethics Committee in Iceland (VSN: 00-063) and National Institute on Ageing Intramural Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

To cite Axelsson GT, Putman RK, Araki T, *et al*. *Thorax* 2018;**73**:884–886.

Received 26 August 2017
Revised 11 December 2017
Accepted 18 December 2017
Published Online First 9 January 2018

Thorax 2018;**73**:884–886.
doi:10.1136/thoraxjnl-2017-210956

REFERENCES

1 Washko GR, Hunninghake GM, Fernandez IE, *et al*. Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011;364:897–906.
2 Hunninghake GM, Hatabu H, Okajima Y, *et al*. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013;368:2192–200.
3 Putman RK, Hatabu H, Araki T, *et al*. Association between interstitial lung abnormalities and all-cause mortality. *JAMA* 2016;315:672–81.
4 Podolanczuk AJ, Oelsner EC, Barr RG, *et al*. High attenuation areas on chest computed tomography in community-dwelling adults: the MESA study. *Eur Respir J* 2016;48:1442–52.
5 American Thoracic Society, European Respiratory Society. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. *Am J Respir Crit Care Med* 2002;165:277–304.
6 Harris TB, Launer LJ, Eiriksdottir G, *et al*. Age, gene/environment susceptibility-Reykjavik study: multidisciplinary applied phenomics. *Am J Epidemiol* 2007;165:1076–87.
7 Siggeirsdottir K, Aspelund T, Jonsson BY, *et al*. Effect of vertebral fractures on function, quality of life and hospitalisation the AGES-Reykjavik study. *Age Ageing* 2012;41:351–7.
8 Doyle TJ, Washko GR, Fernandez IE, *et al*. Interstitial lung abnormalities and reduced exercise capacity. *Am J Respir Crit Care Med* 2012;185:756–62.
9 Doyle TJ, Hunninghake GM, Rosas IO. Subclinical interstitial lung disease: why you should care. *Am J Respir Crit Care Med* 2012;185:1147–53.

Paper II

ORIGINAL ARTICLE
INTERSTITIAL LUNG DISEASE

The associations of interstitial lung abnormalities with cancer diagnoses and mortality

Gisli T. Axelsson[1], Rachel K. Putman[2], Thor Aspelund[1,3], Elias F. Gudmundsson[3], Tomayuki Hida[4,5], Tetsuro Araki[4,5], Mizuki Nishino[4,5], Hiroto Hatabu[4,5], Vilmundur Gudnason[1,3], Gary M. Hunninghake[2,5] and Gunnar Gudmundsson[1,6]

Affiliations: [1]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [2]Pulmonary and Critical Care Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [3]Icelandic Heart Association, Kopavogur, Iceland. [4]Dept of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [5]Center for Pulmonary Functional Imaging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [6]Dept of Respiratory Medicine, Landspitali University Hospital, Reykjavik, Iceland.

Correspondence: Gisli T. Axelsson, Faculty of Medicine, University of Iceland, Saemundargata 2, 102 Reykjavik, Iceland. E-mail: gtha7@hi.is

@ERSpublications
Interstitial lung abnormalities are associated with an increased hazard of lung cancer diagnosis and lung cancer mortality in a general population cohort. Cancers other than lung cancer were not associated with interstitial lung abnormalities. https://bit.ly/3hWdc6m

Cite this article as: Axelsson GT, Putman RK, Aspelund T, *et al*. The associations of interstitial lung abnormalities with cancer diagnoses and mortality. *Eur Respir J* 2020; 56: 1902154 [https://doi.org/10.1183/13993003.02154-2019].

ABSTRACT An increased incidence of lung cancer is well known among patients with idiopathic pulmonary fibrosis. It is not known whether interstitial lung abnormalities, *i.e.* early fibrotic changes of the lung, are a risk factor for lung cancer in the general population.

The study's objective was to assess whether interstitial lung abnormalities were associated with diagnoses of, and mortality from, lung cancer and other cancers. Data from the AGES-Reykjavik study, a cohort of 5764 older Icelandic adults, were used. Outcome data were ascertained from electronic medical records. Gray's tests, Cox proportional hazards models and proportional subdistribution hazards models were used to analyse associations of interstitial lung abnormalities with lung cancer diagnoses and lung cancer mortality as well as diagnoses and mortality from all cancers.

There was a greater cumulative incidence of lung cancer diagnoses ($p<0.001$) and lung cancer mortality ($p<0.001$) in participants with interstitial lung abnormalities than in others. Interstitial lung abnormalities were associated with an increased hazard of lung cancer diagnosis (hazard ratio 2.77) and lung cancer mortality (hazard ratio 2.89) in adjusted Cox models. Associations of interstitial lung abnormalities with all cancers were found in models including lung cancers but not in models excluding lung cancers.

People with interstitial lung abnormalities are at increased risk of lung cancer and lung cancer mortality, but not of other cancers. This implies that an association between fibrotic and neoplastic diseases of the lung exists from the early stages of lung fibrosis and suggests that interstitial lung abnormalities could be considered as a risk factor in lung cancer screening efforts.

This article has supplementary material available from erj.ersjournals.com

Received: 5 Nov 2019 | Accepted after revision: 16 June 2020

Introduction

Interstitial lung abnormalities (ILA) are commonly defined as abnormalities noted on chest computed tomography (CT) scans that are similar in appearance to those noted in patients with interstitial lung disease but occurring in a person without a known diagnosis of interstitial lung disease [1]. There is evidence to suggest that some research participants with ILA may share a common syndrome noted in patients with idiopathic pulmonary fibrosis (IPF) that includes the development of a restrictive lung deficit [1], accelerated lung function decline [2], imaging progression [2, 3], shared genetic determinants [4, 5], poorer subjective health and physical function, and increased rates of mortality [6–8]. ILA have been further categorised into specific subtypes and imaging patterns [1, 3]. Associations with restrictive lung deficits, genetic polymorphisms, imaging progression and mortality have been found to vary between these subtypes and patterns [1, 3, 9].

A number of studies have demonstrated an increased incidence of lung cancer among IPF patients compared with the general population, even when adjusted for confounders such as cigarette smoking [10, 11]. The development of lung cancer in IPF patients has been shown to severely impair their survival [12]. While ILA have been associated with increased prevalence of, and mortality from, lung cancer in cohorts of smokers intended for lung cancer screening [13–16], there is less known about these risks in the general population. In addition, data are scarce regarding whether there is an increased risk of non-pulmonary malignancies among people with ILA.

Thus, the objectives of this study were to explore the associations of ILA with diagnoses of both lung cancer and other cancers and to assess whether ILA were associated with increased mortality from lung cancer and other malignancies.

Methods

Data acquisition and materials

The Age, Gene/Environment Susceptibility-Reykjavik (AGES-Reykjavik) study is a longitudinal birth cohort study, derived from the previous Reykjavik study, in which older individuals were recruited between 2002 and 2006 in an effort to identify the causal factors of diseases and disabilities associated with ageing. Additional details on the study design have been previously published [17].

CT imaging of the thorax was characterised for the presence of ILA in 5320 out of 5764 AGES-Reykjavik participants (92%) by up to three readers, as previously described [1]. ILA were defined as nondependent ground-glass or reticular abnormalities, diffuse centrilobular nodularity, non-emphysematous cysts, honeycombing and traction bronchiectasis that affected >5% of any lung zone [1]. Participants who had focal or unilateral ground-glass attenuation, focal or unilateral reticulation and patchy ground-glass abnormalities present in <5% of any lung zone were regarded as having indeterminate changes [1]. Images from participants with ILA were further classified by the presence of the definite fibrosis imaging pattern, defined as pulmonary parenchymal architectural distortion consistent with a fibrotic lung disease [3].

Data on cancer diagnoses were available in 5270 (99%) of the 5320 AGES-Reykjavik participants previously characterised for ILA.

Participants were followed from their entry into the study (between 2002 and 2006) until their first diagnosis of cancer or until the end of observation (August 31, 2016). Information regarding cancer diagnoses was ascertained from electronic medical records from Landspitali University Hospital, Iceland's largest, and only tertiary care, hospital. Participants' hospital visits with a registered International Classification of Diseases, Tenth Revision (ICD-10) diagnosis ranging from C00 to C97 were defined as cancer diagnoses, with the date of the first such visit defined as the date of first diagnosis. Lung cancer diagnoses were likewise defined from hospital visits with a registered ICD-10 diagnosis starting with C34. Information on mortality and causes of death was obtained from the Icelandic Directorate of Health, with follow-up from study entry until the end of August 2016. Mortality from cancer was defined as having the cause of death registered as C00–C97, coded according to the ICD-10, while mortality from lung cancer was defined as having C34 as the registered cause of death.

Statistical analyses

Comparable to previous studies [4, 6], participants indeterminate for ILA were excluded from analyses of the associations between ILA, cancer diagnoses and cancer-associated mortality. The cumulative incidences of lung cancer diagnoses and diagnoses of other cancers among participants with and without ILA were calculated, with the risk of mortality regarded as a competing risk. Gray's tests were used to assess for differences in these cumulative incidences. The cumulative incidences of mortality from lung cancer, mortality from non-pulmonary cancers and mortality from other causes were calculated and compared between participants with and without ILA using Gray's tests with all risks regarded as competing.

Cox proportional hazards models were used to quantify the associations of ILA and several outcomes: lung cancer diagnoses, diagnoses of all cancers, mortality from lung cancers and mortality from all cancers. The proportional hazards assumption was tested and graphically verified for all models. The covariates included in all adjusted models were age, sex, pack-years of smoking and smoking at the beginning of the study. In addition, models analysing the associations of ILA with diagnoses of all cancers and mortality from all cancers were constructed in which lung cancer diagnoses and lung cancer mortality were excluded from the outcomes. Identical Cox proportional hazards models were created in which participants with ILA were compared with both participants indeterminate for ILA and participants without ILA. Results from these models are shown in the supplementary material.

Proportional subdistribution hazards models [18] were created to verify results from Cox models using regression methods accounting for competing risks. To assess whether lung cancer outcomes differed depending on the presence of the definite fibrosis pattern, adjusted and unadjusted Cox proportional hazards models were created. In these models, the associations of lung cancer diagnoses and mortality from lung cancer were assessed, comparing participants with ILA and definite fibrosis or ILA without definite fibrosis to participants without ILA.

Statistical analyses were done using R, version 3.5.2 (R Project for Statistical Computing, Vienna).

Results

Participants' characteristics

Demographic variables and the incidence of cancer diagnoses in participants stratified by ILA status are included in table 1. Comparable to previous reports [8], participants with ILA were on average older, more likely to be male and more likely to be exposed to tobacco smoke than participants without ILA.

ILA and cancer diagnoses

Subsequent to study entry, participants with ILA were more likely to have received a diagnosis of cancer overall, and lung cancer specifically, than participants without ILA (table 1).

The cumulative incidences of lung cancer diagnoses and other cancer diagnoses are displayed in figure 1. There was a greater cumulative incidence of lung cancer diagnoses among participants with ILA than

TABLE 1 Baseline participant characteristics

	No ILA	Indeterminate for ILA	ILA
Participants n	3183	1712	375
Age years	76.0±5.4	77.4±5.7	77.8±5.6
Women	1887 (59)	953 (56)	170 (45)
BMI kg·m^{-2}	27.2±4.4	26.8±4.4	27.0±4.6
History of smoking	1732 (54)	1013 (59)	269 (72)
Median pack-years (IQR)	0 (0–16)	2.5 (0–23)	11 (0–28)
Current smoker	368 (12)	203 (12)	68 (18)
Days of follow-up to all-cause mortality	3675±1228	3396±1347	2981±1433
Imaging patterns			
Without fibrosis			246 (66)
Definite fibrosis			129 (34)
Participants diagnosed with cancer before beginning of study			
Overall	194 (6.1)	132 (7.7)	32 (8.5)
Participants diagnosed with cancer after beginning of study			
Overall	668 (21)	383 (22)	97 (26)
Lung cancer (C34)	77 (2.4)	58 (3.4)	27 (7.2)
Gastrointestinal cancer (C15–C26)	176 (5.5)	86 (5.0)	20 (5.3)
Skin cancers (C43–C44)	45 (1.4)	30 (1.8)	4 (1.1)
Cancers of breasts and female genitalia (C50–C58)	108 (3.4)	44 (2.6)	8 (2.1)
Cancers of male genitalia (C60–C63)	124 (3.9)	71 (4.1)	20 (5.3)
Urinary tract cancers (C64–C68)	81 (2.5)	49 (2.9)	11 (2.9)
Haematologic malignancies (C81–C96)	57 (1.8)	34 (2.0)	9 (2.4)
Mortality due to cancer during study follow-up			
Cancer overall	388 (12)	232 (14)	63 (17)
Lung cancer (C34)	65 (2.0)	61 (3.6)	25 (6.7)

Data are presented as mean±SD or n (%), unless otherwise stated. ILA: interstitial lung abnormalities; BMI: body mass index; IQR: interquartile range.

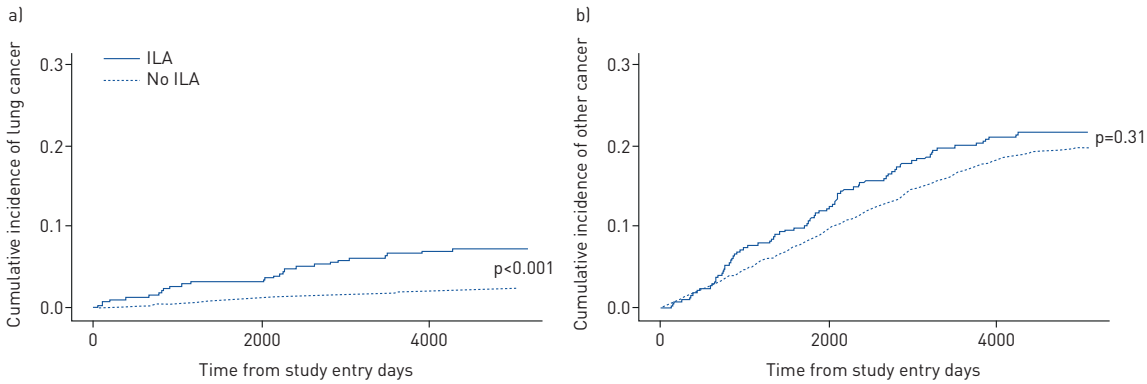

FIGURE 1 Cumulative incidence of cancer diagnoses among participants with and without interstitial lung abnormalities (ILA). p-values are for differences in the cumulative incidence of cancer diagnoses between participants with ILA and participants without ILA for a) lung cancer and b) other cancers using Gray's tests.

among participants without ILA (p<0.001). There were no significant differences in the cumulative incidences of other cancer diagnoses between participants with and without ILA (figure 1).

In Cox proportional hazards models, participants with ILA were at increased risk of lung cancer diagnosis than those without ILA, both in an unadjusted model (hazard ratio (HR) 3.76, 95% CI 2.42–5.84, $p=3.59\times10^{-9}$) and in a model adjusting for age, sex, pack-years of smoking and smoking at the beginning of the study (HR 2.77, 95% CI 1.76–4.36, $p=1.08\times10^{-5}$) (table 2). In adjusted models, the increase in risk of lung cancer diagnosis was statistically significant for both participants with the definite fibrosis imaging pattern (HR=3.95, 95% CI=2.07–7.57, $p=3.32\times10^{-5}$) and without it (HR=2.26, 95% CI=1.29–3.96, p=0.004), although participants with fibrosis were at greater risk (table 3). Participants with ILA were at an increased risk of diagnosis of cancer overall in adjusted models (HR 1.35, 95% CI 1.09–1.68, p=0.006). In contrast, the increase in risk of a diagnosis of all cancers excluding lung cancers was not statistically significant among participants with ILA (HR 1.24, 95% CI 0.98–1.57, p=0.07) (table 2).

ILA and mortality from cancer

The cumulative incidences of mortality from lung cancer and mortality from cancers other than lung cancer are displayed in figure 2. There was greater mortality from lung cancer among participants with ILA than without ILA (p<0.001), as well as greater mortality from causes other than cancer. However, mortality from cancers other than lung cancer was not increased among participants with ILA (figure 2).

