MUC5B, Telomere Length, and Longitudinal Quantitative Interstitial Lung Changes: the MESA Lung Study

John S. Kim, Ani Manichaikul, Eric A. Hoffman, Pallavi Balte, Michaela R. Anderson, Elana J. Bernstein, Purnema Madahar, Elizabeth C. Oelsner, Steven M. Kawut, Artur Wysoczanski, Andrew F. Laine, Ayodeji Adegunsoye, Jennie Z. Ma, Margaret A. Taub, Rasika A. Mathias, Stephen S. Rich, Jerome I. Rotter, Imre Noth, Christine Kim Garcia, R. Graham Barr, Anna J. Podolanczuk

SUPPLEMENTARY APPENDIX

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table of Contents

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Adjudication of events in MESA

A two-member adjudication panel reviewed inpatient medical records and death certificates of participants who were hospitalized with one of the following International Classification of Diseases, Ninth Revision (ICD-9), ILD diagnosis codes (ICD-9 495.XX, 515.XX, or 516.XX) or ICD-10 codes (J60.X-J64.X, J67.X, or J84.X). Each member of the panel reviewed half of the records. If there was uncertainty, the panel members came to a consensus on the diagnosis and event classification after reviewing the case together. A death related to ILD was based on either inpatient records that indicated the cause of death to be respiratory-related or ILD being listed as the primary cause of death in the death certificate.

Longitudinal high attenuation areas analysis

We used linear mixed effects models to examine the associations of independent variables (i.e., *MUC5B* risk allele, cigarettes smoked per day, baseline telomere length) with longitudinal changes in high attenuation areas (HAAs). Linear mixed effects models are used to account for repeated measurements over time, which in this analysis was HAAs. This modeling approach has been employed in several observational cohort studies as well as clinical trials[1-3]. Below is the formula we used in which high attenuation areas is the outcome (Y) and our primary independent variable of interest is the number of *MUC5B* risk alleles an individual carries and we use random intercept and slope. The effect estimate of interest was the term,

"$\beta_{\text{MUC5B risk allele} \times \text{time since initial HAA assesment}}$" which is reported in the tables and manuscript.

$$Y_{HAA} = \beta_0 + \beta_{MUC5B\ \text{risk allele}} + \beta_{\text{time since initial HAA assessment}} +$$
$$\beta_{\text{MUC5B risk allele} \times \text{time since initial HAA assesment}} + \beta_{\text{scanner parameters}} + \beta_{\text{baseline age}} +$$
$$\beta_{\text{baseline age} \times \text{time since initial HAA assessment}} + \beta_{\text{sex}} + \beta_{\text{sex} \times \text{time since initial HAA assessment}} +$$

$$\beta_{\text{baseline smoking status}} + \beta_{\text{baseline smoking status} \times \text{time since initial HAA assessment}} +$$

$$\beta_{\text{baseline cigarette pack-years}} + \beta_{\text{baseline cigarette pack-years} \times \text{time since initital HAA assessment}} +$$

$$\beta_{\text{race/ethnicity}} + \beta_{\text{race/ethnicity} \times \text{time since initial HAA assessment}} + \beta_{\text{percent emphysema}} +$$

$$\beta_{\text{percent emphsyema} \times \text{time since initial HAA assessment}} + \beta_{\text{height (time-varying)}} +$$

$$\beta_{\text{weight (time-varying)}} + \beta_{\text{cigarettes smoked per day (time-varying)}} +$$

$$\beta_{\text{principal components of genetic ancestry}}$$

Event Analysis

We used joint modeling to examine the associations of longitudinal changes in high attenuation areas (HAAs) and our clinical outcomes of interest. Joint modeling is a well-described statistical procedure to examine the relationship between covariate data that is repeatedly measured over time (HAAs for this analysis) and time-to-event data (i.e., overall death and interstitial lung disease-related hospitalization and death).[4] This procedure is considered more rigorous compared with using a Cox regression model in which the primary independent variable of interest is a time-varying covariate. We first generated a linear mixed effects model that examines the change in HAAs over time with adjustments for scanner parameters using the "nlme" package from R code (R statistical, Vienna, Austria) and the formula below:

$$Y_{HAAs} = \beta_0 + \beta_{\text{time since initial HAAs assessment}} + \beta_{\text{scanner parameters}} + \beta_{\text{percent emphysema}} + \beta_{\text{percent emphsyema} \times \text{time since initial HAA assessment}}$$

In the linear mixed effects model, the covariates included time since initial HAAs assessment, scanner parameters, percent emphysema and its interaction with time since initial HAAs assessment.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

