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 Madhwar,2 Elizabeth C Oelsner,2 Steven M Kayut,6,7 Artur Wysoczanski,2 Andrew F Laine,8 Ayodeji Adegunsoye,9 Jennie Z Ma,9 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,1,2 R Graham Barr,2,14 Anna J Podolanczuk15

**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 percent 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 HAA.

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 1.51% to 3.68%) 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 5% increase in mortality risk (rate ratio (RR)=1.10, 95% CI 1.01 to 1.19) for overall death and 16% increase in ILD (RR=1.16, 95% CI 1.01 to 1.33). 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.

**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.\(^\text{4,5}\) 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.\(^\text{6}\) 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.\(^\text{6,9-10}\) 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.\(^\text{11}\)

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.\(^\text{12}\) 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.\(^\text{13}\) Briefly, Illumina HiSeq X Ten instruments were used for WGS that targeted a mean depth of at least 30x (paired-end, 150bp 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.\(^\text{14}\) 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/SPRIMICS protocol.\(^\text{15}\) 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.\(^\text{16}\) 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.\(^\text{17}\) HAA was defined as the percentage of voxels with attenuation values in the range of −600 and −250 Hounsfield units (HU), as previously described.\(^\text{18}\) The Pulmonary Analysis Software Suite at the University of Iowa’s Advanced Pulmonary Physiomic Imaging Laboratory was used by trained technicans for segmentation and correction of the lung.\(^\text{19}\) Per cent emphysema was defined as the percentage of lung voxels below −950 HU.\(^\text{20}\)

**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.\(^\text{21}\) 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.\(^\text{2}\) 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.\(^\text{17,22}\) 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.\(^\text{23,24}\) 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.\(^\text{8}\) 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.\(^\text{25-26}\) 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. Sex was also examined as a potential effect modifier due to the differential associations of sex and pulmonary fibrosis risk. 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.

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. 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 V9.4 (SAS Institute). The ‘nlme’, ‘survival’ and ‘JMbayes2’ packages were used for the joint model analysis in R V3.6.1 (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 4532 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>
<thead>
<tr>
<th>Characteristic</th>
<th>Overall</th>
<th>GG genotype</th>
<th>GT genotype</th>
<th>TT genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>4552</td>
<td>3959</td>
<td>569</td>
<td>24</td>
</tr>
<tr>
<td>Age</td>
<td>61 (10)</td>
<td>61 (10)</td>
<td>61 (10)</td>
<td>60 (9)</td>
</tr>
<tr>
<td>Female sex</td>
<td>51%</td>
<td>51%</td>
<td>53%</td>
<td>50%</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>41%</td>
<td>37%</td>
<td>68%</td>
<td>75%</td>
</tr>
<tr>
<td>Asian</td>
<td>13%</td>
<td>15%</td>
<td>2%</td>
<td>0</td>
</tr>
<tr>
<td>African-American</td>
<td>24%</td>
<td>25%</td>
<td>10%</td>
<td>4%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>22%</td>
<td>23%</td>
<td>20%</td>
<td>21%</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoker</td>
<td>46%</td>
<td>47%</td>
<td>40%</td>
<td>21%</td>
</tr>
<tr>
<td>Former smoker</td>
<td>40%</td>
<td>39%</td>
<td>46%</td>
<td>75%</td>
</tr>
<tr>
<td>Current smoker</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>Cigarette pack-years</td>
<td>12 (22)</td>
<td>11.5 (21.2)</td>
<td>13.9 (23.4)</td>
<td>19.4 (30.0)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>167 (10)</td>
<td>167 (10)</td>
<td>168 (10)</td>
<td>169 (7)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>79 (17)</td>
<td>78 (17)</td>
<td>80 (18)</td>
<td>86 (21)</td>
</tr>
<tr>
<td>Telomere length (kb)</td>
<td>4.4 (0.9)</td>
<td>4.4 (0.9)</td>
<td>4.4 (0.9)</td>
<td>4.1 (0.9)</td>
</tr>
<tr>
<td>Continuous variables presented as mean (SD). Categorical variables presented as percentage.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 6.69% (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
Interstitial lung disease

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

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

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.
suggesting its pathogenic role in pulmonary fibrosis.8 33 Ours is linked to a higher prevalence of ILA and its progression, further inhaled materials and alveolar repair upon injury, and has been expression of MUC5B impairs mucosal host defences, clearance of higher risk of pulmonary fibrosis.5 31

Also adjusted for time-varying covariates height, weight, per cent emphysema and cigarettes smoked per day. **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.

Established relationships 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.20 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

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

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 5D increment of log-transformed telomere length.
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. In our study, shorter absolute telomere length was associated with more HAAs from Examination 1 scans, which is consistent with prior ILA studies. 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.

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. Previous studies investigating this have largely been case-control in design, with samples comprised of clinically diagnosed patients or adults of relatives with disease. 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. However, this association has been suggested to be due to index event bias. 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. Second, one of the concerns with HAAs, and other CT densimetric-based quantitative measures, is potential confounding by body habitus, atelectasis, and artefacts generated from foreign bodies, calcifications, and lung densities unrelated to ILD. 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. 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. 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.

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

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.

Concepts
Conception and design of the study—JSK, AWM and AP. Data acquisition—AWM, SSR, JR, 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 RO1 HL-131565 from the NHLBI. EB was supported by grant K23 ES-075112 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases. The MESA Lung Study was supported by grants R01 HL-077612, R01 HL-093081 and R01 HL-100543 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 5N9200200001, HHSN268201500033, N01-RC-95159, 7SN9200200006, N01-RC-95163, 7SN9200200004, N01-RC-95164, 7SN9200200007, N01-RC-95165, N01-RC-95166, N01-RC-95167, N01-RC-95168, N01-RC-95169, Ul1-TR-000400, Ul1-TR-001079, and Ul1-TR-001420. Also supported in part by the National Center for Advancing Translational Sciences, CTSA grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK083931 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 (35US4HG03067:1351). 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-117626-02S; contract HHSN268201800002). 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 HHSN268201800001). Infrastructure for the CHARGE Consortium is supported in part by the NHLBI grant R01-HL105756.

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_00010361, Johns Hopkins University; IRB# 99-11-057-238, University of California Los Angeles; IRB# 9805M0034, University of Minnesota; IRB# BGO-035, Wake Forest University; IRB# AAA7791, 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.

Supplemental material
This content has been supplied by the author(s), it has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error or omission arising from translation and adaptation or otherwise.

ORCID iDs
John S Kim http://orcid.org/0000-0002-8887-150X
Ani W Manichaikul http://orcid.org/0000-0002-5998-795X
John S Kim 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