In unadjusted Cox proportional hazards models, participants with ILA were at increased risk of death from all cancers (HR 1.81, 95% CI 1.39–2.37, $p=1.23\times10^{-5}$) and from lung cancer specifically (HR 4.19, 95% CI 2.64–6.66, $p=1.27\times10^{-9}$) compared to those without ILA. In models adjusting for age, sex,

TABLE 2 Associations of interstitial lung abnormalities (ILA) with cancer diagnoses

Model	HR (95% CI)	p-value
Lung cancer diagnoses		
Unadjusted	3.76 (2.42–5.84)	3.59×10^{-9}
Adjusted	2.77 (1.76–4.36)	1.08×10^{-5}
Cancer diagnoses of all causes		
Unadjusted	1.57 (1.27–1.95)	3.07×10^{-5}
Adjusted	1.35 (1.09–1.68)	0.006
Cancer diagnoses of all causes excluding lung cancer		
Unadjusted	1.39 (1.11–1.76)	0.005
Adjusted	1.24 (0.98–1.57)	0.07

All models are Cox proportional hazards models of the association of ILA with the specified cancer diagnoses. Adjusted models are adjusted for age, sex, pack-years and smoking at entry. HR: cause-specific hazard ratio.

TABLE 3 Associations of imaging patterns with lung cancer diagnoses and mortality

Model	HR (95% CI)	p-value
Definite fibrosis		
Lung cancer diagnoses		
Unadjusted	5.49 (2.91–10.4)	1.56×10^{-7}
Adjusted	3.95 (2.07–7.57)	3.32×10^{-5}
Mortality from lung cancer		
Unadjusted	8.86 (4.94–15.9)	2.37×10^{-13}
Adjusted	5.98 (3.29–10.9)	4.17×10^{-9}
Without fibrosis		
Lung cancer diagnoses		
Unadjusted	3.10 (1.81–5.32)	3.90×10^{-5}
Adjusted	2.26 (1.29–3.96)	0.004
Mortality from lung cancer		
Unadjusted	2.53 (1.33–4.79)	0.005
Adjusted	1.68 (0.86–3.29)	0.13

All models are Cox proportional hazards models of the association of the specified pattern of interstitial lung abnormalities (ILA) with diagnoses of, or mortality from, lung cancer. All comparisons are made with participants without ILA. Adjusted models are adjusted for age, sex, pack-years and smoking at entry. HR: cause-specific hazard ratio.

pack-years of smoking and smoking at the beginning of the study, the same was true for death from cancer overall (HR 1.47, 95% CI 1.12–1.94, p=0.005) and from lung cancer (HR 2.89, 95% CI 1.80–4.66, $p=1.26\times10^{-5}$). However, the risk of death from all cancers excluding lung cancer was not statistically significantly increased among those with ILA (HR 1.15, 95% CI 0.82–1.61, p=0.43) (table 4). Participants with definite fibrosis were, in adjusted models, at increased risk of death from lung cancer (HR 5.98, 95% CI 3.29–10.9, $p=4.17\times10^{-9}$). This increase was not statistically significant for participants with ILA without definite fibrosis (HR 1.68, 95% CI 0.86–3.29, p=0.13) (table 3).

In proportional subdistribution hazards models adjusted for covariates, ILA was also found to be associated with an increased risk of lung cancer diagnosis (HR 2.63, 95% CI 1.58–4.38, $p=1.9\times10^{-4}$) and mortality from lung cancer (HR 2.55, 95% CI 1.56–4.18, $p=2.1\times10^{-4}$).

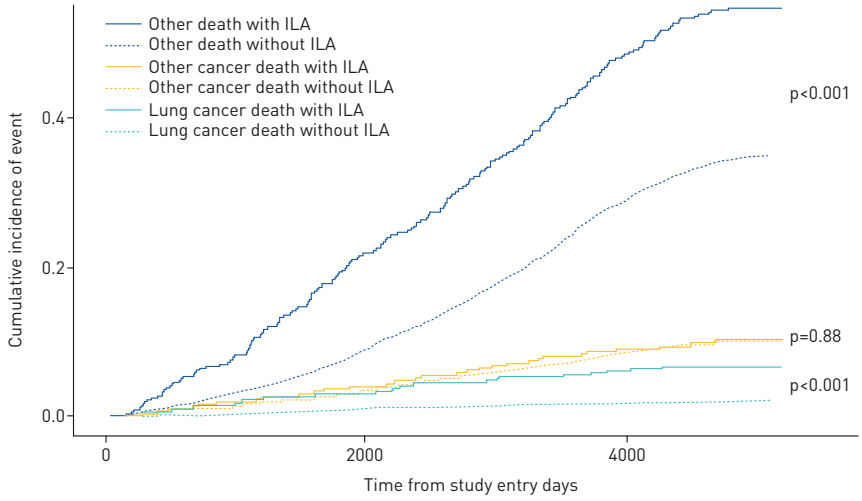

FIGURE 2 Cumulative incidence of cancer mortality among participants with and without interstitial lung abnormalities (ILA). p-values are for differences in the cumulative incidence of mortality due to the specified cause between participants with ILA and participants without ILA using Gray's tests.

TABLE 4 Associations of interstitial lung abnormalities (ILA) with mortality from cancer

Model	HR (95% CI)	p-value
Mortality from lung cancer		
Unadjusted	4.19 (2.64–6.66)	1.27×10^{-9}
Adjusted	2.89 (1.80–4.66)	1.26×10^{-5}
Mortality from all cancers		
Unadjusted	1.81 (1.39–2.37)	1.23×10^{-5}
Adjusted	1.47 (1.12–1.94)	0.005
Mortality from all cancers excluding lung cancer		
Unadjusted	1.32 (0.94–1.85)	0.10
Adjusted	1.15 (0.82–1.61)	0.43

All models are Cox proportional hazards models of the association of ILA with mortality from the specified cancers. Adjusted models are adjusted for age, sex, pack-years and smoking at entry. HR: cause-specific hazard ratio.

Discussion

These results demonstrate that AGES-Reykjavik participants with ILA are at an increased risk of both lung cancer diagnosis and mortality from lung cancer. These associations between ILA, lung cancer and lung cancer-associated mortality were consistent between Cox proportional hazards models and methods accounting for the competing risks of other cancer diagnoses or other causes of mortality. Associations varied with ILA patterns; participants with the definite fibrosis pattern were at greater risk of lung cancer diagnosis than participants without definite fibrosis, and an increase in lung cancer-associated mortality was only found among participants with definite fibrosis. The associations of ILA, which in some cases may represent early fibrotic changes of the lung [19], with pulmonary malignancies are in concordance with the well-established increase in risk of lung cancer among patients with more advanced pulmonary fibrosis such as IPF [10, 11]. These results also support previous findings from lung cancer screening studies of smokers that demonstrate an increased prevalence of ILA among patients with lung cancer [14, 15], as well as increased mortality from lung cancer among participants with ILA [13, 15]. However, associations of ILA with lung cancer diagnoses and lung cancer mortality have not been reported in a population-based cohort.

The mechanisms underlying these associations are yet to be clarified. It is possible that a common pathobiological process exists for fibrotic lung disease and lung cancer. Several studies have explored this possibility with regards to lung cancer and IPF [20–22]. Among similarities noted in the pathogenesis of these diseases are genetic alterations [22–25], epigenetic similarities including in DNA methylation and altered mRNA expression profiles [20, 24, 26], altered cell-to-cell communication, abnormalities in intracellular signalling pathways and overexpression of several signalling molecules [21, 24, 27, 28]. Besides a common biological pathway, it is possible that the results could be explained by a common, unmeasured risk factor *via* residual confounding. However, the mechanisms underlying these results cannot be determined from the cohort data shown here and thus remain a topic of research.

The results oppose the suggestion that participants with ILA are at an increased risk of diagnoses of and mortality from cancers other than lung cancers. While there were small associations between ILA and these outcomes, they did not reach statistical significance when lung cancer diagnoses were excluded. This is supported by the lack of difference in the cumulative incidence of non-pulmonary cancer diagnoses and mortality from non-pulmonary cancers between participants with and without ILA (figures 1 and 2).

The study has several limitations. Cancer diagnoses were obtained from medical records from the National Hospital of Iceland. The outcome data regarding both cancer diagnoses and mortality are separately registered health record data, meaning that the quality of the data is dependent on the quality of clinicians' diagnoses and clinical registration. Among other limitations is the possibility that unknown confounding factors were not adjusted for. This is especially a concern in analyses regarding diagnoses of all cancers because various cancers have different risk factors that were not all adjusted for in these analyses. The association of ILA with mortality from lung cancer was dependent on the presence of the definite fibrosis pattern. That supports the notion that some of the associations presented could be limited to very extensive or progressive abnormalities, similar to changes seen in interstitial lung disease that are known to be associated with lung cancer [10, 11]. Finally, while our findings suggest that ILA preceded the diagnosis of lung cancer in the AGES-Reykjavik study, and we excluded cancer diagnoses that were present on participant entry, we cannot exclude the possibility that some slowly growing lung cancers could have occurred coincident with, or preceded the development of, ILA in some participants.

Despite these limitations, these findings have several implications for further research. The associations presented here between ILA and lung cancer indicate that studies and theories investigating the biological relationship between cancer and fibrotic lung diseases such as IPF could extend their approach to earlier stages of pulmonary fibrosis. In addition, the increased risk of lung cancer among people with ILA could, if replicated in studies of other populations, suggest that early fibrotic changes of the lung such as ILA should be considered as a risk factor in lung cancer screening.

In conclusion, ILA were found to be associated with an increased hazard of lung cancer diagnosis as well as increased mortality from lung cancer. Such associations were not found for non-pulmonary malignancies.

Author contributions: Study design: G.T. Axelsson, R.K. Putman, T. Araki, H. Hatabu, V. Gudnason, G.M. Hunninghake and G. Gudmundsson; acquisition or interpretation of the data: G.T. Axelsson, R.K. Putman, T. Aspelund, E.F. Gudmundsson, T. Hida, T. Araki, M. Nishino, H. Hatabu, V. Gudnason, G.M. Hunninghake and G. Gudmundsson; critical revision of the manuscript for important intellectual content: G.T. Axelsson, R.K. Putman, T. Aspelund, E.F. Gudmundsson, T. Hida, T. Araki, M. Nishino, H. Hatabu, V. Gudnason, G.M. Hunninghake and G. Gudmundsson; statistical analysis: G.T. Axelsson, T. Aspelund, G.M. Hunninghake and E.F. Gudmundsson; obtained funding: G. Gudmundsson, V. Gudnason and G.M. Hunninghake.

Conflict of interest: G.T. Axelsson has nothing to disclose. R.K. Putman reports grants from NIH, during the conduct of the study. T. Aspelund has nothing to disclose. E.F. Gudmundsson has nothing to disclose. T. Hida has nothing to disclose. T. Araki has nothing to disclose. M. Nishino reports personal fees for consultancy from Daiichi Sankyo and AstraZeneca; honoraria from Roche; and grants from Merck, AstraZeneca, Canon Medical Systems and NIH (R01CA203636, U01CA209414, R01HL111024), outside the submitted work. H. Hatabu reports grants from Canon Medical System Inc. and Konica-Minolta Inc., personal fees for consultancy from Mitsubishi Chemical Inc. and personal fees for advisory board work from Canon Medical System Inc., outside the submitted work. V. Gudnason has nothing to disclose. G.M. Hunninghake reports personal fees from Genentech, Boehringer Ingelheim, The Gerson Lehrman Group and Mitsubishi Chemical, outside the submitted work. G. Gudmundsson has nothing to disclose.

Support statement: Supported by National Institutes of Health (NIH) grants K08 HL140087 (R.K. Putman); R01 CA203636 (M. Nishino); and R01 HL111024, R01 HL130974 and R01 135142 (G.M. Hunninghake); National Institute on Aging (NIA) grant 27120120022C (V. Gudnason); the Icelandic Research Fund, project grant 141513-051 (G. Gudmundsson, V. Gudnason and G.M. Hunninghake); Oddur Olafsson Fund; and Landspitali Scientific Fund A-2015-030 and A-2016-023 (G. Gudmundsson). The Age, Gene/Environment Susceptibility-Reykjavik Study was supported by NIH contracts N01-AG-1-2100 and HHSN27120120022C, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association) and the Althingi (the Icelandic Parliament). Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Washko GR, Hunninghake GM, Fernandez IE, *et al.* Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011; 364: 897–906.
2 Araki T, Putman RK, Hatabu H, *et al.* Development and progression of interstitial lung abnormalities in the Framingham Heart Study. *Am J Respir Crit Care Med* 2016; 194: 1514–1522.
3 Putman RK, Gudmundsson G, Axelsson GT, *et al.* Imaging patterns are associated with interstitial lung abnormality progression and mortality. *Am J Respir Crit Care Med* 2019; 200: 175–183.
4 Hunninghake GM, Hatabu H, Okajima Y, *et al.* MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013; 368: 2192–2200.
5 Hobbs BD, Putman RK, Araki T, *et al.* Overlap of genetic risk between interstitial lung abnormalities and idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019; 200: 1402–1413.
6 Axelsson GT, Putman RK, Araki T, *et al.* Interstitial lung abnormalities and self-reported health and functional status. *Thorax* 2018; 73: 884–886.
7 Axelsson GT, Putman RK, Miller ER, *et al.* Interstitial lung abnormalities and physical function. *ERJ Open Res* 2018; 4: 00057-2018.
8 Putman RK, Hatabu H, Araki T, *et al.* Association between interstitial lung abnormalities and all-cause mortality. *JAMA* 2016; 315: 672–681.
9 Putman RK, Gudmundsson G, Araki T, *et al.* The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *Eur Respir J* 2017; 50: 1700537.
10 Le Jeune I, Gribbin J, West J, *et al.* The incidence of cancer in patients with idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Respir Med* 2007; 101: 2534–2540.
11 Hubbard R, Venn A, Lewis S, *et al.* Lung cancer and cryptogenic fibrosing alveolitis. *Am J Respir Crit Care Med* 2000; 161: 5–8.
12 Tomassetti S, Gurioli C, Ryu JH, *et al.* the impact of lung cancer on survival of idiopathic pulmonary fibrosis. *Chest* 2015; 147: 157–164.
13 Hoyer N, Wille MMW, Thomsen LH, *et al.* Interstitial lung abnormalities are associated with increased mortality in smokers. *Respir Med* 2018; 136: 77–82.
14 Wille MM, Thomsen LH, Petersen J, *et al.* Visual assessment of early emphysema and interstitial abnormalities on CT is useful in lung cancer risk analysis. *Eur Radiol* 2016; 26: 487–494.
15 Brown SAW, Padilla M, Mhango G, *et al.* Interstitial lung abnormalities and lung cancer risk in the National Lung Screening Trial. *Chest* 2019; 156: 1195–1203.
16 Nishino M, Cardarella S, Dahlberg SE, *et al.* Interstitial lung abnormalities in treatment-naive advanced non-small-cell lung cancer patients are associated with shorter survival. *Eur J Radiol* 2015; 84: 998–1004.
17 Harris TB, Launer LJ, Eiriksdottir G, *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* 2007; 165: 1076–1087.

18 Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999; 94: 496–509.
19 Miller ER, Putman RK, Vivero M, *et al.* Histopathology of interstitial lung abnormalities in the context of lung nodule resections. *Am J Respir Crit Care Med* 2018; 197: 955–958.
20 Antoniou KM, Tomassetti S, Tsitoura E, *et al.* Idiopathic pulmonary fibrosis and lung cancer: a clinical and pathogenesis update. *Curr Opin Pulm Med* 2015; 21: 626–633.
21 Vancheri C. Common pathways in idiopathic pulmonary fibrosis and cancer. *Eur Respir Rev* 2013; 22: 265–272.
22 Buendía-Roldán I, Mejía M, Navarro C, *et al.* Idiopathic pulmonary fibrosis: clinical behavior and aging associated comorbidities. *Respir Med* 2017; 129: 46–52.
23 Carpagnano GE, Lacedonia D, Soccio P, *et al.* How strong is the association between IPF and lung cancer? An answer from airway's DNA. *Med Oncol* 2016; 33: 119.
24 Ballester B, Milara J, Cortijo J. Idiopathic pulmonary fibrosis and lung cancer: mechanisms and molecular targets. *Int J Mol Sci* 2019; 20: 593.
25 Wang Y, Kuan PJ, Xing C, *et al.* Genetic defects in surfactant protein A2 are associated with pulmonary fibrosis and lung cancer. *Am J Hum Genet* 2009; 84: 52–59.
26 Rabinovich EI, Kapetanaki MG, Steinfeld I, *et al.* Global methylation patterns in idiopathic pulmonary fibrosis. *PLoS One* 2012; 7: e33770.
27 Mercer PF, Woodcock HV, Eley JD, *et al.* Exploration of a potent PI3 kinase/mTOR inhibitor as a novel anti-fibrotic agent in IPF. *Thorax* 2016; 71: 701–711.
28 Vancheri C. Idiopathic pulmonary fibrosis and cancer: do they really look similar? *BMC Med* 2015; 13: 220.

