Online Data Supplement to "The *MUC5B* Promoter Polymorphism and Radiologic Patterns of Interstitial Lung Abnormalities"

Rachel K. Putman MD MPH[1*], Gunnar Gudmundsson MD PhD[2*], Tetsuro Araki MD PhD[3,4], Mizuki Nishino MD MPH[3,4], Sigurdur Sigurdsson BSc MSc[5], Elías F. Gudmundsson MSc[5], Gudny Eiríksdottír MSc[5], Thor Aspelund MSc PhD[5,6], James C. Ross PhD[7,8], Raúl San José Estépar PhD[3,8], Ezra R. Miller MD[1]; Yoshitake Yamada MD PhD[3,4], Masahiro Yamagawa MD PhD[9], Noriyuki Tomiyama MD PhD[9], Lenore J. Launer PhD[10], Tamara B. Harris MD MS[10], Souheil El-Chemaly MD MPH[1], for the COPDGene Investigators, Benjamin A. Raby MD MPH[1,7], Michael H. Cho MD MPH[1,7], Ivan O. Rosas MD[1], George R. Washko MD[1,4], David A. Schwartz MD[10], Edwin K. Silverman MD PhD[1,7], Vilmundur Gudnason MD PhD[5,6], Hiroto Hatabu MD PhD[3,4] and Gary M. Hunninghake MD MPH[1,4]

*These authors contributed equally to this work.

MATERIALS AND METHODS

Chest CT Analysis

All chest CTs were reviewed on AZE VirtualPlace Fujin Raijin workstations (AZE, Tokyo, Japan) using axial images at window level of -700 Hounsfield Units (HU) and a window width of 1500 HU. All chest CT scans were evaluated by up to three readers, (chest radiologists and pulmonologists), all of whom were blind to all participant specific information, using the sequential reading method as previously described (1). The first reader scored the CT scans as follows, no interstitial lung abnormalities (ILA), indeterminate ILA status or ILA. The second reader (blinded to prior ILA scoring) would review and score all scans that were scored as ILA and indeterminate ILA status, along with 20% of the scans that did not have ILA. Finally, the third reader provided a majority opinion on all CT scans with discordant scores. As in previous work, ILA were defined as non-dependent changes affecting >5% of any lung zone including, nondependent ground-glass or reticular abnormalities, diffuse centrilobular nodularity, non-emphysematous cysts, honeycombing, or traction bronchiectasis (1-5).

Next, to provide additional detail on the radiologic patterns of ILA, all the CT scans with ILA reviewed for subtyping using three classification systems. First, all CT scans were reviewed for the presence architectural distortion, including, but not limited to traction bronchiectasis and honeycombing, which defined the category of Definite Fibrosis (3). Next, all scans with ILA were evaluated for the presence of usual interstitial pneumonia (UIP), as defined by the ATS/ERS/JRS/ALAT criteria (6). “UIP pattern contains all four features: subpleural, basal predominance, reticular abnormality, honeycombing with or without traction bronchiectasis, and absence of features that are inconsistent with a UIP pattern. Possible UIP includes the following three features: subpleural, basal predominance, reticular abnormality, and absence of features that are inconsistent with UIP. Inconsistent with a UIP pattern includes any of the

following seven features: upper or mid-lung predominance, peribronchovascular predominance, extensive ground glass abnormality (extent > reticular abnormality), profuse micronodules (bilateral, predominantly upper lobes), discrete cysts (multiple, bilateral, away from areas of honeycombing), diffuse mosaic attenuation/air-trapping (bilateral, in three or more lobes), and consolidation in bronchopulmonary segment(s)/lobe(s)" (6). During this phase all CT scans with ILA were scored as follows: inconsistent with UIP, possible UIP or definite UIP. Finally, all CT scans with ILA were evaluated for the following patterns of ILA: centrilobular (predominantly centrilobular and/or peribronchial ground glass opacities sparing the peripheral lung parenchyma), subpleural (reticular/nodular and/or ground glass opacities in a predominantly subpleural distribution), mixed (mixed centrilobular and subpleural abnormalities), and radiologic ILD (radiologic evidence of interstitial lung disease characterized by extensive changes consistent with firm radiologic evidence of ILD based on ATS guidelines (6), as previously described (2, 5). All subtyping was done by majority opinion of at least three readers.

