Air Pollution and Subclinical Interstitial Lung Disease: Supplementary Appendix

Table of Contents

PARTICIPANT SELECTION

Figure S1: Number of participants included in the analyses of ILA prevalence. The number in the purple shaded box (last row) represents the number of measurements included in the analysis of outdoor $PM_{2.5}$. ILA- interstitial lung abnormalities

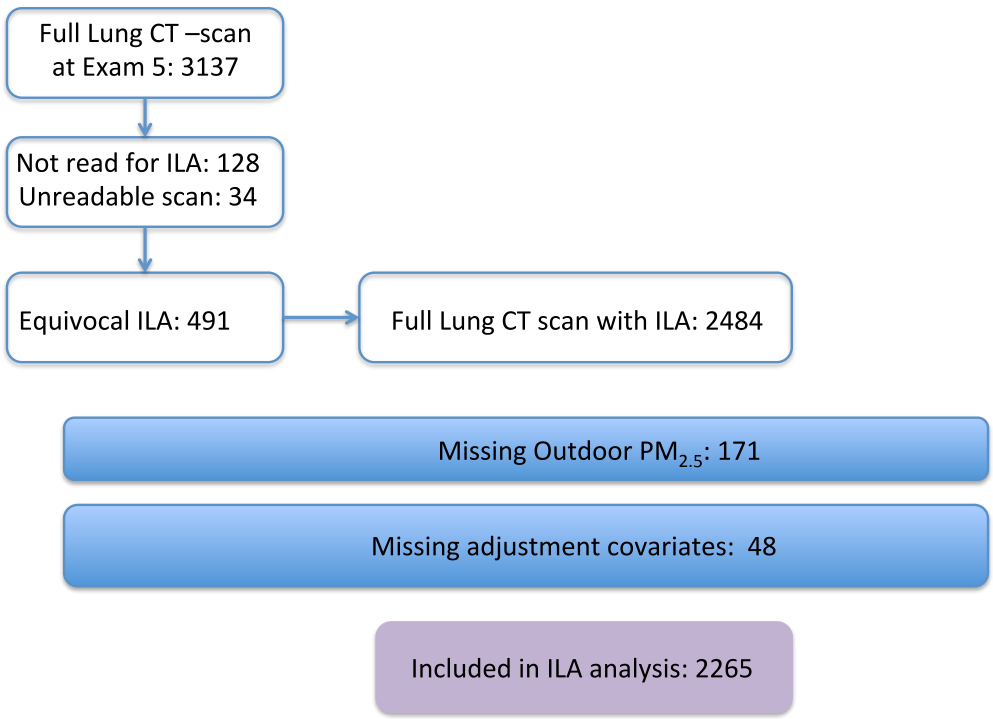
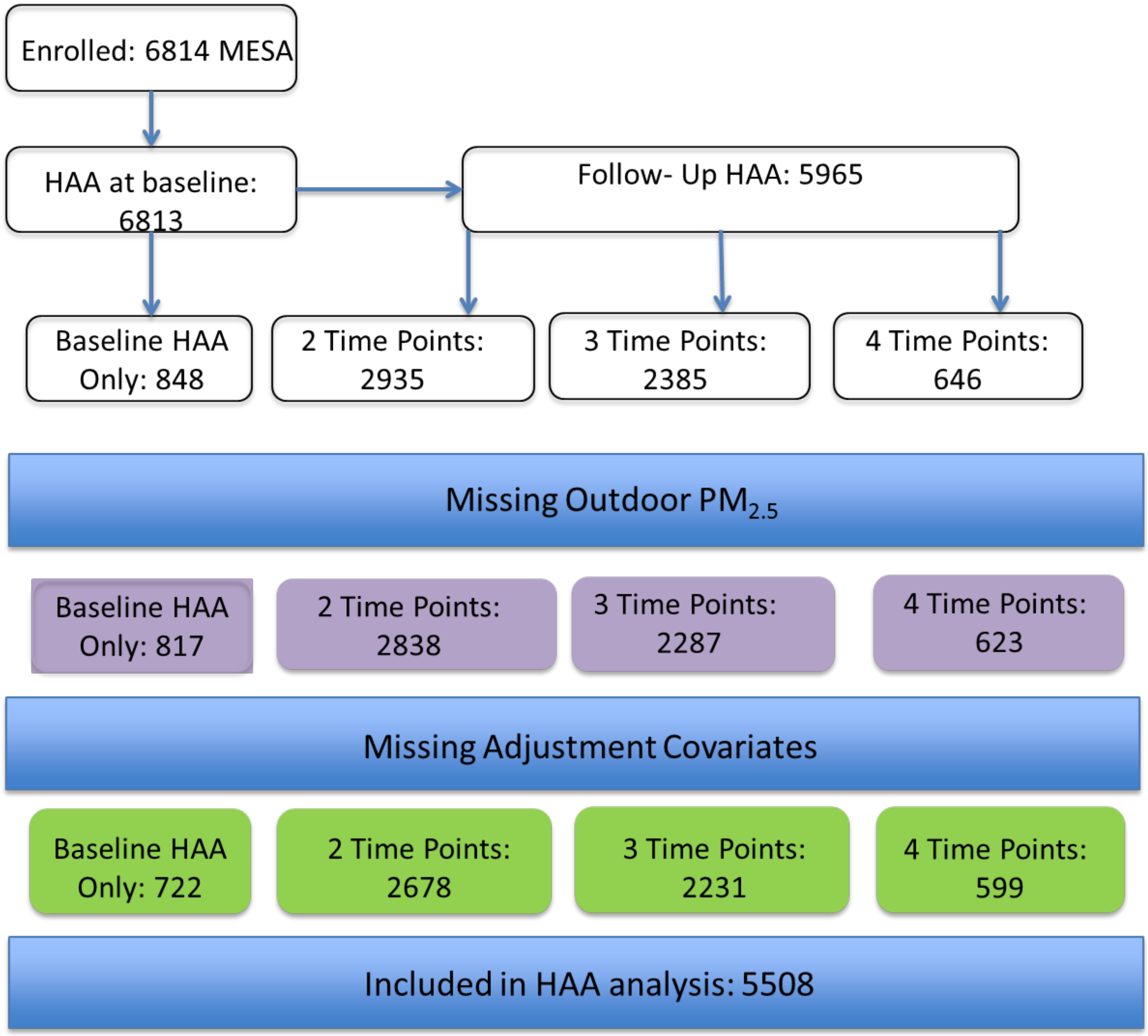