We then generated a Cox regression model to examine associations of pertinent covariates that include confounders with our clinical outcomes of interest which were overall death and a composite outcome of interstitial lung disease-related death/hospitalization. We used the "survival" package from R code.

$$Y_{Death} = \beta_0 + \beta_{\text{baseline age}} + \beta_{\text{sex}} + \beta_{\text{baseline smoking history}} + \beta_{\text{baseline cigarette pack-years}} + \beta_{\text{race/ethnicity}} + \beta_{\text{percent emphysema}} + \beta_{\text{baseline height}} + \beta_{\text{baseline weight}} + \beta_{\text{principal components of genetic ancestry}} + \beta_{\text{baseline total cholesterol level}} + \beta_{\text{baseline high-density lipoprotein level}} + \beta_{\text{baseline systolic blood pressure}} + \beta_{\text{baseline diastolic blood pressure}} + \beta_{\text{diabetes history}} + \beta_{\text{cancer history}} + \beta_{\text{baseline coronary artery calcium Agatston score}} + \beta_{\text{baseline total intentional exercise}}$$

For the Cox regression model for both death and ILD-related events as outcomes, we adjusted for the following: baseline age, cigarette pack-years, smoking history, height, weight, total cholesterol level, high-density lipoprotein cholesterol level, systolic blood pressure, diastolic blood pressure, history of diabetes, history of cancer, coronary artery calcium Agatston score, total intentional exercise (met-min/week), sex, race/ethnicity, and principal components of genetic ancestry. The joint modeling procedure then links these two models and examine the associations of longitudinal change in HAAs with our clinical outcome of interest by incorporating the random effects of HAAs in the Cox regression analysis. We used the "JMbayes2" to perform the joint modeling."

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S1 *MUC5B* (rs35705950) genotypes by self-reported race/ethnicity in MESA

	***MUC5B* (rs35705950) Genotype**		
Race/Ethnicity	**GG**	**GT**	**TT**
Non-Hispanic White (n=1,864)	79%	20%	1%
Asian (n=601)	96%	4%	<1%
African-American (=1,066)	94%	5%	<1%
Hispanic (n=1,021)	89%	11%	<1%

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S2 Study sample characteristics by age-adjusted telomere length percentile

	Age-adjusted telomere length percentiles					
	<5th Percentile	**≥5th Percentile**	**<10th Percentile**	**≥10th Percentile**	**<25th Percentile**	**≥25th Percentile**
No. Participants	221	4,267	446	4,042	1,119	3,369
Telomere length (kb)	2.87 (0.23)	4.46 (0.87)	3.05 (0.26)	4.53 (0.84)	3.35 (0.34)	4.72 (0.78)
Age	61 (10)	61 (10)	61 (10)	61 (10)	61 (10)	61 (10)
Female sex	45%	52%	44%	52%	46%	53%
Race/ethnicity						
Non-Hispanic white	41%	41%	42%	41%	44%	40%
Asian	7%	14%	10%	14%	10%	14%
African-American	30%	23%	25%	23%	23%	24%
Hispanic	22%	22%	23%	22%	23%	22%
Smoking Status						
Never smoker	42%	46%	40%	46%	42%	47%
Former smoker	40%	41%	45%	40%	42%	40%
Current smoker	18%	13%	15%	14%	16%	13%
Cigarette pack-years	12 (19)	12 (22)	13 (21)	12 (22)	13 (22)	11 (21)
Height (cm)	169 (11)	167 (10)	168 (10)	167 (10)	168 (10)	167 (10)
Weight (kg)	81 (17)	78 (17)	81 (17)	78 (17)	80 (17)	78 (17)

Continuous variables presented as mean (standard deviation)
Categorical variables presented as percentage

Table S3 Number of CT scans performed at each exam for high attenuation areas longitudinal analysis

Exam	Overall	Non-Hispanic White	Asian	African-American	Hispanic
1 (2000-2002)	4,552	1,864	601	1,066	1,021
2 (2002-2004)	1,994	796	284	454	460
3 (2004-2005)	2,101	867	287	475	472
4 (2005-2007)	888	356	118	158	256
5 (2010-2012)	2,029	769	326	472	462
6 (2016-2018)	1,858	738	275	433	412

CT scans from Exams 1 through 5 are cardiac. Exam 6 scans are full-lung.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S4 Associations of cigarette smoking with longitudinal changes in high attenuation areas stratified by sex and smoking history