Paper III

Cite: Axelsson G.T., Gudmundsson G., Pratte K.A., Aspelund T., Putman R.K., Sanders J.L., . . . Gudnason V. (2022). The Proteomic Profile of Interstitial Lung Abnormalities. *American Journal of Respiratory and Critical Care Medicine, 206(3),* 337-346

The American Journal of Respiratory and Critical Care Medicine is an official journal of the American Thoracic Society.

ORIGINAL ARTICLE

The Proteomic Profile of Interstitial Lung Abnormalities

Gisli Thor Axelsson[1,2], Gunnar Gudmundsson[1,3], Katherine A. Pratte[4], Thor Aspelund[1,2], Rachel K. Putman[5], Jason L. Sanders[5], Elias F. Gudmundsson[2], Hiroto Hatabu[6,7], Valborg Gudmundsdottir[1,2], Alexander Gudjonsson[2], Takuya Hino[7], Tomoyuki Hida[7,8], Brian D. Hobbs[5,9], Michael H. Cho[5,9], Edwin K. Silverman[5,9], Russell P. Bowler[4,10], Lenore J. Launer[11], Lori L. Jennings[12], Gary M. Hunninghake[5,7], Valur Emilsson[2], and Vilmundur Gudnason[1,2]

[1]Faculty of Medicine, University of Iceland, Reykjavik, Iceland; [2]Icelandic Heart Association, Kopavogur, Iceland; [3]Department of Respiratory Medicine, Landspitali University Hospital, Reykjavik, Iceland; [4]National Jewish Health, Denver, Colorado; [5]Pulmonary and Critical Care Division, [6]Department of Radiology, and [7]Center for Pulmonary Functional Imaging, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; [8]Department of Clinical Radiology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; [9]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts; [10]School of Medicine, University of Colorado, Aurora, Colorado; [11]Laboratory of Epidemiology and Population Sciences, Intramural Research Program, National Institute on Aging, Bethesda, Maryland; and [12]Novartis Institutes for Biomedical Research, Cambridge, Massachusetts

ORCID ID: 0000-0002-7156-9080 (G.T.A.).

Abstract

Rationale: Knowledge on biomarkers of interstitial lung disease is incomplete. Interstitial lung abnormalities (ILAs) are radiologic changes that may present in its early stages.

Objectives: To uncover blood proteins associated with ILAs using large-scale proteomics methods.

Methods: Data from two prospective cohort studies, the AGES-Reykjavik (Age, Gene/Environment Susceptibility–Reykjavik) study ($N = 5{,}259$) for biomarker discovery and the COPDGene (Genetic Epidemiology of COPD) study ($N = 4{,}899$) for replication, were used. Blood proteins were measured using DNA aptamers, targeting more than 4,700 protein analytes. The association of proteins with ILAs and ILA progression was assessed with regression modeling, as were associations with genetic risk factors. Adaptive Least Absolute Shrinkage and Selection Operator models were applied to bootstrap data samples to discover sets of proteins predictive of ILAs and their progression.

Measurements and Main Results: Of 287 associations, SFTPB (surfactant protein B) (odds ratio [OR], 3.71 [95% confidence interval (CI), 3.20–4.30]; $P = 4.28 \times 10^{-67}$), SCGB3A1 (Secretoglobin family 3A member 1) (OR, 2.43 [95% CI, 2.13–2.77]; $P = 8.01 \times 10^{-40}$), and WFDC2 (WAP four-disulfide core domain protein 2) (OR, 2.42 [95% CI, 2.11–2.78]; $P = 4.01 \times 10^{-36}$) were most significantly associated with ILA in AGES-Reykjavik and were replicated in COPDGene. In AGES-Reykjavik, concentrations of SFTPB were associated with the rs35705950 *MUC5B* (mucin 5B) promoter polymorphism, and SFTPB and WFDC2 had the strongest associations with ILA progression. Multivariate models of ILAs in AGES-Reykjavik, ILAs in COPDGene, and ILA progression in AGES-Reykjavik had validated areas under the receiver operating characteristic curve of 0.880, 0.826, and 0.824, respectively.

Conclusions: Novel, replicated associations of ILA, its progression, and genetic risk factors with numerous blood proteins are demonstrated as well as machine-learning–based models with favorable predictive potential. Several proteins are revealed as potential markers of early fibrotic lung disease.

Keywords: interstitial lung abnormalities; interstitial lung disease; idiopathic pulmonary fibrosis; proteomics; biomarkers

Many biologically active proteins have been proposed as potential biomarkers for advanced interstitial lung diseases (ILDs) that can result in pulmonary fibrosis, such as idiopathic pulmonary fibrosis (IPF). These are indicators of alveolar epithelial cell damage, proteins involved in extracellular remodeling, proteins involved in immune response, adhesion molecules, and growth factors (1–5). Still, novel ILD biomarkers are needed (1).

Interstitial lung abnormalities (ILAs) are chest computed tomography (CT) abnormalities resembling the radiologic appearance of ILD (6, 7). They are associated with risk factors common in patients with IPF, such as age, smoking, restrictive lung deficits, and certain genetic polymorphisms, most notably the rs35705950 promoter polymorphism of the *MUC5B* gene (8).

(Received in original form October 9, 2021; accepted in final form April 18, 2022)

Am J Respir Crit Care Med Vol 206, Iss 3, pp 337–346, Aug 1, 2022

Originally Published in Press as DOI: 10.1164/rccm.202110-2296OC on April 19, 2022
Internet address: www.atsjournals.org

At A Glance Commentary

Scientific Knowledge on the Subject: Knowledge of protein biomarkers of interstitial lung disease is insufficient, especially in its early forms.

What This Study Adds to the Field: We present a large-scale proteomic study, the first such study of interstitial lung abnormalities, that suggests several potential biomarkers of changes suggestive of early pulmonary fibrosis.

In addition, some patients with ILA have histopathological evidence of pulmonary fibrosis (6, 8–10), progression of ILAs has been reported (11), and ILAs are associated with increased mortality (12). Patterns of ILA have been classified, among which associations with progression and mortality vary (11, 13). Interest in ILAs stems from an ambition to detect an early stage of pulmonary fibrosis before advanced architectural remodeling develops (7). Blood biomarkers of ILA could aid with identifying those at greatest risk of progression to pulmonary fibrosis and improving understanding of the pathogenesis of early disease stages (7, 14).

Blood-based proteomics methods are emerging as an effective way of uncovering accessible biomarkers of human disease (15). Although such methods have been applied in small cohorts of patients with IPF and their relatives (2–5), no study exists in which methods of proteomics are applied to early stages of fibrotic lung disease.

Therefore, the objective of this study was to apply large-scale proteomics methods to identify biomarkers of ILA and their progression.

Methods

Study Design

The AGES-Reykjavik (Age, Gene/Environment Susceptibility–Reykjavik) study was designed to explore risk factors of disease among the elderly with a multidisciplinary approach (16). The 5,764 participants underwent a range of examinations, including CT imaging of the thorax. Five years later, 3,167 participants had a follow-up examination and CT.

The COPDGene (Genetic Epidemiology of COPD) study is a multicenter cohort study of non-Hispanic White and African American individuals, 99% of whom were smokers, designed to investigate the genetics and epidemiology of chronic obstructive pulmonary disease (COPD) (6). Subjects with significant interstitial lung disease at enrollment were ineligible for COPDGene. Data were used from 5,339 participants who participated in the 5-year follow-up (phase 2) visit, who had chest CT imaging and fresh-frozen plasma samples available (17).

Definitions of ILA, ILA Subtypes, and ILA Progression

Images from initial and follow-up examinations were visually assessed for the presence of ILAs in AGES-Reykjavik (6, 11), as were images from the COPDGene 5-year follow-up visit. ILAs were defined per recent Fleischner Society guidelines (14). Changes present in <5% of any lung zone were deemed indeterminate (6). Images from participants with ILAs were classified with regard to the presence of definite fibrosis and the usual interstitial pneumonia (UIP) pattern (6, 11). For participants who had ILA present at initial examination or at follow-up, CT scans of AGES-Reykjavik participants were simultaneously compared for the development and progression (heretofore referred to as "progression") of ILA (11).

Protein Profiling

Proteomic measurements were performed coincident with the initial chest CT in AGES-Reykjavik and with the 5-year follow-up visit chest CT in COPDGene. Single-stranded DNA aptamers designed to recognize target proteins, termed Slow-Off rate Modified Aptamers (SOMAmers), were used for

Supported by National Institutes of Health grants K08 HL140087 (R.K.P.); K08 HL136928 (B.D.H.); R01 HL135142 and R01 HL137927 (M.H.C.); U01 HL089856, R01 HL113264, R01 137927, R01 133135, and P01 HL114501 (E.K.S.); R01 HL137995 and R01 HL152735 (R.P.B.); R01 HL111024, R01 HL130974, and R01 135142 (G.M.H., H.H.); and R01CA203636 and 5U01CA209414-03 (H.H.). Supported by National Institute on Aging grant 27120120022C (V. Gudnason). Supported by the Icelandic Centre for Research, project grants 141513-051 (G.G., V. Gudnason, and G.M.H.); 195761-051 (V.E.); and 184845-051 and 206692-051 (V. Gudmundsdottir). Supported by Landspítali Háskólasjúkrahús grants A-2019-029, A-2019-030, A-2020-018, and A-2020-017 (G.G.), University of Iceland Research Fund 2021, and the Eimskip University Fund (G.T.A.). The Age, Gene/Environment Susceptibility-Reykjavik Study was supported by National Institutes of Health contracts N01-AG-1-2100 and HHSN27120120022C, the National Institute on Aging Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). COPDGene: The COPDGene study (NCT00608764) is also supported by the COPD Foundation through contributions made to an Industry Advisory Committee that has included AstraZeneca, Bayer Pharmaceuticals, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, and Sunovion. The project described was supported by National Heart, Lung, and Blood Institute grants U01 HL089897, U01 HL089856, R01 HL137995, and R01 HL129937. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health. SCCOR: This project was supported by National Heart, Lung, and Blood Institute grants P50HL084948 and R21HL129917 and Pennsylvania Department of Health CURE SAP 4100062224. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute. Funding sources of the study had no role in the collection, analysis, or interpretation of the data, writing of the manuscript, or the decision to submit for publication.

Author Contributions: Conceptualization, methodology, and design of the study: G.T.A., G.G., J.L.S., B.D.H., M.H.C., E.K.S., R.P.B., L.J.L., L.L.J., G.M.H., V.E., and V.Gudnason. Data acquisition: R.K.P., E.F.G., H.H., V. Gudmundsdottir, A.G., T. Hino, T. Hida, B.D.H., E.K.S., R.P.B., G.M.H., V.E., and V. Gudnason. Analysis of the data: G.T.A., G.G., K.A.P., T.A., V. Gudmundsdottir, and A.G. Administration and supervision: E.K.S., R.P.B., L.J.L., L.L.J., G.M.H., V.E., and V. Gudnason. Drafting of the initial manuscript: G.T.A., G.G., G.M.H., and V. Gudnason. All authors edited the manuscript for scientific content. All the authors agree to be accountable for the work with regard to accuracy and integrity.

Correspondence and requests for reprints should be addressed to Gisli Thor Axelsson, M.D., Faculty of Medicine, University of Iceland, Vatnsmyrarvegur 16, 101 Reykjavik, Iceland. E-mail: gtha7@hi.is.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

protein detection and measurements. Serum samples from 5,457 participants were incubated with a mixture of 5,034 SOMAmers, creating SOMAmer–protein complexes. After washout of unbound proteins and SOMAmers, these enriched SOMAmers were quantified using a hybridization array. Median intraassay and interassay coefficients of variation were found to be <5%. SOMAmer specificity was validated by cross-platform validation, and the specificity of 779 SOMAmers was directly confirmed using mass spectrometry techniques. Details of this process are previously described (18). Human proteins were targeted by 4,782 SOMAmers. Because some proteins were targeted by multiple SOMAmers and some were annotated to multiple genes, these SOMAmers targeted 4,137 human proteins with 4,115 unique genetic targets. For statistical analyses, one SOMAmer per genetic target was used, selecting the SOMAmer with the stronger association with ILA at baseline for proteins targeted by multiple SOMAmers.

In the COPDGene study, proteomic measurements were conducted on ethylenediaminetetraacetic acid plasma samples collected during the phase 2 clinical visit ($n = 6{,}018$). Samples were stored at −80°C until the time of assaying by SomaLogic on their SomaScan version 4.0 (5K) assay for human plasma. This version of SomaScan contained 5,285 SOMAmers, of which 4,979 target human proteins, representing 4,776 unique proteins with 4,720 unique Uniprot numbers. SomaLogic standardized the SomaScan data per their protocol. It consisted of within-plate hybridization to control for variability across array signals, median signal normalization to control for technical variability of replicates within a run, plate scaling and calibration of SOMAmers to control for interassay variation between analytes, and batch differences between plates. Finally, median normalization to a reference using adaptive normalization by maximum likelihood is applied within dilution group to quality control (QC) replicates and individual samples to remove edge effects and technical variance. After data cleaning ($n = 101$), removing those with lung reduction or transplant surgery ($n = 19$), and samples that failed QC ($n = 228$), 5,670 results were available for analysis. The QC coefficient of variation at the 10%, 50%, and 90% percentiles were 3.1%, 5.0%, and 9.8%, respectively. SomaScan results were reported in relative fluorescence units.

Statistical Analysis

A flow chart of study design is shown in Figure E1 in the online supplement. Because of variability in measurement values, protein data were transformed using a variant of the Box-Cox transformation, providing results per SD (19). Extreme outliers (0.2% of values) were removed because of likelihood of measurement error and imputed before analysis using K-nearest neighbor imputing. All logistic regression models were adjusted for age, sex, pack-years, and smoking status at the beginning of the study. Bonferroni-corrected P values < 0.05 were considered significant for single-point analyses. Actual P values are shown throughout the paper. Statistical analyses were performed using R.

Analyses of ILA in AGES-Reykjavik

Of 5,764 AGES-Reykjavik participants, 5,259 (91%) had data on ILA status at baseline and protein measurements. Logistic regression models of the associations of single proteins with ILA were fitted for all proteins. For comparison, associations of single proteins with indeterminate ILA were also assessed with logistic regression; otherwise, participants with indeterminate ILA status were excluded from all analyses. The 1,609 proteins with suggestive associations ($P < 0.05$) with ILA in single-protein models were explored using adaptive LASSO (Least Absolute Shrinkage and Selection Operator) modeling of 200 bootstrap data samples (20). The associations of the eight proteins that occurred in all 200 LASSO models were analyzed using multivariate logistic regression models. The areas under the receiver operating characteristic curves (AUROCs) were calculated for these models and validated using resampling methods (21). The variance inflation factor was calculated for the multivariate regression model. Further methodological details, comparative analyses of proteins occurring in fewer LASSO models, tissue expression analyses, and functional enrichment analyses based on the GTEx, KEGG, WikiPathways, TRANSFAC, CORUM, and HPO databases are described in online supplemental methods.

Replication of ILA Analyses in COPDGene

In the COPDGene study, SOMAmer data were available for 4,899 participants (92% of participants with CT data available). After exclusion of those indeterminate for ILA and with missing covariate data, data were available for 2,974 participants. Single-protein logistic regression models using the three proteins with the strongest associations with ILA in AGES-Reykjavik (SFTPB [surfactant protein B], SCGB3A1 [Secretoglobin family 3A member 1], and WFDC2 [WAP four-disulfide core domain protein 2]) were fitted, as well as the eight-protein multivariate logistic regression model based on adaptive LASSO modeling in AGES-Reykjavik. The AUROCs of these models were calculated and validated, with methods identical to those in AGES-Reykjavik. Because of their association with variable standardization, models in COPDGene were additionally adjusted for white blood cell and platelet count and study center.