Figure S2: Number of measurements included in the analysis of HAA progression. The numbers in the green boxes (last row) represent the number of measurements included in the analysis of outdoor $PM_{2.5}$. HAA- high attenuation areas



ANALYSIS APPROACH:

ILA Logistic Regression Model:

We used logistic regression to estimate adjusted odds ratios for ILA for the different pollutants. We chose to estimate odds ratios instead of relative risk as part of the cohort (those with equivocal scans) were excluded from analysis. Models were adjusted for age, gender, ethnicity, tobacco use (current smoking status and pack years) and site.

HAA Linear Mixed Model:

We modeled HAA as continuous dependent variable, expressed as the log transformed percentage of imaged lung volume with an attenuated value between -600 and -250 HU. To interpret effect sizes from the linear mixed model, we back-transformed β-coefficients and 95% CIs to get the percent change in HAA.

The linear mixed model we used in this analysis is similar to other health effects models developed by the MESA Air group and described in detail elsewhere.[1,2] The model is designed to jointly model the cross-sectional and longitudinal relationships between air pollution and HAA. It is comprised of three parts:

$$Y_{iv} = [x_{i0}\alpha_1 + \alpha_i] + [t_{iv}\beta_0 + W_{iv}t_{iv}\beta_1 + t_{iv}b_i] + [U_{iv}\gamma_1 + \epsilon_{iv}]$$

where:

Y_{iv} = Outcome measurement for subject i at v^{th} follow-up exam (log-transformed HAA in our analyses)

x_{i0} = time-invariant cross-sectional confounders and risk factors at Exam 1 for subject i, including mean air pollution exposure during the year 2000. Also includes site indicator.