Model	No. Participants	% Longitudinal change in HAAs over 10 years (95% CI) per 10 cigarettes smoked per day	P-value
Overall	4,552	4.69 (2.24 to 7.15)	<0.001

Abbreviations: CI=confidence intervals; HAAs=high attenuation areas

Overall model is adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, and self-reported race/ethnicity were also adjusted for including their interaction terms with "time since initial HAAs assessment." Time-varying covariates height, weight, and percent emphysema also adjusted for in the model. Results are reported per 10 cigarettes smoked per day on average.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S5 Associations of *MUC5B* (rs35705950) risk allele with longitudinal changes in high attenuation areas stratified by sex and smoking history

Model	No. Participants	% Longitudinal change in HAAs over 10 years (95% CI) per *MUC5B* (rs35705950) risk allele (T)	P-value for interaction
Sex			0.70
Female	2,343	3.05 (0.04 to 6.06)	
Male	2,209	2.18 (-1.11 to 5.47)	
Smoking History			0.91
Never	2,081	2.49 (-1.01 to 6.01)	
Ever	2,471	2.75 (-0.12 to 5.63)	

Abbreviations: CI=confidence intervals; HAA=high attenuation areas

Model is adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, self-reported race/ethnicity, smoking status, cigarette pack-years were also adjusted for including their interaction terms with "time since initial HAAs assessment." Time-varying covariates height, weight, percent emphysema, and cigarettes smoked per day were also adjusted for in the model. Three-way interaction term of effect modifier, *MUC5B* risk allele, and time since initial HAAs assessment also included (e.g., "sex × *MUC5B* risk allele × time since initial HAAs assessment").

All results are reported per risk allele (T) of the *MUC5B* (rs35705950) promoter variant.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S6 Associations of *MUC5B* (rs35705950) risk allele with longitudinal changes in high attenuation areas by self-reported race/ethnicity

Model	Non-Hispanic White			Asian		
	No. Participants	% Longitudinal change in HAAs (95% CI) per *MUC5B* (rs35705950) risk allele (T)	P-value	No. Participants	% Longitudinal change in HAAs (95% CI) per *MUC5B* (rs35705950) risk allele (T)	P-value
Overall	1,864	3.05 (0.64 to 5.46)	0.01	601	5.89 (-6.72 to 18.65)	0.36
Stratified						
Telomere Length						
Fifth percentile cutoff						
Below 5th percentile*	90	15.31 (12.78 to 17.83)	0.02	27	N/A	N/A
Above 5th percentile	1,747	2.21 (-7.94 to 12.47)		568	5.75 (-6.84 to 18.50)	
Tenth percentile cutoff						
Below 10th percentile	189	5.48 (2.89 to 8.08)	0.44	57	N/A	N/A
Above 10th percentile	1,648	2.54 (-4.42 to 9.55)		538	5.89 (-7.01 to 18.34)	
Twenty-fifth percentile cutoff						
Below 25th percentile	457	7.02 (4.91 to 9.13)	0.004	145	5.74 (-21.54 to 33.78)	0.98
Above 25th percentile	1,380	1.49 (-1.68 to 4.67)		450	6.24 (-7.91 to 20.59)	
Sex			0.23			0.78
Male	912	1.46 (-1.79 to 4.73)		302	3.39 (-14.31 to 21.40)	
Female	952	4.43 (0.87 to 8.01)		299	6.99 (-9.76 to 24.02)	
Smoking History			0.59			0.91
Ever	1,115	3.61 (0.59 to 6.64)		177	7.37 (-13.57 to 28.76)	
Never	749	2.24 (-1.72 to 6.21)		424	5.83 (-9.92 to 21.84)	

Abbreviations: CI=confidence intervals; HAA=high attenuation areas

Model is adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, smoking status, cigarette pack-years were also adjusted for including their interaction terms with "time since initial HAAs assessment." Time-varying covariates height, weight, percent emphysema, and cigarettes smoked per day were also adjusted for in the model. Three-way interaction term of effect modifier, *MUC5B* risk allele, and time since initial HAAs assessment also included (e.g., "sex × *MUC5B* risk allele × time since initial HAAs assessment").

P-values for stratified analysis represent p-values for interaction.