Analyses of Pulmonary Fibrosis–associated SNPs

Genotyping was done for 5,656 AGES-Reykjavik participants, as previously described (22). For the 5,368 participants with both genotyping and protein measurements available, linear regression analyses adjusted for age, sex, pack-years of smoking, and smoking at study entry were performed to evaluate the associations of previously reported pulmonary fibrosis–related SNPs (SNPs previously associated with either IPF or ILA [9]) with all 4,782 human SOMAmers.

Analyses of ILA Imaging Patterns

The 287 proteins that were significantly associated with ILA in AGES-Reykjavik were assessed by logistic regression models for associations with ILA imaging patterns (i.e., the definite fibrosis pattern and the UIP pattern). Logistic regression models of the associations of these proteins with each pattern were constructed. Comparisons were made with participants without ILA, excluding participants with other imaging patterns. For analyses involving the UIP pattern, participants with probable or definite UIP were regarded as having UIP, and participants with no or indeterminate UIP were not.

Analyses of ILA Progression

Included in analyses of ILA progression in AGES-Reykjavik were the 223 participants who had ILA progression at follow-up examination and the 1,425 who did not have ILA at either baseline or follow-up. Participants with definite or probable progression on follow-up examinations were compared with participants with no ILA in

Table 1. Overview of the Study Participants with Available Protein Measurements and Characterization of Interstitial Lung Abnormalities

	AGES-Reykjavik			COPDGene		
Participants	**No ILAs (*n* = 3,187)**	**Indeterminate for ILAs (*n* = 1,703)**	**ILAs (*n* = 329)**	**No ILAs (*n* = 2,484)**	**Indeterminate for ILAs (*n* = 1,891)**	**ILAs (*n* = 524)**
Age, mean (SD)	75.9 (5.4)	77.3 (5.6)	78.1 (5.6)	64.0 (8.3)	66.4 (8.7)	69.0 (8.8)
Women, *n* (%)	1,889 (59)	950 (56)	149 (45)	1,221 (49)	955 (50)	249 (48)
BMI, mean (SD)	27.2 (4.4)	26.8 (4.4)	27.2 (4.7)	29.0 (6.4)	28.8 (6.4)	29.6 (6.1)
History of smoking, *n* (%)	1,742 (55)	1,015 (60)	237 (72)	2,448 (99)	1,872 (99)	518 (99)
Pack-years, median (IQR)	0 (0–17)	2.9 (0–23)	11 (0–28)	38.0 (25.1–51.0)	41.4 (28.2–57.0)	43.9 (28.4–59.1)
Current smoker, *n* (%)	371 (12)	203 (12)	54 (16)	911 (37)	762 (40)	199 (38)
Definite fibrosis pattern, *n* (%)						
Without fibrosis	—	—	206 (63)	—	—	469 (90)
Definite fibrosis	—	—	123 (37)	—	—	55 (10)
UIP pattern, *n* (%)						
No UIP	—	—	104 (32)	—	—	129 (25)
Indeterminate for UIP	—	—	131 (40)	—	—	346 (66)
Probable UIP	—	—	77 (23)	—	—	46 (9)
Definite UIP	—	—	17 (5)	—	—	3 (0.6)

Definition of abbreviations: AGES-Reykjavik = Age, Gene/Environment Susceptibility–Reykjavik; BMI = body mass index; COPDGene = Genetic Epidemiology of COPD; ILA = interstitial lung abnormality; IQR = interquartile range; UIP = usual interstitial pneumonia.

both examinations (11). ILA progression was not assessed in COPDGene, as data collection is not yet complete.

The associations of single proteins with the progression of ILA were tested with logistic regression. As in analyses of ILA at baseline, the proteins suggestively associated with progression ($P < 0.05$) in single-protein logistic regression models (1,562 proteins) were used to create 200 bootstrap data samples. An adaptive LASSO regression model was created in each sample. A multivariate regression model was constructed, using proteins in all 200 LASSO models with the AUROC calculated and validated. Comparative analyses and further details are found in the online supplement.

Data Availability

The custom-design Novartis SOMAscan is available through a collaboration agreement with the Novartis Institutes for BioMedical Research (lori.jennings@novartis.com). Data from the AGES Reykjavik study are available through collaboration (AGES_data_request@hjarta.is) under a data usage agreement with the IHA in accordance with participants' informed consent.

Results

Participant characteristics at baseline in AGES-Reykjavik and in phase 2 of COPDGene are shown in Table 1. Representative ILA cases are shown in Figure E2.

Associations between Serum Proteins and ILA in AGES-Reykjavik

Of the 4,137 proteins tested, 287 were associated with ILA after Bonferroni adjustment ($P < 1.22 \times 10^{-5}$) (Figure 1A and Tables E1 and E2). The association of SFTPB was the most significant ($P = 4.28 \times 10^{-67}$) and was associated with the greatest odds increase of ILA (odds ratio [OR], 3.71 [95% confidence interval [CI], 3.20–4.30]). Among other significant associations of ILA at baseline were associations with SCGB3A1 (OR, 2.43 [95% CI, 2.13–2.77]; $P = 8.01 \times 10^{-40}$) and WFDC2 (OR, 2.42 [95% CI, 2.11–2.78]; $P = 4.01 \times 10^{-36}$). The distributions of these three proteins grouped by ILA status are shown in Figure E3. Proteins associated with ILA were enriched for lung-tissue specificity when compared with all proteins in the genome but not when compared with all proteins with available SOMAmer measurements in AGES-Reykjavik (Figure E4). Full results of analyses of indeterminate ILAs, tissue expression, and functional enrichment, and direct comparisons with previously published ILA biomarkers are shown in online supplement results, Figures E4 and E5, and Tables E2–E4.

Eight proteins featured in all 200 adaptive LASSO models of ILAs (Table E5). In a multivariate logistic regression model exploring the association of these proteins with ILA, WFDC2 (OR, 3.15 [95% CI, 2.55–3.89]; $P = 2.81 \times 10^{-26}$) and SFTPB (OR, 3.14 [95% CI, 2.66–3.71]; $P = 1.78 \times 10^{-41}$) had the largest effect on ILAs. SCGB3A1 (OR, 1.62 [95% CI, 1.37–1.91]; $P = 1.84 \times 10^{-8}$) and CBLN4 (Cerebellin 4 Precursor; OR, 1.26 [95% CI, 1.09–1.45]; $P = 0.0019$) were also positively associated with ILAs, whereas WFIKKN2 (WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 2; OR, 0.42 [95% CI, 0.35–0.51]; $P = 3.83 \times 10^{-19}$), ADAM Metallopeptidase Domain 9 (ADAM9) (OR, 0.59 [95% CI, 0.50–0.70]; $P = 6.93 \times 10^{-10}$), and Annexin A9 (ANXA9) (OR, 0.69 [95% CI, 0.59–0.81]; $P = 6.39 \times 10^{-6}$) were negatively associated (Figure 2A and Table E6). The validated AUROC of this model, based on 200-fold resampling, was 0.880, compared with 0.670 for a model with only age, sex, pack-years, and smoking at beginning of the study and 0.749, 0.760, and 0.826 for single-protein models with SCGB3A1, WFDC2, and SFTPB, respectively, added to these demographic factors (Figure 3A). The variance inflation factor was less than two for all components of the eight-protein multivariate model. Results of comparative models based on proteins used in a lower number of LASSO models are seen in Figure E6.

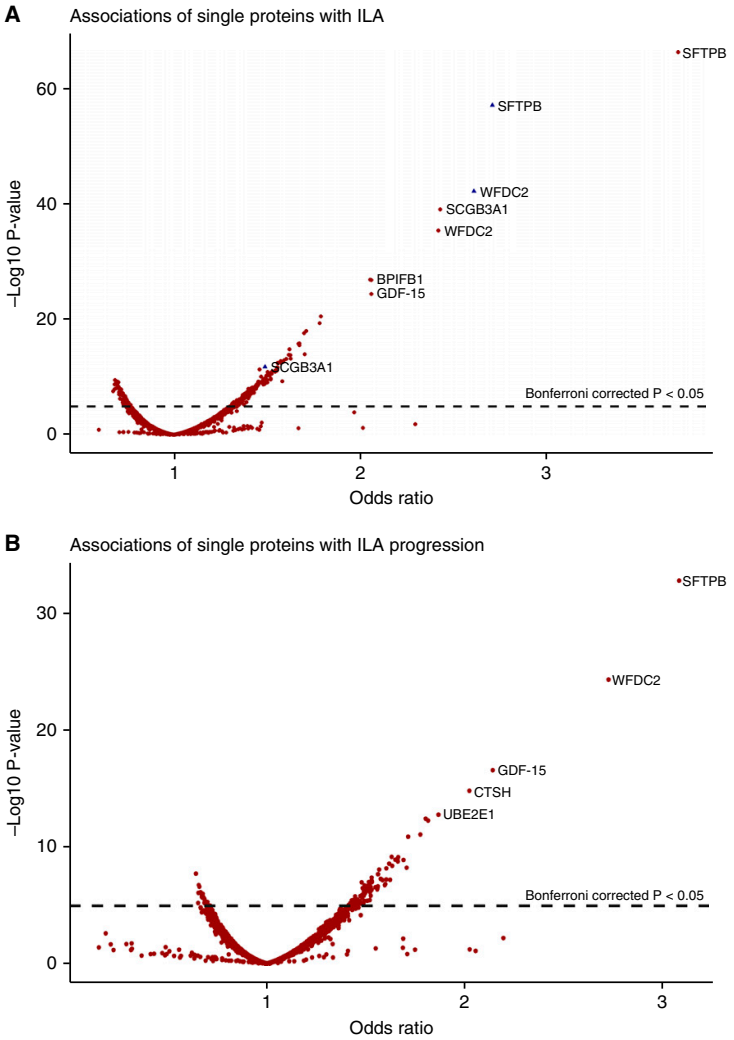

Figure 1. The associations of single proteins with interstitial lung abnormalities (ILAs) at baseline and progression of ILAs. Shown in red circles are results from the AGES-Reykjavik (Age, Gene/Environment Susceptibility–Reykjavik) study. Results from COPDGene (Genetic Epidemiology of COPD) are shown in blue triangles. Models are (*A*) logistic regression models of a single protein with ILAs at baseline, and (*B*) progression of ILAs, adjusted for age, sex, pack-years, and smoking at study entry. Models in COPDGene are additionally adjusted for white blood cell count, platelet count, and study center. BPIFB1 = BPI fold containing family B member 1; CTSH = cathepsin H; GDF-15 = growth differentiation factor 15; SCGB3A1 = secretoglobin family 3A member 1; SFTPB = surfactant protein B; UBE2E1 = ubiquitin conjugating enzyme E2 E1; WFDC2 = WAP four-disulfide core domain protein 2.

Replication Analyses of ILAs in COPDGene

In models of single proteins, SFTPB (OR, 2.70 [95% CI, 2.39–3.05]; $P = 1.13 \times 10^{-57}$), WFDC2 (OR, 2.61 [95% CI, 2.27–2.99]; $P = 5.89 \times 10^{-43}$), and SCGB3A1 (OR, 1.49 [95% CI, 1.33–1.67]; $P = 1.77 \times 10^{-12}$) were all associated with ILA in the COPDGene cohort (Figure 1A and Table E7). The validated AUROCs of these models were 0.787, 0.763, and 0.709, respectively, compared with 0.692 for demographic factors only (Figure 3B). In a multivariate model with the eight proteins selected using LASSO modeling, SFTPB (OR, 2.33 [95% CI 2.04–2.66]; $P = 9.33 \times 10^{-36}$), WFDC2 (OR, 2.82 [95% CI, 2.39–3.33]; $P = 2.33 \times 10^{-34}$), WFIKKN2 (OR, 0.50 [95% CI, 0.43–0.57]; $P = 1.64 \times 10^{-22}$), and CBLN4 (OR, 1.15 [95% CI, 1.02–1.31]; $P = 0.03$) were associated with ILAs (Figure 2B and Table E7), and the validated AUROC of this model was 0.826 (Figure 3B).

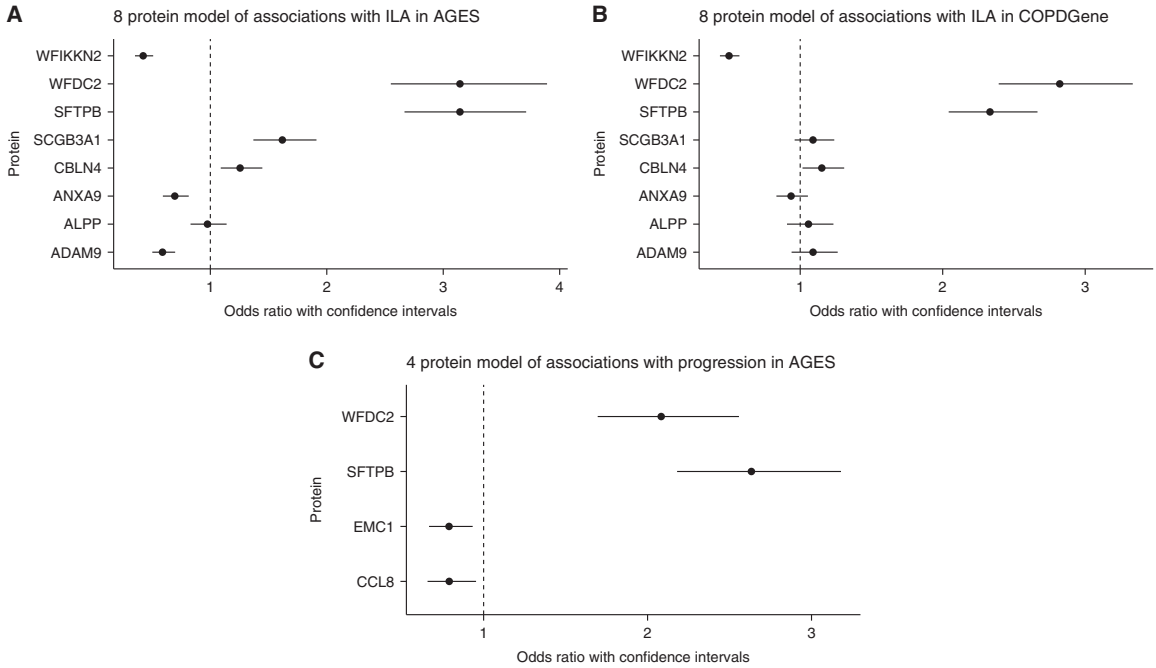

Figure 2. Multivariate logistic regression models of the association of proteins with interstitial lung abnormalities (ILAs) and ILA progression. (*A*) A model of the associations of the eight proteins selected for 200 adaptive Least Absolute Shrinkage and Selection Operator (LASSO) models with ILAs in the AGES-Reykjavik (Age, Gene/Environment Susceptibility–Reykjavik) cohort. (*B*) A model of the associations of the eight proteins selected for 200 adaptive LASSO models with ILA in the COPDGene (Genetic Epidemiology of COPD) cohort. (*C*) A model of the associations of the four proteins selected for 200 adaptive LASSO models with progression of ILA in the AGES-Reykjavik cohort. Models in AGES-Reykjavik are logistic regression models, adjusted for age, sex, pack-years, and smoking at study entry. The model in COPDGene is additionally adjusted for white blood cell count, platelet count, and study center. ADAM9 = ADAM metallopeptidase domain 9; ALPP = alkaline phosphatase, placental; ANXA9 = annexin A9; CBLN4 = cerebellin 4 precursor; CCL8 = C-C motif chemokine ligand 8; EMC1 = ER membrane protein complex subunit 1; SCGB3A1 = secretoglobin family 3A member 1; SFTPB = surfactant protein B; WFDC2 = WAP four-disulfide core domain protein 2; WFIKKN2 = WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 2.