W_{iv} = possibly time-varying longitudinal confounders and risk factors at exam v for subject i, including mean air pollution exposure during the time period between baseline (v = 0) and v^{th} follow-up exam, rounded to the nearest whole year

U_{iv} = time-varying variables to adjust measurements at exam v for subject i, primarily CT scanner

t_{iv} = time in years from baseline (v = 0) to the v^{th} follow-up exam for subject i

β_0 = Outcome progression (annual rate of change) in average participants in the reference group

β_1 = coefficients for interaction between risk factors and time; this includes the air pollution by time interaction which is interpreted as a rate (association between air pollution and annual progression)

α_1 = coefficients for cross-sectional associations between baseline outcome measurements and risk factors (including year 2000 air pollution exposure)

γ_1 = coefficients for cross-sectional associations between time-varying variables and HAA measurements at all exams

a_i = subject-specific random intercept, which is nested within a neighborhood specific intercept

b_i = subject-specific random slope

ε_{iv} = error associated with Y_{iv}

1) Cross sectional term: The cross-sectional term models and reduces any confounding effect from the relationship between estimated baseline measurements and HAA. As done in prior work with the MESA cohort, we chose year 2000 as a proxy for chronic exposure to air pollution prior to study enrollment.

2) Rate of change term: The longitudinal relationship used to model the rate of change in HAA. For pollution, the coefficient is interpreted as the rate of change in log-transformed HAA per year in a subject holding all other covariates constant. Long-term average pollution estimates were calculated by averaging time- and location-specific two-week predictions for each individual between baseline exam and follow up exam.

3) time varying covariates: Adjust for variables relevant to specific measurements

Adjustment Variable Details:

The questionnaires administered to MESA participants have been well described. Tobacco use was categorized as non-smoker versus current-smoker and pack years were included as a continuous variable. Of note, in stratified analyses assessing for effect modification, tobacco use was dichotomized as ever smoker versus never smoker to improve power. Educational attainment was collapsed into three categories: high school or less; some college or technical school; bachelor's degree or more. Occupational exposures to vapors, gas, dust and fumes were quantified according to a job exposure matrix (JEM) previously created by the National Institute for Occupational Safety and Health (NIOSH).[3] Demographic questionnaires ascertained current occupation if employed or the occupation where last employed if retired. These reported occupations were coded using Bureau of Census 2002 occupational codes by trained staff from NIOSH. An industrial hygienist then assigned an exposure score based on occupational codes that reflected the likelihood and severity of VGDF exposure (low, moderate, high). Three NIOSH industrial hygienists then reached a final consensus on the exposure scores.

SUPPLEMENTARY TABLES:

This section contains tables and descriptive statistics that may be of interest.
Tables S1 and S2 present the correlation between the pollutant exposures in primary ILA and HAA analyses, respectively. Table S3 presents results of the cross sectional association between year 2000 pollution estimates and progression of HAA, derived from the main effect of mixed linear model.

TABLE S1: 10 year Estimates of Pollution Exposure Correlations used in ILA analysis

	$PM_{2.5}$	NO_x	NO_2	O_3
$PM_{2.5}$	1.0			
NO_x	0.69	1.0		
NO_2	0.69	0.96	1.0	
O_3	-0.46	-0.87	-0.87	1.0

TABLE S2: Estimates of Air Pollution Exposure Correlations for HAA Analysis

	Year 2000					**Long-Term***			
	$PM_{2.5}$	NO_x	NO_2	O_3		$PM_{2.5}$	NO_x	NO_2	O_3
$PM_{2.5}$	1.0				**$PM_{2.}$**	1.0			
NO_x	0.65	1.0			**NO_x**	0.70	1.0		
NO_2	0.63	0.93	1.0		**NO_2**	0.70	0.96	1.0	
O_3	-0.45	-0.86	-0.88	1.0	**O_3**	-0.46	-0.87	-0.87	1.0

*Based on long-term average over the time rounded to the nearest whole year between baseline and the last HAA measurements for participants