All results are reported per risk allele (T) of the *MUC5B* (rs35705950) promoter variant

*Due to lower number of Asian participants, unable to perform stratified analysis of age-adjusted telomere 5% and 10% percentile cutoffs.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S6 Associations of *MUC5B* (rs35705950) risk allele with longitudinal changes in high attenuation areas by self-reported race/ethnicity (continued)

Model	African-American No. Participants	African-American % Longitudinal change in HAAs (95% CI) per *MUC5B* (rs35705950) risk allele (T)	African-American P-value	Hispanic No. Participants	Hispanic % Longitudinal change in HAAs (95% CI) per *MUC5B* (rs35705950) risk allele (T)	Hispanic P-value
Overall	1,066	0.32 (-7.68 to 8.39)	0.94	1,021	2.52 (-2.78 to 7.82)	0.35
Stratified						
Telomere Length						
Fifth percentile cutoff						
Below 5th percentile*	51	0.48 (-64.57 to 70.05)	0.98	48	18.87 (-4.51 to 42.79)	0.17
Above 5th percentile	1,005	1.23 (-7.00 to 9.52)		952	1.82 (-3.68 to 7.36)	
Tenth percentile cutoff						
Below 10th percentile	106	-7.20 (-31.57 to 17.78)	0.48	98	11.12 (-8.16 to 30.77)	0.38
Above 10th percentile	950	2.16 (-6.48 to 10.88)		902	1.99 (-3.59 to 7.61)	
Twenty-fifth percentile cutoff						
Below 25th percentile	262	-3.27 (-19.35 to 13.06)	0.55	247	10.23 (-2.24 to 22.86)	0.19
Above 25th percentile	794	2.42 (-7.00 to 11.93)		753	0.91 (-5.02 to 6.88)	
Sex			0.44			0.35
Male	499	4.50 (-9.36 to 18.55)		496	4.99 (-2.41 to 12.45)	
Female	567	-2.15 (-11.92 to 7.72)		525	-0.04 (-7.57 to 7.54)	
Smoking History			0.30			0.57
Ever	645	2.43 (-7.19 to 12.14)		534	0.96 (-6.68 to 8.65)	
Never	421	-6.94 (-21.65 to 8.00)		487	4.05 (-3.25 to 11.41)	

Abbreviations: CI=confidence intervals; HAAs=high attenuation areas

Model is adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, smoking status, cigarette pack-years were also adjusted for including their interaction terms with "time since initial HAAs assessment." Time-varying covariates height, weight, percent emphysema, and cigarettes smoked per day were also adjusted for in the model. Three-way interaction term of effect modifier, *MUC5B* risk allele, and time since initial HAAs assessment also included (e.g., "sex × *MUC5B* risk allele × time since initial HAAs assessment").

P-values for stratified analysis represent p-values for interaction.

All results are reported per risk allele (T) of the *MUC5B* (rs35705950) promoter variant

*Due to lower number of Asian participants, unable to perform stratified analysis of age-adjusted telomere 5% and 10% percentile cutoffs.

Table S7 Baseline telomere length with high attenuation areas stratified analysis

Model	No. Participants	Mean percent change in Exam 1 HAAs (95% CI)	P-value for interaction	% Longitudinal change in HAAs over 10 years (95% CI)	P-value for interaction
Sex			0.94		0.76
Female	2,317	-6.38 (-11.41 to -1.06)		1.04 (-5.69 to 7.81)	
Male	2,171	-6.65 (-11.89 to -1.09)		-0.50 (-7.66 to 6.72)	
Smoking Status			0.81		0.27
Never smoker	2,435	3.08 (-2.87 to 9.38)		3.00 (-4.09 to 10.16)	
Ever smoker	2,053	-11.51 (-16.32 to -6.43)		-2.43 (-9.21 to 4.40)	
Race/ethnicity			0.99		0.86
Non-Hispanic White	1,837	-0.98 (-1.98 to 0.04)		0.01 (-1.30 to 1.31)	
Asian	595	-1.21 (-2.60 to 0.19)		0.50 (-1.19 to 2.20)	
African-American	1,056	-1.22 (-2.53 to 0.10)		-0.38 (-2.02 to 1.26)	
Hispanic	1,000	-1.15 (-2.94 to 0.67)		0.58 (-1.52 to 2.69)	

Abbreviations: CI=confidence intervals; HAAs=high attenuation areas

*Reported per standard deviation increment of log-transformed telomere length

Exam 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 percent emphysema. Two-way interaction term of effect modifier and telomere length also included (e.g., "sex × telomere length").

Longitudinal HAAs model: Adjusted for scanner parameters and principal components of genetic ancestry. Baseline age, sex, self-reported race/ethnicity, smoking status, cigarette pack-years were also adjusted for including their interaction terms with "time since initial HAAs assessment." Time-varying covariates height, weight, percent emphysema, and cigarettes smoked per day were also adjusted for in the model. Three-way interaction term of effect modifier, telomere length, and time since initial HAAs assessment also included (e.g., "sex × telomere length × time since initial HAAs assessment").