Statistical Analyses

Prediction Models

Models were built using baseline clinical information, that would be available to a primary care provider prior to pulmonary consultation and CT imaging results and covariate selection was based on prior work (3-5). The AGES-Reykjavik cohort was used for the derivation of the prediction model and COPDGene was used for replication of the selected model. For all prediction models, all participants other than those with ILA (or the ILA subtypes) that were the outcome of the prediction model were included in the model as no ILA. The initial model included age, sex, body-mass-index (BMI), pack-years of smoking and current smoking status. Using Hosmer and Lemeshow Goodness-of-Fit test, the P-value was 0.94. Next, to try and the

most parsimonious model and based on prior significant associations with ILA (4), first BMI was removed from the model, with a slight decrease in P-value to 0.87. Next, current smoking status was removed from the model with an improvement in the P-value from the Goodness-of-Fit test to 0.97. Next, pack-years of smoking was removed from the model with a subsequent decrease in p-value to 0.77, then pack-years of smoking was added back with sex removed and the p-value decreased even further to 0.35, the final model tested included just age with a p-value of 0.76. Based on these results, the model including age, sex and pack-years of tobacco smoking was used as the final model for prediction. In addition to c-statistics, Net Reclassification Indices were calculated using the same prediction model, using risk categories in increments of 10% from 0% to 100%) (7).

RESULTS

Table S1. Baseline characteristics of participants in COPDGene stratified by race and Interstitial Lung Abnormality (ILA) status.*

	African-Americans			Non-Hispanic Whites		
	No ILA (N=1728)	Indeterminate (N=1207)	ILA (N=223)	No ILA (N=3667)	Indeterminate (N=1982)	ILA (N=485)
Age – yrs	54 ± 7	55 ± 7	55 ± 8	61 ± 9	64 ± 9	64 ± 10
Gender – no. female(%)	705 (41)	559 (46)	126 (57)	1752 (48)	918 (46)	219 (45)
Body Mass Index	29 ± 7	29 ± 7	29 ± 7	29 ± 6	29 ± 6	29 ± 6
Pack Years Smoking, median (IQR)†	34 (22, 46)	36 (24, 48)	35 (24, 47)	40 (29, 56)	44 (32, 62)	45 (34, 63)
Current Smokers – no. yes (%)	1393 (81)	984 (82)	183 (82)	1402 (38)	807 (41)	256 (53)
History of COPD‡ – no. (%)	381 (22)	325 (27)	54 (25)	1474 (40)	946 (48)	167 (35)
Percentage of the lung less than -950 Hounsfield Units, median (IQR)	1.1 (0.4, 3.3)	0.7 (0.2, 2.4)	0.7 (0.2, 2.4)	3 (0.9, 9)	2.8 (0.8, 9.6)	1.4 (0.5, 4.8)

***ILA** stands for interstitial lung abnormality. ± values are means and standard deviations.
†IQR is interquartile range
‡COPD is chronic obstructive pulmonary disease

Table S2. Results of ILA Radiologic Subtyping into major categories by cohort*

	Centrilobular	Subpleural	Mixed	Radiologic ILD†
AGES-Reykjavik				
Inconsistent with UIP‡	39 (1)§	21 (8)	84 (20)	0
Possible UIP	0	216 (82)	0	0
UIP	0	0	0	17 (17)
Non-Hispanic Whites COPDGene				
Inconsistent with UIP	132	35 (8)	91 (4)	3 (2)
Possible UIP	0	197 (65)	2 (1)	11 (9)
UIP	0	0	0	12 (12)
African-Americans COPDGene				
Inconsistent with UIP	64	13 (1)	74 (6)	2 (2)
Possible UIP	0	62 (12)	4 (1)	0
UIP	0	0	0	3 (3)

*ILA is interstitial lung abnormality
†ILD is interstitial lung disease
‡UIP is usual interstitial pneumonia
§Numbers in ( ) are the number of cases with definite fibrosis (evidence of pulmonary parenchymal architectural distortion) in that category

Table S3. Association between Interstitial Lung Abnormalities (ILA) and the *MUC5B* promoter polymorphism when limited to participants over the age of 70*