Associations with Pulmonary Fibrosis–associated SNPs in AGES-Reykjavik

Results of linear regression analyses modeling the relationship of pulmonary fibrosis–related SNPs with human proteins are shown in Table 2, for proteins with associations reaching genome-wide significance ($P < 5 \times 10^{-8}$). Also shown are the associations of these proteins with ILA and the previously reported associations of the SNPs with ILAs (9). Four proteins were found to be associated with these SNPs, and three of these were themselves associated with ILAs. These SNPs were the rs35705950 *MUC5B* promoter polymorphism associated with SFTPB, the rs2736100 *TERT* (Telomerase reverse transcriptase) polymorphism associated with thrombopoietin, and the rs4727443 polymorphism associated with Paired Immunoglobin Like Type 2 Receptor Alpha (PILRA). Of those, only the *MUC5B* promoter polymorphism was significantly associated with ILAs in the cohort. This association was consistent when conducted with stratification based on ILA status (Table E8). The associations of the *MUC5B* promoter polymorphism with proteins are shown graphically in Figure E7 and numerically in Table E9.

Associations with ILA Imaging Patterns

Results of models of the associations of ILA-associated proteins with imaging patterns are shown in Tables E10 and E11. The 20 proteins most significantly associated with ILAs had significant associations regardless of the presence or absence of specific imaging patterns. All twenty proteins were associated with greater odds of ILA with definite fibrosis than ILA without definite fibrosis. Similar results were seen for the UIP pattern; all proteins were associated with ILAs with and without UIP, but ORs were higher for ILA with UIP (Table E10).

Analyses of ILA Progression in AGES-Reykjavik

In single-protein models of ILA progression, 121 proteins were significant after Bonferroni correction ($P < 1.22 \times 10^{-5}$) (Figure 1B and Tables E12 and E13). The protein associated with the greatest odds of ILA progression was SFTPB (OR, 3.08 [95% CI, 2.56–3.69]; $P = 1.59 \times 10^{-39}$). Other strongly associated proteins were WFDC2 (OR, 2.72 [95% CI, 2.25–3.29]; $P = 4.90 \times 10^{-25}$), GDF-15 (Growth Differentiation Factor 15; OR, 2.14 [95% CI, 1.79–2.55]; $P = 3.01 \times 10^{-17}$), and CTSH (Cathepsin H; OR, 2.02 [95% CI, 1.70–2.40]; $P = 1.72 \times 10^{-15}$).

The proteins used in all 200 adaptive LASSO models of ILA progression (Table E14) were all associated with ILA

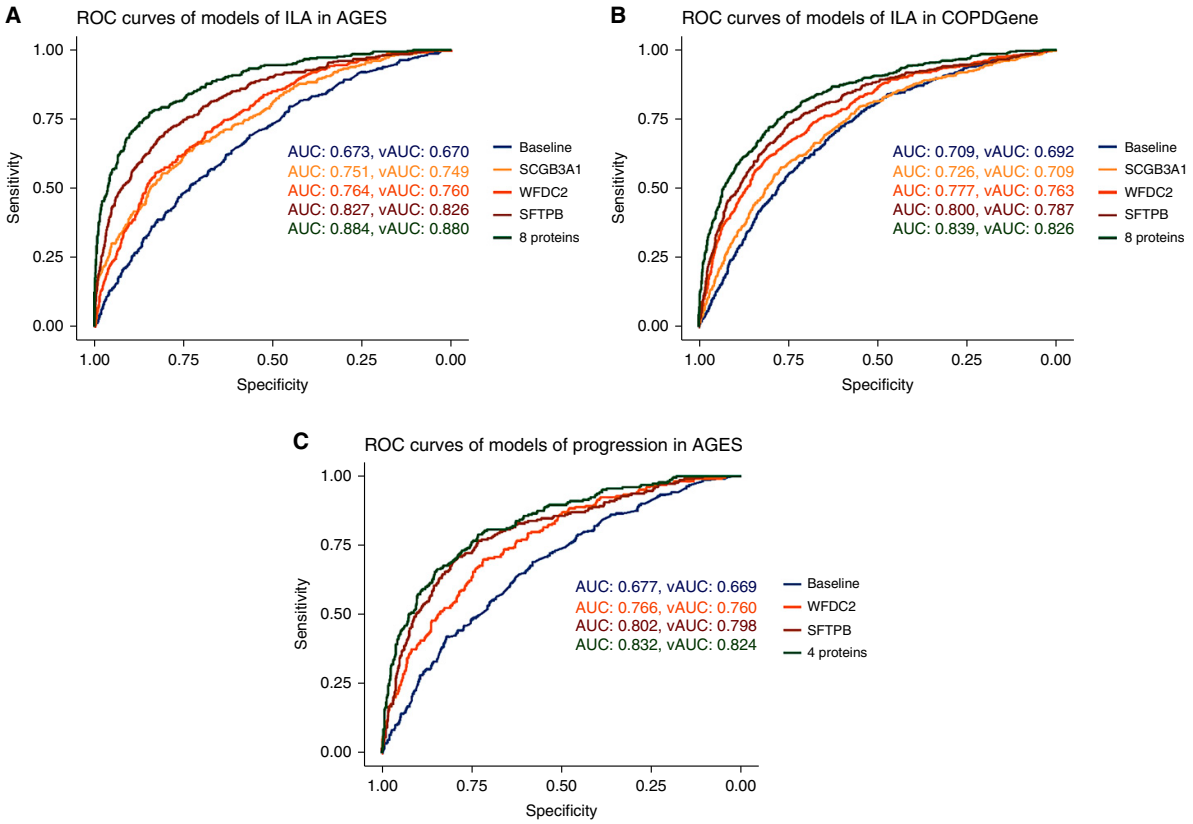

Figure 3. Receiver operating characteristic (ROC) curves of models of the associations of proteins with interstitial lung abnormalities (ILAs) at baseline and ILA progression. ROC curves for the specified logistic regression models. (*A*) Curves for models in the AGES-Reykjavik (Age, Gene/Environment Susceptibility–Reykjavik) cohort with ILAs at baseline as the outcome. (*B*) Curves for models in the COPDGene (Genetic Epidemiology of COPD) cohort with ILAs at baseline as the outcome. (*C*) Curves for models in the AGES-Reykjavik cohort with progression of ILAs as the outcome. Baseline: A model with age, sex, pack-years, and smoking at study entry, for which all other models are adjusted. Names of proteins refer to single-protein models of that protein with ILAs at baseline or at progression. Eight proteins (*A* and *B*): A model with proteins used in 200 adaptive Least Absolute Shrinkage and Selection Operator (LASSO) models of ILAs at baseline. Four proteins (*C*): A model with proteins used in 200 adaptive LASSO models of ILA progression. AUC = area under curve; SCGB3A1 = secretoglobin family 3A member 1; SFTPB = surfactant protein B; WFDC2 = WAP four-disulfide core domain protein 2; vAUC = area under curve, validated with 200-fold resampling.

progression in multiprotein models. The most significant associations were for SFTPB (OR, 2.63 [95% CI, 2.17–3.18]; $P = 1.96 \times 10^{-23}$) and WFDC2 (OR, 2.08 [95% CI, 1.69–2.55]; $P = 3.05 \times 10^{-12}$), whereas ER Membrane Protein Complex Subunit 1 (EMC1) (OR, 0.79 [95% CI, 0.67–0.94]; $P = 0.0063$) and C-C Motif Chemokine Ligand 8 (CCL8) (OR, 0.79 [95% CI, 0.66–0.95]; $P = 0.015$) were also associated with progression (Figure 2C and Table E15). The validated AUROC of this model was 0.824, compared with 0.669 for a model with only demographic factors and 0.760 and 0.798 for single-protein models with WFDC2 and SFTPB (Figure 3C).

Discussion

This study represents the first proteomic assessment of ILAs and is the largest blood proteomic assessment of ILD or pulmonary fibrosis to date (5, 23). In addition to a comprehensive assessment of thousands of proteins, uncovering hundreds of associations with ILAs and ILA progression, these results provide replicable evidence for the association between protein measures, alone (e.g., SFTPB, WFDC2, and SCGB3A1) or in machine-learning–based models, and ILAs across independent populations. The magnitudes of the associations of single proteins are strong compared with known risk factors (6, 8), and multiprotein models demonstrate replicable ability to improve risk prediction of ILA and its progression over demographic risk factors. The proteins most strongly associated with ILA consistently had larger associations with imaging patterns that are correlated with unfavorable outcomes, such as progression and mortality (11).

These findings greatly expand on the small but growing number of independent studies demonstrating that peripheral blood protein measures can help detect ILAs (24–27). These proteins were all suggestively associated with ILAs based on an unadjusted *P* value, and most were associated with ILAs and/or ILA progression after multiple testing

Table 2. Associations between Previously Identified Pulmonary Fibrosis–associated Genetic Loci, Single-Protein Measurements, and Interstitial Lung Abnormalities

SNPs and ILAs (Previously Reported)					SNPs and Proteins			Proteins and ILAs	
rsID	Chromosomal Location	Nearest Gene	OR	*P* Value	Protein	β (SE)	*P* Value	OR	*P* Value
rs73199442	3q13	*FCF1P3*	1.68 (1.39–2.02)	5×10^{-8}	–	–	–	–	–
rs6886640	5q12	*IPO11*	1.28 (1.18–1.41)	4×10^{-8}	–	–	–	–	–
rs7744971	6q15	*HTR1E*	1.26 (1.16–1.37)	1×10^{-7}	–	–	–	–	–
rs35705950	11p15	*MUC5B*	1.97 (1.74–2.22)	3×10^{-27}	SFTPB	0.26 (0.030)	8×10^{-18}	3.71	4.28×10^{-67}
rs2609255	4q22	*FAM13A*	1.18 (1.07–1.29)	5×10^{-4}	–	–	–	–	–
rs2076295	6p24	*DSP*	1.14 (1.05–1.2)	0.001	–	–	–	–	–
rs2034650	15q15	*IVD*	1.08 (0.99–1.17)	0.07	–	–	–	–	–
rs12610495	19p13	*DPP9*	1.14 (1.03–1.26)	0.01	N/A	N/A	N/A	N/A	N/A
rs6793295	3q26	*LRRC34*	1.06 (0.97–1.15)	0.2	N/A	N/A	N/A	N/A	N/A
rs1981997	17q21	*MAPT*	1.16 (1.03–1.30)	0.01	–	–	–	–	–
rs2736100	5p15	*TERT*	1.03 (0.95–1.12)	0.44	THPO	0.11 (0.019)	4×10^{-9}	0.77*	1.39×10^{-4}
								1.19*	3.93×10^{-3}
rs11191865	10q24	*OBFC1*	1.03 (0.95–1.12)	0.46	–	–	–	–	–
rs1278769	13q34	*ATP11A*	1.04 (0.95–1.15)	0.37	F7	−0.13 (0.020)	6×10^{-11}	0.91	0.17
rs62025270	15q25	*AKAP13*	1.09 (0.99–1.20)	0.08	–	–	–	–	–
rs4727443	7q22	*LOC100128334/LOC105375423*	0.95 (0.87–1.03)	0.19	PILRA	−0.50 (0.019)	2×10^{-151}	0.97*	0.66
								0.97*	0.56
								1.30*	1.20×10^{-5}
								1.21*	1.69×10^{-3}
								1.07*	0.26

Definition of abbreviations: AGES-Reykjavik = Age, Gene/Environment Susceptibility–Reykjavik; CI = confidence interval; ILA = interstitial lung abnormality; OR = odds ratio; PILRA = Paired Immunoglobin Like Type 2 Receptor Alpha; SFTPB = surfactant protein B; SOMAmers = Slow-Off rate Modified Aptamers; THPO = thrombopoietin.

SNPs and ILAs: the associations between the listed SNPs and ILAs as previously reported (8). SNPs and proteins: the associations between the listed SNPs and proteins in the AGES-Reykjavik cohort, adjusted for age, sex, pack-years of smoking, and smoking at study entry. Proteins with an association with a P value $< 5 \times 10^{-8}$ are shown. Proteins and ILAs: the associations between shown proteins and ILAs, calculated with logistic regression adjusted for age, sex, pack-years of smoking, and smoking at study entry. N/A indicates data for SNP not available in AGES-Reykjavik.

*Data shown for multiple SOMAmers binding to the same protein.

(Tables E1, E2, E4, E12, and E13). Although magnitudes of associations are hard to directly compare owing to differences in the units in which they are provided, the directionality of associations was consistent for all previously published protein biomarkers of ILA (Table E4). In addition to providing an independent replication of these prior findings, this reproducibility across different platforms provides some assurance that other associations are not platform specific, as protein measurements for these previously published studies were mainly done using ELISA (24–27).

Although some of the proteins most significantly associated with ILAs in the present findings have previously been associated with IPF (1), many associations are novel. SFTPB is a small hydrophobic protein that is an essential component of the regulation and function of pulmonary surfactant (28). Its concentration in plasma is normally low, and its elevation is believed to represent a breakdown of the alveolar–capillary membrane, which likely contributes to its elevation in other conditions (29, 30). Rare genetic variants in multiple surfactant proteins (31) have been implicated in IPF. Common variants in SFTPB have been previously reported in association with IPF (32), although these findings have not been confirmed in larger studies (33), and elevated concentrations have been reported in patients with pulmonary fibrosis in some (34), but not all (35), studies. This reflects the protein's different isoforms and the challenges of reliably measuring it in plasma (29). Although WFDC2 (also referred to as HE-4 [human epididymis protein 4]) is highly expressed in epithelial cells and submucosal glands in the human respiratory tract, its function remains incompletely characterized (36). Elevation of WFDC2 in plasma has been associated with IPF (37). SCGB3A1 (also referred to as UGRP2 [uteroglobin-related protein 2] or HIN-1 [high in normal 1]) is a tumor-suppressor gene known to be secreted throughout the conducting airways (38, 39). To the best of our knowledge, it has not been previously studied in IPF. WFIKKN2 has a replicable negative association with ILAs. This protein is an antagonist of GDF-8 and GDF-11, implicated in various aging-related processes and diseases, and interacts with other members of the Transforming Growth Factor-β (TGF-β) superfamily (40, 41). The TGF-β family of proteins, a member of which is the previously mentioned GDF-15, is central to the tissue-injury response and plays a pivotal role in the pathogenesis of IPF as it is currently understood (42). The proposed role of cytokine signaling in ILD pathogenesis extends well beyond that of TGF-β, possibly explaining the enrichment of gene ontology terms related to such signaling among ILA-associated proteins (42).

The potential of SFTPB as a biomarker of ILAs and ILA progression is supported by the association of the rs35705950 promoter polymorphism of *MUC5B*, the best-established genetic risk factor for pulmonary fibrosis (33). The mechanism of this

association is unknown but of great interest. Although it is possible that this finding is a result of two separate strong associations with ILAs, no other proteins had associations with *MUC5B* that reached genome-wide significance. Therefore, it is conceivable that this association represents an unknown biological mechanism, considering that the distal airways are a major site of MUC5B expression, that MUC5B is known to be coexpressed with other surfactant proteins in distal airways (43, 44), and the suggestion that an appropriate ratio of mucins relative to surfactant may be necessary for normal physiology of certain airway zones (44). Another notable finding was the association of *TERT*, an established genetic risk factor of IPF (42), with THPO (thrombopoietin), concentrations of which were associated with ILAs. Although data on the role of thrombopoietin in pulmonary fibrosis are scarce, an increase in its amount in lungs of mice with medication-induced lung fibrosis has been documented (45). Still, the meaning of this finding is unclear, because *TERT* is not associated with ILAs. This could be because THPO is more strongly associated with ILAs than *TERT* or because the association of THPO with ILAs is independent of its association with *TERT*.

This study has several important limitations. First, although this is the largest proteomic analysis of ILAs, and any form of ILD, to date, it is possible that larger sample sizes of those with ILAs will be needed to uncover additional important proteins and pathways associated with the early stages of pulmonary fibrosis. Second, although we demonstrate replicable evidence for associations of proteins and multiprotein models with ILAs across independent populations, it is likely that associations of some of these proteins are not specific to ILAs alone. Third, not all the proteins selected for replication from the AGES-Reykjavik cohort were validated in COPDGene. It is possible that demographic differences between these two cohorts (e.g., age, smoking history, and racial background) could explain some of these discrepancies. Fourth, the validation of results is based on SOMAmer technology, a novel measurement platform. The specificity of this method compared with other methods has been extensively studied, and a minority of proteins have potential cross-reactivity with protein isoforms or related proteins with high amino acid homology (46). Although validation data from recent studies based on both mass spectrometry and antibody-based technology, shown for the proteins highlighted in the results (Table E16), were reassuring, not all proteins had such data available (18, 47). Therefore, as the SOMAmer technology is novel and the validation of SOMAmer specificity is ongoing, future replication of the findings with standard platforms would be useful. Fifth, the preferred method of selecting multiprotein models for biomarker discovery has not been established in proteomics studies to date. The method of machine learning by adaptive LASSO regression of bootstrap data samples provides sets of proteins that jointly predict ILAs. This selection method is based on statistical correlations but not biological mechanisms, possibly explaining why not all proteins selected had significant associations when externally validated. The decision to use only proteins that featured in all 200 adaptive LASSO models can be questioned. However, as shown in Figure E6, the gain in prediction by using more inclusive models is slim. Finally, effects of residual confounding (e.g., with chronic diseases that could coincide with ILAs) cannot be excluded for the multitude of associations presented.