P-values for stratified analysis represent p-values for interaction.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Table S8 Associations of longitudinal changes in high attenuation areas with mortality stratified analysis

Model	No. 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 for interaction
***MUC5B* (rs35705950)**						0.85
GG	3,928	74,415	722	111.8 (103.8 to 120.1)	1.08 (1.02 to 1.14)	
GT/TT	589	64,602	105	107.0 (88.0 to 129.0)	3.94 (2.00 to 7.79)	
Telomere Length						
<5th percentile	223	3,618	49	135.5 (101.3 to 177.6)	4.54 (0.61 to 33.62)	0.40
≥5th percentile	4,230	69,740	771	110.6 (103.0 to 118.6)	1.09 (1.03 to 1.16)	
<10th percentile	444	7,213	97	134.5 (109.6 to 163.3)	14.25 (8.38 to 24.23)	0.70
≥10th percentile	4,009	66,144	723	109.3 (101.6 to 117.5)	1.09 (1.02 to 1.15)	
<25th percentile	1,114	18,139	228	125.7 (110.2 to 142.8)	9.94 (5.61 to 17.62)	0.98
≥25th percentile	3,339	55,218	592	107.2 (98.8 to 116.1)	1.12 (1.05 to 1.20)	
Sex						0.78
Female	2,322	38,795	354	91.3 (82.1 to 101.1)	1.08 (1.01 to 1.16)	
Male	2,195	35,620	473	132.8 (121.2 to 145.2)	1.08 (1.02 to 1.15)	
Smoking history						0.84
Never smoker	2,056	34,118	317	92.9 (83.1 to 103.6)	15.40 (10.71 to 22.14)	
Ever smoker	2,461	40,297	510	126.6 (115.9 to 137.9)	1.08 (0.99 to 1.16)	

Abbreviations: CI=confidence intervals; HAAs=high attenuation areas

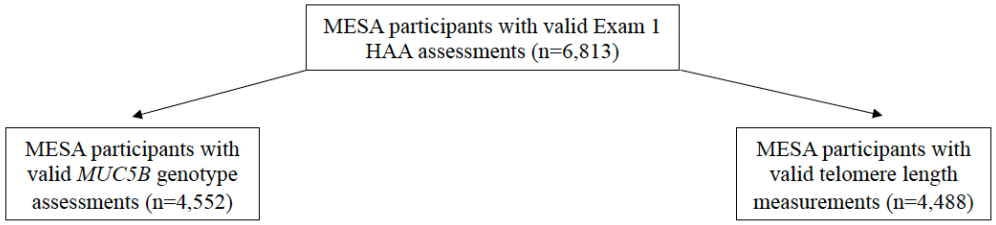
Models adjusted for sex, self-reported race/ethnicity, baseline age, smoking status, cigarette pack-years, height, weight, systolic and diastolic blood pressures, total cholesterol, high-density lipoprotein cholesterol, diabetes history, cancer history, coronary artery calcium score, percent emphysema, and total intentional exercise (met-min/week).

Figure Legend

Figure S1. Flow chart of MESA participants with valid Exam 1 *MUC5B* (rs35705950) assessments, telomere length, and high attenuation area assessments at Exam 1.

Kim JS, *et al. Thorax* 2023; 78:566–573. doi: 10.1136/thorax-2021-218139

Figure S1


MESA participants with valid Exam 1 HAA assessments (n=6,813)
MESA participants with valid *MUC5B* genotype assessments (n=4,552)
MESA participants with valid telomere length measurements (n=4,488)

References

1. Martinez FJ, Yow E, Flaherty KR, et al. Effect of Antimicrobial Therapy on Respiratory Hospitalization or Death in Adults With Idiopathic Pulmonary Fibrosis: The CleanUP-IPF Randomized Clinical Trial. *JAMA* 2021;325(18):1841-51. doi: 10.1001/jama.2021.4956 [published Online First: 2021/05/12]
2. Flaherty KR, Wells AU, Cottin V, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. *N Engl J Med* 2019;381(18):1718-27. doi: 10.1056/NEJMoa1908681 [published Online First: 2019/10/01]
3. Menon AA, Putman RK, Sanders JL, et al. Interstitial Lung Abnormalities, Emphysema and Spirometry in Smokers. *Chest* 2021 doi: 10.1016/j.chest.2021.10.034 [published Online First: 2021/11/08]
4. 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(2) doi: 10.1183/13993003.03206-2020 [published Online First: 2020/11/28]