Original research

MUC5B, telomere length and longitudinal quantitative interstitial lung changes: the MESA Lung Study

John S Kim ,[1,2] Ani W Manichaikul ,[3,4] Eric A Hoffman,[5] Pallavi Balte,[2] Michaela R Anderson,[2] Elana J Bernstein,[2] Purnema Madahar,[2] Elizabeth C Oelsner,[2] Steven M Kawut,[6,7] Artur Wysoczanski,[2] Andrew F Laine,[8] Ayodeji Adegunsoye ,[9] Jennie Z Ma,[3] Margaret A Taub,[10] Rasika A Mathias,[10,11] Stephen S Rich,[3,4] Jerome I Rotter,[12,13] Imre Noth,[1] Christine Kim Garcia ,[2] R Graham Barr,[2,14] Anna J Podolanczuk [15]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/thorax-2021-218139).

For numbered affiliations see end of article.

Correspondence to
Dr John S Kim, Medicine, University of Virginia School of Medicine, Charlottesville, VA 22903, USA; jk6jb@virginia.edu

Received 26 August 2021
Accepted 11 July 2022
Published Online First 5 August 2022

To cite: Kim JS, Manichaikul AW, Hoffman EA, *et al*. *Thorax* 2023;**78**:566–573.

ABSTRACT

Background The *MUC5B* promoter variant (rs35705950) and telomere length are linked to pulmonary fibrosis and CT-based qualitative assessments of interstitial abnormalities, but their associations with longitudinal quantitative changes of the lung interstitium among community-dwelling adults are unknown.

Methods We used data from participants in the Multi-Ethnic Study of Atherosclerosis with high-attenuation areas (HAAs, Examinations 1–6 (2000–2018)) and *MUC5B* genotype (n=4552) and telomere length (n=4488) assessments. HAA was defined as the per cent of imaged lung with attenuation of −600 to −250 Hounsfield units. We used linear mixed-effects models to examine associations of *MUC5B* risk allele (T) and telomere length with longitudinal changes in HAAs. Joint models were used to examine associations of longitudinal changes in HAAs with death and interstitial lung disease (ILD).

Results The *MUC5B* risk allele (T) was associated with an absolute change in HAAs of 2.60% (95% CI 0.36% to 4.86%) per 10 years overall. This association was stronger among those with a telomere length below an age-adjusted percentile of 5% (p value for interaction=0.008). A 1% increase in HAAs per year was associated with 7% increase in mortality risk (rate ratio (RR)=1.07, 95% CI 1.02 to 1.12) for overall death and 34% increase in ILD (RR=1.34, 95% CI 1.20 to 1.50). Longer baseline telomere length was cross-sectionally associated with less HAAs from baseline scans, but not with longitudinal changes in HAAs.

Conclusions Longitudinal increases in HAAs were associated with the *MUC5B* risk allele and a higher risk of death and ILD.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The *MUC5B* (rs35705950) risk allele (T) and telomere length are associated with pulmonary fibrosis and qualitative assessments of interstitial lung changes on CT. There are fewer studies that have examined their relationship with longitudinal interstitial lung changes among community-dwelling adults.

WHAT THIS STUDY ADDS

⇒ More high-attenuation areas over time on CT were associated with the *MUC5B* risk allele, particularly those with shorter baseline telomere length, and a higher risk of overall death and interstitial lung disease.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our study suggests that quantitative CT assessments and their longitudinal measures may be a potential tool to identify earlier stages of interstitial lung disease.

INTRODUCTION

Interstitial lung disease (ILD) is a group of chronic respiratory disorders characterised by inflammation and fibrosis of the lung parenchyma.[1] Non-fibrosing ILDs can become fibrotic which may lead to rapid deterioration, chronic respiratory failure and death. By the time of diagnosis, patients often have significant physiological impairments and decreased exercise capacity, with a poor median survival.[2 3]

An increasing number of studies have focused on visual, qualitative assessment of CT scans for changes suggestive of ILD, termed interstitial lung abnormalities (ILAs), when identified in research studies or incidentally in the clinical context.[4] These studies have yielded important insights into novel risk factors for progressive fibrosing ILD. A recent Fleischner Society position paper defining the radiological criteria for ILA also highlighted the need for quantitative CT methods for evaluation of disease extent and progression as a key research priority.[4]

One of the strongest known risk factors for pulmonary fibrosis is the *MUC5B* promoter polymorphism (rs35705950); it is associated with fourfold higher odds of idiopathic pulmonary fibrosis (IPF). Recent studies have shown that this

gain-of-function promoter variant is linked to other types of fibrosing ILDs like chronic hypersensitivity pneumonitis and rheumatoid arthritis-related ILD.[5–7] The polymorphism is associated with ILA among the general population and in cohorts of smokers, suggesting that carriers of this allele may be at higher risk of developing fibrosing ILDs.[8] Shorter telomere length, a marker of accelerated cellular senescence, is another intrinsic risk factor that is associated with a higher risk of pulmonary fibrosis, more severe disease, mortality and a higher prevalence of ILA.[6 9 10] The few prior studies that have examined the *MUC5B* polymorphism and telomere length in IPF and chronic hypersensitivity pneumonitis did not show evidence of interaction between these genetic factors, suggesting distinct mechanisms by which they influence pathogenesis. However, study of the *MUC5B* promoter polymorphism and telomere length in other ILDs and in earlier stages of ILD is lacking.[6 11]

Whether these genetic risk factors (and their interactions) are linked to a quantitative measure of lung parenchymal changes and its progression over time among a racially/ethnic diverse sample of community-dwelling adults remains unclear. We hypothesised that the *MUC5B* promoter polymorphism (rs35705950) would be associated with longitudinal increases in high-attenuation areas (HAAs) with effect modification by telomere length over time in the Multi-Ethnic Study of Atherosclerosis (MESA). We also examined whether longitudinal changes in HAAs are associated with death and ILD-related clinical events.

METHODS

Study participants

MESA in an ongoing prospective cohort with the original intent of investigating subclinical cardiovascular disease.[12] There were 6814 US adults between the ages of 45 and 84 years without clinical cardiovascular disease recruited at Examination 1 (2000–2002) and followed longitudinally, with repeat examinations that included additional data collection. Participants who provided consent for genetic analyses were included in this study.