In this first proteomic assessment of ILAs and ILA progression, replicable associations of several proteins, notably SFTPB, WFDC2, SCGB3A1, and WFIKKN2, are presented. In conjunction with prior studies of IPF, these findings demonstrate the utility of proteomics in generating reproducible models of ILA that may help to detect patients at risk for pulmonary fibrosis. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank other investigators, staff, and particularly the participants of the AGES-Reykjavik and the COPDGene studies.

References

1. Guiot J, Moermans C, Henket M, Corhay J-L, Louis R. Blood biomarkers in idiopathic pulmonary fibrosis. *Lung* 2017;195:273–280.
2. Mathai SK, Cardwell J, Metzger F, Powers J, Walts AD, Kropski JA, *et al.* Preclinical pulmonary fibrosis circulating protein biomarkers. *Am J Respir Crit Care Med* 2020;202:1720–1724.
3. O'Dwyer DN, Norman KC, Xia M, Huang Y, Gurczynski SJ, Ashley SL, *et al.* The peripheral blood proteome signature of idiopathic pulmonary fibrosis is distinct from normal and is associated with novel immunological processes. *Sci Rep* 2017;7:46560.
4. Niu R, Liu Y, Zhang Y, Zhang Y, Wang H, Wang Y, *et al.* iTRAQ-based proteomics reveals novel biomarkers for idiopathic pulmonary fibrosis. *PLoS One* 2017;12:e0170741.
5. Todd JL, Neely ML, Overton R, Durham K, Gulati M, Huang H, *et al.*; IPF-PRO Registry investigators. Peripheral blood proteomic profiling of idiopathic pulmonary fibrosis biomarkers in the multicentre IPF-PRO Registry. *Respir Res* 2019;20:227.
6. Washko GR, Hunninghake GM, Fernandez IE, Nishino M, Okajima Y, Yamashiro T, *et al.*; COPDGene Investigators. Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011;364:897–906.
7. Hunninghake GM. Interstitial lung abnormalities: erecting fences in the path towards advanced pulmonary fibrosis. *Thorax* 2019:74:506–511.
8. Hunninghake GM, Hatabu H, Okajima Y, Gao W, Dupuis J, Latourelle JC, *et al.* MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013;368:2192–2200.
9. Hobbs BD, Putman RK, Araki T, Nishino M, Gudmundsson G, Gudnason V, *et al.* Overlap of genetic risk between interstitial lung abnormalities and idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;200: 1402–1413.
10. Miller ER, Putman RK, Vivero M, Hung Y, Araki T, Nishino M, *et al.* Histopathology of interstitial lung abnormalities in the context of lung nodule resections. *Am J Respir Crit Care Med* 2018;197:955–958.
11. Putman RK, Gudmundsson G, Axelsson GT, Hida T, Honda O, Araki T, *et al.* Imaging patterns are associated with interstitial lung abnormality progression and mortality. *Am J Respir Crit Care Med* 2019;200:175–183.
12. Putman RK, Hatabu H, Araki T, Gudmundsson G, Gao W, Nishino M, *et al.*; Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators; COPDGene Investigators. Association between interstitial lung abnormalities and all-cause mortality. *JAMA* 2016;315:672–681.
13. Putman RK, Gudmundsson G, Araki T, Nishino M, Sigurdsson S, Gudmundsson EF, *et al.* The *MUC5B* promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *Eur Respir J* 2017;50:1700537.
14. Hatabu H, Hunninghake GM, Richeldi L, Brown KK, Wells AU, Remy-Jardin M, *et al.* Interstitial lung abnormalities detected incidentally on

CT: a position paper from the Fleischner Society. *Lancet Respir Med* 2020;8:726–737.
15. Emilsson V, Gudnason V, Jennings LL. Predicting health and life span with the deep plasma proteome. *Nat Med* 2019;25:1815–1816.
16. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* 2007;165: 1076–1087.
17. Wan ES, Fortis S, Regan EA, Hokanson J, Han MK, Casaburi R, *et al.*; COPDGene Investigators. Longitudinal phenotypes and mortality in preserved ratio impaired spirometry in the COPDGene study. *Am J Respir Crit Care Med* 2018;198:1397–1405.
18. Emilsson V, Ilkov M, Lamb JR, Finkel N, Gudmundsson EF, Pitts R, *et al.* Co-regulatory networks of human serum proteins link genetics to disease. *Science* 2018;361:769–773.
19. Yeo IK, Johnson RA. A new family of power transformations to improve normality or symmetry. *Biometrika* 2000;87:954–959.
20. Zou H. The adaptive lasso and its oracle properties. *J Am Stat Assoc* 2006;101:1418–1429.
21. Efron B, Tibshirani RJ. An introduction to the bootstrap. New York: Chapman & Hall; 1993.
22. Gudjonsson A, Gudmundsdottir V, Axelsson GT, Gudmundsson EF, Jonsson BG, Launer LJ, *et al.* A genome-wide association study of serum proteins reveals shared loci with common diseases. *Nat Commun* 2022;13:480.
23. Alqalyoobi S, Adegunsoye A, Linderholm A, Hrusch C, Cutting C, Ma S-F, *et al.* Circulating plasma biomarkers of progressive interstitial lung disease. *Am J Respir Crit Care Med* 2020;201:250–253.
24. Ho JE, Gao W, Levy D, Santhanakrishnan R, Araki T, Rosas IO, *et al.* Galectin-3 is associated with restrictive lung disease and interstitial lung abnormalities. *Am J Respir Crit Care Med* 2016; 194:77–83.
25. Armstrong HF, Podolanczuk AJ, Barr RG, Oelsner EC, Kawut SM, Hoffman EA, *et al.*; MESA (Multi-Ethnic Study of Atherosclerosis). Serum matrix metalloproteinase-7, respiratory symptoms, and mortality in community-dwelling adults. *Am J Respir Crit Care Med* 2017;196: 1311–1317.
26. McGroder CF, Aaron CP, Bielinski SJ, Kawut SM, Tracy RP, Raghu G, *et al.* Circulating adhesion molecules and subclinical interstitial lung disease: the Multi-Ethnic Study of Atherosclerosis. *Eur Respir J* 2019; 54:19900295.
27. Sanders JL, Putman RK, Dupuis J, Xu H, Murabito JM, Araki T, *et al.* The association of aging biomarkers, interstitial lung abnormalities, and mortality. *Am J Respir Crit Care Med* 2021; 203:1149–1157.
28. Martínez-Calle M, Olmeda B, Dietl P, Frick M, Pérez-Gil J. Pulmonary surfactant protein SP-B promotes exocytosis of lamellar bodies in alveolar type II cells. *FASEB J* 2018;32:4600–4611.
29. Leung JM, Mayo J, Tan W, Tammemagi CM, Liu G, Peacock S, *et al.* Plasma pro-surfactant protein B and lung function decline in smokers. *Eur Respir J* 2015;45:1037–1045.
30. Sin DD, Tammemagi CM, Lam S, Barnett MJ, Duan X, Tam A, *et al.* Pro-surfactant protein B as a biomarker for lung cancer prediction. *J Clin Oncol* 2013;31:4536–4543.
31. Ley B, Brown KK, Collard HR. Molecular biomarkers in idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2014;307: L681–L691.
32. Selman M, Lin H-M, Montaño M, Jenkins AL, Estrada A, Lin Z, *et al.* Surfactant protein A and B genetic variants predispose to idiopathic pulmonary fibrosis. *Hum Genet* 2003;113:542–550.
33. Allen RJ, Guillen-Guio B, Oldham JM, Ma SF, Dressen A, Paynton ML, *et al.* Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2020;201:564–574.
34. Kahn N, Rossler A-K, Hornemann K, Muley T, Grünig E, Schmidt W, *et al.* C-proSP-B: a possible biomarker for pulmonary diseases? *Respiration* 2018;96:117–126.
35. Papaioannou AI, Kostikas K, Manali ED, Papadaki G, Roussou A, Spathis A, *et al.* Serum levels of surfactant proteins in patients with combined pulmonary fibrosis and emphysema (CPFE). *PLoS One* 2016;11:e0157789.
36. Bingle L, Armes H, Bingle C. Expression and function of murine WFDC2 in the respiratory tract. *Eur Respir J* 2020;56:2775.
37. Raghu G, Richeldi L, Jagerschmidt A, Martin V, Subramaniam A, Ozoux M-L, *et al.* Idiopathic pulmonary fibrosis: prospective, case-controlled study of natural history and circulating biomarkers. *Chest* 2018;154: 1359–1370.
38. Reynolds SD, Reynolds PR, Pryhuber GS, Finder JD, Stripp BR. Secretoglobins SCGB3A1 and SCGB3A2 define secretory cell subsets in mouse and human airways. *Am J Respir Crit Care Med* 2002;166: 1498–1509.
39. Mazumdar J, Hickey MM, Pant DK, Durham AC, Sweet-Cordero A, Vachani A, *et al.* HIF-2alpha deletion promotes Kras-driven lung tumor development. *Proc Natl Acad Sci USA* 2010;107:14182–14187.
40. Szláma G, Kondás K, Trexler M, Patthy L. WFIKKN1 and WFIKKN2 bind growth factors TGFβ1, BMP2 and BMP4 but do not inhibit their signalling activity. *FEBS J* 2010;277:5040–5050.
41. Frohlich J, Vinciguerra M. Candidate rejuvenating factor GDF11 and tissue fibrosis: friend or foe? *Geroscience* 2020;42:1475–1498.
42. Wolters PJ, Collard HR, Jones KD. Pathogenesis of idiopathic pulmonary fibrosis. *Annu Rev Pathol* 2014;9:157–179.
43. Hancock LA, Hennessy CE, Solomon GM, Dobrinskikh E, Estrella A, Hara N, *et al.* Muc5b overexpression causes mucociliary dysfunction and enhances lung fibrosis in mice. *Nat Commun* 2018;9:5363.
44. Okuda K, Chen G, Subramani DB, Wolf M, Gilmore RC, Kato T, *et al.* Localization of secretory mucins MUC5AC and MUC5B in normal/ healthy human airways. *Am J Respir Crit Care Med* 2019;199: 715–727.
45. Zhou Y, Zhang B, Li C, Huang X, Cheng H, Bao X, *et al.* Megakaryocytes participate in the occurrence of bleomycin-induced pulmonary fibrosis. *Cell Death Dis* 2019;10:648.
46. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, *et al.* Genomic atlas of the human plasma proteome. *Nature* 2018;558: 73–79.
47. Pietzner M, Wheeler E, Carrasco-Zanini J, Kerrison ND, Oerton E, Koprulu M, *et al.* Synergistic insights into human health from aptamer- and antibody-based proteomic profiling. *Nat Commun* 2021; 12:6822.

Paper IV

Interstitial lung abnormalities are associated with decreased mean telomere length

Rachel K. Putman[1], Gisli Thor Axelsson[2,3], Samuel Y. Ash[1], Jason L. Sanders[1], Aravind A. Menon[1], Tetsuro Araki[4,5], Mizuki Nishino[4,5], Masahiro Yanagawa[6], Elías F. Gudmundsson[3], Dandi Qiao[7], Raúl San José Estépar[4], Josée Dupuis[8], George T. O'Connor[9], Ivan O. Rosas[10], George R. Washko[1], Souheil El-Chemaly[1], Benjamin A. Raby[1,11], Vilmundur Gudnason[3], Dawn L. DeMeo[7], Edwin K. Silverman[7], Hiroto Hatabu[4,5], Immaculata DeVivo[7,12], Michael H. Cho[1,7], Gunnar Gudmundsson[3,13] and Gary M. Hunninghake[1]

[1]Pulmonary and Critical Care Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [2]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [3]Icelandic Heart Association, Kopavogur, Iceland. [4]Dept of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [5]Center for Pulmonary Functional Imaging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [6]Dept of Radiology, Graduate School of Medicine, Osaka University, Osaka, Japan. [7]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA. [8]Dept of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [9]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA. [10]Pulmonary and Critical Care Division, Baylor University Medical Center, Houston, TX, USA. [11]Division of Pulmonary and Respiratory Diseases, Boston Children's Hospital, Boston, MA, USA. [12]Harvard T.H. Chan School of Public Health, Boston, MA, USA. [13]Dept of Respiratory Medicine, Landspitali – The National University Hospital of Iceland, Reykjavik, Iceland.

Corresponding author: Rachel K. Putman (rputman@bwh.harvard.edu)

Shareable abstract (@ERSpublications)
Interstitial lung abnormalities are associated with decreased telomere length https://bit.ly/3qo5pV3

Cite this article as: Putman RK, Axelsson GT, Ash SY, *et al*. Interstitial lung abnormalities are associated with decreased mean telomere length. *Eur Respir J* 2022; 60: 2101814 [DOI: 10.1183/13993003.01814-2021].

Received: 27 June 2021
Accepted: 29 Dec 2021

Abstract

Background Interstitial lung abnormalities (ILA) share many features with idiopathic pulmonary fibrosis; however, it is not known if ILA are associated with decreased mean telomere length (MTL).

Methods Telomere length was measured with quantitative PCR in the Genetic Epidemiology of Chronic Obstructive Pulmonary Disease (COPDGene) and Age Gene/Environment Susceptibility Reykjavik (AGES-Reykjavik) cohorts and Southern blot analysis was used in the Framingham Heart Study (FHS). Logistic and linear regression were used to assess the association between ILA and MTL; Cox proportional hazards models were used to assess the association between MTL and mortality.

Results In all three cohorts, ILA were associated with decreased MTL. In the COPDGene and AGES-Reykjavik cohorts, after adjustment there was greater than twofold increase in the odds of ILA when comparing the shortest quartile of telomere length to the longest quartile (OR 2.2, 95% CI 1.5–3.4, p=0.0001, and OR 2.6, 95% CI 1.4–4.9, p=0.003, respectively). In the FHS, those with ILA had shorter telomeres than those without ILA (−767 bp, 95% CI 76–1584 bp, p=0.03). Although decreased MTL was associated with chronic obstructive pulmonary disease (OR 1.3, 95% CI 1.1–1.6, p=0.01) in COPDGene, the effect estimate was less than that noted with ILA. There was no consistent association between MTL and risk of death when comparing the shortest quartile of telomere length in COPDGene and AGES-Reykjavik (HR 0.82, 95% CI 0.4–1.7, p=0.6, and HR 1.2, 95% CI 0.6–2.2, p=0.5, respectively).

Conclusion ILA are associated with decreased MTL.

Introduction

There is growing evidence of the similarities between specific patterns of interstitial lung abnormalities (ILA) and pulmonary fibrosis (PF). These similarities include decrements in diffusion capacity for carbon monoxide and total lung capacity [1, 2], decreased exercise capacity [3, 4], imaging progression [5–7], decline in forced vital capacity [5] and an increased risk of death [8]. There is also evidence that some forms of ILA and idiopathic pulmonary fibrosis (IPF) have overlapping genetic risk loci [9, 10], including

increased prevalence of the rs35705950 *MUC5B* promoter polymorphism [1, 9–15]. While IPF has been associated with reduced telomere length [16, 17], and reduced telomere length has recently been associated with early stages of PF in high-risk relatives of patients with PF [13, 14], less is known about the associations between telomere length and ILA at the population level.