	AGES-Reykjavik (N=3192)		**COPDGene Non-Hispanic Whites (N=784)**		**COPDGene African-Americans (N=69)**	
ILA Subtype	Adjusted† Odds Ratio (95% CI)	P-Value	Adjusted† Odds Ratio (95% CI)	P-Value	Adjusted† Odds Ratio (95% CI)	P-Value
ILA	2.7 (2.2, 3.3)	$2x10^{-21}$	2.1 (1.4, 3.1)	0.0001	--	--
ILA without Fibrosis‡	2.4 (1.9, 3.1)	$3x10^{-12}$	2.1 (1.3, 3.3)	0.001	--	--
Definite Fibrosis	3.2 (2.4,4.4)	$9x10^{-14}$	2.0 (1.2, 3.6)	0.01	--	--
Centrilobular	1.3 (0.6, 2.7)	0.46	2.1 (0.4, 9.9)	0.34	--	--
Subpleural	2.9 (2.3, 3.7)	$2x10^{-17}$	2.1 (1.4, 3.3)	0.0005	--	--
Mixed	2.4 (1.6, 3.5)	$1x10^{-5}$	0.85 (0.25, 2.9)	0.80	--	--
Radiologic ILD§	4.3 (2.1, 9.1)	$1x10^{-4}$	4.7 (1.8, 12.3)	0.001	--	--
Inconsistent with UIP‖	2.3 (1.7, 3.2)	$8x10^{-8}$	1.6 (0.8, 3.1)	0.16	--	--
Possible UIP	2.7 (2.1, 3.5)	$7x10^{-15}$	2.2 (1.4, 3.4)	0.0007	--	--
UIP	4.3 (2.1, 9.1)	$1x10^{-4}$	7.6 (1.8, 32.8)	0.006	--	--

*ILA is interstitial lung abnormality. Analyses of *MUC5B* genotype were performed using additive genetic models, odds ratios are per copy o*f MUC5B* minor allele.

†Models are adjusted for age, sex and tobacco exposure

‡ Fibrosis is evidence of pulmonary parenchymal architectural distortion

§ ILD is interstitial lung disease. Analysis was not done in African-Americans; no participants with radiologic ILD had the *MUC5B* genotype

‖UIP is usual interstitial pneumonia. Analysis was not done in African-Americans; no participants with UIP had the *MUC5B* genotype

Table S4. Association between Interstitial Lung Abnormalities (ILA) and the *MUC5B* promoter polymorphism when limited to participants under the age of 70*

	AGES-Reykjavik (N=394)		**COPDGene Non-Hispanic Whites (N=3409)**		**COPDGene African-Americans (N=1884)**	
ILA Subtype	Adjusted† Odds Ratio (95% CI)	P-Value	Adjusted† Odds Ratio (95% CI)	P-Value	Adjusted† Odds Ratio (95% CI)	P-Value
ILA	--	--	1.5 (1.2, 1.9)	0.001	1.3 (0.6, 2.5)	0.49
ILA without Fibrosis‡	--	--	1.4 (1.04, 1.8)	0.02	1.0 (0.45, 2.1)	0.96
Definite Fibrosis	--	--	2.2 (1.3, 3.7)	0.003	3.9 (1.2, 12.8)	0.02
Centrilobular	--	--	0.6 (0.3, 1.0)	0.06	1.2 (0.4, 3.8)	0.79
Subpleural	--	--	2.6 (1.9, 3.6)	1×10^{-8}	1.6 (0.6, 4.4)	0.36
Mixed	--	--	1.1 (0.6, 1.9)	0.74	1.05 (0.3, 3.3)	0.94
Radiologic ILD§	--	--	4.8 (2.1, 11.2)	0.0003	--	--
Inconsistent with UIP‖	--	--	0.8 (0.6, 1.2)	0.33	1.0 (0.4, 2.4)	0.95
Possible UIP	--	--	2.8 (2.0, 3.9)	2×10^{-8}	1.8 (0.7, 5.1)	0.24
UIP	--	--	3.7 (1.3, 10.6)	0.02	--	--

*ILA is interstitial lung abnormality. Analyses of *MUC5B* genotype were performed using additive genetic models, odds ratios are per copy o*f MUC5B* minor allele. Many of the analyses include less than 10 (at times less than 5) cases with at least one copy of the *MUC5B* minor allele, limiting the ability to draw conclusions.