MUC5B genotype and telomere length

MUC5B (rs35705950) genotypes were obtained from whole-genome sequencing (WGS) that was performed on DNA samples from whole blood in MESA as part of the National Heart, Lung and Blood Institute TOPMed programme. Procedures related to the TOPMed programme have been previously described in detail.[13] Briefly, Illumina HiSeq X Ten instruments were used for WGS that targeted a mean depth of at least 30× (paired-end, 150 bp reads). PCR-free library preparation kits from KAPA Biosystems were used for all sequencing. Using blood samples from MESA Examination 1, telomere length was extracted from the TOPMEd WGS using TelSeq and has been previously described. Briefly, counts of sequencing reads that contained fixed number of repeats of the telomeric nucleotide motif 'TTAGGG' were used to estimate individual telomere lengths and accounted for batch corrections. TelSeq estimates of telomere length have a Pearson correlation of 0.68–0.72 with Southern blot data in TOPMed.[14] There were no repeat telomere length measurements from WGS in MESA.

High-attenuation area

MESA participants underwent cardiac CT scans at Examinations 1 (2000–2002), 2 (2002–2004), 3 (2004–2005), 4 (2005–2007) and 5 (2010–2012), and full-lung scans at Examination 6 (2016–2018) using the MESA/SPIROMICS protocol.[15] Sixty-six per cent of the lung volume from the carina to lung bases was captured by cardiac scans. The intraclass correlation coefficient for lung density measures between cardiac and full-lung scans is 0.93.[16] HAAs from Examination 6 full-lung scans were assessed in the lower two-thirds lung fields to be consistent with the previous cardiac scans; this approach has been used previously for other densitometry-based measures like per cent emphysema.[17] HAA was defined as the percentage of voxels with attenuation values in the range of −600 and −250 Hounsfield units (HU), as previously described.[18] The Pulmonary Analysis Software Suite at the University of Iowa's Advanced Pulmonary Physiomic Imaging Laboratory was used by trained technicians for segmentation and correction of the lung.[19] Per cent emphysema was defined as the percentage of lung voxels below −950 HU.[20]

Mortality and ILD

Vital status of participants was determined by the contact of each MESA participant or family member by interviewers every 9–12 months. The National Death Index was used to supplement this review and ensure complete follow-up. Mortality was adjudicated up until 31 December 2018. Hospitalisation data were complete as of 31 December 2013. Inpatient medical records were reviewed by an adjudication panel and International Classification of Diseases (ICD), 9th Revision (495.XX, 515.X or 516.XX) and ICD, 10th Revision (J60.X–J64.X, J67.X or J84.X) were used to determine whether participants who were hospitalised or died had ILD as previously described.[21] Further details are in the online supplemental file 1.

Statistical analysis

We used linear mixed-effects models with random intercept and slope to examine associations of the *MUC5B* risk allele (T) with the longitudinal change in HAAs over time (Examinations 1 through 6, 2000–2018). We used a genetic additive model to present our results.[5] We used mixed-effects model to account for within-subject correlations among repeated measurements of HAAs over time for each participant. Such modelling approach has been used in other studies.[17 22] Measurements of HAAs from all six examinations were used in our linear mixed-effects model. Additional information is provided in the online supplemental file 1.

The longitudinal analysis was adjusted for scanner parameters, principal components of genetic ancestry, baseline age, sex, race/ethnicity, smoking status and cigarette pack-years, and time-varying covariates of height, weight, cigarettes smoked per day, and per cent emphysema. We included interaction terms of covariates with the time since initial HAA assessment to evaluate the impact of covariates on HAA trajectories over time. We performed an analysis of the overall cohort and race/ethnic-specific analyses due to the heterogeneity of *MUC5B* (rs35705950) genotype frequencies.[23 24] Positive and negative beta estimates from the term, '*MUC5B* risk allele×time since initial HAA assessment', were interpreted as more rapid progression and slower progression in HAAs over time, respectively. We used natural log-transformed HAAs as the outcome in all of our regression models. To improve interpretability, we exponentiated the regression coefficients of our log-transformed HAA analysis to report our outcomes as per cent change in HAAs.

We examined whether telomere length modified the association between *MUC5B* and HAAs based on a previous study that examined interactions between *MUC5B* and telomere length in pulmonary fibrosis.[6] Studies have used different percentile cut-offs for telomere length as there is uncertainty of what is a relevant cut-off in general population-based cohorts.[25–28] Therefore,

stratified analyses were performed using age-adjusted percentile cut-offs for telomere length at 5%, 10% and 25%. We accounted for baseline age in generating each of these cut-offs by using the percentile cut-off for each of the age groups: <50 years, 50–60, 60–70 and >80 years.

We examined associations of baseline (Examination 1) telomere length with Examination 1 HAAs using linear regression models with adjustment for baseline scanner parameters, age, sex, race/ethnicity, smoking history, cigarette pack-years, height, weight, per cent emphysema and principal components of genetic ancestry. For the longitudinal HAA analysis, we used the same approach as our *MUC5B* analysis. We report results per 1 SD increment in natural log-transformed telomere length.

We performed prespecified stratified analyses by smoking status (never vs ever) based on past studies that suggest overproduction of mucin 5B may impair clearance of cigarette particles that leads to lung injury.[29] Sex was also examined as a potential effect modifier due to the differential associations of sex and pulmonary fibrosis risk.[1] For associations between telomere length and HAAs, we performed a stratified analysis by race/ethnicity. The log-likelihood ratio was used to test for effect modification in cross-sectional analyses. We used the F-test of the three-way interaction term, 'effect modifier×primary exposure variable of interest×time since initial HAA assessment', to determine effect modification in longitudinal analyses.[17]

Joint modelling was used to examine the association of longitudinal changes of HAAs with time-to-event data (ie, overall death and ILD-related events). This approach uses linear mixed-effects models to model longitudinal changes of a time-varying covariate of interest (ie, HAAs) and examine its association with a time-to-event outcome using a Cox regression model.[30] For the ILD analysis, we used a composite binary outcome of ILD-related death or hospitalisation. Due to the time when events were adjudicated, we used HAAs from Examinations 1 through 5 (2010–2012) for the joint model. Results are reported as rate ratios (RRs) per 1% increment in HAAs per year. A description of this method and the covariates we adjusted for in the models are in the supplement.