We hypothesised that research participants with ILA would have relative reductions in their measures of telomere length. To test this hypothesis, we evaluated the associations between ILA and measures of relative telomere length in the Genetic Epidemiology of Chronic Obstructive Pulmonary Disease (COPDGene) and Age Gene/Environment Susceptibility Reykjavik (AGES-Reykjavik) cohorts measured by quantitative PCR, and telomere length measured by Southern blot in the Framingham Heart Study (FHS). Based on these results we evaluated the associations between telomere length and 1) ILA subtypes in COPDGene and AGES-Reykjavik and 2) quantitative assessments of interstitial features in COPDGene. To provide context on the effect estimates for these associations we provide a comparison of the associations between measures of telomere length and chronic obstructive pulmonary disease (COPD) in COPDGene. Finally, based on prior associations between reduced telomere length and mortality among IPF patients [17], we evaluated the associations between measures of telomere length and mortality among those with ILA.

Methods

Study populations

Protocols for participant enrolment in COPDGene, AGES-Reykjavik and the FHS have been previously reported [18, 19]. Briefly, COPDGene is a multicentre, longitudinal study of smokers that was designed to determine the epidemiological and genetic risk factors for COPD [19]. The AGES-Reykjavik study is a cohort derived from the Reykjavik Study, which was established in 1967 and includes men and women who were born in Reykjavik, Iceland, from 1907 to 1935 who are followed by the Icelandic Heart Association [18]. The FHS is a longitudinal study that began in 1948 and was originally designed to identify epidemiological risk factors for cardiovascular disease; the current investigation includes the offspring cohort [20], and participants were included in this analysis if they had chest computed tomography (CT) scans from the FHS Multidetector Computed Tomography 2 study and telomere length measured by Southern blot. Participants in COPDGene and the FHS were not selected for telomere length measurement on the basis of ILA status, while AGES-Reykjavik participants with ILA were oversampled to provide adequate comparison. This study was approved by the Icelandic Bioethics Committee (VSN: 00-063) and the Institutional Review Board of the Brigham and Women's Hospital.

Telomere length measurements

In COPDGene and the AGES-Reykjavik studies, mean telomere length (MTL) was measured using genomic DNA that was extracted from peripheral blood leukocytes. The measurements were done using a modified, high-throughput version of a quantitative real-time PCR-based telomere assay, as previously described [21]. The relative MTL was then calculated as a ratio of telomere repeat copy number to single gene copy number; each sample was run in triplicate and the value reported is the mean of those values. The MTL is then reported as the exponentiated ratio of telomere repeat copy number to single gene copy number corrected for a reference sample. The coefficient of variation for the triplicate assays was 6.4%. Given the large number of samples assayed and the time required to complete the assays, to account for potential batch effect, z-scores were calculated for the exponentiated relative telomere length by dividing by the mean per 384-well plate. Telomere lengths were measured using Southern blot analysis in the FHS as previously described [22, 23]; these measurements were performed from 1995 to 1998, ~13 years prior to the ILA assessment. The MTL measurements in COPDGene and AGES-Reykjavik were performed on samples that were collected at the time of the chest CT imaging. The difference in timing of the sample measurements was based on availability in each of the cohorts.

ILA evaluation

First, chest CT scans were evaluated for ILA using a sequential reading method, as previously described [1, 2, 24], by up to three readers (radiologists and pulmonologists on a Canon Medical Systems Inc. workstation), who were blind to all participant-specific information. ILA were initially defined as nondependent changes affecting >5% of any lung zone. These abnormalities included ground glass, reticular abnormalities, diffuse centrilobular nodularity, non-emphysematous cysts, traction bronchiectasis or honeycombing. Chest CT scans with either focal or unilateral ground glass attenuation, focal or unilateral reticulation, patchy ground glass abnormalities (bilateral or unilateral) or changes that affected <5% of the lung were indeterminate for ILA [1, 2]. The definition of ILA for this manuscript uses the updated definition of ILA adopted by the Fleischner Society [25]. This definition excludes those with ILA with centrilobular nodules alone, based on imaging, genetic and longitudinal outcome data, demonstrating

that this ILA subset should be viewed as a distinct phenotype [6, 11]. As a result, individuals with only centrilobular nodules were considered to have indeterminate disease and were excluded from the analyses. Further ILA subtyping (*e.g.* the identification of definite fibrosis [1]) was performed by a consensus of at least three readers as previously described [2, 11].

Identification of interstitial features

The objective identification of interstitial features on chest CT scans in COPDGene has been previously described in detail [26, 27]. Briefly, a previously trained, local histogram-based, machine-learning classifier was used to measure the percentage of total lung volume occupied by interstitial features (reticulation, honeycombing, centrilobular nodules, linear scar, nodular opacities, linear scar and ground glass), emphysema and normal tissue. Using this method, every part of the lung tissue was then classified as interstitial, normal or emphysema, and the total volumes were summed and then expressed as a percentage of the total lung volume.

Statistical analysis

In various analyses as indicated, MTL was analysed as a continuous variable and divided into quartiles (and deciles), with comparisons made to the quartile (or decile) with the longest telomere length. Analyses for the association between MTL and ILA were performed using logistic regression. The multivariable analyses were adjusted for age, sex, body mass index, pack-years of smoking, current smoking status and race in COPDGene. For analyses assessing the relationship between telomere length and ILA in the FHS, generalised estimating equations to account for familial correlation were used as previously described [28]. Logistic regression was also used to assess the association between MTL and COPD. To evaluate the association of MTL and mortality among those with ILA, only data from COPDGene and AGES-Reykjavik were used; the FHS subset lacked adequate statistical power for proper assessment due to the small sample size. Cox proportional hazards models were used to assess the association between MTL and time-to-mortality. All variables were assessed, and none violated the proportional hazards assumption. Reported p-values were two-sided and those <0.05 were considered statistically significant. SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) and R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria) were used for the analyses.

Results

ILA characterisation and MTLs were available for 4452 participants in COPDGene and for 498 participants from AGES-Reykjavik. ILA characterisation and telomere length were available for 439 participants from the FHS. In this subset from COPDGene, 240 (5%) had ILA, 2606 (59%) did not have ILA and 1503 (34%) were indeterminate for ILA status; 103 (2%) were excluded for having ILA characterised by centrilobular abnormalities only. In AGES-Reykjavik, 163 had ILA (33%), 243 (49%) did not have ILA, 88 (18%) were indeterminate for ILA and four (1%) were excluded for having only centrilobular abnormalities. In FHS, 44 (10%) had ILA, 204 (46%) did not have ILA and 191 (44%) were indeterminate for ILA; no participants had only centrilobular abnormalities on chest CT. Baseline characteristics of participants stratified by cohort and ILA status are included in table 1, and baseline

TABLE 1 Baseline characteristics of participants stratified by ILA status

	COPDGene		p-value	AGES-Reykjavik		p-value	Framingham Heart Study		p-value
	No ILA	ILA		No ILA	ILA		No ILA	ILA	
Subjects, n	2606	240		243	163		204	44	
Age, years	58±9	63±10	<0.0001	75±5	78±6	<0.0001	66±7	75±8	<0.0001
Sex (% female)	1221 (47)	102 (43)	0.20	132 (54)	67 (41)	0.01	117 (57)	23 (52)	0.54
Body mass index, kg·m^{-2}	29±6	30±7	0.01	27±4	27±4	0.96	29±5	29±6	0.71
Race (% white)	1886 (72)	171 (71)	0.71	--	--	--	--	--	--
Pack-years smoking (median (IQR))	37 (25–50)	42 (33–60)	<0.0001	0 (0–22)	14 (0–30)	<0.0001	5 (0–18)	12 (1–26)	0.01
Smoking status			0.28			<0.0001			0.55
Current	1401 (54)	138 (58)		27 (11)	28 (17)		6 (3)	2 (5)	
Former	1205 (46)	102 (42)		101 (42)	98 (60)		131 (64)	31 (70)	
Never	--	--		115 (47)	37 (23)		67 (33)	11 (25)	

Data are presented as mean±SD or n (%), unless otherwise stated. ILA: interstitial lung abnormalities; COPDGene: Genetic Epidemiology of Chronic Obstructive Pulmonary Disease; AGES: Age Gene/Environment Susceptibility; IQR: interquartile range.

characteristics by quartile of telomere length can be found in supplementary table S1. As has been previously demonstrated, participants with ILA were older and had increased tobacco exposure, and in the AGES-Reykjavik cohort participants with ILA were more likely to be men.

Interstitial abnormalities and telomere length

In the COPDGene and AGES-Reykjavik cohorts, after adjusting for covariates, measures of MTL were associated with ILA status (table 2). For example, in COPDGene, after adjusting for age, sex, body mass index, pack-years of smoking, current smoking status, race and Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) stage of COPD, participants in the shortest quartile of MTL had a 2.2-fold increase (95% CI 1.5–3.4, p=0.0001) in their odds of having ILA compared to those in the longest quartile of MTL (table 2). In AGES-Reykjavik, after adjusting for age, sex, body mass index, pack-years of smoking and current smoking status, participants in the shortest quartile of MTL had a 2.6-fold increase (95% CI 1.4–4.9, p=0.003) in their odds of having ILA compared to those in the longest quartile of MTL (table 2). Similar results were noted in analyses between ILA and continuous measures of MTL (table 2).

Comparably, in the FHS, after adjusting for age, sex, body mass index, pack-years of smoking and current smoking status, participants with ILA had decreased telomere length (767 bp, 95% CI 76–1584 bp, p=0.03) compared to those without ILA (table 2).

Interstitial lung abnormality subtypes and telomere length

We examined whether there were differences in MTL between different subtypes or radiological patterns of ILA. First, the relationship between definite fibrosis (evidence of parenchymal architectural distortion) and MTL was explored. In COPDGene, when compared to those without ILA, both fibrotic and non-fibrotic ILA were associated with shorter MTL (OR 11.6, 95% CI 2.7–49.7, p=0.009, and OR 1.7, 95% CI 1.1–2.6, p=0.02, respectively) for the comparison of the shortest quartile of MTL to the longest quartile. Among those with ILA, ILA with definite fibrosis was associated with a shorter MTL than those with ILA without fibrosis (OR 6.8, 95% CI 1.4–33.3, p=0.02 for the comparison of the shortest quartile of MTL to the longest quartile). In AGES-Reykjavik there was no difference in MTL between those with ILA with and without evidence of definite fibrosis (OR 1.2, 95% CI 0.4–3.3, p=0.80). Similar results were seen when evaluating the association between MTL and consistency with a usual interstitial pneumonia (UIP) pattern (supplementary material). Additional analyses limited to those with centrilobular abnormalities on chest CT in COPDGene are available in supplementary table S2.

Quantitative measure of interstitial features and MTL

Next, we assessed whether a quantitative measure of interstitial features on chest CT, which is associated with *MUC5B* promoter polymorphism in COPDGene [27], was associated with telomere length. After adjustment, the shortest quartile of MTL was associated with a mean±SE increase of 0.67±0.20% (p=0.001) in the amount of interstitial features when compared to the quartile with the longest telomere length. There was also a significant association when the MTL was evaluated continuously: per standard deviation decrease in MTL, there was a 1.13±0.33% (p=0.0007) increase in the percentage of interstitial features seen on chest CT (figure 1). When narrowed to those participants with ILA on chest CT, there was an increase in the effect size of this association. Among those with ILA, there was a 2% increase in the

TABLE 2 Association between telomere length and interstitial lung abnormalities

qPCR measurement	Continuous telomere length		Shortest quartile of telomere length[#]	
	OR (95% CI)[¶]	p-value	OR (95% CI)	p-value
AGES-Reykjavik	15.4 (3.8–62.5)	0.0001	2.7 (1.4–5.1)	0.002
COPDGene	4.2 (2.1–8.5)	<0.0001	2.2 (1.5–3.4)	0.0001
Southern blot measurement	**Length difference (95% CI)**		**p-value**	
Framingham Heart Study	767 bp (76–1584 bp)		0.03	

qPCR: quantitative polymerase chain reaction; AGES: Age Gene/Environment Susceptibility; COPDGene: Genetic Epidemiology of Chronic Obstructive Pulmonary Disease. [#]: comparison is to the longest quartile of telomere length; [¶]: analyses are adjusted for age, sex, body mass index, pack-years smoking, current smoking status and in COPDGene also adjusted for race.

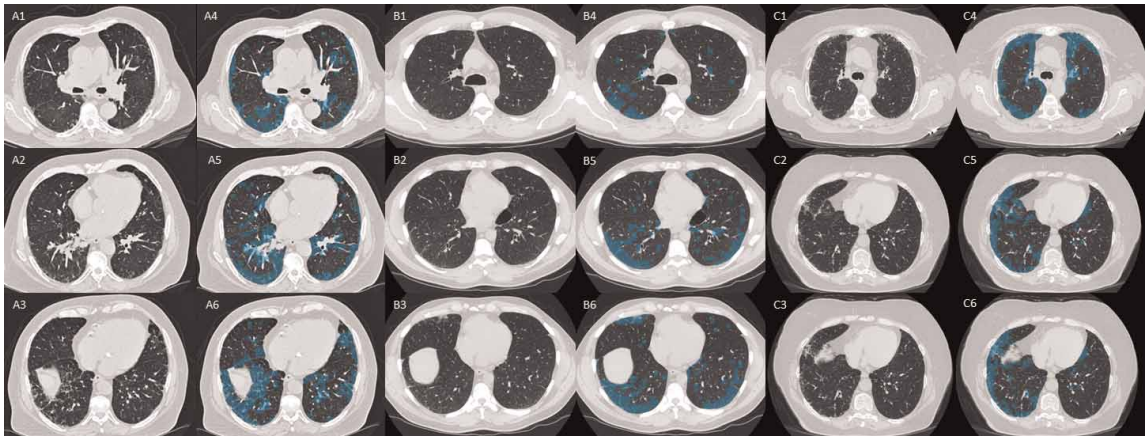

FIGURE 1 Axial computed tomography (CT) images from three COPDGene participants (A, B and C) with interstitial lung abnormalities. Images labelled 1, 2 and 3 are axial images at the level of the carina (1), right inferior pulmonary vein (2) and base (3). Images 4, 5 and 6 are the same axial images but with the overlay of the local histogram used to detect interstitial features. Blue represents areas of interstitial features.

amount of interstitial features present on chest CT in the shortest quartile of telomere length compared to longest quartile. We then assessed which components of the interstitial features were associated with decreased telomere length. After adjustment, the shortest quartile of MTL was associated with an increase in linear scar, reticular markings and subpleural line (supplementary table S3).

COPD and MTL

To provide a comparison to the results for ILA, and to follow up on prior reports of an association between decreased telomere length and measures of COPD [29], we analysed the associations between COPD and MTL in COPDGene. Although there was evidence for an association between MTL and COPD in COPDGene, the effect estimate was smaller than that noted with ILA. A diagnosis of COPD (as defined by spirometry and GOLD stage 2 and higher) was associated with the lowest quartile of MTL (OR 1.3, 95% CI 1.1–1.6, p=0.01); similar results were seen for the lowest quartile when those with ILA were removed from the analysis (OR 1.3, 95% CI 1.1–1.6, p=0.01).

MTL and mortality

Finally, we evaluated the association between telomere length and the risk of death among those with ILA in the COPDGene and AGES-Reykjavik cohorts. Although most associations between MTL and mortality among those with ILA in the COPDGene and AGES-Reykjavik cohort were not statistically significant, positive associations were found in some analyses in AGES-Reykjavik. For example, in both the COPDGene and AGES-Reykjavik cohorts there was no evidence for an association between the lowest quartile of MTL and risk of death among those with ILA (HR 0.82, 95% CI 0.4–1.7, p=0.6, and HR 1.2, 95% CI 0.6–2.2, p=0.5, respectively). While there was no evidence for an increased risk of death among those with ILA in the lowest 10th percentile of MTL in COPDGene (HR 1.3, 95% CI 0.9–1.8, p=0.14), there was evidence for an increased risk of death among those with ILA in the lowest 10th percentile of MTL in AGES-Reykjavik (HR 2.0, 95% CI 1.2–3.4, p=0.007). The evidence for an association between an increased risk of death among those with ILA in the lowest 10th percentile of MTL in AGES-Reykjavik was even greater when ILA was limited to those with probable UIP and UIP patterns (HR 3.5, 95% CI 1.7–7.4, p=0.0009); however, there was not a positive association among those with ILA limited to those with probable UIP and UIP patterns in COPDGene (HR 1.2, 95% CI 0.8–1.9, p=0.3).