TABLE S3: Estimates of Year 2000 Pollution Exposure and Progression of HAA

	Percent change in HAA per 5 mcg/m^3 increment in $PM_{2.5}$ (95% CI)	p	Percent change in HAA per 40 ppb increment in NO_x (95%CI)	p	Percent change in HAA per 10 ppb increment in NO_2 (95% CI)	p	Percent change in HAA per 10 ppb increment in O_3 (95% CI)	p
Overall	0.44 (-2.37 to 3.35)	0.76	-0.96 (-2.86 to 0.98)	0.33	-0.078 (-1.88 to 1.71)	0.93	0.61 (-2.86 to 4.19)	0.73
Sex								
Female	2.22 (-1.78 to 6.40)	0.29	-0.78 (-3.44 to 2.02)	0.57	0.87 (-1.69 to 3.56)	0.51	-1.78 (-6.67 to 1.03)	0.50
Male	-1.09 (-4.69 to 2.74)	0.58	-0.88 (-3.34 to 1.61)	0.49	-0.52 (-2.86 to 1.92)	0.67	0.64 (-3.92 to 5.34)	0.78
Race								
White	-4.40 (-7.87 to -0.77)	0.02	-2.96 (-6.01 to 0.17)	0.06	-2.47 (-4.97 to 0.22)	0.07	-1.39 (-5.54 to 3.04)	0.54
Asian	-1.39 (-9.43 to 17.4)	0.66	-1.09 (-5.82 to 3.87)	0.66	-0.06 (-4.30 to 4.39)	0.98	-7.60 (-18 to 4.2)	0.20
Black	1.82 (-4.30 to 8.32)	0.56	-1.59 (-3.82 to 0.59)	0.15	1.71 (-1.98 to 5.55)	0.37	2.0 (-4.69 to 9.09)	0.56
Hispanic	-1.49 (-8.15 to 5.65)	0.68	-4.30 (-8.15 to -0.32)	0.03	-4.40 (-8.70 to 0.14)	0.06	10.5 (0.08 to 22.1)	0.05
Tobacco Use								
Never-smoker	-1.69 (-1.88 to 5.65)	0.34	-2.27 (-5.07 to 0.65)	0.13	-1.19 (-3.82 to 1.61)	0.41	2.43 (-3.25 to 8.32)	0.41
Ever-smoker	1.82 (-5.82 to 2.63)	0.43	-0.40 (-1.78 to 2.12)	0.75	0.57 (-1.78 to 3.05)	0.64	-1.09 (-5.45 to 3.36)	0.62

Table S3: Estimates of Year 2000 Pollution Exposure and Progression of HAA. Cross-sectional association of fine particulate matter ($PM_{2.5}$), nitrogen oxides (NO_x), nitrogen dioxide (NO_2) and ozone (O_3) with percent change in high attenuation areas (HAA), from linear mixed models adjusted for age, gender, race/ ethnicity, educational attainment, JEM, height, BMI, waist circumference, smoking status, cigarette pack-years, glomerular filtration rate, total volume of image lung, percent emphysema on CT scan, CT scanner type and study site

Additional Sensitivity Analyses:

Table S4 depicts estimated OR from a multiplicative model of the interaction between smoking status and pollutant exposure. Table S5 shows the results of the ILA analysis, including participants with scans initially read as "equivocal" for ILA. Scans were labelled "equivocal" if they included a solitary focus of ground-glass attenuation, reticulation or multifocal ground-glass abnormality in less than 5% of the lung. In the main analysis, these scans were excluded from further analysis, as previously described. In this analysis, indeterminate scans were coded as "no ILA." Table S6 shows the results from ILA analysis adjusting for additional confounder of occupational exposures (JEM). Results from analyses of different pollutant exposure periods and the odds of ILA are presented in Table S7. 20 and 30-year pollutant estimates were only available for $PM_{2.5}$ and NO_2 (estimates for NO_2 not pictured).