MUC5B and telomere length analyses were performed using SAS V.9.4 (SAS Institute). The 'nlme', 'survival' and 'JMbayes2' packages were used for the joint model analysis in R V.3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

There were 4552 MESA participants with *MUC5B* genotype and Examination 1 HAA assessments (online supplemental figure S1). Table 1 shows the baseline characteristics of the study sample. Over 50% of the cohort were women, Asian, African-American or Hispanic, and had a history of smoking. There was a higher proportion of self-identified non-Hispanic white participants with the GT or TT genotype compared with other races/ethnicities (online supplemental table S1). Participant characteristics among those with longer versus shorter telomere length are shown in online supplemental table S2.

MUC5B and HAAs

Online supplemental table S3 summarises the number of CT scans performed at each examination between 2000 and 2018 for our longitudinal analysis. There were 4552 MESA participants with the *MUC5B* genotype and at least two repeat HAA measurements between 2000 and 2018. Representative CT images of increasing HAAs between Examinations 1 and 6 are shown in figure 1.

Table 1 Study sample characteristics

Characteristic	Overall	*MUC5B* genotype		
		GG	GT	TT
Number of participants	4552	3959	569	24
Age	61 (10)	61 (10)	61 (10)	60 (9)
Female sex	51%	51%	53%	50%
Race/ethnicity				
Non-Hispanic white	41%	37%	68%	75%
Asian	13%	15%	2%	0
African-American	24%	25%	10%	4%
Hispanic	22%	23%	20%	21%
Smoking status				
Never smoker	46%	47%	40%	21%
Former smoker	40%	39%	46%	75%
Current smoker	14%	14%	14%	4%
Cigarette pack-years	12 (22)	11.5 (21.2)	13.9 (23.4)	19.4 (30.0)
Height (cm)	167 (10)	167 (10)	168 (10)	169 (7)
Weight (kg)	79 (17)	78 (17)	80 (18)	86 (21)
Telomere length (kb)	4.4 (0.9)	4.4 (0.9)	4.4 (0.9)	4.1 (0.9)

Continuous variables presented as mean (SD).
Categorical variables presented as percentage.

Each *MUC5B* risk allele (T) was associated with an increase in HAAs of 2.60% (95% CI 0.36% to 4.86%) per 10 years (table 2 and figure 2). In comparison, smoking 10 cigarettes per day on average was associated with an increase in HAAs of 4.69% (95% CI 2.24% to 7.15%) per 10 years (online supplemental table S4). The association between *MUC5B* and HAAs was stronger among those with a telomere length below an age-adjusted cut-off of 5% (p value for interaction=0.008). There was no significant effect modification with a telomere length cut-off of 10% and 25%. There was not a significant interaction with sex or smoking (online supplemental table S5). The *MUC5B* and HAA association was strongest among non-Hispanic white participants with an increase in HAAs of 3.05% (95% CI 0.64% to 5.46%) per 10 years, and was significant after applying a Bonferroni corrected p value of 0.0125 (online supplemental table S6). Associations between *MUC5B* and HAAs were in the same direction but weaker and not significant in the other race/ethnic subgroups (online supplemental table S6).

Telomere length and HAAs

There were 4488 MESA participants with telomere length and Examination 1 HAA assessments (online supplemental figure S1). Shorter baseline telomere length was associated with more HAAs on Examination 1 cardiac CT. A 1 SD decrease in log-transformed telomere length was associated with an HAA of 6.51% (95% CI 2.62% to 10.24%) (table 3). Participants with a telomere length below the 25th percentile had more baseline HAAs compared with those with longer telomeres (1.56%, 95% CI 0.04% to 3.11%). Associations were not statistically significant using 5th and 10th percentile cut-offs. Baseline telomere length was not associated with a longitudinal change in HAAs overall and in stratified analyses (table 3 and online supplemental table S7).

HAAs, mortality and ILD

Among the 4552 participants, there were 4517 (99%) who had adjudicated death status and complete covariate data. There was

Figure 1A

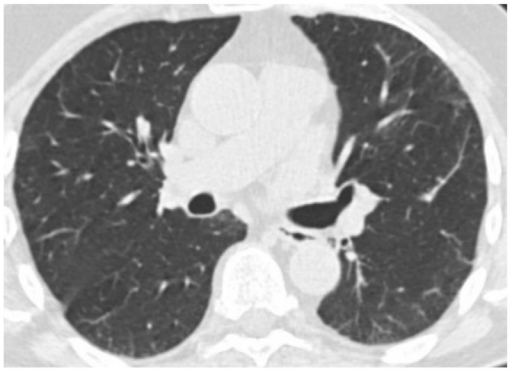

Figure 1B

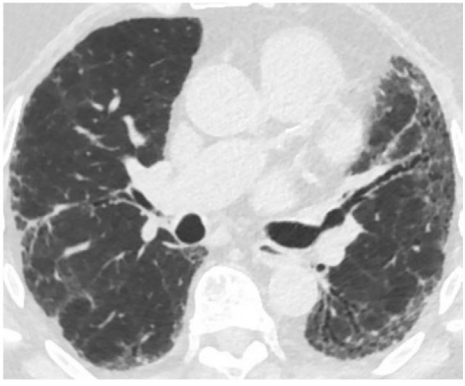

Figure 1C

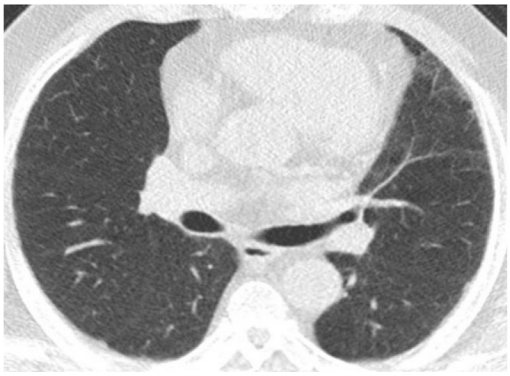

Figure 1D

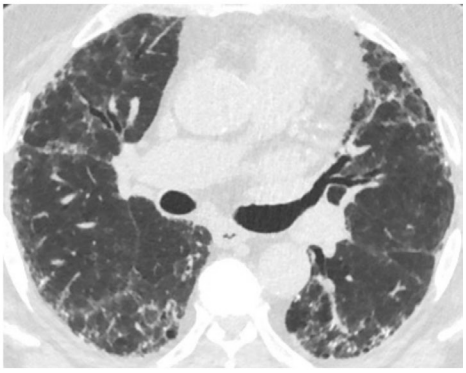

Figure 1 CT axial images that show an increase in high-attenuation areas over time between Examinations 1 (2000–2002) and 6 (2016–2018). (A,B) CT images of one participant with high-attenuation areas per cent of 4.34 at Examination 1 and 12.21 at Examination 6. (C,D) CT images of another participant with high-attenuation areas per cent of 7.38 at Examination 1 and 17.42 at Examination 6.