†Models are adjusted for age, sex and tobacco exposure

‡ Fibrosis is evidence of pulmonary parenchymal architectural distortion

§ ILD is interstitial lung disease. Analysis was not done in African-Americans; no participants with radiologic ILD had the *MUC5B* genotype

ǁUIP is usual interstitial pneumonia. Analysis was not done in African-Americans; no participants with UIP had the *MUC5B* genotype

Table S5. Net Reclassification Results by ILA Subtype and Cohort*

	Clinical Data + *MUC5B* Minor Allele		
	Net Reclassification Index	Standard Error	**P-Value**
AGES-Reykjavik			
ILA	12%	3%	0.0001
Subpleural & Radiologic ILD†	15%	3%	<.0001
Definite Fibrosis	9%	3%	0.008
Possible & Definite UIP‡	13%	3%	<.0001
Non-Hispanic Whites - COPDGene			
ILA	6%	2%	0.01
Subpleural & Radiologic ILD	13%	3%	<.0001
Definite Fibrosis	4%	2%	0.05
Possible & Definite UIP	10%	3%	0.0004
African-Americans – COPDGene			
ILA	0.2%	1%	0.86
Subpleural & Radiologic ILD	2%	2%	0.20
Definite Fibrosis	7%	5%	0.17
Possible & Definite UIP	3%	2%	0.18

*ILA is interstitial lung abnormality
†ILD is interstitial lung disease
‡UIP is usual interstitial pneumonia

REFERENCES

1. Washko GR, Lynch DA, Matsuoka S, Ross JC, Umeoka S, Diaz A, Sciurba FC, Hunninghake GM, San Jose Estepar R, Silverman EK, Rosas IO, Hatabu H. Identification of early interstitial lung disease in smokers from the COPDGene Study. *Acad Radiol* 2010; 17: 48-53.
2. Washko GR, Hunninghake GM, Fernandez IE, Nishino M, Okajima Y, Yamashiro T, Ross JC, Estepar RS, Lynch DA, Brehm JM, Andriole KP, Diaz AA, Khorasani R, D'Aco K, Sciurba FC, Silverman EK, Hatabu H, Rosas IO, Investigators CO. Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011; 364: 897-906.
3. Hunninghake GM, Hatabu H, Okajima Y, Gao W, Dupuis J, Latourelle JC, Nishino M, Araki T, Zazueta OE, Kurugol S, Ross JC, San Jose Estepar R, Murphy E, Steele MP, Loyd JE, Schwarz MI, Fingerlin TE, Rosas IO, Washko GR, O'Connor GT, Schwartz DA. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013; 368: 2192-2200.
4. Putman RK, Hatabu H, Araki T, Gudmundsson G, Gao W, Nishino M, Okajima Y, Dupuis J, Latourelle JC, Cho MH, El-Chemaly S, Coxson HO, Celli BR, Fernandez IE, Zazueta OE, Ross JC, Harmouche R, Estepar RS, Diaz AA, Sigurdsson S, Gudmundsson EF, Eiriksdottir G, Aspelund T, Budoff MJ, Kinney GL, Hokanson JE, Williams MC, Murchison JT, MacNee W, Hoffmann U, O'Donnell CJ, Launer LJ, Harrris TB, Gudnason V, Silverman EK, O'Connor GT, Washko GR, Rosas IO, Hunninghake GM, Evaluation of CLtIPSEI, Investigators CO. Association Between Interstitial Lung Abnormalities and All-Cause Mortality. *JAMA* 2016; 315: 672-681.
5. Araki T, Putman RK, Hatabu H, Gao W, Dupuis J, Latourelle JC, Nishino M, Zazueta OE, Kurugol S, Ross JC, San Jose Estepar R, Schwartz DA, Rosas IO, Washko GR, O'Connor GT, Hunninghake GM. Development and Progression of Interstitial Lung Abnormalities in the Framingham Heart Study. *Am J Respir Crit Care Med* 2016.
6. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Carvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE, Jr., Kondoh Y, Myers J, Muller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schunemann HJ, Fibrosis AEJACoIP. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788-824.
7. Nakahira K, Kyung SY, Rogers AJ, Gazourian L, Youn S, Massaro AF, Quintana C, Osorio JC, Wang Z, Zhao Y, Lawler LA, Christie JD, Meyer NJ, Mc Causland FR, Waikar SS, Waxman AB, Chung RT, Bueno R, Rosas IO, Fredenburgh LE, Baron RM, Christiani DC, Hunninghake GM, Choi AM. Circulating mitochondrial DNA in patients in the ICU as a marker of mortality: derivation and validation. *PLoS Med* 2013; 10: e1001577; discussion e1001577.