TABLE S4: 10 Year Estimates of Pollution Exposure and Risk of Interstitial Lung Abnormalities with Effect Modification by Smoking

Pollutants	Main Pollutant Effect OR (95%CI)	Pollutant* smoking OR (95%CI)	p for interaction
$PM_{2.5}$	1.61 (0.76 to 3.39)	0.59 (0.29 to 1.16)	0.12
NO_x	1.89 (0.99 to 3.59)	0.71 (0.42 to 1.2)	0.21
NO_2	1.35 (0.85 to 2.12)	0.79 (0.57 to 1.09)	0.29
O_3	0.54 (0.22 to 1.30)	1.45 (0.75 to 2.80)	0.27

Table S4: Sensitivity analysis for effect modification by smoking status. Cross sectional association of fine particulate matter (per 5mcg/m^3 increments of $PM_{2.5}$), nitrogen oxides (per 40ppb increments in NO_x), nitrogen dioxide (per 10ppb increments in NO_2) and ozone (per 10ppm increments in O_3) with the odds of interstitial lung abnormalities (ILA). Estimates are from multivariable logistic regression models with an interaction term for pollutant exposure and smoking status, adjusted for age, gender, ethnicity, site, pack years and occupational exposures. The p-value for the interaction was derived from the likelihood ratio test comparing nested models with the main effect and multiplicative interaction to test for statistical significance.

TABLE S5: 10 Year Estimates of Pollution Exposure and Risk of Interstitial Lung Abnormalities, including Indeterminate Scans

	OR per 5 mcg/m³ increment in $PM_{2.5}$ (95% CI)	p	OR per 40 ppb increment in NO_x (95%CI)	p	OR per 10 ppb increment in NO_2 (95% CI)	p	OR per 10 ppb increment in O_3 (95% CI)	P
Overall	1.43 (0.81 to 2.54)	0.22	1.77 (1.06 to 2.95)	0.03	1.28 (0.90 to 1.82)	0.16	0.63 (0.31 to 1.28)	0.20
Sex								
Female	1.15 (0.54 to 2.44)	0.72	1.55 (0.79 to 3.05)	0.20	1.27 (0.80 to 2.01)	0.30	0.62 (0.24 to 1.57)	0.32
Male	1.96 (0.82 to 4.66)	0.13	2.08 (0.94 to 4.58)	0.07	1.31 (0.75 to 2.28)	0.33	0.60 (0.20 to 1.83)	0.37
Race								
White	1.37 (0.61 to 3.08)	0.45	1.84 (0.77 to 4.41)	0.17	1.29 (0.75 to 2.20)	0.36	0.51 (0.19 to 1.41)	0.20
Asian	6.79 (0.85 to 54.0)	0.07	2.93 (0.69 to 12.5)	0.15	1.82 (0.71 to 4.64)	0.21	1.72 (0.18 to 16.2)	0.64
Black	0.84 (0.22 to 3.16)	0.79	1.03 (0.38 to 2.78)	0.96	0.69 (0.35 to 1.38)	0.30	0.97 (0.25 to 3.82)	0.97
Hispanic	1.24 (0.32 to 4.84)	0.76	2.33 (0.70 to 7.73)	0.17	1.89 (0.75 to 4.80)	0.18	0.23 (0.03 to 1.93)	0.18
Tobacco Use								
Never-smoker	1.74 (0.73 to 4.16)	0.22	2.89 (1.35 to 6.16)	0.006	2.08 (1.17 to 3.69)	0.01	0.30 (0.10 to 0.93)	0.04
Ever-smoker	1.04 (0.40 to 2.76)	0.93	1.05 (0.45 to 2.43)	0.91	0.81 (0.46 to 1.43)	0.46	1.44 (0.52 to 4.01)	0.49

Table S5: 10 Year Estimates of Pollution Exposure and Risk of Interstitial Lung Abnormalities, including equivocal scans which have been recoded as "no ILA." sectional association of fine particulate matter (PM2.5), nitrogen oxides (NOx), nitrogen dioxide (NO2) and ozone (O3) with the odds of interstitial lung abnormalities (ILA), from multivariable logistic regression models adjusted for age, gender, ethnicity, tobacco use (current smoking status and pack years) and site.