a total of 827 deaths with an event rate of 111.1 (95% CI 103.8 to 118.9) per 10 000 person-years. A 1% increment in HAAs per year was associated with a 7% higher risk of overall death (RR 1.07, 95% CI 1.02 to 1.12) after adjustment for covariates (table 4). The association between longitudinal changes in HAAs and mortality was not significantly different by *MUC5B* risk allele carrier status, shorter telomere length, sex or smoking history (p values for interaction >0.40) (online supplemental table S8). There were 4443 participants with adjudicated ILD-related outcomes with a total of 29 events and an event rate of 5.5 (95% CI 3.7 to 7.7) per 10 000 person-years. A 1% increment in HAAs per year was associated with a 34% higher risk of ILD (ie, ILD-related hospitalisation or death) (RR 1.34, 95% CI 1.20 to 1.50).

DISCUSSION

The *MUC5B* (rs35705950) promoter variant was associated with more HAAs over time among community-dwelling adults. This association was stronger among those with shorter telomere

Table 2 Associations of *MUC5B* (rs35705950) risk allele with longitudinal changes in high-attenuation areas (HAAs)

Model	Number of participants	% longitudinal change in HAAs per 10 years (95% CI) per *MUC5B* (rs35705950) risk allele (T)	P value*
Overall	4552	2.60 (0.36 to 4.86)	0.02
Stratified by telomere length percentile cut-off			
Below 5th percentile	221	15.13 (5.60 to 24.75)	0.008
Above 5th percentile	4267	1.93 (–0.40 to 4.27)	
Below 10th percentile	446	5.26 (–1.35 to 11.91)	0.41
Above 10th percentile	4042	2.30 (–0.11 to 4.71)	
Below 25th percentile	1119	5.03 (0.83 to 9.25)	0.19
Above 25th percentile	3369	1.72 (–0.94 to 4.38)	

Overall model is adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, self-reported race/ethnicity, smoking status and cigarette pack-years were also adjusted for including their interaction terms with 'time since initial HAA assessment'. Time-varying covariates height, weight, per cent emphysema and cigarettes smoked per day were also adjusted for in the model.
Stratified models include their respective three-way interaction term (eg, 'telomere length percentile below 5%×*MUC5B* risk allele×time since initial HAA assessment').
All results are reported per risk allele (T) of the *MUC5B* (rs35705950) promoter variant.
*P values for stratified analysis represent p values for interaction.

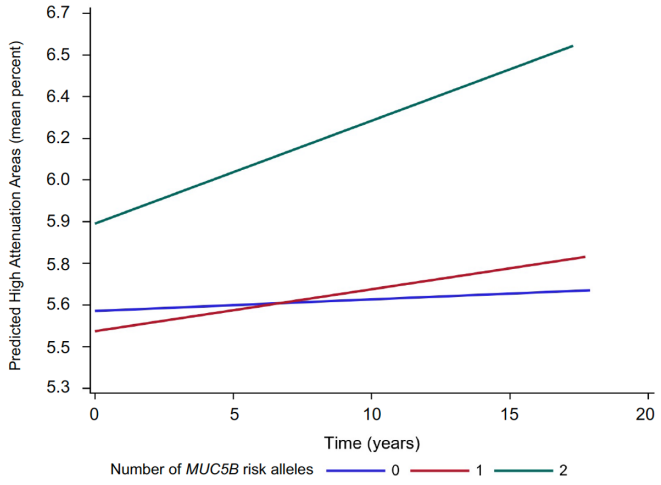

Figure 2 Predicted high-attenuation areas (HAAs) over time by number of *MUC5B* (rs35705950) risk alleles (T). Linear mixed-effects model with random intercept and slope adjusted for scanner parameters and principal components of genetic ancestry. Baseline sex, self-reported race/ethnicity, age, smoking status, cigarette pack-years and their interaction term with 'time since initial HAA assessment'. Also adjusted for time-varying covariates height, weight, per cent emphysema and cigarettes smoked per day. Y-axis is log-scale.

length using an age-adjusted 5th percentile cut-off and non-Hispanic white individuals. Shorter baseline telomere length was associated with more HAAs at baseline, but not with changes in HAAs over time. Longitudinal increases in HAAs were associated with worse survival and a higher risk of ILD-related events.

Prior studies have identified a strong association between the *MUC5B* (rs35705950) gain-of-function promoter variant and a higher risk of pulmonary fibrosis.[5 31] *MUC5B* encodes the protein mucin 5B and is highly expressed in the lungs of adults with pulmonary fibrosis, particularly in the distal airways.[32] Overexpression of *MUC5B* impairs mucosal host defences, clearance of inhaled materials and alveolar repair upon injury, and has been linked to a higher prevalence of ILA and its progression, further suggesting its pathogenic role in pulmonary fibrosis.[8 33] Ours is the first study to report on the association of the *MUC5B* polymorphism with longitudinal quantitative changes in CT density (that could represent earlier stages of ILD) in a more general population cohort.