Discussion

Our study presents the first comprehensive assessment of telomere length and ILA in the general population and in large populations of smokers and demonstrates several important findings. First, in three cohorts, including two general population samples, and with different types of measures, we demonstrate that ILA are associated with reduced telomere length. These findings imply that reduced telomere length, or processes strongly correlated with reduced telomere length, may be associated with the early

developmental stages of PF. In addition, while we present evidence that there may be some correlation between COPD and measures of reduced telomere length, this association is weaker than the association with ILA. Finally, our varying findings of associations with reduced telomere length and mortality among those with ILA suggest that this correlation may be more limited to those with the shortest telomere lengths, with more advanced imaging findings (*e.g.* patients with IPF) [17] and with additional risk factors for an increased rate of mortality (*e.g.* advanced age).

Our findings add to a growing body of evidence that measures of reduced telomere length overall not only contribute to the presence [16, 17] and morbidity [17] of advanced stages of PF, and early stages of PF in high-risk relatives [13, 14], but also likely play a role in increasing the risk of early undetected stages of PF in smokers and in the general population. However, it is important to note that reduced telomere length has also been associated with various forms of interstitial lung disease, including chronic hypersensitivity pneumonitis, interstitial pneumonia with autoimmune features, connective tissue disease-associated interstitial lung disease and unclassifiable patterns, among others [30, 31]. The spectrum of fibrotic lung disease associated with reduced telomere length may help to explain the consistent associations we demonstrate with ILA, which encompass a variety of imaging features and patterns, and likely represent disorders not limited to IPF [9].

Telomeres are regions of repetitive nucleotide sequences at the end of chromosomes that protect against the loss of genetic information during cell division. They shorten over time with repeated cell divisions and, once they reach a critical length, cellular senescence is induced that ultimately progresses to apoptosis or cell cycle arrest [29, 32, 33]. In adults the effect of telomere length shortening most commonly affects low turnover tissues, including the lung, and it has also been shown the telomeres shorten during the ageing of human fibroblasts [34]. It is possible that telomerase mutations accelerate this shortening and subsequent senescence in these cells, potentially explaining why PF is the most frequent manifestation of telomerase-associated disease [35]. Future studies should look for evidence of the induction of biological processes implicated by reduced telomere length in those with early developing stages of this disease.

The associations between telomere length and ILA and COPD deserve further consideration. Our findings are consistent with prior studies that have demonstrated associations between decreased telomere length and COPD [36, 37]. While confirming these prior associations, it is important to note that the size of this effect is much less than that seen with ILA (there is a 30% increase in the odds of developing COPD in the quartile with the shortest telomere length compared to a 120–170% increase in the odds of developing ILA).

Clusters of rare *TERT* mutations have previously been reported in a small number of research participants (three of 292) with severe emphysema [29]. In addition to the genetic association, it has been shown that participants with COPD have decreased telomere length compared to controls [36, 37]. By contrast, a study of Mendelian randomisation using polygenic risk scores demonstrated that IPF was associated with mutations in telomere-regulating genes, but COPD was not [38]. Further work is needed to determine if these findings are generalisable across the COPD spectrum or if there are severe COPD phenotypes that may have different genetic associations, as has been previously demonstrated [29].

Our study has several limitations. First, although we present associations between reduced telomere length and ILA in multiple cohorts, and with different types of measures, many of our findings of association are with MTL, which limits our ability to determine if specific thresholds of reduced telomere length contribute to our findings in some cohorts. Second, the small numbers of participants with telomere length measurements in the FHS limits our ability to evaluate the association between telomere length and important outcomes, including risk of death. In addition, the length of time between telomere measurement and ILA assessment in the FHS may limit our ability to evaluate outcomes. Third, our findings of association with reduced telomere length and mortality were inconsistent across populations. Although we speculate that these differences may be due to the distinct characteristics of the underlying populations (*e.g.* increased age in the AGES-Reykjavik cohort), we cannot rule out the possibility that they are due to technical limitations in telomere length measurement or missing data in the AGES-Reykjavik cohort, or that some of our positive findings between ILA and mortality are spurious. Additionally, despite the consistency of the association between MTL and ILA, the oversampling of ILA cases in the AGES-Reykjavik cohort may have led to bias in our results. Finally, although our analyses are adjusted for covariates including age and smoking history, we cannot rule out the possibility that some other unmeasured confounding variable could be contributing to our findings.

In conclusion, our study demonstrates that reduced telomere length is associated with ILA. In addition, our study also provides evidence for the differential association between telomere length and ILA, and

telomere length and measures of COPD. Future studies are needed to help determine the role of telomere length in the progression of early stages of PF, and whether efforts to intervene in this pathway could help prevent progression to more advanced stages of fibrotic lung disease.

Author contributions: Study design: G. Gudmundsson, G.M. Hunninghake and R.K. Putman; acquisition, analysis or interpretation of the data: R.K. Putman, H. Hatabu, T. Araki, M. Nishino, M. Yanagawa, E.F. Gudmundsson, J. Dupuis, I. DeVivo and G.M. Hunninghake; critical revision of the manuscript for important intellectual content: R.K. Putman, G.T. Axelsson, J.L. Sanders, S.Y. Ash, A.A. Menon, T. Araki, M. Yanagawa, E.F. Gudmundsson, D. Qiao, J. Dupuis, G.T. O'Connor, I.O. Rosas, G.R. Washko, S. El-Chemaly, B.A. Raby, V. Gudnason, D.L. DeMeo, E.K. Silverman, I. DeVivo, M.H. Cho, G. Gudmundsson, G.M. Hunninghake and R. San José Estépar; statistical analysis: R.K. Putman, J. Dupuis and G.M. Hunninghake; obtained funding: G.M. Hunninghake, M.H. Cho, R.K. Putman and G. Gudmundsson.

Conflict of interest: R.K. Putman reports grants from NIH/NHLBI, during the conduct of the study. G.T. Axelsson has nothing to disclose. S.Y. Ash reports grants from NHLBI (K08HL145118) and Pulmonary Fibrosis Foundation, during the conduct of the study; and is the owner of Quantitative Imaging Solutions, outside the submitted work. J.L. Sanders has commenced employment at Vertex Pharmaceuticals Inc. since the study was completed. A.A. Menon has nothing to disclose. T. Araki has nothing to disclose. M. Nishino reports personal fees for consultancy from Daiichi Sankyo and AstraZeneca, and grants from the Merck investigator studies programme, Canon Medical Systems, AstraZeneca and Daiichi Sankyo, outside the submitted work. M. Yanagawa reports travel expenses from Canon Medical Systems, outside the submitted work. E.F. Gudmundsson has nothing to disclose. D. Qiao has nothing to disclose. R. San José Estépar reports grants from Boehringer Ingelheim, Lung Biotechnology and Insmed, payment for lectures and travel expenses from Chiesi, and stock and/or stock options from Quantitative Imaging Solutions, outside the submitted work. J. Dupuis reports grants from NIH, during the conduct of the study. G.T. O'Connor reports grants from NIH, during the conduct of the study. I.O. Rosas has a patent 62/849,644 pending. G.R. Washko reports grants from Boehringer Ingelheim, consultancy for Pulmonx, Janssen Pharmaceuticals, Novartis and Vertex, and is founder and co-owner of Quantitative Imaging Solutions, outside the submitted work. S. El-Chemaly has nothing to disclose. B.A. Raby reports grants from NIH, outside the submitted work. V. Gudnason has nothing to disclose. D.L. DeMeo reports grants from Novartis and Bayer, during the conduct of the study. E.K. Silverman reports grants from GlaxoSmithKline and Bayer, during the conduct of the study. H. Hatabu reports grants from Canon Medical Systems Inc. and Konica-Minolta Inc., and personal fees for consultancy from Mitsubishi Chemical Inc. and Canon Medical Systems Inc., outside the submitted work. I. DeVivo has nothing to disclose. M.H. Cho reports grants from GlaxoSmithKline and Bayer, and consultancy and/or speaker fees from Genentech, AstraZeneca and Illumina, outside the submitted work. G. Gudmundsson has nothing to disclose. G.M. Hunninghake reports personal fees from Gerson Lehrman Group and Boehringer Ingelheim, outside the submitted work.

Support statement: The COPDGene project was supported by Award Number U01 HL089897 and Award Number U01 HL089856 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health. COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprising AstraZeneca, Boehringer Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, Siemens and Sunovion. R.K. Putman is supported by NIH grant K08 HL140087. J.L. Sanders is supported by NIH grant T32 HL007633. S.Y. Ash is supported by NIH grant K08 HL145118. G. Gudmundsson is supported by project grant 141513-051 from the Icelandic Research Fund and Landspitali Scientific Fund A-2019-029 and A-2018-024. M. Nishino is supported by NIH grant R01 CA203636. D. Qiao is supported by NIH grant K01 HL129039. R. San José Estépar is supported by NIH grants R21HL140422 and R01HL149877. I.O. Rosas is supported by NIH grants U01 HL133232 and R01 HL130974. G.R. Washko is supported by NIH grants R01 HL116473 and R01 HL122464. The Age, Gene/Environment Susceptibility Reykjavik Study was supported by NIH contracts N01-AG-1-2100 and HHSN27120120022C, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association) and the Althingi (the Icelandic Parliament). V. Gudnason is supported by NIA grant 27120120022C and project grant 141513-051 from the Icelandic Research Fund. D.L. DeMeo is supported by NIH grant P01 HL132825. H. Hatabu is supported by NIH grant 5U01CA209414-03. M.H. Cho is supported by NIH grants HL137927, HL147148, HL149861 and HL135142. G.M. Hunninghake and this work were supported by NIH grants R01 HL111024, R01 HL130974 and R01 135142, and project grant 141513-051 from the Icelandic Research Fund. The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute in collaboration with Boston University (contract numbers N01-HC-25195, HHSN268201500001I and 75N92019D00031). The COPDGene study is supported by NIH grants U01 HL089897 and U01 HL089856. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Hunninghake GM, Hatabu H, Okajima Y. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013; 368: 2192–2200.
2 Washko GR, Hunninghake GM, Fernandez IE, *et al.* Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011; 364: 897–906.
3 Doyle TJ, Washko GR, Fernandez IE, *et al.* Interstitial lung abnormalities and reduced exercise capacity. *Am J Respir Crit Care Med* 2012; 185: 756–762.
4 Axelsson GT, Putman RK, Miller ER, *et al.* Interstitial lung abnormalities and physical function. *ERJ Open Res* 2018; 4: 00057-2018.
5 Araki T, Putman RK, Hatabu H, *et al.* Development and progression of interstitial lung abnormalities in the Framingham Heart Study. *Am J Respir Crit Care Med* 2016; 194: 1514–1522.
6 Putman RK, Gudmundsson G, Axelsson GT, *et al.* Imaging patterns are associated with interstitial lung abnormality progression and mortality. *Am J Respir Crit Care Med* 2019; 200: 175–183.
7 Hino T, Lee KS, Han J, *et al.* Spectrum of pulmonary fibrosis from interstitial lung abnormality to usual interstitial pneumonia: importance of identification and quantification of traction bronchiectasis in patient management. *Korean J Radiol* 2021; 22: 811–828.
8 Putman RK, Hatabu H, Araki T, *et al.* Association between interstitial lung abnormalities and all-cause mortality. *JAMA* 2016; 315: 672–681.
9 Hobbs BD, Putman RK, Araki T, *et al.* Overlap of genetic risk between interstitial lung abnormalities and idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019; 200: 1402–1413.
10 Allen RJ, Guillen-Guio B, Oldham JM, *et al.* Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2020; 201: 564–574.
11 Putman RK, Gudmundsson G, Araki T, *et al.* The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *Eur Respir J* 2017; 50: 1700537.
12 Mathai SK, Humphries S, Kropski JA, *et al. MUC5B* variant is associated with visually and quantitatively detected preclinical pulmonary fibrosis. *Thorax* 2019; 74: 1131–1139.
13 Salisbury ML, Hewlett JC, Ding G, *et al.* Development and progression of radiologic abnormalities in individuals at risk for familial ILD. *Am J Respir Crit Care Med* 2020; 201: 1230–1239.
14 Hunninghake GM, Quesada-Arias LD, Carmichael NE, *et al.* Interstitial lung disease in relatives of patients with pulmonary fibrosis. *Am J Respir Crit Care Med* 2020; 201: 1240–1248.
15 Seibold MA, Wise AL, Speer MC, *et al.* A common *MUC5B* promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011; 364: 1503–1512.
16 Newton CA, Batra K, Torrealba J, *et al.* Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive. *Eur Respir J* 2016; 48: 1710–1720.
17 Stuart BD, Lee JS, Kozlitina J, *et al.* Effect of telomere length on survival in patients with idiopathic pulmonary fibrosis: an observational cohort study with independent validation. *Lancet Respir Med* 2014; 2: 557–565.
18 Harris TB, Launer LJ, Eiriksdottir G, *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* 2007; 165: 1076–1087.
19 Regan EA, Hokanson JE, Murphy JR, *et al.* Genetic epidemiology of COPD (COPDGene) study design. *COPD* 2010; 7: 32–43.
20 Splansky GL, Corey D, Yang Q, *et al.* The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007; 165: 1328–1335.
21 McGrath M, Wong JY, Michaud D, *et al.* Telomere length, cigarette smoking, and bladder cancer risk in men and women. *Cancer Epidemiol Biomarkers Prev* 2007; 16: 815–819.
22 Okuda K, Bardeguez A, Gardner JP, *et al.* Telomere length in the newborn. *Pediatr Res* 2002; 52: 377–381.
23 Vasan RS, Demissie S, Kimura M, *et al.* Association of leukocyte telomere length with circulating biomarkers of the renin-angiotensin-aldosterone system: the Framingham Heart Study. *Circulation* 2008; 117: 1138–1144.
24 Xu JF, Washko GR, Nakahira K, *et al.* Statins and pulmonary fibrosis: the potential role of NLRP3 inflammasome activation. *Am J Respir Crit Care Med* 2012; 185: 547–556.
25 Hatabu H, Hunninghake GM, Richeldi L, *et al.* Interstitial lung abnormalities detected incidentally on CT: a position paper from the Fleischner Society. *Lancet Respir Med* 2020; 8: 726–737.
26 Ash SY, Harmouche R, Ross JC, *et al.* The objective identification and quantification of interstitial lung abnormalities in smokers. *Acad Radiol* 2017; 24: 941–946.
27 Ash SY, Harmouche R, Putman RK, *et al.* Clinical and genetic associations of objectively identified interstitial changes in smokers. *Chest* 2017; 152: 780–791.
28 Uh HW, Wijk HJ, Houwing-Duistermaat JJ. Testing for genetic association taking into account phenotypic information of relatives. *BMC Proc* 2009; 3: Suppl. 7, S123.
29 Stanley SE, Chen JJ, Podlevsky JD, *et al.* Telomerase mutations in smokers with severe emphysema. *J Clin Invest* 2015; 125: 563–570.

30 Ley B, Newton CA, Arnould I, *et al.* The MUC5B promoter polymorphism and telomere length in patients with chronic hypersensitivity pneumonitis: an observational cohort-control study. *Lancet Respir Med* 2017; 5: 639–647.

31 Newton CA, Oldham JM, Ley B, *et al.* Telomere length and genetic variant associations with interstitial lung disease progression and survival. *Eur Respir J* 2019; 53: 1801641.

32 Aviv A, Valdes AM, Spector TD. Human telomere biology: pitfalls of moving from the laboratory to epidemiology. *Int J Epidemiol* 2006; 35: 1424–1429.

33 Armanios M. Telomeres and age-related disease: how telomere biology informs clinical paradigms. *J Clin Invest* 2013; 123: 996–1002.

34 Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature* 1990; 345: 458–460.

35 Armanios M. Telomerase and idiopathic pulmonary fibrosis. *Mutat Res* 2012; 730: 52–58.

36 Cordoba-Lanus E, Cazorla-Rivero S, Espinoza-Jimenez A, *et al.* Telomere shortening and accelerated aging in COPD: findings from the BODE cohort. *Respir Res* 2017; 18: 59.

37 Rode L, Bojesen SE, Weischer M, *et al.* Short telomere length, lung function and chronic obstructive pulmonary disease in 46,396 individuals. *Thorax* 2013; 68: 429–435.

38 Duckworth A, Gibbons MA, Allen RJ, *et al.* Telomere length and risk of idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease: a Mendelian randomisation study. *Lancet Respir Med* 2021; 9: 285–294.