TABLE S6: 10 Year Estimates of Pollution Exposure and Risk of Interstitial Lung Abnormalities, Adjusting for Occupational Exposures

	OR per 5 mcg/m^3 increment in $PM_{2.5}$ (95% CI)	p	OR per 40 ppb increment in NO_x (95%CI)	p	OR per 10 ppb increment in NO_2 (95% CI)	p	OR per 10 ppb increment in O_3 (95% CI)	p
Overall	1.30 (0.66 to 2.57)		1.64 (0.90 to 2.97)		1.20 (0.78 to 1.85)		0.65 (0.29 to 1.45)	
Sex		0.14		0.29		0.24		0.42
Female	1.22 (0.48 to 3.08)		1.67 (0.75 to 3.76)		1.32 (0.75 to 2.34)		0.54 (0.18 to 1.63)	
Male	1.52 (0.55 to 4.19)		1.52 (0.62 to 3.7)		1.04 (0.54 to 2.02)		0.77 (0.23 to 2.58)	
Race		0.13		0.20		0.16		0.72
White	1.85 (0.63 to 5.43)		3.13 (1.00 to 9.86)		1.67 (0.79 to 3.54)		0.43 (0.13 to 1.47)	
Asian	9.84 (0.99 to 97.4)		4.35 (0.85 to 22.4)		2.08 (0.74 to 5.82)		1.01 (0.08 to 12.4)	
Black	0.48 (0.10 to 2.24)		0.68 (0.21 to 2.17)		0.45 (0.19 to 1.09)		1.16 (0.27 to 4.97)	
Hispanic	0.67 (0.14 to 3.22)		1.19 (0.31 to 4.54)		1.17 (0.42 to 3.26)		0.17 (0.02 to 2.06)	
Tobacco Use		0.12		0.21		0.29		0.27
Never-smoker	1.58 (0.62 to 4.0)		2.44 (1.09 to 5.46)		1.93 (1.05 to 3.55)		0.33 (0.10 to 1.09)	
Ever-smoker	0.97 (0.35 to 2.66)		0.96 (0.39 to 2.36)		0.72 (0.39 to 1.32)		1.19 (0.39 to 3.60)	

Table S6: 10 Year Estimates of Pollution Exposure and Risk of Interstitial Lung Abnormalities. Cross sectional association of fine particulate matter (PM2.5), nitrogen oxides (NOx), nitrogen dioxide (NO2) and ozone (O3) with the odds of interstitial lung abnormalities (ILA), from multivariable logistic regression models adjusted for age, gender, ethnicity, tobacco use (current smoking status and pack years), site and occupational exposures. P-values in the subgroup analyses were obtained from a likelihood ratio test comparing nested models with a main effect and multiplicative interaction with potential effect modifier. A total of 2150 participants were included in analysis.

TABLE S7: Estimates of $PM_{2.5}$ Exposures over Different Time Periods and Risk of ILA

Years Prior to Full Lung CT Scan	Exposure Model	OR per 5μg/m^3 increment in $PM_{2.5}$ (95%CI)	P
1	Spatio-temporal	0.75 (0.17 to 3.22)	0.69
10	Spatio-temporal	1.34 (0.75 to 2.38)	0.32
20	Historic	1.37 (0.57 to 2.71)	0.48
30	Historic	1.16 (0.53 to 2.56)	0.71

Table S7: Sensitivity analysis for different $PM_{2.5}$ exposure periods prior to full lung CT scan at Exam 5. Cross-sectional association of fine particulate matter ($PM_{2.5}$) with the odds of ILA. Exposure estimates were estimated from spatio-temporal models with land use regression or historic models based on annual average exposure as specified. Odds ratios are derived from multivariable logistic regression models adjusted for age, gender, ethnicity, site, smoking status and occupational exposures.

REFERENCES:

1. Kaufman JD, Adar SD, Barr RG, Budoff M, Burke GL, Curl CL, et al. Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the Multi-Ethnic Study of Atherosclerosis and Air Pollution): a longitudinal cohort study. Lancet. 2016.
2. Gassett AJ, Sheppard L, McClelland RL, Olives C, Kronmal R, Blaha MJ, et al. Risk Factors for Long-Term Coronary Artery Calcium Progression in the Multi-Ethnic Study of Atherosclerosis. Journal of the American Heart Association. 2015;4(8):e001726.
3. Doney B, Hnizdo E, Graziani M, Kullman G, Burchfiel C, Baron S, et al. Occupational risk factors for COPD phenotypes in the Multi-Ethnic Study of Atherosclerosis (MESA) Lung Study. Copd. 2014;11(4):368-80.