Our longitudinal analysis shows that the *MUC5B* (rs35705950) risk allele (T) was associated with increasing HAAs over time. This is consistent with prior studies that have identified associations between the *MUC5B* risk allele and other quantitative CT-based interstitial lung measures. Those previous studies have largely been cross-sectional and used cohorts comprised of individuals with a significant smoking history or first-degree relatives of adults with pulmonary fibrosis.[34 35] This study extends the prior work to community-dwelling adults sampled without regard for respiratory symptoms or smoking history. Notably, the association between *MUC5B* and HAAs was not affected by smoking history. Previous studies did not find an association between *MUC5B* and HAAs assessed from a single time point; the focus of our study was to examine longitudinal changes in HAAs.[36 37] Another difference from prior work was that we accounted for changes over time in scanner models, height and weight. Our longitudinal analysis over an 18-year span with repeated CT measurements supports the potential unidirectional relationship between the *MUC5B* variant and interstitial lung changes that may represent the early stages of ILD.[8]

Race/ethnic-specific analyses revealed that the association between *MUC5B* (rs35705950) and HAAs was stronger in non-Hispanic white participants compared with other groups. Consistent with prior literature, the genotypes GT and TT were rare in the Asian and African-American subgroups and less common in the Hispanic group in MESA compared with the non-Hispanic white group.[23] Thus, we may have been underpowered to detect differences in the smaller racial/ethnic groups. There has been one prior study demonstrating an association between *MUC5B* (rs35705950) and ILA in a cohort of adults from Mexico.[38] Few other studies of genetic risk factors for early ILD/ILA have focused on under-represented racial and ethnic groups, highlighting a pressing need for research in this area. African-American patients with ILD have been shown to have a younger age at diagnosis and longer survival time, underscoring the importance of research into the determinants of racial/ethnic differences in disease risk and progression.[39] The recent identification of variants related to obstructive lung disease and ILD in Hispanic/Latinx cohorts of admixed populations and African ancestry demonstrates that such studies are feasible.[40–42]

Table 3 Associations of baseline telomere length with high-attenuation areas (HAAs)

Model	Number of participants	Mean per cent change in Examination 1 HAAs (95% CI)	P value	% longitudinal change in HAAs per 10 years (95% CI)	P value
Absolute telomere length*	4488	−6.51 (−10.24 to −2.62)	0.001	0.32 (−4.68 to 5.35)	0.90
Age-adjusted telomere length percentile cut-offs					
Above 5th percentile	4267	0.0 (ref)		0.0 (ref)	
Below 5th percentile	221	0.44 (−2.55 to 3.52)	0.78	0.87 (−2.82 to 4.56)	0.65
Above 10th percentile	4042	0.0 (ref)		0.0 (ref)	
Below 10th percentile	446	0.16 (−2.00 to 2.36)	0.88	1.41 (−1.28 to 4.10)	0.30
Above 25th percentile	3369	0.0 (ref)		0.0 (ref)	
Below 25th percentile	1119	1.56 (0.04 to 3.11)	0.04	0.11 (−1.74 to 1.96)	0.91

Examination 1 HAAs model: adjusted for scanner parameters, principal components of genetic ancestry, and baseline age, sex, self-reported race/ethnicity, smoking status, cigarette pack-years height, weight, and per cent emphysema.
Longitudinal HAAs model: adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, self-reported race/ethnicity, smoking status and cigarette pack-years were also adjusted for including their interaction terms with 'time since initial HAA assessment'. Time-varying covariates height, weight, per cent emphysema and cigarettes smoked per day were also adjusted for in the model.
*Reported per SD increment of log-transformed telomere length.

Table 4 Associations of longitudinal changes in high-attenuation areas (HAAs) with mortality and interstitial lung disease

Model	Number of participants	Total person-years	Events	Event rate per 10 000 person-years (95% CI)	Rate ratio per 1% Increment in HAAs per year (95% CI)	P value
Overall mortality	4517	74 415	827	111.1 (103.8 to 118.9)	1.07 (1.02 to 1.12)	<0.001
Interstitial lung disease	4443	53 173	29	5.5 (3.7 to 7.7)	1.34 (1.20 to 1.50)	<0.001

Models adjusted for sex, self-reported race/ethnicity, baseline age, smoking status, cigarette pack-years, height, weight, systolic and diastolic blood pressures, total cholesterol, high-density lipoprotein cholesterol, diabetes history, cancer history, coronary artery calcium score, per cent emphysema and total intentional exercise (MET-min/week).
Interstitial lung disease events defined as hospitalisation or death related to interstitial lung disease using ICD-9 and ICD-10 codes (ICD-9 codes 495.XX, 515.XX or 516.XX, and ICD-10 codes J60.X–J64.X, J67.X or J84.X).
ICD-9/10, International Classification of Diseases, 9th Revision/10th Revision; MET, metabolic equivalent of task.

In addition to the *MUC5B* polymorphism, telomere length has been strongly implicated in pulmonary fibrosis. Adults with IPF and other fibrosing ILDs who have shorter telomere lengths have an accelerated decline in their lung function and reduced survival.[9 43] In our study, shorter absolute telomere length was associated with more HAAs from Examination 1 scans, which is consistent with prior ILA studies.[10 44] However, when using different telomere length thresholds, only telomere lengths below the 25th percentile showed a significant association with HAAs. Our findings highlight an important knowledge gap. More research is needed to determine what is a relevant percentile cut-off for telomere length and its relationship to chronic diseases in the general population.

In contrast to the cross-sectional analyses, baseline telomere length was not associated with longitudinal changes in HAAs. A potential reason may be the dynamic nature of telomeres, with lengths that can change over time. In addition to telomerase-related genetic variants (eg, *TERC*, *PARN*), environmental factors influence telomere length, including socioeconomic status (ie, neighbourhood status, home ownership), nutrition, cigarette smoking, dietary supplementation and physical activity.[45–50] We did not have repeat measures of telomere length at each examination in MESA and thus were not able to examine whether longitudinal changes in telomere length and its rate of attrition may be more informative in how cellular senescence relates to longitudinal changes in lung density.

We also found that individuals with the *MUC5B* risk allele and a baseline telomere length below an age-adjusted 5th percentile cut-off had more HAAs over time. It has been speculated that individuals with a genetic predisposition (ie, *MUC5B* risk allele carrier, telomerase mutations, etc), in combination with other environmental risk factors, may increase one's risk of developing pulmonary fibrosis.[51] Previous studies investigating this have largely been case–control in design, with samples comprised of clinically diagnosed patients or adults of relatives with disease.[6 44] It is intriguing to consider the possibility that our finding of an interaction between *MUC5B* and telomere length on progression of HAA, in a prospective, population-based cohort of community-dwelling individuals, may support this hypothesis. However, the subgroup nature of this analysis is exploratory as replication in independent cohorts and investigation of the potential intersection of biological pathways of *MUC5B* and telomere biology are needed.

More consequential may be our finding that increased HAAs over time were associated with worse overall mortality and a higher risk of ILD-related clinical events. This finding suggests that an increase in HAAs may indicate ongoing pathological processes in the lung, which have clinical implications. Increases in HAAs among carriers of the *MUC5B* risk allele had a higher risk of overall death compared with non-carriers. However, this interaction was not statistically significant. In contrast, among patients with IPF, the *MUC5B* risk allele was shown to be associated with improved survival.[52] However, this association has been suggested to be due to index event bias.[53] Further research is needed to understand the complex relationship between *MUC5B*, disease progression and survival in patients with fibrosing ILD.

This study had several limitations. First, histopathological correlations of HAAs are unknown. However, HAAs are associated with ILA, biomarkers of extracellular matrix remodelling and ILD-related outcomes, suggesting that they have construct validity as a measure of parenchymal lung injury and fibrosis.[18 21] Second, one of the concerns with HAAs, and other CT densiometric-based quantitative measures, is potential confounding by body habitus, atelectasis, and artefacts generated from foreign bodies, calcifications, and lung densities unrelated to ILD.[36] Although we cannot completely rule out residual confounding, we demonstrate that one of the strongest genetic risk factors for pulmonary fibrosis (ie, *MUC5B* promoter variant), which does not appear to be influenced by adiposity or has a role in adipogenesis, is strongly associated with more HAAs. Notably, the effect estimate of the association between *MUC5B* and HAAs was comparable with that of cigarette smoking, which is a universally accepted risk factor for ILD.[54] This provides further validity for HAAs as a risk factor for ILD.

Third, we did not have repeated measurements of WGS-extracted telomere length. Although telomere length from WGS correlates with laboratory-based measurements, further research is needed to improve and refine the techniques to extract telomere length from WGS, and obtain consensus on the platforms and relevant cut-offs for broader clinical use. Fourth, we may have been underpowered to detect a significant association in the under-represented racial/ethnic subgroups.[23 55] Fifth, we acknowledge the possibility of survival bias (an issue common to longitudinal population cohort studies), which may have influenced our findings. Finally, there was a very low number of ILD-related events, which is not unexpected for a population cohort without known cardiovascular disease at baseline, and adjudicated ILD diagnoses from outpatient records were not available in this study population. The increasing incidence of ILD that is likely due to growing awareness of this condition and the adjudication of ILD-related outcomes by more cohort studies will be important to determine whether radiological tools can help identify at-risk individuals.[56]

In conclusion, longitudinal increases in HAAs were associated with the *MUC5B* (rs35705950) promoter variant and a higher risk of death and ILD-related events in a population-based cohort of community-dwelling adults. This study demonstrates the feasibility and potential utility of using a quantitative assessment of lung parenchymal changes to investigate intrinsic risk factors for pulmonary fibrosis in a diverse cohort of community-dwelling adults. Quantitative CT assessments, especially more

sophisticated texture-based methods that can more definitively distinguish ILD-specific characteristics from artefact, hold promise for the detection of ILD at its earlier stages and track its progression. Future studies that develop and refine quantitative methods to detect early interstitial lung changes and expand to more diverse study populations will be important in identifying novel factors that contribute to the development and progression of ILD.

Author affiliations
[1]Department of Medicine, University of Virginia School of Medicine, Charlottesville, Virginia, USA
[2]Department of Medicine, Columbia University Irving Medical Center, New York, New York, USA
[3]Department of Public Health Sciences, University of Virginia School of Medicine, Charlottesville, Virginia, USA
[4]Center for Public Health Genomics, University of Virginia School of Medicine, Charlottesville, Virginia, USA
[5]Department of Radiology, University of Iowa, Iowa City, Iowa, USA
[6]Department of Medicine, Perelman School of Medicine University of Pennsylvania, Philadelphia, Pennsylvania, USA
[7]Department of Biostatistics and Epidemiology, Perelman School of Medicine University of Pennsylvania, Philadelphia, Pennsylvania, USA
[8]Department of Biomedical Engineering, Columbia University, New York, New York, USA
[9]Department of Medicine, University of Chicago, Chicago, Illinois, USA
[10]Department of Biostatistics, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
[11]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[12]Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, California, USA
[13]The Institute for Translational Genomics and Population Sciences, The Lundquist Institute, Harbor-UCLA Medical Center, Torrance, California, USA
[14]Department of Epidemiology, Columbia University Mailman School of Public Health, New York, New York, USA
[15]Division of Pulmonary and Critical Care, Weill Cornell Medical College, New York, New York, USA

Twitter John S Kim @jskim8223, Ayodeji Adegunsoye @drdayjee and Anna J Podolanczuk @AnnaPodolanczuk

Acknowledgements The authors thank the other investigators, staff and participants of the MESA Study for their contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. We gratefully acknowledge the studies and participants who provided biological samples and data for MESA and TOPMed.

Contributors Conception and design of the study—JSK, AWM and AP. Data acquisition—AWM, SSR, JR, MT, RM, EAH and RGB. Analysis of the data—JSK and AP. JSK drafted the initial manuscript. All authors contributed to the data interpretation and edited the manuscript for important scientific content. All of the authors agree to be accountable for all aspects of the work in regard to accuracy and integrity. JSK accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Funding JSK was supported by the Pulmonary Fibrosis Foundation Scholars Award and grant K23-HL-150301 from the National Heart, Lung and Blood Institute (NHLBI). MRA was supported by grant K23-HL-150280, AP was supported by grant K23-HL-140199, SK was supported by grant K24-HL-103844, and AWM was supported by R01-HL131565 from the NHLBI. EB was supported by grant K23-AR-075112 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases. The MESA Lung Study was supported by grants R01-HL077612, R01-HL093081 and RC1-HL100543 from the NHLBI. The MESA Lung Fibrosis Study was funded by grants R01-HL-103676 from the NHLBI. The MESA Project is conducted and supported by the NHLBI in collaboration with MESA investigators. Support for MESA is provided by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420. Also supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. Whole-genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) programme was supported by the NHLBI. WGS for 'NHLBI TOPMed: Multi-Ethnic Study of Atherosclerosis (MESA)' (phs001416.v1.p1) was performed at the Broad Institute of MIT and Harvard (3U54HG003067-13S1). Core support including centralised genomic read mapping and genotype calling, along with variant quality metrics and filtering, were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1; contract HHSN268201800002I). Core support including phenotype harmonisation, data management, sample-identity QC and general programme coordination were provided by the TOPMed Data Coordinating Center (R01HL-120393; U01HL-120393; contract HHSN268201800001I). Infrastructure for the CHARGE Consortium is supported in part by the NHLBI grant R01-HL105756.

Competing interests EAH is the founder of VIDA Diagnostics, a company that is commercialising lung image analysis software developed, in part, at the University of Iowa. SK reports receipt of grant support from the NIH and Cardiovascular Medical Research and Education Fund and payments/honoria from Children's Hospital of Philadelphia, University of California, San Francisco, University of Minnesota, PVRI and Stanford University, and support for attending meetings from the Pulmonary Hypertension Association, American Thoracic Society study section, Duke University, World Symposium on Pulmonary Hypertension, PVRI, Children's Hospital of Philadelphia, University of Miami, Wake Forest University, Aspen Lung Conference and NIH Study Section. SK reports participation on a data safety monitoring board for United Therapeutics and editorial board for the European Respiratory Journal. IN reports receipt of personal fees and other support from Boehringer Ingelheim, HLR/Genentech, Sanofi Aventis, Global Blood Therapeutics and Veracyte outside the scope of this work. CKG reports support from the Department of Defense, Boehringer Ingelheim and AstraZeneca outside the scope of this work. RGB reports receipt of grants from the Alpha-1 Foundation and the COPD Foundation outside the scope of this work. AP reports grant support from the NHLBI and American Lung Association and consulting fees from Regeneron and Boehringer Ingelheim outside the submitted work.

Patient consent for publication Not required.

Ethics approval This study involves human participants and institutional review board (IRB) approval was obtained at all participating study sites (IRB# STU00021057, Northwestern University; IRB# NA_00030361_AM00004951, Johns Hopkins University; IRB# 99-11-057-23B, University of California Los Angeles; IRB# 9805M00034, University of Minnesota; IRB# BG00-035, Wake Forest University; IRB-AAAA7791, Columbia University). All MESA participants provided written informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data available upon request if study approved by MESA.

ORCID iDs
John S Kim http://orcid.org/0000-0002-8887-150X
Ani W Manichaikul http://orcid.org/0000-0002-5998-795X
Ayodeji Adegunsoye http://orcid.org/0000-0002-7015-9610
Christine Kim Garcia http://orcid.org/0000-0002-0771-1249
Anna J Podolanczuk http://orcid.org/0000-0002-9559-1485

REFERENCES

1 Wijsenbeek M, Cottin V. Spectrum of fibrotic lung diseases. *N Engl J Med* 2020;383:958–68.
2 Raghu G, Chen S-Y, Yeh W-S, *et al*. Idiopathic pulmonary fibrosis in US medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001-11. *Lancet Respir Med* 2014;2:566–72.
3 du Bois RM, Weycker D, Albera C, *et al*. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *Am J Respir Crit Care Med* 2011;184:1382–9.
4 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;8:726–37.
5 Seibold MA, Wise AL, Speer MC, *et al*. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503–12.
6 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;5:639–47.
7 Juge P-A, Lee JS, Ebstein E, *et al*. MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. *N Engl J Med* 2018;379:2209–19.

8 Hunninghake GM, Hatabu H, Okajima Y, *et al*. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013;368:2192–200.
9 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;2:557–65.
10 Putman RK, Axelsson GT, Ash SY, *et al*. Interstitial lung abnormalities are associated with decreased mean telomere length. *Eur Respir J* 2022. doi:10.1183/13993003.01814-2021. [Epub ahead of print: 03 Feb 2022].
11 Dressen A, Abbas AR, Cabanski C, *et al*. Analysis of protein-altering variants in telomerase genes and their association with MUC5B common variant status in patients with idiopathic pulmonary fibrosis: a candidate gene sequencing study. *Lancet Respir Med* 2018;6:603–14.
12 Bild DE, Bluemke DA, Burke GL, *et al*. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 2002;156:871–81.
13 Taliun D, Harris DN, Kessler MD, *et al*. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. *Nature* 2021;590:290–9.
14 Taub MA, Conomos MP, Keener R, *et al*. Genetic determinants of telomere length from 109,122 ancestrally diverse whole-genome sequences in TOPMed. *Cell Genom* 2022;2. doi:10.1016/j.xgen.2021.100084. [Epub ahead of print: 12 Jan 2022].
15 Sieren JP, Newell JD, Barr RG, *et al*. SPIROMICS protocol for multicenter quantitative computed tomography to phenotype the lungs. *Am J Respir Crit Care Med* 2016;194:794–806.
16 Hoffman EA, Jiang R, Baumhauer H, *et al*. Reproducibility and validity of lung density measures from cardiac CT scans--the multi-ethnic study of atherosclerosis (MESA) lung study. *Acad Radiol* 2009;16:689–99.
17 Wang M, Aaron CP, Madrigano J, *et al*. Association between long-term exposure to ambient air pollution and change in quantitatively assessed emphysema and lung function. *JAMA* 2019;322:546–56.
18 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.
19 Hoffman EA, Reinhardt JM, Sonka M, *et al*. Characterization of the interstitial lung diseases via density-based and texture-based analysis of computed tomography images of lung structure and function. *Acad Radiol* 2003;10:1104–18.
20 Barr RG, Bluemke DA, Ahmed FS, *et al*. Percent emphysema, airflow obstruction, and impaired left ventricular filling. *N Engl J Med* 2010;362:217–27.
21 Podolanczuk AJ, Oelsner EC, Barr RG, *et al*. High-Attenuation areas on chest computed tomography and clinical respiratory outcomes in community-dwelling adults. *Am J Respir Crit Care Med* 2017;196:1434–42.
22 Waxman A, Restrepo-Jaramillo R, Thenappan T, *et al*. Inhaled treprostinil in pulmonary hypertension due to interstitial lung disease. *N Engl J Med* 2021;384:325–34.
23 Peljto AL, Selman M, Kim DS, *et al*. The MUC5B promoter polymorphism is associated with idiopathic pulmonary fibrosis in a Mexican cohort but is rare among Asian ancestries. *Chest* 2015;147:460–4.
24 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. doi:10.1183/13993003.00537-2017. [Epub ahead of print: 11 09 2017].
25 Alder JK, Chen JJ-L, Lancaster L, *et al*. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proc Natl Acad Sci U S A* 2008;105:13051–6.
26 Alter BP, Rosenberg PS, Giri N, *et al*. Telomere length is associated with disease severity and declines with age in dyskeratosis congenita. *Haematologica* 2012;97:353–9.
27 Newton CA, Zhang D, Oldham JM, *et al*. Telomere length and use of immunosuppressive medications in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;200:336–47.
28 Schneider CV, Schneider KM, Teumer A. Association of telomere length with risk of disease and mortality. *JAMA Intern Med* 2022.
29 Yang IV, Fingerlin TE, Evans CM, *et al*. MUC5B and idiopathic pulmonary fibrosis. *Ann Am Thorac Soc* 2015;12 Suppl 2:S193–9.
30 Chen Y, Postmus D, Cowie MR, *et al*. Using joint modelling to assess the association between a time-varying biomarker and a survival outcome: an illustrative example in respiratory medicine. *Eur Respir J* 2021;57:2003206.
31 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;201:564–74.
32 Nakano Y, Yang IV, Walts AD, *et al*. MUC5B promoter variant rs35705950 affects MUC5B expression in the distal airways in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2016;193:464–6.
33 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;200:1402–13.
34 Mathai SK, Humphries S, Kropski JA, *et al*. MUC5B variant is associated with visually and quantitatively detected preclinical pulmonary fibrosis. *Thorax* 2019;74:1131–9.
35 Ash SY, Harmouche R, Putman RK, *et al*. Clinical and genetic associations of objectively identified interstitial changes in smokers. *Chest* 2017;152:780–91.
36 Kliment CR, Araki T, Doyle TJ. A comparison of visual and quantitative methods to identify interstitial lung abnormalities. *BMC Pulm Med* 2015;15:134.
37 Manichaikul A, Wang X-Q, Sun L, *et al*. Genome-wide association study of subclinical interstitial lung disease in MESA. *Respir Res* 2017;18:97.
38 Buendía-Roldán I, Fernandez R, Mejía M, *et al*. Risk factors associated with the development of interstitial lung abnormalities. *Eur Respir J* 2021;58. doi:10.1183/13993003.03005-2020. [Epub ahead of print: 05 08 2021].
39 Adegunsoye A, Oldham JM, Bellam SK, *et al*. African-merican race and mortality in interstitial lung disease: a multicentre propensity-matched analysis. *Eur Respir J* 2018;51. doi:10.1183/13993003.00255-2018. [Epub ahead of print: 14 06 2018].
40 Zhao X, Qiao D, Yang C, *et al*. Whole genome sequence analysis of pulmonary function and COPD in 19,996 multi-ethnic participants. *Nat Commun* 2020;11:5182.
41 Burkart KM, Sofer T, London SJ, *et al*. A genome-wide association study in hispanics/latinos identifies novel signals for lung function. The hispanic community health study/study of latinos. *Am J Respir Crit Care Med* 2018;198:208–19.
42 Zhang D, Povysil G, Newton CA, *et al*. Genome-wide enrichment of TERT rare variants in IPF patients of Latino ancestry. *Am J Respir Crit Care Med* 2022. doi:10.1164/rccm.202203-0622LE. [Epub ahead of print: 06 Jun 2022].
43 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.
44 Salisbury ML, Hewlett JC, Ding G, *et al*. Development and progression of radiologic abnormalities in individuals at risk for familial interstitial lung disease. *Am J Respir Crit Care Med* 2020;201:1230–9.
45 Garrett-Bakelman FE, Darshi M, Green SJ, *et al*. The NASA twins study: a multidimensional analysis of a year-long human spaceflight. *Science* 2019;364. doi:10.1126/science.aau8650. [Epub ahead of print: 12 04 2019].
46 Stuart BD, Choi J, Zaidi S, *et al*. Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet* 2015;47:512–7.
47 Needham BL, Carroll JE, Diez Roux AV, *et al*. Neighborhood characteristics and leukocyte telomere length: the multi-ethnic study of atherosclerosis. *Health Place* 2014;28:167–72.
48 Kiecolt-Glaser JK, Epel ES, Belury MA, *et al*. Omega-3 fatty acids, oxidative stress, and leukocyte telomere length: a randomized controlled trial. *Brain Behav Immun* 2013;28:16–24.
49 Cherkas LF, Hunkin JL, Kato BS, *et al*. The association between physical activity in leisure time and leukocyte telomere length. *Arch Intern Med* 2008;168:154–8.
50 Astuti Y, Wardhana A, Watkins J, *et al*. Cigarette smoking and telomere length: a systematic review of 84 studies and meta-analysis. *Environ Res* 2017;158:480–9.
51 Wolters PJ, Collard HR, Jones KD. Pathogenesis of idiopathic pulmonary fibrosis. *Annu Rev Pathol* 2014;9:157–79.
52 Peljto AL, Zhang Y, Fingerlin TE, *et al*. Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013;309:2232–9.
53 Dudbridge F, Allen RJ, Sheehan NA, *et al*. Adjustment for index event bias in genome-wide association studies of subsequent events. *Nat Commun* 2019;10:1561.
54 Baumgartner KB, Samet JM, Stidley CA, *et al*. Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1997;155:242–8.
55 Wang C, Zhuang Y, Guo W, *et al*. Mucin 5B promoter polymorphism is associated with susceptibility to interstitial lung diseases in Chinese males. *PLoS One* 2014;9:e104919.
56 Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, *et al*. Trends and patterns of differences in chronic respiratory disease mortality among US counties, 1980-2014. *JAMA* 2017;318:1136–49.